Italian National Agency for New Technologies,
Energy and Sustainable Economic Development

High Performance Computing on CRESCO infrastructure: research activities and results 2024

December 2025

Italian National Agency for New Technologies,
Energy and Sustainable Economic Development

High Performance Computing on CRESCO Infrastructure: research activity and results 2024

December 2025

High Performance Computing on CRESCO Infrastructure: research activity and results 2024

Contributions provided by a selection of users of the CRESCO infrastructure
Scientific Editor: Beatrice Calosso, ENEA, TERIN-ICT-HPC and coeditor Filippo Palombi, ENEA, TERIN-ICT-HPC
Cover: Amedeo Trolese, ENEA, TERIN-ICT, CR Frascati
ISBN: 978-88-8286-532-0

Index

FOREWORD

During the year 2024, the CRESCO high performance computing clusters have provided 122 million hours of "core" computing time, at a high availability rate, to about 200 users, supporting ENEA research and development activities in many relevant scientific and technological domains. In the framework of joint programs with ENEA researchers and technologists, computational services have been provided also to academic and industrial communities. In this use of the CRESCO clusters, Cresco7 and XCRESCO are not included, as they are still in a pre-production phase with usage monitoring not yet available. This report, the sixteenth of a series started in 2008, is a collection of 30 papers illustrating the main results obtained during 2024 using the CRESCO/ENEAGRID HPC facilities. The significant number of contributions proves the importance of the HPC facilities in ENEA for the research community. The topics cover various fields of research, such as materials science applied to energy research such as renewable sources and battery technologies, efficient combustion, climate research, nuclear technologies, molecular dynamic, artificial intelligence applied to cultural and software heritage. The report shows the wide spectrum of applications of high-performance computing, which has become an all-round enabling technology for science and engineering. Since 2008, the main ENEA computational resources are located near Naples, in Portici Research Centre. This is a result of the CRESCO Project (Computational Centre for Research on Complex Systems), co-funded, in the framework of the 2001-2006 PON (European Regional Development Funds Program), by the Italian Ministry of Education, University and Research (MIUR).The CRESCO Project provided the financial resources to set up the first HPC x86_64 Linux cluster in ENEA; a major computing installation for both the Italian and the International context: it ranked 126 in the HPC Top 500 June 2008 world list, with 17.1 Tflops and 2504 cpu cores. It was later decided to keep CRESCO as the name for all the Linux clusters in the ENEAGRID infrastructure, which integrates all ENEA scientific computing systems, and is currently distributed in six Italian sites. CRESCO computing resources were later upgraded in the framework of PON 2007-2013 with the project TEDAT and the cluster CRESCO4, 100 Tflops computing power. In 2020 the ENEAGRID computational resources consist of 25000 computing cores and a raw data storage of about 5 PB.

In 2015 ENEA and CINECA, the main HPC institution in Italy, signed a collaboration agreement to promote joint activities and projects in HPC. In this framework, CINECA and ENEA are providing HPC services for the nuclear fusion researcher community until 2023. In this framework, the new system MARCONI-FUSION started the operations in July 2016 with 1 Pflops computation power and they have been extended until 2029 with PITAGORA that has a power peak of 14 PFlops of conventional processors AMD Turin and 20 PFlops of accelerated GPU NVIDIA H100. The extension of the HPC services for EUROfusion has allowed to deploy new HPC systems in Portici: CRESCO7 and in Frascati: XCRESCO started in operation in 2023.

The ENEA-CINECA agreement was the key basis for ENEA HPC developments in the past. The CRESCO6 cluster has been installed in 2018 and its own 1.4 PFlops peak computing power, ranked 420th in November 2018 Top500 list. CRESCO6 has been a challenge in HPC co-design system thanks to implement a multi-fabric network able for working InfiniBand and Omni-Path on a single GPFS cluster using the same storage systems of CRESCO data centre. Within the framework of the NextGeneration EU funding plan the new ENEA Tier-0, CRESCO8 supercomputer has been deployed in 2024 to support the plasma physics modelling and new materials research relevant for nuclear fusion reactors.CRESCO8 supercomputer is essentially based on LENOVO technological solutions that

provides an aggregated peak performance of 9 PFlops for the compute partitions based on conventional CPUs and accelerated GPUs, with 6 PFlops of peak performance delivered by the CPU-only partition.

The new cluster brings Lenovo Neptune, a direct liquid cooling technology, to ENEA for the first time. Direct liquid cooling is more efficient than air cooling and helps to reduce energy use while supporting the Cresco8 HPC that has a 700kW of electrical power consuming. A Vertiv dry cooler of 1 MW with inlet temperature of 45 °C guarantee a Power Usage Effectiveness (PUE) of 1.045.

CRESCO8 provides the high-performance, scalable platform needed for advanced DTT plasma modelling, showing excellent strong and weak scaling with fusion-relevant codes and supporting reliable, production-level simulations.

CRESCO's strong results highlight the crucial role of HPC in supporting ENEA's scientific work and collaborations, and the ongoing infrastructure upgrades will help maintain this role in the coming years.

Department of Energy Technologies and Renewable Energy, Division for the Development of Information and Communication Technology Systems - CRESCO Team

CRESCO8: THE NEW ENEA HPC FACILITY FOR NUCLEAR FUSION RESEARCH

Francesco Iannone* and *CRESCO team*:
L.Acampora, D.Alderuccio, F.Ambrosino, G.Baldassarre, T.Bastianelli, G.Bracco, L.Bucci, F.Buonocore, M.Caiazzo, B.Calosso, G.Cannataro, M.Caporicci, G.Caretto, M.Casa, G.Cascone, M.Catillo, M.Celino, M.Chinnici, R.Clemente, D. De Chiara, M.De Rosa, D.Di Mattia, S.D'Onofrio, M.Faltelli, G.Ferro, M.Fois, N.Fonso, M.Fratarcangeli, A.Funel, F.Genovesi, S.Giusepponi, G.Guarnieri, M.Gusso, W.Lusani, M.Marano, S.Marchio, A.Mariano, M.Mongelli, P.Palazzari, F.Palombi, F.Pascarella, S.Pecoraro, S.Pierattini, G.Ponti, E.Rossi, G.Santomauro, A.Scalise, F.Simoni, M.Steffè.

Energy Technologies & Renewable Sources Department - Information Communication Technologies, Lungotevere Thaon di Revel, 76, 00196 Rome Italy

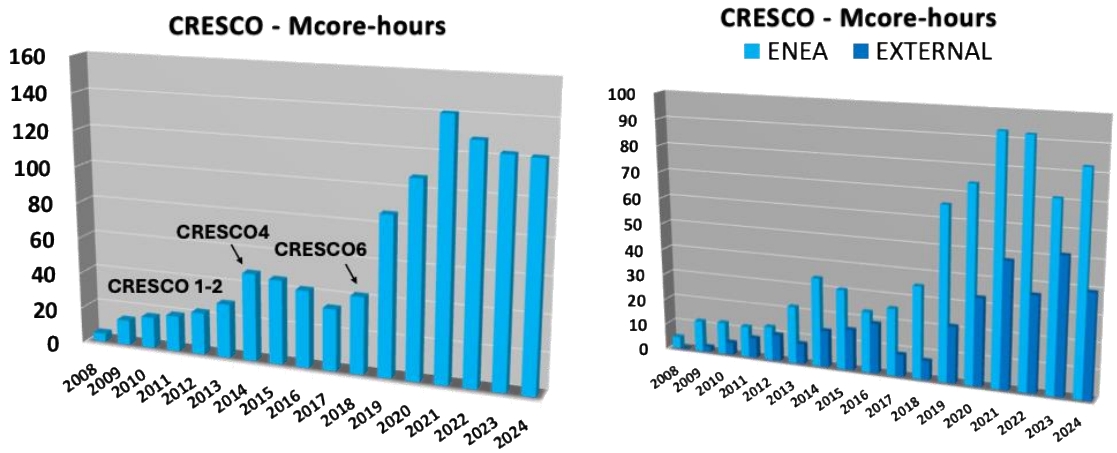
* Corresponding author. E-mail: *francesco.iannone@enea.it*

Abstract. The CRESCO8 supercomputing facility at ENEA represents a strategic national infrastructure supporting advanced research in nuclear fusion. Designed to deliver high computational throughput and scalable performance, CRESCO8 enables large-scale simulations, data analysis, and multi-physics modelling essential for next-generation fusion experiments. In particular, the system plays a crucial role in supporting the Divertor Tokamak Test (DTT) facility in Frascati, providing the computational backbone required to study plasma behavior, optimize divertor configurations, and assess power exhaust solutions under reactor-relevant conditions. Through its integration of high-performance computing resources, tailored software environments, and dedicated user support, CRESCO8 accelerates scientific discovery and engineering innovation for the Italian and international fusion community. This article presents the architecture, operational model, and key workflows enabled by CRESCO8, highlighting its pivotal contribution to the DTT program and to fusion research at large.

Introduction

The Cresco project [1],[2] launched in 2008, marked the beginning of ENEA's journey into High-Performance Computing (HPC) systems, utilizing an architecture based on multicore and multiprocessor nodes connected by low-latency, high-bandwidth networks. This architectural approach has been consistently maintained across various stages of scaling ENEA's HPC systems, with advancements closely aligned with the evolution of CPU technology. As a result of this strategic growth, a robust user community, both within ENEA and externally, has steadily expanded, leading to a significant increase in the demand for computing time. Figure 1(a) illustrates the growth in millions of core-hours over approximately a decade, beginning in 2008 with the first HPC systems, Cresco1-2, which delivered a peak performance of 25 Tflops. This was followed by the installation of Cresco4 in 2015, achieving a peak power of 100 Tflops, and later, Cresco6, which reached a peak of 1.4 Pflops in 2018. In Figure 1(b), the same upward trend is shown for Cresco usage, divided between internal ENEA

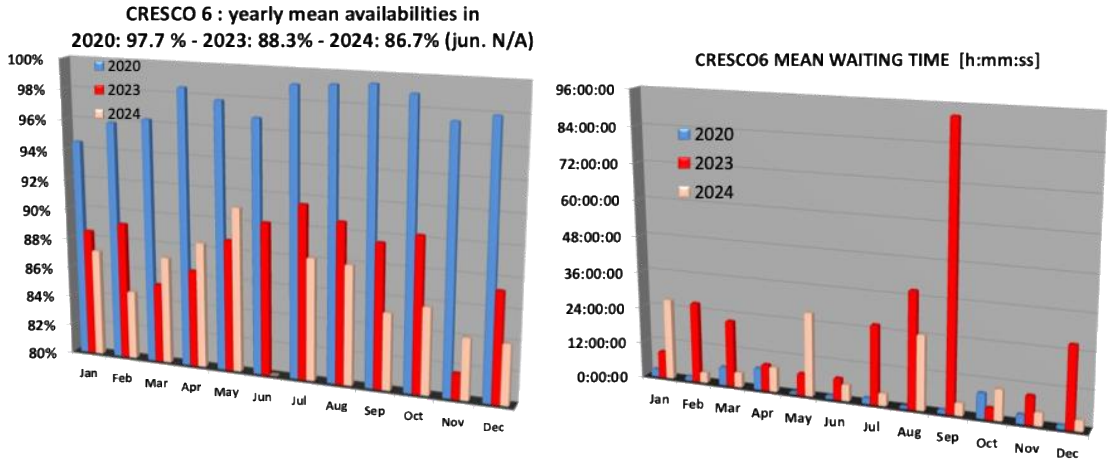
users and external users. By 2024, the utilization of Cresco HPC systems reached around 122 million core-hours.



(a) (b)

Fig.1: (a) CRESCO usage since 2008; (b) CRESCO usage since 2008 break down in ENEA internal and External Users.

To complete the statistics of the Cresco6 operations in 2024, the figure 2(a) shows the availabilities of TIER0 Cresco6 in 2020, in 2023 and 2024 respectively 97.7, 88.3 % and as annual mean, whilst the figure 2(b) shows the wait main time of the jobs in the TIER0 Cresco6 queues respectively in 2020, 2023 and 2024.



(a) (b)

Fig.2: (a) CRESCO6 availability in 2020, 2023 and 2024; (b) CRESCO6 mean waiting time in the queues in 2020 and 2023.

The Fig. 2(a) illustrates the monthly mean availability of the TIER0 Cresco6 HPC cluster over a three-year period, comparing data from 2020, 2023 and 2024. The metric of interest is the percentage of time the HPC system was operational and available for computational tasks. In 2020, the availability was notably high, maintaining an operational status for approximately 98% as yearly mean. This high level of availability reflects the system's relative newness, with minimal hardware-related failures and

downtime. Routine maintenance and software updates were the primary causes of any observed interruptions.

By 2023, however, a marked decline in availability is evident, with the system operational for only 88% as yearly mean. This reduction, continued in 2024 with an availability of 86.7%, can be attributed to the aging hardware, which has increasingly led to unplanned outages due to component failures, decreased system efficiency, and longer repair times. As infrastructure ages, the risk of hardware degradation and failure increases, contributing to extended downtimes and reduced reliability. Additionally, the increasing complexity of managing aging systems may have further compounded this decrease in availability. The data underscores the critical impact that hardware aging can have on the reliability and performance of HPC infrastructures, emphasizing the need for timely upgrades and maintenance strategies to mitigate availability loss over time.

The Fig.2(b) presents the monthly average queue waiting times in the Cresco6 HPC cluster for the years 2020, 2023 and 2024. The data reveals a significant increase in the waiting times over this three-year period, pointing to changes in system usage and resource availability. In 2020, the average waiting times in the queue remained relatively low, with most months recording monthly mean waiting times under 3 hours. This reflects a balanced system load, where computational demand was well-aligned with available resources. The system's hardware was still operating efficiently, and maintenance was adequately managed, ensuring optimal performance and minimal delays. However, by 2023, a clear upward trend in queues waiting times is observed, with averages reaching as high 24 hours in some months of 2023 and 27 hours in some months of 2024. This substantial increase is attributed to two primary factors. First, there was a notable rise in computational workload, driven by increased demand from users, which exceeded the system's capacity to efficiently distribute jobs. Second, the reduced availability of computational resources due to aging hardware significantly impacted the system’s performance. The degradation of hardware components, coupled with the absence of a proper maintenance service, led to more frequent downtimes and slower recovery times, further exacerbating the backlog of queued jobs. This combination of higher demand and lower resource availability illustrates the critical challenges faced by HPC infrastructures as they age, especially when maintenance is not adequately sustained. The graph highlights the importance of both capacity planning and timely hardware upgrades to ensure the continued efficiency of HPC resources.

The two new HPC clusters: CRESCO7 and XCRESCO, installed in 2023, are still in a pre-production phase and the availability and usage data are not available yet.

New Tier-0 ENEA CRESCO8

Within the framework of the NextGeneration EU funding plan a new ENEA supercomputer has been deployed in 2024 to support the plasma physics modelling and new materials research relevant for nuclear fusion reactors. The new ENEA supercomputer, CRESCO8 (Fig.3), shall provide HPC resources for DTT, Divertor Tokamak Test facility: a fusion experiment under construction at the Frascati ENEA Research Center.

Fig.3: New Tier-0 ENEA CRESCO8

CRESCO8 supercomputer is essentially based on LENOVO technological solutions that provides an aggregated peak performance of 9 PFlops for the compute partitions based on conventional CPUs and accelerated GPUs, with 6 PFlops of peak performance delivered by the CPU-only partition.

Furthermore it integrates a high-performance storage system with a raw capacity of no less than 10 PBytes and I/O performance exceeding 50 GBytes/s, and a high-performance interconnect network based on Infiniband NDR 200 Gbps technology, with a sufficient number of switches to interconnect 880 nodes/servers in a 2:1 blocking fat-tree topology. Finally, for the management network is composed of a sufficient number of Ethernet switches with 1 Gbps ports and 10 Gbps uplinks, while for the cluster network is composed of Ethernet switches with 25 Gbps ports and 40 Gbps uplinks, to be connected to the ENEA central hub. CRESCO8 is a X86_64 CPU architecture based on 5th Gen Intel® Xeon® Platinum 8592+ processors, with built-in accelerators including Intel® Advanced Vector Extensions 512 (Intel® AVX-512) and Intel® Advanced Matrix Extensions (Intel® AMX). For acceleration, ENEA is deploying 30 additional 5th Gen Intel® Xeon® Platinum 8592+ processors with 60 Intel® Data Center GPU Max 1550 accelerators attached. The cluster also includes 32 Intel® Xeon® Max 9480 processors with high-bandwidth memory (HBM) for bandwidth-constrained workloads. OneAPI enables researchers to run the same application code across the heterogeneous architecture, spanning CPUs, CPUs with HBM, and GPUs. The operative environments are based on Redhat 9 as operating, Kerberos/Openafs as Single-Sign-On and network filesystem, Slurm as resource management system and 10 PBytes of DDN/Lustre ES400NVX2 as high performance parallel filesystem with a throughput of 50 GB/s. The three partitions are described in the Tab. 1.

Partition	#nodes	CPU	Memory	Network
Lenovo EMR	758	2x Intel 8592+ Emerald Rapids (64c @1.9GHz)	16x 32GB (DIMM @5600MHz)	nVidia NDR400 SharedIO
Lenovo SPR HBM	16	2x CPU Intel Max 9480 (56c @1.9GHz)	2x 64GB HBM 16x 64GB (DIMM @4800MHz)	nVidia NDR400 SharedIO
Lenovo PVC	15	2x Intel 8592+ Emerald Rapids (64c @1.9GHz) + 4x Intel Max 1550	16 x 32GB (DIMM @5600 MHz)	nVidia NDR400 SharedIO

Tab.1: CRESCO8 hardware partitions.

For the Cresco8 HPC cluster, ENEA chose a solution proposed by Lenovo. The new cluster brings Lenovo Neptune, a direct liquid cooling technology, to ENEA for the first time. Direct liquid cooling is more efficient than air cooling and helps to reduce energy use while supporting the Cresco8 HPC that has a

700kW of electrical power consuming. A Vertiv dry cooler of 1 MW with inlet temperature of 45 °C guarantee a Power Usage Effectiveness (PUE) of 1.045.

ENEA carried out testing using several of its application benchmarks, measuring the performance of the latest generation CPU against ENEA's CRESCO6 and XCRESCO. The results showed in Tab.2, provide performance accelerations of up to 8x with the same number of nodes.

Scientific Application (EMR)	**#nodes**	**CRESCO6 Time to solution [sec.]**	**CRESCO8 Time to solution [sec.]**	**Boost**
QE	16	3186.95	376.4	Up to 8x
GROMACS	16	1659.224	264.9	Up to 6x
ANSYS FLUENT	4	379	87	Up to 4x
OPENFOAM	16	193.06	72.2	Up to 2x
LAMMPS	16	1160	255.9	Up to 4x
Scientific Application (PVC)	**#nodes**	**XCRESCO Time to solution [sec.]**	**CRESCO8 Time to solution [sec.]**	**Boost**
CP2K	8	26818	1892.2	> 14x
LAMMPS	8	5316	128	> 41x

Tab.2: CRESCO8 benchmark results

DTT software suite

The CRESCO8 high-performance computing (HPC) system hosts a comprehensive software suite designed to support advanced numerical modeling, data analysis, and visualization activities in the field of nuclear fusion research. The installed and configured environment integrates the software stack required for both experimental data interpretation and numerical simulation workflows associated with ITER, DTT (Divertor Tokamak Test facility), and the broader EU fusion research community. The installed software includes:

- Intel OneApi and MPI libraries for distributed and hybrid parallelization.
- Plasma physics codes for kinetic, fluid, and hybrid modelling (e.g., GENE, ORB5, GYSELA, WARPX and STARWALL).
- Data processing and machine learning frameworks (IMAS, Python, NumPy, SciPy, TensorFlow, PyTorch) tailored for large-scale diagnostic data analysis from fusion experiments.
- Visualization and post-processing tools (Paraview, VisIt, Matplotlib) for 3D field and particle data inspection.
- Scientific workflow and environment management tools (Kepler, Singularity, Spack, SLURM, HDF5, NetCDF).

The configuration ensures full compatibility with existing fusion community workflows, including ITER Integrated Modelling & Analysis Suite (IMAS) [3] standards, enabling interoperability with external data repositories and experimental diagnostics.

Overall, the CRESCO8 software environment provides a robust, scalable, and reproducible computational ecosystem designed to advance simulation-based research and data-driven analysis in support of ongoing and future magnetic confinement fusion experiments.

IMAS: Integrating Modelling Analysis Suite

CRESCO8 provides a fully configured installation of the Integrated Modelling & Analysis Suite (IMAS), the standardized data and software framework developed for ITER and adopted by EUROfusion. The IMAS, installed on CRESCO8, offers a unified data model and a set of APIs enabling seamless access, manipulation, and storage of experimental and simulation data. The installation includes:

- IMAS Core Framework and Data Dictionary (standard ITER schema);

- Data Access Layer (DAL) supporting MDSplus and HDF5 backends;
- Python, C/C++, and Fortran interfaces for integration with existing simulation codes;

Fusion Data Management Tools

To ensure compatibility and data exchange within the European fusion community, the following tools are also deployed:

- MDSplus for experimental data acquisition and hierarchical data storage;
- HDF5 and NetCDF for structured numerical output and interoperability;
- Data staging and transfer tools optimized for parallel I/O and remote access through the CRESCO8 data mover.

Fusion Data Visualization and Workflows

CRESCO8 includes a wide range of software environments supporting high-end visualization and data analysis in both 2D and 3D:

- ParaView and VisIt for large-scale scientific visualization (parallel rendering, in-situ visualization);
- Matplotlib, NumPy, SciPy, and Pandas for data analysis and numerical post-processing;
- JupyterLab and JupyterHub for interactive workflows and collaborative analysis;
- OpenGL-based visualization libraries for embedded and in-situ rendering in simulation codes.
- Kepler for scientific workflow

Gyrokinetic and Plasma Physics Codes

Several advanced computational models and simulation codes describing gyrokinetic plasma behavior are installed and optimized for execution on the CRESCO8 CPU-enabled architecture. These codes contribute to understanding the microturbulent transport processes and confinement properties in magnetically confined plasmas relevant to next-generation fusion reactors. The suite includes:

- GENE [4] (Gyrokinetic Electromagnetic Numerical Experiment) for nonlinear gyrokinetic turbulence simulations;
- ORB5 [5] is a global, gyrokinetic, electromagnetic, multi-species code based on a Lagrangian variational description.
- GYSELA [6] (Gyrokinetic SEmi-LAgrangian) codes for modelling on-scale turbulence and transport in tokamak plasmas;
- WARPX [7] for particle-in-cell (PIC) simulations of plasma micro instabilities
- STARWALL [8] based on JOREK-STARWALL model, it is a simulation tool for studying plasma behavior in fusion reactors.

This integrated environment on CRESCO8 provides researchers with a robust, interoperable, and high-performance ecosystem for fusion plasma modeling, data management, and visualization. It enables end-to-end workflows—from raw data acquisition and integrated modeling to physics interpretation and large-scale numerical simulation—fully aligned with the computational standards of the ITER, DTT and EUROfusion programs.

Benchmarks

Scalability benchmarks were carried out for three gyrokinetic codes: GENE, ORB5 and GYSELA. A comparison between two different versions has been done for the first two codes: GENE and ORB5 whilst for GYSELA the last stable release was under test.

The GENE scalability tests were carried out in order to compare the performance between v.3.0 and v.3.1. The code runs with 2 MPI tasks/node and 64 OpenMP threads/task using from 16 to 512 compute nodes. The benchmarks result report the time step in second as depicted in Fig.4.

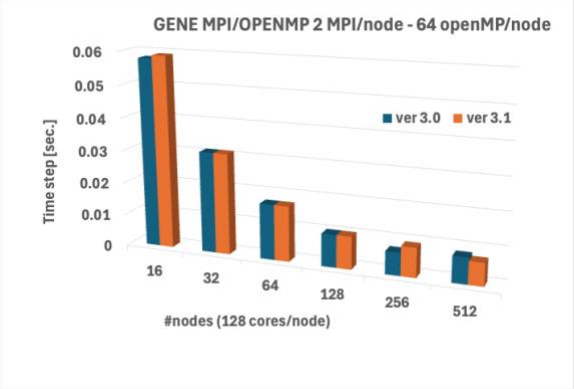


Fig.4: GENE scalability tests on CRESCO8 CPU partition

The ORB5 scalability tests were carried out in order to compare the performance between v.4.0.0 and v.4.9.4. The code runs with 8 MPI tasks/node and 16 OpenMP threads/task using from 16 to 512 compute nodes. The benchmarks result report the time loop wall clock in second as depicted in Fig.5.

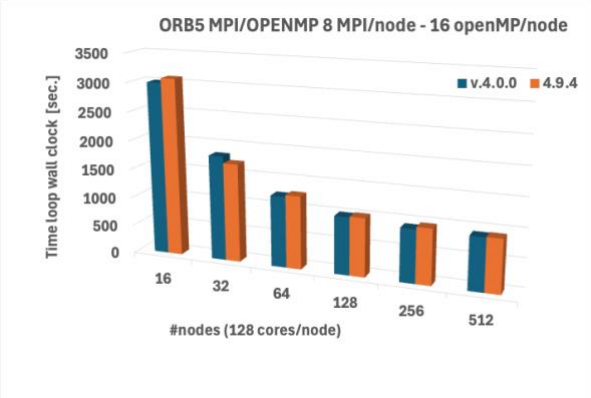


Fig.5: ORB5 scalability tests on CRESCO8 CPU partition

The GYSELA scalability tests were carried out with 2 MPI tasks/node and 64 OpenMP threads/task using from 16 to 256 compute nodes. The benchmarks result report the time elapsed in second as depicted in Fig.6.

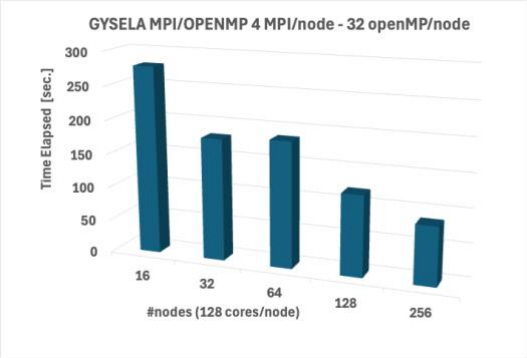


Fig.6: GYSELA scalability tests on CRESCO8 CPU partition

Conclusions

The ENEA CRESCO8 computing infrastructure provides the high-performance capabilities required for the advanced modelling activities of the DTT experiment in Frascati. Its architecture, based on energy-efficient multicore processors and a high-bandwidth interconnect, enables large-scale simulations essential for plasma physics and engineering studies.

Benchmark results obtained with gyrokinetic plasma codes relevant to fusion research confirm excellent strong and weak scaling across the system. The parallel algorithms employed by these codes fully exploit the multicore architecture, demonstrating high efficiency even at large core counts.

Overall, CRESCO8 proves to be a robust and scalable platform for the computational demands of the DTT programme, supporting reliable, production-level simulations and providing a solid foundation for future high-performance modelling activities.

Acknowledgements

This work was granted by the European Commission NextGenerationEU within the framework of the National Recovery and Resilience Plan (PNRR) (Mission 4 "Education and Research", Component 2, Investment 3.1), Project “Fund for the realisation of an integrated research and innovation infrastructure system”, code IR0000001.

References

[1] G.Ponti et al. *The role of medium size facilities in the HPC ecosystem: The case of the new CRESCO4 cluster integrated in the ENEAGRID infrastructure*. Proceedings of the 2014 International Conference on High Performance Computing and Simulation, HPCS 2014 6903807, pp. 1030-1033

[2] F.Iannone et alt. *CRESCO ENEA HPC clusters: A working example of a multifabric GPFS Spectrum Scale layout.t* Open Access International Conference on High Performance Computing and Simulation, HPCS 2019 9188135, pp. 1051-1052

[3] Imbeaux F., Pinches S.D., Lister J.B., et al. “Design and first applications of the ITER integrated modelling & analysis suite (IMAS)”, Nuclear Fusion, 55 (2015) 123006. DOI: 10.1088/0029-5515/55/12/123006.

[4] F. Jenko, W. Dorland, M. Kotschenreuther, B. N. Rogers (2000) Electron temperature gradient driven turbulence Physics of Plasmas 7, 1904 (2000) DOI: 10.1063/1.874014.

[5] Lanti, E., Ohana, N., Tronko, N., Hayward-Schneider, T., Bottino, A., McMillan, B. F., Mishchenko, A., Scheinberg, A., Biancalani, A., Angelino, P., Brunner, S., Dominski, J., Donnel, P., Gheller, C., Hatzky, R., Jocksch, A., Jolliet, S., Lu, Z. X., Martin Collar, J. P., Novikau, I., Sonnendrücker, E., Vernay, T., Villard, L. (2020) ORB5: A global electromagnetic gyrokinetic code using the PIC approach in toroidal geometry, Computer Physics Communications, 251, 107072. DOI:10.1016/j.cpc.2019.107072.

[6] Grandgirard, Virginie & Abiteboul, Jeremie & Bigot, Julien & Cartier-Michaud, T. & Crouseilles, Nicolas & Erhlacher, Charles & Damien, Esteve & Dif-Pradalier, G. & Garbet, Xavier & Ghendrih, Philippe & Latu, Guillaume & Mehrenberger, Michel & Norscini, Claudia & Passeron, Chantal & Rozar, Fabien & Sarazin, Yanick & Strugarek, Antoine & Sonnendrücker, Eric & Zarzoso, David. (2015). A 5D gyrokinetic full- global semi-Lagrangian code for flux-driven ion turbulence simulations. Computer Physics Communications. 207. 10.1016/j.cpc.2016.05.007.

KNOWLEDGE ASYMMETRIES IN LLMS: ENHANCING EFFICIENCY AND SAFETY THROUGH CULTURAL AWARENESS

Daniela Alderuccio*, Rossana Cotroneo[2]

[1]*ENEA – TERIN-ICT-HPC*[1]

[2]*ENEA – TTEC-GOT-KTM*[2]

ABSTRACT. This paper focuses on Knowledge Asymmetries in Large Language Models (LLMs). LLMs abilities and performances are the results of the interplay of *Parametric* and *Contextual* Knowledge. *Parametric Knowledge* is the internal, knowledge, implemented by initial training data. *Contextual Knowledge* is the external knowledge from private repositories. Asymmetries origin from the variety of Natural Language but this diversity generates disparities in the digital representation of linguistic knowledge, thus influencing LLM training, data storage, latency and service costs. Languages are connoted as 'low' vs 'high' resourced in relation to disparity in data, research, and performance of Language Technologies across languages. Disparities play a critical role in digital competitiveness and in economic development. Emerging technologies such as Artificial Intelligence (AI) can also widen the digital divide leading to an AI divide, strengthening power dynamics inequity, and exacerbating the economic divide in the accessibility of NLP technology. Speakers of under-represented languages experience the technology exclusion from technological advances (ranging from lack of computing resources and basic computational tools to exclusion from pre-trained language models). While this technology exclusion prevents only them from taking advantage of NLP innovations, Knowledge Asymmetries in Parametric Knowledge implemented in LLMs impact on the Efficiency and Safety of all users. Embracing linguistic and cultural differences (by introducing Cultural Sensitivity, Fairness, ethical Training Data and Safety Protocols), is a first step toward reducing Knowledge Asymmetries in Multilingual Large Language Models and reaching Digital Language Equity.

Keyword: Knowledge Asymmetry, Artificial Intelligence, multilingual LLM, Cultural Awareness, Fairness, Ethics, Efficiency, Safety.

Introduction

Languages are not created equal. In the World there are 7,159 languages [1], adopting 6 different writing systems, denominated scripts. Asymmetries origin both from the variety of Natural Language, and from the uneven presence/ distribution of languages in the Digital World (in terms of digital encoding and volume of digital contents). The most fundamental issue is the representation disparity in data, i.e. not all cultures are equally represented. Furthermore, access to the Internet is unevenly distributed often within each country. Even if you were to include the entire Internet in all languages, large sections of humanity would still not be represented in the resulting training dataset. If we consider that Language

1 Corresponding author. E-mail: Daniela.alderuccio@enea.it

and Culture are intertwined and Culture is also the lens through which people understand linguistic messages [2] when contributing to the Internet, the resulting text usually reflects socio-cultural values and identity. When we use those texts to train a language model that makes stochastic decisions based on the training datasets, we often see a reflection of embedded values in generated outputs. Values shift from people, to training data, to models, to generated outputs [3].

Knowledge in LLMs

LLMs parameters encode a representation of the world's knowledge. This Knowledge is implicitly embedded within the model's architecture, and it can be retrieved and manipulated through the computational processes of the neural network [4]. LLMs abilities/performances are the results of the interplay of *Parametric* and *Contextual Knowledge*. *Parametric Knowledge* is the internal, static, general, and dated knowledge, implemented by initial training data. Parametric Knowledge in LLMs is encoded in models' parameters through vast amounts of text data during pre-training. *Contextual Knowledge* is external, dynamic, domain-specific, and up-dated knowledge from private repositories. Contextual Knowledge for LLMs refers to the information (or tools) augmented in LLMs context window to enhance generation (RAG), which hopefully supplements LLMs internal PK [5]. Parametric knowledge enables LLMs to deeply compress and integrate information, allowing them to generalize and apply this knowledge across various contexts. Parametric knowledge within LLMs is opaque, often encountering challenges such as interpretability issues, outdated information, hallucinations, and security concerns. Linguistic Knowledge is considered parametric knowledge.
Blasi et al (2022) [6] found that systematic performance drops accord to dimension such as language varieties. The performances of language technology are sensitive to diverse aspects of the language under study, including morphology, word order, writing script or data availability. Out of the over 7.000 languages spoken or signed in the world today, only a handful are systematically represented in academia and industry.

Asymmetries

After English, a handful of Western Languages dominate the field, as well as even fewer non-Indo-European Languages, primarily Chinese, Japanese, and Arabic. This gap between overrepresented languages as English and underrepresented languages is evident also in Europe: the best supported languages are English, followed by Spanish, German and French; the least supported among the EU official: Irish and Maltese. Also, the other European languages are massively under-resourced in terms of data, tools, and technologies. From 2012 to 2022 the gap between English and the next cluster of languages is getting bigger instead of smaller and EU aim is to seek balance between European languages in the digital realm [7]. Giagkou (2023) [8] defined the concept of Digital Language Equality (DLE) as the state of affairs in which all languages have the technological support and situational context necessary for them to continue to exist and to prosper as living languages in the digital age. Data scarcity is usually adopted as a criterion to identify low vs high-resource languages as there is an asymmetry in the availability of amount of data (no. of content: labelled, unlabelled, internet content, websites) for all languages. Nevertheless, asymmetry already arise in the digital format of language representation. Language diversity generates disparities in the digital representation of linguistic knowledge: ASCII characters are sufficient for English and require 1 byte; other Latin scripts character (Greek, Cyrillic, Coptic, Armenian, Hebrew, Arabic Syrian) require 2 bytes; Chinese, Korean, Japanese require 3 bytes; languages with consonant and diacritics more bytes. Asymmetry in the digital format of language representation, then result in tokenization dis-parity for LLMs. Tokenization transforms text into processable tokens: LLM (such as ChatGPT) process and generate text sequences by first splitting the text into smaller units called tokens, short and common words such as "you" have its own token, whereas longer or less common words are broken into smaller sub-words. Languages and relative tokenization are not equal. Morphologically complex languages (such as Arabic, Finnish) necessitate a greater number of tokens to convey the same meaning as English, resulting in inefficiencies, particularly in low-resource languages. This disparity increases computational costs and diminishes fluency text

generation. These asymmetries in multilingual tokenization methods incurs varying costs across languages. Furthermore, in OpenAI's GPT-4, Arabic requires over three times as many tokens as English, leading to slower inference times and a reduction in output quality, which also illustrates the importance and urgency of fair tokenization [9]. The unfair digital representation of linguistic knowledge leads to Tokenization disparity not only among different scripts but also in languages sharing the same scripts (Standard vs language varieties (dialects); English-centric vs Western Languages (see Irish and Maltese); Western Colonial Languages vs creole languages; Latin-alphabet vs other alphabets or other scripts. Inequities/Asymmetries in tokenization happens within: (i) the same language and is related to differences in lengths, commonness/rarity of words; (ii) fonts (many of the languages e.g. Arabic, Hindi, Korean, etc., cannot be rendered using a single font so requiring more tokens; (iii) within languages with different script (see English vs. Burmese, Shan, Amharic). *"The English "you" has three characters but a single token" - "as an example "ၵူၺ်း", one of the Shan words for "you". This word is constructed from one consonant and three diacritics. As the diacritics are encoded separately, there are 4 Unicode codepoints for this Shan character, resulting in 9 tokens (It is tokenized by ChatGPT and GPT-4 as ◌ၵ◌ူ◌ၺ◌်◌း: 25870 247 - 157 224 224 - 25870 118 - 25870 116)* [10]; (iv) within different languages sharing same script, same sentence, with same meaning and same number of words: i.e. *The dog eats the apple (en – 5 tokens) - Il cane mangia la mela (it – 7 tokens - El perro come la manzana (sp – 8 tokens)* [11] ; (v) within different languages, same sentence, with same meaning: the English (ISO EN-US) sentence *"what will the weather be next week"* has 7 token; in Spanish (es-ES) 8 tokens; in Korean (ISO ko-KR) 12 tokens, in Burmese 61 tokens in Amharic (am-ET) has 69 tokens [12]. Ahuja 2023 [13] investigated the Tokenization Fertility (defined as the average number of sub-words produced per tokenized word (higher means worse quality) of the tokenizers used by different models. Ahuja (2023) observed that the tokenizers for the OpenAI models are substantially worse for low-resource, non-Latin script languages: - where the fertility for languages like Malayalam and Tamil is so high (∼10) that the tokenizer essentially operates as a byte-level tokenizer for these languages. Note that this means that for low-resource languages, substantially larger number of tokens are needed to encode the inputs as well as for generation, which results in a significant additional API cost. Ahia et al. (2023) [14] discusses how this phenomenon leads to large socio-economic disparities for speakers of underrepresented languages. Furthermore, tokenization disparities influence context length: due to the limited-context-sized window in LLM, this result in less informative context for low-represented languages than English-centric context.

Asymmetry Impact

The optimal performance on LLMs requires balancing computational efficiency, desired performance, and the size of the context window. In an LLM, a context window is the maximum number of tokens that can be accommodated or consulted by the model at any point. These tokens constitute the basic building blocks of text from which LLMs derive meaning. Large, bigger context windows permit models to catch long-range connections among texts and enhance their ability to understand and generate sequences of coherent text parts. However, larger context windows also come at the expense of computational and memory resources, thus affecting scalability and efficiency in deploying models [15]. Asymmetries in character length representation led to increased context-window. In Large Language Models (LLMs), longer training corpus and increased context window size, result in higher machine learning capabilities but generate higher carbon footprints. To ensure an environmentally sustainable growth of AI we must focus on efficiency optimization holistically across the entire AI system stock: data, algorithms and models, systems, and infrastructures at scale [16]. As far as Environmental Impact: Each query generates 4.3 grams of CO_2, over 20× the carbon cost of a typical Google search [17], whereas a bottle of water per email is the hidden environmental costs of using AI chatbot [18]. As far as energy efficiency is concerned, the average ChatGPT query uses about 0.34 watt-hours (about what an oven would use in a little over one second, or a high-efficiency lightbulb would use in a couple of minutes). It also uses about 0.000085 gallons of water (roughly one fifteenth of a teaspoon) [19]. The economic sustainability of models like OpenAI o3 could also become a practical

limit to their large-scale diffusion. Although OpenAI o3 is a more capable model, the imbalance between efficiency and performance is huge (no. of token per inference[2]): R1-Zero uses only an average of 11K token per inference, while OpenAI o3 (high compute) requires 57M. OpenAI has better results but with enormously higher computational and inference costs (OpenAI o3 (high) has an inference cost of $ 3400; OpenAI o1 (low) $ 0.43 – four times higher than r1 zero $ 0.11)[20] [11]. Linguistic diversity has an impact not only on efficiency. Asymmetry also endangers safety. Linguistic inequality of safety training data enhances cross-lingual vulnerability: translating unsafe English inputs into low-resource languages successfully circumvents LLM's safeguard. Cross-lingual vulnerability mainly applies to low-resource languages (79% of the time LLM engages with the unsafe translated inputs and provides actionable items that can get users towards their harmful goals. LLMS perform poorer with low-resource languages: asymmetry in safety training data and benchmark in under-represented languages impact on safety. Limited training on low-resource languages causes technological disparities for native speakers, but cross-lingual vulnerability poses risks to all LLM users. Focusing on existing English-centric benchmarks may give false sense of security while the model remains susceptible to attacks in languages out of distribution of the safety training data. Translating English inputs into low-resource languages is sufficient to bypass safeguards and elicit harmful responses, increasing the chance to bypass GPT-4's safety filter from 1% to 79%. Other high/mid-resource languages have significantly lower attack success rates. This is due to the fact that safety mechanism do not generalize to low resource languages, Safety measures are circumvented and harmful responses are elicited nearly half of the time - by translating unsafe inputs into low-resource languages like Zulu or Scot Gaelic; combining different low-resource languages increases the jailbreaking success rate to around 79%, whereas the original English inputs have less than 1% of success rate.; individually all high/mid-resource languages have less than 15% attack success rate. To be truly safe, LLMs need safety mechanism like training and red-teaming to apply to a wide range of languages. [9].

Conclusions

Awareness of linguistic diversity and knowledge asymmetry in LLM highlights their impacts on results/outputs, on efficiency and safety, as well as on geo-political, economic, social dimensions and equity in real-world. The underestimation of risks from linguistic inequality generates both: (i) technological disparities primarily affecting speakers of low-resource languages (lack of infrastructure and basic computational tools; exclusion from pre-trained language models; limited training data); (ii) and safety vulnerabilities posing a risk to all LLMs users. It is not only a matter of epistemic or algorithmic justice, but a new holistic approach is also needed to develop wide and fair language coverage, cultural sensitivity, ethical training data, and safety protocols. AI is one of the most transformative technologies influencing digital competitiveness. Limited access to the basic component of AI: electricity, connectivity and affordable services contributes to widening digital divide and inequalities [21]. Digital competitiveness has become fundamental for economic growth. The digital divide remains a significant challenge, both across and within countries. Wealthier economies or regions with better infrastructure can sustain their levels of digital competitiveness, while others lag. The role of emerging technologies is also crucial for long-term digital competitiveness, particularly in relation to widening digital disparities. *"In human groups, Knowledge Asymmetry between the holder of knowledge and know-how (practitioner or expert) and the beneficiary of the end-product (observer) is overcome by trust between themselves.(...) Transparency and Trust are the key to success, otherwise Knowledge Asymmetries will raise ethical concerns and create uneven power dynamics at social-*

[2] AI inference costs* have rapidly declined due to innovation https://www.bain.com/insights/deepseek-a-game-changer-in-ai-efficiency/# - Inference costs refer to the expense of querying a trained model (and are typically measured in USD per million tokens. - Training costs: Foundation models have a high training costs influenced by training duration, hardware type, quantity and utilization rate. There is a direct correlation between the training costs of AI models and their computational requirements, Models with greater computational trainings needs, cost more to train.

economic and political level.” [22]. After Trust, next steps are (i) gaining awareness of asymmetries, (ii) then paying Respect and giving recognition to Diversity: these are the key elements needed for a new paradigm shifted towards the plurality of knowledges [23].

References

[1] Ethnologue 2025 - https://www.ethnologue.com/ethnoblog/welcome-28th-edition/
[2] Herschovich D., Challenges and Strategies in cross-cultural NLP in 60th annual meeting of the Association for Computational Linguistics, volume 1, pages 6997-7013, 2022 - https://aclanthology.org/2022.acl-long.482/
[3] Johnson R.L., The Ghost in the Machine has an American accent: value conflict in GPT-3. https://arxiv.org/pdf/2203.07785 - https://doi.org/10.48550/arXiv.2203.07785
[4] Wang M. et al.,Knowledge Mechanism in Large Language Models: Survey & Perspective” arXiv:2407.15017v4) 2024
[5] Cheng S. et al., Understanding the Interplay between Parametric and Contextual Knowledge for Large Language Models - https://arxiv.org/pdf/2410.08414 - 2024
[6] Blasi d. et al. Systematic Inequalities in Language Technology Performance across the World’s Languages in Proceedings of the 60th Meeting of the ACL vol. 1, pp.5486-5505 (2022) - https://aclanthology.org/2022.acl-long.376/

[7] Giagkou M., Digital Language Equality in Europe - https://www.europarl.europa.eu/cmsdata/257076/Giagkou.pdf
[8] Giagkou M. et al. - European Language Equality: A Strategic Agenda for Digital Language Equality – DOI: 10.1007/978-3-031-28819-7_4 https://link.springer.com/chapter/10.1007/978-3-031-28819-7_4 - https://link.springer.com/content/pdf/10.1007/978-3-031-28819-7_38.pdf
[9] Qin L. et al. (2025) A survey of multilingual large language models, in Patterns 6, Review Volume 6, 2025 - https://www.cell.com/patterns/fulltext/S2666-3899(24)00290-3 - doi: 10.1016/j.patter.2024.101118.
[10] Petrov et al. - Language Model Tokenizers Introduce Unfairness Between Languages - https://arxiv.org/pdf/2305.15425
[11] De Baggis M. - Puliafito A., E poi arrivò DeepSeek, 2025, ISBN: 9788850339266 statistics https://arcprize.org/blog/r1-zero-r1-results-analysis
[12] Yennie Jun (2023) in https://medium.com/data-science/all-languages-are-not-created-tokenized-equal-cd87694a97c1 -

[13] Ahuja et al., MEGA: Multilingual Evaluation of Generative AI https://arxiv.org/abs/2303.12528
[14] Ahia et al., Do All Languages Cost the Same? Tokenization in the Era of Commercial Language Models - https://arxiv.org/abs/2305.13707

[15] Williamson T., How Context Window Size affect LLM performance - https://www.deepchecks.com/question/how-does-context-window-size-affect-llm-performance/
[16] Wuet et al., Efficiency Optimization helps bend the ever-increasing resource demand, cited in Wuet et al. 2024, Beyond Efficiency: Scaling AI Sustainability
[17] Global ChatGPT Usage Trends by Country and Age (2022–2025) (browsercat.com)
[18] Washington Post https://www.washingtonpost.com/technology/2024/09/18/energy-ai-use-electricity-water-data-centers/
[19] Sam Altman post https://blog.samaltman.com/the-gentle-singularity
[20] Qwen (cited in [11]) Technical Report https://arxiv.org/pdf/2309.16609 - https://lunabot.ai/it/models/qwen-2.5-max Qwen2 https://kili-technology.com/large-language-models-llms/qwen-2-guide-what-can-we-learn-about-alibaba-s-multilingual-llm#multilingual-capabilities
[21] UN AI Advisory board, 2023, Interim Report: Governing AI for Humanity, UN AI Advisory Board , https://www.un.org/sites/un2.un.org/files/ai_advisory_body_interim_report.pdf

[22] Kenza Arab, “Knowledge Asymmetry” 43 Int. Conference on Information Systems, Copenhagen, 2022

[23] Melber H. 2019, Knowledge, Asymmetric Power Relations and Us. In https://www.developmentresearch.eu/?p=364

STATISTICAL PROCESS CONTROL OF PROFILES USING MIXTURE REGRESSION MODELS

Fiorenzo Ambrosino[1], Christian Capezza[2], Fabio Centofanti[2], Davide Forcina[2], Antonio Lepore[2*], Biagio Palumbo[2]

[1]*ENEA - Italian National Agency for New Technologies, Energy and Sustainable Economic Development, P.le E. Fermi, 1 (Loc. Granatello), 80055, Portici (NA), Italy*

[2]*University of Naples Federico II, Department of Industrial Engineering, P.le V. Tecchio 80, 80125, Naples, Italy*

ABSTRACT. Modern industrial processes often operate under diverse conditions, giving rise to multiple in-control states that challenge traditional statistical process control (SPC) techniques. In this work, we propose a new method (FunMixReg), a novel approach for monitoring functional quality characteristics influenced by covariates in the presence of heterogeneous operating regimes. The method models the in-control process as a finite mixture of functional linear models (FLMs), capturing multimodal behavior and enhancing detection performance. Deviations from the in-control distribution are identified using a likelihood ratio test. The methodology is implemented in R and evaluated through extensive Monte Carlo simulations, which compare FunMixReg with existing control charts. Simulations are performed on the CRESCO6 high-performance computing cluster, allowing parallel execution of over 100 scenarios and sensitivity analyses. Results show that FunMixReg consistently achieves superior detection power. The method will be made publicly available through the funcharts R package, promoting reproducibility and future adoption.

Introduction

Statistical Process Control (SPC) is a cornerstone of quality monitoring in manufacturing and industrial systems [9]. Classical SPC tools, often based on multivariate statistics, rely on scalar summaries of observed data. However, with the increasing complexity of modern processes and the availability of high-frequency sensors, quality characteristics are frequently observed as functions—trajectories over time, space, or other continuous domains. Profile monitoring [10], a subfield of SPC, is thus better suited to this setting, as it directly models such functional data.

Nevertheless, an important challenge persists: real-world processes often operate under varying environmental or production conditions that generate heterogeneous, multimodal in-control (IC) behavior [7]. In this setting, standard functional control charts, such as those assuming a single functional linear model (FLM), are no longer adequate, as they fail to represent multiple IC states, leading to misclassifications and compromised performance.

To overcome this limitation, we propose a novel methodology based on functional mixture regression (FunMixReg) [2,3,4,5]. This control chart models the IC distribution as a finite mixture of FLMs, where each component captures a distinct subpopulation characterized by a specific functional relationship

* Corresponding author. E-mail: antonio.lepore@unina.it.

between a response and a covariate. This approach accounts for both functional heterogeneity and covariate effects, offering a significant advancement over existing methods.
The methodology is evaluated via extensive Monte Carlo simulations, tailored to assess the behavior of the proposed control chart under different scenarios of heterogeneity and process shifts. Data are simulated under multiple regimes varying by the nature and severity of the shifts introduced (in intercept functions, regression coefficients, or both), and across several levels of cluster dissimilarity.
The entire approach is situated within the broader research domain of intelligent manufacturing, statistical learning for high-dimensional data, and functional data analytics.

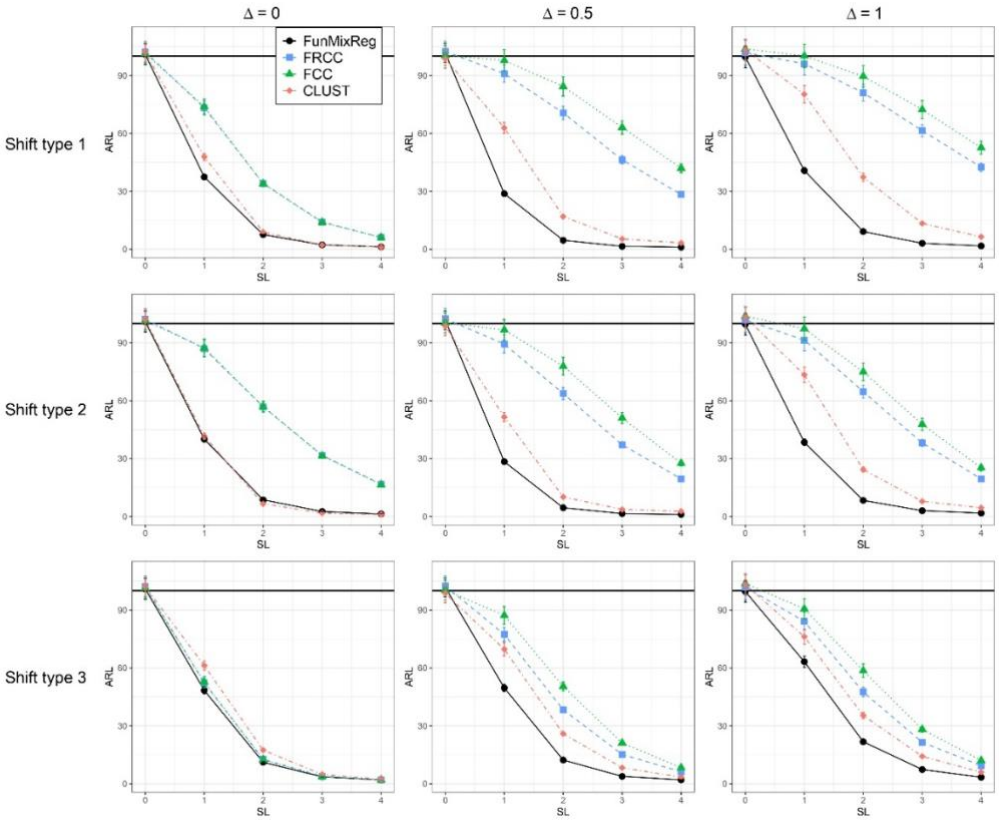


Fig. 1 Mean ARL achieved in Phase II by FunMixReg, FRCC, FCC, and CLUST for each combination of Δ and shift types as a function of the severity level (SL).

Simulation study

Simulation settings

The simulation study required for evaluating the FunMixReg methodology is computationally intensive, involving 100 independent Monte Carlo experiments across a wide range of settings. These settings involve the comparison of the proposed method to existing methods in the literature, namely: the functional regression control chart (FRCC) [6], a functional control chart (FCC) based on Hotelling's T2 and squared prediction error (SPE) applied to principal component scores of the response, and a clustering-based approach (CLUST) that first clusters the data and then applies the FCC within each identified cluster, following ideas from [7], under different setting of heterogeneity and shift types. Moreover, sensitivity analyses are also conducted to quantify the sensitivity of FunMixReg to the choice of hyperparameters and initialization strategies in terms of performance. For illustrative purposes, Figure 1 shows some simulation results for three heterogeneity levels, indexed by the parameter Δ, and

three shift types. Each Monte Carlo experiment corresponds to the evaluation of Average Run Length (ARL) performance across multiple severity levels (SL).

Use of HPC resources

The programming language used throughout the project is R [11] (version 3.6.1). No compiled code, external compilers, or low-level programming languages were involved, and therefore, no specific compilers were used. The software does not rely on external numerical libraries such as LAPACK or BLAS, nor does it make use of GPU acceleration. The parallelization strategy is implemented using R's built-in parallel package, specifically via the mclapply function. This approach leverages embarrassingly parallel computation, as each Monte Carlo simulation scenario is independent of the others. Each Monte Carlo experiment is parallelized internally using 48 cores. Moreover, simulations are distributed across available nodes on the CRESCO6 high-performance computing cluster located in Portici. Every job simulates an entire Monte Carlo experiment, each lasting about 1 hour when run on 48 cores. Considering 5 SL values and 9 panels represented in Figure 1, the total computing time without parallelization would have exceeded 4500 hours. Thanks to the HPC infrastructure, this was reduced to approximately 2 hours by leveraging multiple nodes concurrently. The total storage requirement was relatively low, estimated at 1–2 GB, with no GPU boards used. The ability to perform simultaneous jobs and full-core utilization per node was essential to completing the study efficiently and reproducibly.

Software and Code Development

The simulations relied on a combination of in-house code specifically developed for this project and routines available in the funcharts R package [1], which was developed and maintained by the authors. The funcharts package provided essential tools for functional data simulation, dimension reduction, preprocessing, and visualization. Custom code was written to implement the mixture regression model, the estimation algorithm, and the likelihood ratio test used in the control chart.

Simulation Results

The FunMixReg control chart consistently outperformed existing approaches across a wide range of simulated scenarios. It achieved excellent false alarm control with ARL_0 values meeting or exceeding nominal thresholds, while also providing superior sensitivity to OC shifts, particularly in the presence of heterogeneity in regression relationships. The advantage of FunMixReg was most evident when clusters shared similar intercepts but differed in regression coefficients. In such cases, traditional methods like FRCC and FCC were unable to distinguish between subpopulations, resulting in poor detection performance. The CLUST method, while competitive in simpler settings, struggled under complex configurations or when cluster separability was low. Beyond the results presented in this manuscript, the methodology has been progressively developed and disseminated through various academic venues [2,3,4,5].

Conclusions

The FunMixReg control chart offers an effective solution for monitoring complex functional processes with heterogeneous IC structures. By modeling the IC distribution as a mixture of FLMs, the proposed approach accommodates multiple process states and leverages covariate information for improved detection. Extensive simulations show that FunMixReg consistently outperforms existing methods, particularly in scenarios with subtle or complex heterogeneity.

The availability of high-performance computing resources was essential to the timely and thorough validation of the methodology. The integration of FunMixReg into the funcharts package will ensure future accessibility and impact across academic and applied domains.

Acknowledgments

This research activity was carried out within the MICS (Made in Italy – Circular and Sustainable) Extended Partnership and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.3 – D.D. 1551.11-10-2022, PE00000004).
The research activity of B. Palumbo was carried out within the MOST - Sustainable Mobility National Research Center and received funding from the European Union Next-GenerationEU (PIANO NAZIONALE DI RIPRESA E RESILIENZA (PNRR) – MISSIONE 4 COMPONENTE 2, INVESTIMENTO 1.4 – D.D. 1033.17-06-2022, CN00000023).
This work reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them.
The computing resources and related technical support used for this work have been provided by the CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [8]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see \url{http://www.cresco.enea.it/english} for information.

References

[1] Capezza, C., Centofanti, F., Lepore, A., Menafoglio, A., Palumbo, B., & Vantini, S. (2023). funcharts: control charts for multivariate functional data in R. *Journal of Quality Technology*, *55*(5), 566-583.

[2] Capezza, C., Centofanti, F., Forcina, D., Lepore, A., Palumbo, B., (2023) "Resistance Spot Welding Process Monitoring Through Mixture Function-on-scalar Regression Analysis." Proceedings of the 23rd Annual Conference of the European Network for Business and Industrial Statistics (ENBIS), Valencia, Spain, 10–14 September

[3] Capezza, C., Centofanti, F., Forcina, D., Lepore, A., Palumbo, B., (2023) "Control Charts for Functional Data Based on Functional Mixture Regression." Proceedings of the 16th International Conference of the ERCIM WG on Computational and Methodological Statistics (CMStatistics), Berlin, Germany, 16–18 December

[4] Capezza, C., Centofanti, F., Forcina, D., Lepore, A., & Palumbo, B. (2024). Functional mixture regression control chart. arXiv preprint arXiv:2410.20138.

[5] Capezza, C., Centofanti, F., Forcina, D., Lepore, A., Palumbo, B. (2025) A Mixture Regression Framework for Profile Monitoring. In the Third Conference of the Statistics and Data Science Group of the Italian Statistical Society

[6] Centofanti, F., Lepore, A., Menafoglio, A., Palumbo, B., Vantini, S.: Functional regression control chart. Technometrics 63(3), 281–294 (2021)

[7] Grasso, M., Colosimo, B.M., Tsung, F.: A phase I multi-modelling approach for profile monitoring of signal data. International Journal of Production Research 55(15), 4354–4377 (2017)

[8] Iannone, F., Ambrosino, F., Bracco, G., De Rosa, M., Funel, A., Guarnieri, G., Migliori, S., Palombi, F., Ponti, G., Santomauro, G., Procacci, P.: Cresco enea hpc clusters: a working example of a multifabric gpfs spectrum scale layout. In: 2019 International Conference on High Performance Computing Simulation (HPCS), pp. 1051–1052 (2019)

[9] Montgomery, D.C.: Introduction to Statistical Quality Control. Wiley (2012)

[10] Noorossana, R., Saghaei, A., Amiri, A.: Statistical analysis of profile monitoring, vol. 865. John Wiley & Sons (2011)

[11] R CORE TEAM (2021). R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria.

SENSITIVITY OF THE WRF MODEL TO SUB-GRID PHYSICS PARAMETERIZATIONS

Marta Antonelli[1*], Emanuela Pichelli[1,2], Maria Vittoria Struglia[1,2]

[1] *ENEA Italian National Agency for New Technologies, Energy and Sustainable Economic Development, Rome, 00123, Italy*

[2]*ICSC Italian Research Center on High-Performance Computing, Big Data and Quantum Computing, Casalecchio di Reno (BO), 40033, Italy*

**Corresponding author email: marta.antonelli@enea.it*

ABSTRACT. This work presents the numerical experiments carried out to investigate the sensitivity to sub-grid schemes of microphysics and turbulence of the regional climate model wrf. the aim is to implement an optimal configuration at the convection-permitting scale, needed over the mediterranean region for a realistic representation of its climate and of its climatic extremes, strongly related to local and complex interactions.

Introduction

Regional Climate Models (RCMs) are numerical tools for investigating the Earth climate over limited regions of the planet at resolution finer (~10 km) than the Global Climate Models (GCMs, ~100 km). The Earth's climate is characterized by a wide range of spatial and temporal scales (~nm-Mm; ms-100years). Once the grid-step of a model is set, the processes evolving on finer sub-grid scales are not explicitly resolved and must be approximated through parameterizations, which are inherently source of uncertainty. Last generation of RCM application is at the so called “convection-permitting” scale, i.e. at scales of 4 km or finer, which allow the deep-convection to be solved explicitly rather than parameterized, thus turning-off one of the major sources of uncertainty [1], [2], due to incomplete understanding of processes and uncertainties of scheme parameters [3]. The convection has the crucial role of redistributing vertically heat and moisture, triggering convective precipitation and influencing mesoscale dynamics by changing vertical stability and affecting surface heating and radiation trough clouds. The interest in correctly simulating convection is also related to its high relevance in driving extreme events such as heavy precipitation, windstorms and floods. In this study we present the sensitivity tests performed to assess the best configuration of the WRF model at convection-permitting scale in terms of turbulence and microphysics for climatic application over Italy.

Model Setup

A Convection Permitting version of the regional model WRF (CPM) has been implemented over the Italian region (Fig.1). The adopted grid-step of 3 km allows to resolve the deep-convection explicitly. The CPM domain is nested within a regional simulation at 15 km of resolution over Europe in turn driven by GCM data (ERA5, [4]).

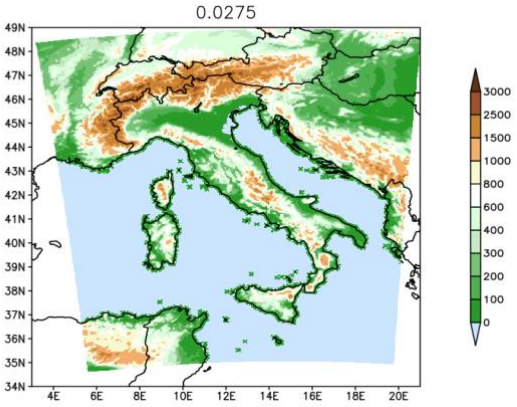


Fig.1: The 3km domain covered by the WRF model in its convection-permitting configuration, topography (m)

The domain size is 426x501x54 grid points. The CPM configuration has the same base version of the model adopted in [5] (their Table 1), but the cumulus parameterization is switched off, while the urban and lake schemes are active. A series of sensitivity tests of the CMP to sub-grid processes schemes have been carried out to assess an optimal model configuration, varying the microphysics and planetary boundary layer (PBL) parameterizations (Table 1). Each experiment is two years long (1980-1981), but the intercomparison among the tests is done for the year 1981 only, being the 1980 model's spin-up time. Tests 1 to 3 were meant to define modelling strategy aspects, other than the sub-grid physics, and have not been included in the analysis. The tests are carried out by changing one of the sub-grid processes at a time.

Table 1: Sensitivity test configurations.

Test-num	microphysics	PBL	Surface layer
Test 4	Thompson (2 moment rain)	MYNN	Monin-Obukhov (Janjic) scheme
Test 5	Thompson	YSU	MM5 Monin-Obukhov scheme
Test 6	Morrison (2 moments)	MYNN	Monin-Obukhov (Janjic) scheme
Test 7	WSM5 (1 moment)	MYNN	Monin-Obukhov (Janjic) scheme

Results and conclusions

The analysis has been performed comparing daily precipitation indices of mean, heavy precipitation (99^{th} percentile of the daily precipitation distribution, p99), wet-days (>1 mm/day) intensity and wet-days frequency. Results of the experiments were compared to the observations EURO4M (5 km) [6]. Fig.2 shows the summer (JJA) maps of all precipitation indices (from left to right as listed above) over the Alpine Region (ALP) covered by EURO4M. The analysis was also performed for other seasons, but only the JJA is shown as representative of almost purely convective precipitation. The tests show similar statistics, with the exception of test 5 that has a strong wet bias during winter and test 6 that overestimates the mean precipitation during spring (not shown). We can speculate that such similarities among the different configurations might be due to the short simulation length, as well as to the influence of the domain size. The domain is small along the longitude axis, which may allow the boundary conditions signal to influence the simulation more than the independent solution from the sub-grid physics.

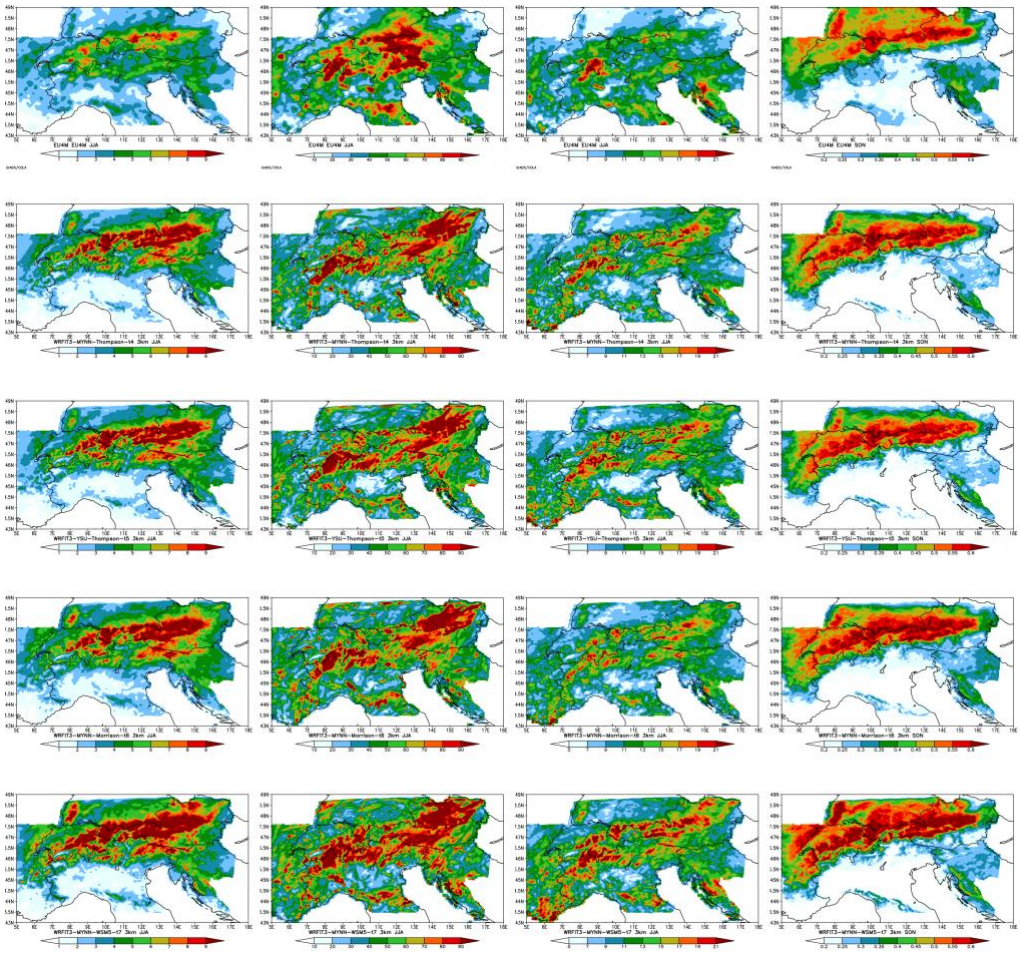


Fig.2: Daily precipitation indexes for JJA season: from left to right: mean, p99, wet days intensity, wet days frequency. From top to bottom: EURO4M, test 4, test5, test 6, test 7. The shaded area is where EURO4M is available

An analysis of the mean vertical profiles of temperature, specific humidity and winds over the ALP region was also carried out (not shown). Sensitivity tests, compared to ERA5 reanalysis profiles, show non-significant differences in the lower atmosphere (below 850 hPa, <0.5 °C, 0.3 g/Kg, 1 m/s respectively). The only notable difference between tests 4 and 5 was found in the meridional wind. This demonstrates that using a non-local PBL (test 5) in winter, summer and autumn allows for better agreement with the ERA5. Conversely, using a single-moment microphysics scheme (test 7) provides some benefits in terms of moisture content from spring to autumn. However, neither scheme offers an evident advantage over the others for improving significantly the statistics.

In conclusion, the analysis of the WRF CPM sensitivity to sub-grid physics schemes revealed minimal differences among the tests performed. Therefore, we opted for the same configuration of test 4 for a climatic run at the convection-permitting scale. This new, cutting-edge application of the ENEA WRF CPM will be used to gain new insights into the effect of global warming on extreme hazards. It will also serve as benchmark for the 5 km experiment described in [5] (5Km-Hindcast), having the same configuration, apart from the cumulus scheme being switched off and the grid-step (3 and 5 km). This will allow us to investigate the model's performance in the so-called "grey zone" for cumulus schemes. This is poorly documented in literature due to the risk of algorithm violations at such scales. This could be key to saving computing resources while avoiding the underrepresentation of physics and interactions at the local scale in the next generation of RCM applications, but a robust evaluation is required.

Acknowledgments
The computing resources and the technical support to computing facilities provided by CRESCO/ENEA-GRID High Performance Computing infrastructure and staff. We acknowledge for funding RETURN Extended Partnership that received funding from the European Union Next-GenerationEU (National Recovery and Resilience Plan – NRRP, Mission 4, Component 2, Investment 1.3 – D.D. 1243 2/8/2022, PE0000005) and ICSC Italian Research Center on High-Performance Computing, Big Data and Quantum Computing that received funding from the European Union Next-GenerationEU (National Recovery and Resilience Plan – NRRP, Mission 4, Component 2, Investment 1.4 – D.D: 3138 16/12/2021, CN00000013)

References

[1] Foley, A. M. Uncertainty in regional climate modelling: A review. P*rogress in Physical Geography*, **34**, pp. 647–670, (2010). https://doi.org/10.1177/0309133310375654
[2] N. Ban, J. Schmidli, C. Schär. Evaluation of the new convective-resolving regional climate modeling approach in decade-long simulations. *Journal of Geophysical Research*, **119**, pp. 7889–7907, (2014), https://doi.org/10.1002/2014JD021478
[3] S. Li, S. Zhang, Z. Liu, X. Yang, A. Rosati, J. Golaz, and M. Zhao. The Role of Large-Scale Feedbacks in Cumulus Convection Parameter Estimation. *J. Climate*, **29**, 4099–4119, (2016) https://doi.org/10.1175/JCLI-D-15-0117.1.
[4] Hersbach H, Bell B, Berrisford P, et al (2020) The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society **146**:1999–2049. https://doi.org/10.1002/qj.3803
[5] M. V. Struglia, A. Anav, M. Antonelli, S. Calmanti, F. Catalano, A. Dell'Aquila, E. Pichelli, G. Pisacane. Impact of spatial resolution on CMIP6-driven Mediterranean climate simulations: a focus on precipitation distribution over Italy, *EGUsphere [preprint],* https://doi.org/10.5194/egusphere-2025-387, 2025.
[6] F. Isotta et al. The climate of daily precipitation in the Alps: development and analysis of a high-resolution grid dataset from pan-Alpine rain-gauge data, *International Journal of Climatology*, **34**, pp. 1657-1675, (2014), doi: 10.1002/joc.3794

MATCH: Computational Approaches for Antibody Structure Analysis Using ENEA CRESCO HPC Infrastructure

Caterina Arcangeli[1], Valerio Chiarini[2]. Matteo Fois[3,] Simone Giusepponi[3], Sara Marchio[3*], Mariateresa Mancuso[1], Giuseppe Roscilli[2]

[1]*ENEA, SSPT-BIOTEC, Rome, Italy*[4]

[2]*Takis Biotech s.r.l. Rome, Italy*

[3]*ENEA TERIN-ICT, Rome, Italy*

Abstract. Within the ENEA-Takis MATCH Joint Lab (PNRR Rome Technopole) framework, ENEA's HPC resources have been employed to explore computational methods for antibody structure prediction and conformational sampling. These activities supported the testing of software tools and computational workflows aimed at improving the analysis of antibody-antigen interactions.

Introduction

The MATCH (Monoclonal Antibody developmenT and CHaracterization) Joint Lab, developed within the Rome Technopole PNRR framework, aims to advance computational methodologies for the design, structural optimization, and evaluation of monoclonal antibodies. The project brings together expertise in molecular modelling, high-performance computing (HPC), and artificial intelligence (AI) to establish a scalable, automated pipeline supporting the *in silico* phase of antibody development.

The MATCH pipeline integrates conformational sampling, predictive modelling, and structural scoring into a decision-oriented workflow capable of filtering and prioritizing candidate antibody sequences before experimental validation. This approach significantly reduces downstream testing costs and accelerates the discovery timeline. The ENEA HPC environment, comprising CPU-based clusters and GPU-accelerated platforms, provides the computational infrastructure necessary to run large-scale simulations and AI-enhanced analyses in parallel.

Software and Hardware Infrastructure Setup

In the initial phase of the project, considerable effort was devoted to configuring a scalable and secure computational environment to support the automated workflows developed within the MATCH framework.

User environments and secure access credentials were created across both the AFS (Andrew File System) and GPFS (IBM's General Parallel File System), enabling distributed, high-throughput data management across the ENEAGRID infrastructure. A dedicated 10 TB storage area was allocated to the project on GPFS to accommodate simulation data, modelling outputs, and software assets.

* Corresponding author. E-mail: sara.marchio@enea.it.

The software architecture was implemented across two high-performance computing systems:

- CRESCO6 cluster (CPU-based), used primarily for simulation-intensive and parallelized molecular analysis tasks;
- XCRESCO GPU cluster, equipped with IBM Power9 processors and NVIDIA V100 GPUs, used for AI-enhanced modelling and deep learning operations.

All required software tools were deployed to support structure prediction, molecular simulations, data post-processing, and visualization. Visualization tasks and molecular system inspection were carried out using VMD [1] and Chimera [2], while core simulation workflows relied on integrated molecular modelling packages capable of supporting enhanced sampling and force-field-based dynamics.

To streamline the execution of these workflows, a set of automation scripts was developed, enabling reproducible configuration of molecular systems and batch execution across the HPC platforms. The architecture was designed to allow dynamic allocation of jobs across CPU and GPU resources, ensuring scalability and high-throughput screening capacity.

The combination of CPU- and GPU-resident capabilities allowed for efficient distribution of workloads based on their computational demands, ensuring high-throughput, scalable processing of antibody candidates in parallel. Computational tasks were efficiently distributed across the available high-performance computing infrastructure. Classical CPU-based calculations were executed on the CRESCO6 cluster located in Portici ENEA center, while machine learning and GPU-intensive tasks were performed on the XCRESCO cluster that is equipped with IBM Power9 processors and NVIDIA V100 SXM2 GPUs and is located in Frascati ENEA center. The peak computational performance of XCRESCO reaches approximately 1.9 Pflops, providing sufficient resources to support deep learning and large-scale parallel simulations.

Computational Workflow

The computational activities carried out within the project focused on the application and testing of diverse *in silico* methods for antibody structural modeling, conformational sampling, and interaction analysis. The main goal was to evaluate the performance and scalability of selected bioinformatics and molecular modeling software when executed on ENEA's HPC infrastructure (CRESCO6 and XCRESCO clusters). The workflow included several standard steps typically used in antibody research, such as generation of molecular models based on input sequences, exploration of conformational space under different conditions, and preliminary docking analyses to estimate the antibody's ability to interact with potential targets.

To facilitate this, a combination of CPU- and GPU-based computational tasks was executed, leveraging both traditional simulation techniques and modern machine learning approaches. These steps were carried out independently using established open-source and/or licensed software tools, configured to run efficiently in parallel environments.

Results

Testing the Ab optimization pipeline

The computational resources were tested using publicly available antibody-antigen complexes from the Protein Data Bank (PDB) to evaluate performance and scalability. During the test run, the computational procedures executed without interruption and allowed parallel processing of multiple cases, demonstrating good stability and resource utilization across both CPU and GPU nodes.

Each case generated several gigabytes of data and demonstrated that the combination of molecular modeling and docking tools available on the ENEA HPC infrastructure can be used to explore antibody-antigen interactions *in silico*.

Figure 1 shows an example result from a test system (PDB-ID: 4U6V), where a predicted antibody binding pose was compared to the original crystallographic structure. This illustrates the ability of the tested computational environment to support structure-based analysis of antibody-antigen complexes.

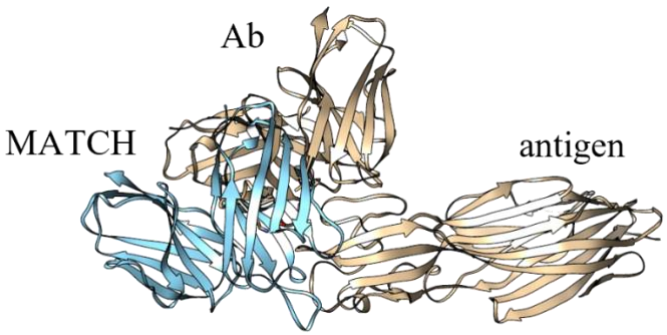


Fig. 1: Example output from a test run (PDB-ID: 4U6V). The antigen and the antibody as in the original PDB structure are depicted in "tan". The binding pose of the same antibody predicted as result of the MATCH pipeline is in light blue.

Conclusion

The MATCH Joint Lab, established within the Rome Technopole framework, demonstrated the potential of ENEA's HPC infrastructure to support advanced computational approaches in antibody research. The activities carried out confirmed the suitability of available resources for large-scale structure-based simulations and exploratory modeling. These preliminary evaluations represent a solid foundation for the future development of high throughput in silico workflows aimed at supporting preclinical studies.

Acknowledgments

This work was supported and funded by the Project ECS 0000024 Rome Technopole – CUP B83C22002820006, National Recovery and Resilience Plan (NRRP), Mission 4, Component 2 Investment 1.5, funded from the European Union – NextGenerationEU. FP7, Joint Lab MATCH. Authors acknowledge Guido Guarnieri, Nicolò Colistra and Paolo Roberto Saraceni from ENEA for their help with the preliminary part of the study.

References

[1] W. Humphrey et al. VMD: visual molecular dynamics. Journal of molecular graphics 14, pp. 33-38, (1996). DOI:10.1016/0263-7855(96)00018-5

[2] E. F. Pettersen et al. UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry 25, pp. 1605-1612, (2004). DOI:10.1002/jcc.20084

COMPUTATIONAL SCREENING OF C_2-C_{10} ALIPHATIC. MOIETIES ON HYDROGENATED SILICON (111) SURFACE

Francesco Buonocore[15], Sara Marchio[1], Simone Giusepponi[1], Massimo Celino[1]

[1]Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA) – C. R. Casaccia, Via Anguillarese 301, 00123 Rome, Italy

ABSTRACT. Silicon is widely used in electronics, sensors, and solar cells due to its flexibility as a semiconductor. In this work, we used first-principles simulations based on density functional theory, performed on Cresco6, to study how simple organic groups (alkyl, alkenyl, and alkynyl) interact with hydrogen-terminated silicon (111) surfaces. We focused on how these chemical bonds influence the electronic properties at the interface, such as the Schottky barrier and energy level alignment, which are important for charge transport. The results offer useful insights for improving the design of silicon-based electronic devices.

Introduction

Silicon, thanks to its excellent semiconductor properties, is a key material in many modern technologies, including integrated circuits, sensors, and photovoltaic cells [1–13]. The performance and versatility of silicon-based devices can be greatly improved by modifying their surfaces, which allows for fine-tuning of their electronic, optical, and chemical characteristics. However, during fabrication, the formation of silicon oxide can degrade device performance or even prevent proper operation, as the oxide acts as an insulating layer. To avoid this, one widely used method is the functionalization of hydrogen-terminated silicon (H–Si) surfaces with molecular monolayers [14].

There is growing interest in developing alternative functionalization strategies that prevent oxide formation while preserving the conductive properties of silicon. Since the first demonstration of Si–C alkyl monolayers [15], it has been shown that the Si–C bond is more stable than the Si–H bond. In this context, aliphatic chains are a promising option to protect the silicon surface from oxidation, forming a direct interface that maintains good electrical properties. This approach also enables the customization of surface features for specific applications and has been proposed as a way to enhance the performance of low-cost polycrystalline semiconductors [16]. However, a detailed understanding of how such functionalization affects the electronic structure and interface behavior is still needed [17–18].

Short aliphatic chains are known to form more ordered monolayers than longer ones [19–20], but synthesizing monolayers from short alkenes and alkynes using wet chemistry is more challenging. For this reason, alternative methods using gaseous aliphatic molecules at higher pressures have been explored [20].

In this study, we use ab initio density functional theory (DFT) calculations to investigate how single aliphatic moieties (alkyl (C_nH_{2n+1}), 1-alkenyl (C_nH_{2n-1}), and 1-alkynyl (C_nH_{2n-3}), with carbon numbers from 2 to 10) adsorb on H–Si(111) surfaces. While previous studies have shown that functionalization with 1-alkenes and 1-alkynes leads to the formation of alkyl and alkenyl monolayers, respectively [21], our work takes a broader view. We focus on how the presence of single, double, or triple C–C bonds

[5] Corresponding author. E-mail: francesco.buonocore@enea.it

near the surface in individual molecules affects the energetics and electronic properties of the silicon/organic interface.
We simulate the replacement of a surface hydrogen atom with an aliphatic group, forming a Si–C bond through chemisorption. This process significantly alters the interface's electronic behaviour, unlike physisorption, where interactions are weak and mainly due to dispersion forces. By analysing charge distribution, dipole formation, and energy level alignment at the Si–C interface, our goal is to better understand how these modifications influence key properties such as Schottky and tunnelling barriers. These insights are valuable for improving the design and performance of silicon-based electronic devices. This study also sets the stage for future investigations on full monolayers and their collective effects.

Theoretical Methods

The calculations were carried out using density functional theory (DFT) and the plane-wave pseudopotential method, implemented in the PWSCF code of the QUANTUM-ESPRESSO package [22-23], and compiled with Intel Fortran compiler, Math Kernel Library (MKL) and Message Passing Interface (MPI) parallelization on the Cresco6 cluster, with a good scalability up to hundreds of cores. To include Van der Waals interactions, we used the nonlocal functional VdW-df-cx [24], which has already been applied in studies of charge transfer at silicon–organic interfaces [25].

We performed the calculations using ultrasoft pseudopotentials [26]. Structural relaxations were done with a kinetic energy cutoff of 40 Ry for the wavefunctions and a 2×2×1 Monkhorst-Pack k-point grid. The total energy was converged below 0.0001 Ry per atom, and atomic forces were relaxed until they were smaller than 0.001 Ry/Å. For the electronic density of states (DOS), we used a denser 8×8×1 k-point mesh.

We studied the adsorption of aliphatic moieties with carbon chains from C_2 to C_{10}, focusing on three types: alkyl (C_nH_{2n+1}) with only single C–C bonds, 1-alkenyl (C_nH_{2n-1}) with a double bond near the surface, and 1-alkynyl (C_nH_{2n-3}) with a triple bond near the surface. The most stable configurations were used to build a slab model of the Si(111) surface with eight bilayers (Figures 1a–b), to evaluate how the dipole changes compared to the fully hydrogenated surface after relaxation.

The silicon slab was 23 Å thick, with a vacuum region of 65 Å in the direction perpendicular to the surface (L_z), and lateral dimensions of 15 Å ($L_x = L_y$). This setup represents a compromise between computational cost and accuracy. The slab was asymmetric, requiring a correction for dipole self-interaction effects [27]. Both the top and bottom surfaces were passivated with hydrogen atoms, with the organic moiety adsorbed on the top side. The two central bilayers were kept fixed to mimic bulk silicon.

The calculations have been run on Cresco6, exploiting two distinct levels of MPI parallelization, as made available by the QUANTUM-ESPRESSO package. We launched the jobs on 12 nodes, each equipped with 48 processors, resulting in a total of 576 processors. The first level of parallelism involves partitioning into pools, each responsible for a subset of k-points. 4 pools were selected, with an allocation of 1 pool per k-point, thus assigning 3 nodes for each pool. The second level of parallelization is PW parallelization. Whitin this scheme, the orbitals represented in the plane-wave (PW) basis set, along with charges and density in either reciprocal or real space, are distributed across multiple processors. The code employs the 3D Fast Fourier Transform (FFT) to convert electronic wave functions between reciprocal and real space. To parallelize this process, the 3D real-space grid is partitioned into planes and distributed across multiple processors. Similarly, in reciprocal space, columns of G-vectors are allocated to processors to efficiently execute the FFT. Parallelization of PWs demonstrates a good scaling, particularly when the number of processors in a pool divides evenly into the number of planes defined by the FFT mesh for charge density and wavefunction (respectively nr3 and nr3s). Here nr3 has been set at 576. Therefore, each pool has 144 cores, and 4 planes are assigned to each core.

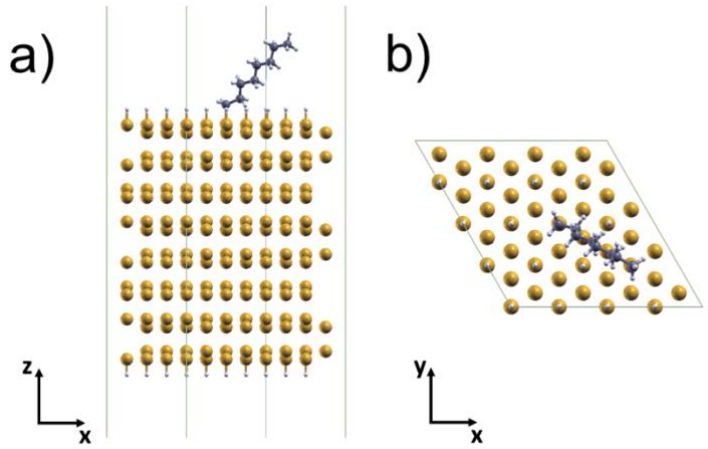


Fig. 1. a) Side and b) top view of the eight bilayers H–Si(111) slab model with adsorption of CN moiety on the top surface. Yellow, white, and grey spheres are Si, H and C atoms, respectively.

Results and Discussion

Screening

We explored a wide set of possible configurations for the adsorption of molecules with single, double, and triple carbon–carbon bonds on the H–Si(111) surface and optimized the structure of each case. For each type of molecule, we identified the most stable geometry, i.e., the one with the lowest adhesion energy, and used it for the analysis of electronic properties [28].

The adhesion strength was quantified by calculating the dissociation energy E_d using the formula:

$$E_d = (E_m + E_s) - E_{sm}, \qquad (1)$$

where E_{sm} is the total energy of the system with the molecule attached to the surface, and E_m and E_s are the energies of the isolated molecule and the clean surface, respectively. Details on the screening workflow and the structural parameters of the tested configurations are available in the dataset referenced in [29]. Among the three types of moieties, 1-alkynyl groups showed the highest dissociation energies, exceeding 5.5 eV, indicating strong binding to the surface. 1-alkenyl groups followed, with dissociation energies ranging from 4.2 eV (C_2H_3) to 4.5 eV ($C_{10}H_{11}$). Alkyl groups had the lowest dissociation energies, around 3.8 eV, suggesting weaker interaction with the silicon surface.

Schottky barriers

A surface dipole is generated when the aliphatic chain is adsorbed on the silicon surface; the Si–C bond is formed, accompanied by charge transfer between the adsorbate and the surface. In the following we

demonstrate that the effect of the Si–C bond formation on the whole chain must be considered to find the correct dipole. The dipole formed at the interface affects the thermionic transport at the silicon/organic molecule/metal heterojunction, because of the Schottky Barrier (SB). Indeed, the SB depends on the surface dipole as follows [10,16]:

$$SB_e = F - c - y \quad (2)$$

$$SB_h = E_g - F + c + y \quad (3)$$

where F is the work function of the metal, c is the electronic affinity of the semiconductor (Si), y is the energy step of the surface dipole, and E_g is the band gap of the semiconductor. SB_e and SB_h are the SB at zero bias of electrons and holes, respectively. In Figure 2 we show how the electron and hole SBs of the Si–H/molecule/metal heterojunction change by varying adsorbed moiety when the metal is Hg by assuming F_{Hg} = 4.5 eV, c_{Si} =4.06 eV [30] and E_g = 1.12 eV. The qualitative variation of the SB for alkyl and 1-alkenyl moieties (i.e., SB_e decreases by increasing the length of the molecular chain) agrees with the experimental findings of Har-Lavan and co-authors [31].

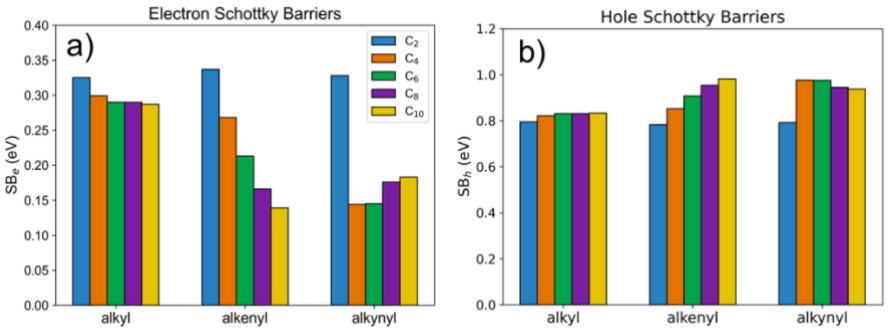


Fig. 2. a) Electron and b) hole Schottky barriers of the alkyl, 1-alkenyl and 1-alkynyl modified Hg/H–Si(111) heterojunction calculated for the one moiety adsorbed per supercell model.

Interface energetics and tunnelling barriers

We determined the bands edges and the frontier orbital levels to investigate the energetics of the H–Si(111)/aliphatic chain heterointerface. We estimated the vacuum level of each adsorption configuration as the constant plane-averaged electrostatic energy in the vacuum gap far away from the top atomic layer. The valence and conduction band edges of the bulk silicon have been aligned to that of the silicon slab by overlapping the plane-averaged electrostatic energy in the central layers. We determined the HOMO and LUMO energies of the aliphatic moieties as the peaks of the projected DOS (PDOS) of the atomic orbitals associated to the organic molecule near the band gap. We found that only 1-alkenyl moieties show a monotonic variation of the tunnelling barrier with respect to the carbon chain length (N) The energy shift of HOMO and LUMO with respect to valence and conduction band edges (E_V and E_C, respectively) of the 1-alkenyl moieties adsorbed on the H–Si(111) surface are shown in Table 1 and indicated as DHOMO (= E_V-HOMO) and DLUMO (= LUMO-E_C), respectively. DHOMO and DLUMO are the energy barrier for the tunnelling of holes and electrons, respectively. Although nonlocal functionals still underestimate the energy gap of molecules and semiconductors, it is interesting to observe that the electron tunnelling barrier of C_9 1-alkenyl moiety in the experiments is found to be around 1 eV [32], in nice agreement with the values of 1.2 and 1.0 eV for C_8 and C_{10} 1-alkenyl moieties, respectively, found in the present report.

Table 1: Ab initio tunnelling barriers for holes (DHOMO) and electrons (DLUMO) of the 1-alkenyl moieties adsorbed on the H–Si(111) surface.

Chain	DHOMO (eV)	DLUMO (eV)
C_2	1.63	2.59
C_4	1.05	1.84
C_6	0.65	1.49
C_8	0.41	1.21
C_{10}	0.21	1.01

Conclusions

We investigated a wide range of possible configurations for the adsorption of alkyl, 1-alkenyl, and 1-alkynyl moieties on the H–Si(111) surface, identifying and analysing the most stable geometries. The type of carbon–carbon bond near the surface—single, double, or triple—has a significant effect on the interface energetics. Among the three, 1-alkynyl groups showed the strongest binding (highest dissociation energy), followed by 1-alkenyl and then alkyl groups.

The surface dipole plays a key role in determining the Schottky barrier (SB) at the semiconductor/organic/metal interface, influencing thermionic charge transport depending on the metal work function and the semiconductor's electron affinity. Notably, 1-alkenyl moieties showed a consistent and significant modulation of the SB as the carbon chain length (N) increased.

To explore tunnelling transport, we calculated the alignment of the frontier molecular orbitals with the silicon band edges. The 1-alkenyl moieties produced the lowest tunnelling barriers, which decreased with longer chains. In particular, the C_{10} 1-alkenyl group reduced the tunnelling barrier to about 1 eV for electrons and 0.2 eV for holes, suggesting it could significantly enhance tunnelling efficiency. These values are in good agreement with available experimental data.

Beyond electronic transport, our findings are relevant for other applications, such as optimizing charge transport layers in photovoltaic devices or tuning interface energetics in water-splitting systems. The approach demonstrated here, supported by calculations on the Cresco6, provides a solid foundation for designing functionalized silicon surfaces tailored to specific technological needs.

Acknowledgments

This work was supported by the European Union through the Next Generation EU funds through the Italian MUR National Recovery and Resilience Plan, Mission 4 Component 2 – Investment 1.4 – National Center for HPC, Big Data and Quantum Computing – CUP I33C22001270007. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff[33]. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research projects, see http://www.cresco.enea.it/english for information. The authors acknowledge the extensive use of the ENEA FARO facility and the support of its management team[34]. The authors thank prof. Mohammed Y. Bashouti for useful discussions.

References

[1] R. Har-Lavan, I. Ron, F. Thieblemont, D. Cahen. Toward metal-organic insulator-semiconductor solar cells, based on molecular monolayer self-assembly on n-Si. *Appl. Phys. Lett.* **94**, 043308 (2009). https://doi.org/10.1063/1.3076115.

[2] S. Maldonado, D. Knapp, N. S. Lewis. Near-Ideal Photodiodes from Sintered Gold Nanoparticle Films on Methyl-Terminated Si(111) Surfaces. *J. Am. Chem. Soc.* **130**, 3300 (2008). https://doi.org/10.1021/ja800603v.
[3] D. Gal, E. Sone, R. Cohen, G. Hodes, J. Libman, A. Shanzer, H. W. Schock, D. Cahen. Engineering the interface energetics of solar cells by grafting molecular properties onto semiconductors. *Proc. Indian Acad. Sci. Chem. Sci.* **109**, 487 (1997).
[4] I. Visoly-Fisher, A. Sitt, M. Wahab, D. Cahen. Molecular Adsorption-Mediated Control over the Electrical Characteristics of Polycrystalline CdTe/CdS Solar Cells. *ChemPhysChem* **6**, 277 (2005). https://doi.org/10.1002/cphc.200400392.
[5] M. Page, O. Niitsoo, Y. Itzhaik, D. Cahen, G. Hodes. Copper sulfide as a light absorber in wet-chemical synthesized extremely thin absorber (ETA) solar cells. *Energy Environ. Sci.* **2**, 220 (2009). https://doi.org/10.1039/B813740D.
[6] D. G. Wu, G. Ashkenasy, D. Shvarts, R. V. Ussyshkin, R. Naaman, A. Shanzer, D. Cahen. Novel NO Biosensor Based on the Surface Derivatization of GaAs by "Hinged" Iron porphyrins. *Angew. Chem. Int. Ed.* **39**, 4496 (2000).
[7] E. Stern, R. Wagner, F. J. Sigworth, R. Breaker, T. M. Fahmy, M. A. Reed. Importance of the Debye Screening Length on Nanowire Field Effect Transistor Sensors. *Nano Lett.* **7**, 3405 (2007). https://doi.org/10.1021/nl071792z.
[8] A. Cattani-Scholz, D. Pedone, M. Dubey, S. Neppl, B. Nickel, P. Feulner, J. Schwartz, G. Abstreiter, M. Tornow. Organophosphonate-Based PNA-Functionalization of Silicon Nanowires for Label-Free DNA Detection. *ACS Nano* **2**, 1653 (2008). https://doi.org/10.1021/nn800136e.
[9] G. Peng, U. Tisch, H. Haick. Detection of Nonpolar Molecules by Means of Carrier Scattering in Random Networks of Carbon Nanotubes: Toward Diagnosis of Diseases via Breath Samples. *Nano Lett.* **9**, 1362 (2009). https://doi.org/10.1021/nl8030218.
[10] A. Vilan, D. Cahen. Chemical Modification of Semiconductor Surfaces for Molecular Electronics. *Chem. Rev.* **117**, 4624 (2017). https://doi.org/10.1021/acs.chemrev.6b00746.
[11] M. Y. Bashouti, M. Pietsch, G. Brönstrup, V. Sivakov, J. Ristein. Heterojunction Based Hybrid Silicon Nanowire Solar Cell: Surface Termination, Photoelectron and Photoemission Spectroscopy Study. *Prog. Photovolt: Res. Appl.* **22**, 1050 (2014). https://doi.org/10.1002/pip.2315.
[12] A. Shalabny, F. Buonocore, M. Celino, L. Zhang, K. Sardashti, M. Härth, D. W. Schubert, M. Y. Bashouti. Enhancing the Electronic Properties of VLS-Grown Silicon Nanowires by Surface Charge Transfer. *Appl. Surf. Sci.* **599**, 153957 (2022). https://doi.org/10.1016/j.apsusc.2022.153957.
[13] A. Shalabny, F. Buonocore, M. Celino, G. Shalev, L. Zhang, W. Wu, P. Li, J. Arbiol, M. Y. Bashouti. Semiconductivity Transition in Silicon Nanowires by Hole Transport Layer. *Nano Lett.* **20**, 8369 (2020). https://doi.org/10.1021/acs.nanolett.0c03543.
[14] S. Harilal, S. Sadhujan, K. Zhang, A. Shalabny, F. Buonocore, B. Ferrucci, S. Giusepponi, M. Celino, M. Y. Bashouti. Uniform Tendency of Surface Dipoles Across Silicon Doping Levels and Types of H-Terminated Surfaces. *Adv. Electron. Mater.* **10**, 2300873 (2024). https://doi.org/10.1002/aelm.202300873.
[15] M. R. Linford, C. E. D. Chidsey. Alkyl monolayers covalently bonded to silicon surfaces. *J. Am. Chem. Soc.* **115**, 12631 (1993). https://doi.org/10.1021/ja00079a071.
[16] A. Vilan, O. Yaffe, A. Biller, A. Salomon, A. Kahn, D. Cahen. Molecules on Si: Electronics with Chemistry. *Adv. Mater.* **22**, 140 (2010). https://doi.org/10.1002/adma.200901834.
[17] X. Li et al. Recent progress on surface chemistry I: Assembly and reaction. *Chinese Chem. Lett.* **35**, 110055 (2024). https://doi.org/10.1016/j.cclet.2024.110055.
[18] X. Li et al. Recent progress on surface chemistry II: Property and characterization. *Chinese Chem. Lett.* **36**, 110100 (2025). https://doi.org/10.1016/j.cclet.2024.110100.
[19] M. Y. Bashouti, C. A. Garzuzi, M. de la Mata, J. Arbiol, J. Ristein, H. Haick, S. Christiansen. Role of Silicon Nanowire Diameter for Alkyl (Chain Lengths C1–C18) Passivation Efficiency through Si–C Bonds. *Langmuir* **31**, 2430 (2015). https://doi.org/10.1021/la5047244.
[20] S. P. Pujari, A. D. Filippov, S. Gangarapu, H. Zuilhof. High-Density Modification of H-Terminated Si(111) Surfaces Using Short-Chain Alkynes. *Langmuir* **33**, 14599 (2017). https://doi.org/10.1021/acs.langmuir.7b03683.

[21] L. Scheres, M. Giesbers, H. Zuilhof. Self-Assembly of Organic Monolayers onto Hydrogen-Terminated Silicon: 1-Alkynes Are Better Than 1-Alkenes. *Langmuir* **26**, 10924 (2010). https://doi.org/10.1021/la100858q.
[22] P. Giannozzi et al. QUANTUM ESPRESSO: A Modular and Open-Source Software Project for Quantum Simulations of Materials. *J. Phys.: Condens. Matter* **21**, 395502 (2009). https://doi.org/10.1088/0953-8984/21/39/395502.
[23] P. Giannozzi et al. Advanced Capabilities for Materials Modelling with QUANTUM ESPRESSO. *J. Phys.: Condens. Matter* **29**, 465901 (2017). https://doi.org/10.1088/1361-648X/aa8f79.
[24] K. Berland, P. Hyldgaard. Exchange Functional That Tests the Robustness of the Plasmon Description of the van der Waals Density Functional. *Phys. Rev. B* **89**, 035412 (2014). https://doi.org/10.1103/PhysRevB.89.035412.
[25] X. Wang, K. Esfarjani, M. Zebarjadi. First-Principles Calculation of Charge Transfer at the Silicon–Organic Interface. *J. Phys. Chem. C* **121**, 15529 (2017). https://doi.org/10.1021/acs.jpcc.7b03275.
[26] D. Vanderbilt. Soft Self-Consistent Pseudopotentials in a Generalized Eigenvalue Formalism. *Phys. Rev. B* **41**, 7892 (1990). https://doi.org/10.1103/PhysRevB.41.7892.
[27] L. Bengtsson. Dipole Correction for Surface Supercell Calculations. *Phys. Rev. B* **59**, 12301 (1999). https://doi.org/10.1103/PhysRevB.59.12301.
[28] F. Buonocore, S. Marchio, S. Giusepponi, M. Celino. Computational Insights into the Energetics of Single C2–C10 Aliphatic Moieties Adsorbed on the Hydrogenated Silicon (111) Surface. *Langmuir* **41**, 9706 (2025). https://doi.org/10.1021/acs.langmuir.4c05103.
[29] S. Marchio, F. Buonocore, S. Giusepponi, M. Celino. Density Functional Theory Study of Silicon Nanowires Functionalized by Grafting Organic Molecules. *Materials Cloud Archive* **81** (2024). https://doi.org/10.24435/materialscloud:15-fs.
[30] A.-E. Haj-Yahia, O. Yaffe, T. Bendikov, H. Cohen, Y. Feldman, A. Vilan, D. Cahen. Substituent Variation Drives Metal/Monolayer/Semiconductor Junctions from Strongly Rectifying to Ohmic Behavior. *Adv. Mater.* **25**, 702 (2013).
[31] R. Har-Lavan, O. Yaffe, P. Joshi, R. Kazaz, H. Cohen, D. Cahen. Ambient Organic Molecular Passivation of Si Yields Near-Ideal, Schottky-Mott Limited Junctions. *AIP Adv.* **2**, 012164 (2012).
[32] A. Aragonès, N. Darwish, S. Ciampi, F. Sanz, J. J. Gooding, I. Dìez-Pérez. Single-Molecule Electrical Contacts on Silicon Electrodes Under Ambient Conditions. *Nat. Commun.* **8**, 15056 (2017). https://doi.org/10.1002/adma.201203028.
[33] F. Iannone et al. CRESCO ENEA HPC Clusters: A Working Example of a Multifabric GPFS Spectrum Scale Layout. *2019 International Conference on High Performance Computing & Simulation (HPCS)*, 1051 (2019). https://doi.org/10.1109/HPCS48598.2019.9188135.
[34] A. Mariano et al. Fast Access to Remote Objects 2.0: A Renewed Gateway to ENEAGRID Distributed Computing Resources. *Future Gener. Comput. Syst.* **94**, 920 (2019). https://doi.org/10.1016/j.future.2018.12.032.

DENSITY OF STATES AND ABSORPTION SPECTRA OF SP^2/SP^3 HYBRID STRUCTURES

Emiliano Burresi[1*]

[1]*ENEA, SSPT Department, Brindisi Research Centre, Brindisi (Br)*

ABSTRACT. Domains of carbon atoms with sp^2 hybridization inside a sp3 carbons matrix are always present in structure like graphene oxide (GO) and reduced graphene oxide (rGO) but their specific features have never been deeply explored. Their ability to stabilize these structures was investigated through density of states and absorption spectra calculations.

Introduction

Graphitic or Graphene-like domains can exist also in more complex structures composed by mixing of sp^2/sp^3 carbon atoms, as example in graphene oxide, a 2D carbon single layer decorates with several oxygen functional groups (epoxy, hydroxyl, carboxyl) on the surface and on sheet edges. The opto-electronic properties are mainly determined by the π states of the sp^2 sites. The π and π* electronic levels of the sp^2 clusters lie within the bandgap of s and s* states of the sp^3 matrix. From experimental point of view, three absorption peaks were found for GO and they were assigned to the n®s, π® π* and π®n*[1]. The investigation of the Density of States (DOS) and the π orbitals around the gap could add important information on the stability of these mixed structures and their optical properties. In this work, density of states was calculated for different sp^2 structures which can be particularly stable for specific aromatic effects already showed in a previous report [2]. The π-gap values of these structures obtained in this work confirm quite well these predictions.

Computational Details

The geometry of the all structures were fully optimized with the PWSCF (Plane Wave Self Consistent Field) code of the QUANTUM ESPRESSO (QE) [3] suite to investigate structural parameters, total energies and band gaps. Plane-wave basis functions, exchange correlation functional PBE generalized-gradient approximation and norm-conserving pseudopotentials were used for PWSCF calculations. The Brillouin zone is sampled using the Gamma-point only and the cutoff energy for the wave function was set at 80 Ry. A cubic box of 40 Å was used to apply the periodic boundary conditions and the system was placed in the center of the box. The Liouville-Lanczos recursive method implemented in TURBO_TDDFT code of the QE suite was used for the calculations of the Lanczos coefficients and the TURBO_SPECTRA code for obtaining the strength of the oscillator in function of the absorption energy.

Results and Discussion

The condensed sp^2 carbons in GR_10, GR_18, GR_22, GR_27 (showed in green in figure 1a) were obtained by removing specific hydrogen atoms from GR_0 that is a structure with hydrogens and carbons arranged like a graphane structure and showed in figure 1b. The distance distribution function revealed two aromatic and aliphatic components of GR_10, centered around 1.410 Å and 1.530 Å

respectively, where the former is closer to the typical sp^2 carbon-carbon bonds in aromatic compounds or for typical C-C bond in graphene layer (1.420 Å). So that in GR_10 It is reasonable to attribute the C-C bond for the atoms C12, C18, C19, C48, C49, C55 an aromatic nature of the sp^2 bonds.

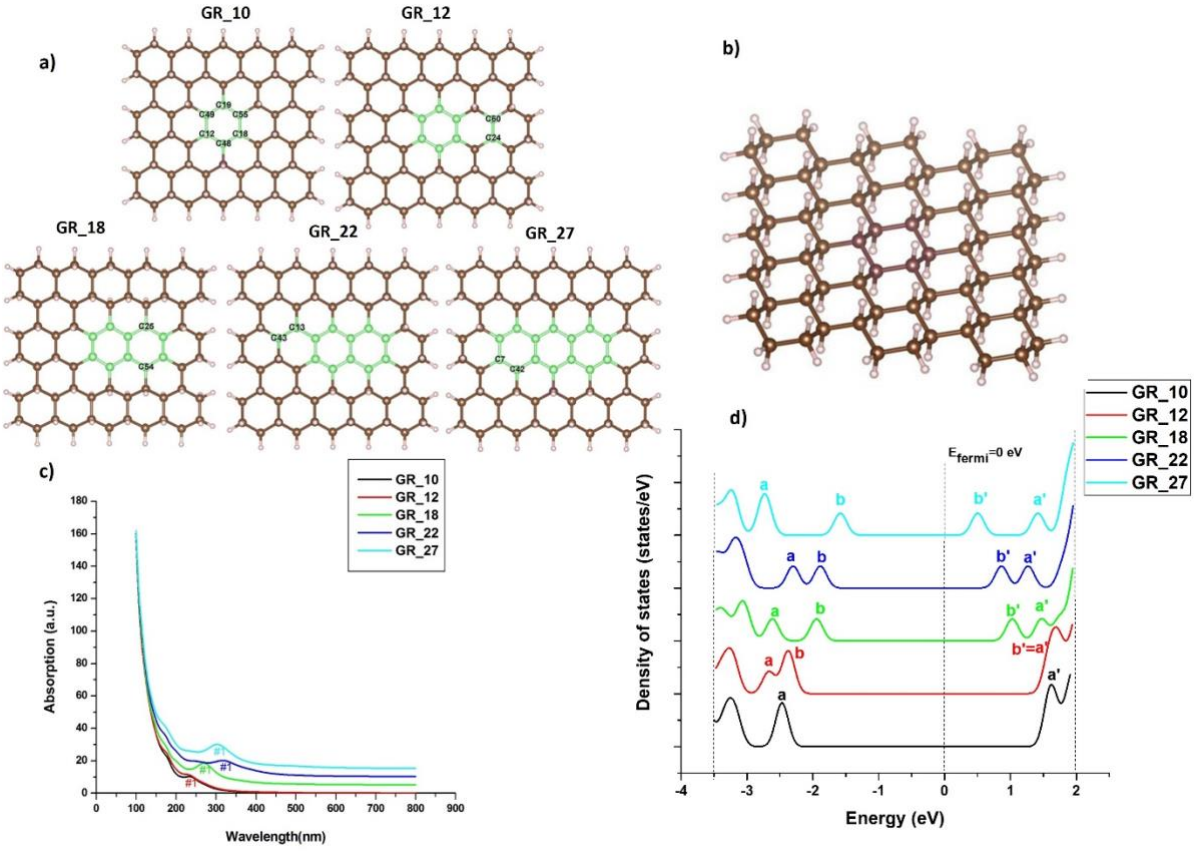


Fig. 1. Figure shows the arrangements of the atoms in GR_10, GR_12, GR_18, GR_22, GR_27 (a), GR_0 (b), absorption spectra (d) and configuration electronic around Fermi (c).

Generally, it can observe as the effect of the aromatic component of GR_10 yields a further symmetry broken of the sp^3 structure activating an optical transition around 237 nm (assigned as aromatic π→π*) and one around 178 nm which is reasonable due again to the transition between electronic states π→π*. The transition revealed around 178 nm is quite close to the transition $A_{1g} \rightarrow E_{1g}$ (experimentally recorded around 180 nm) of the benzene as well as the transition around 237 nm is included in the transition $A_{1g} \rightarrow B_{2u}$. The introduction of further insaturations beyond GR_10 (GR_12 GR_18 and GR_22) lead a read shift of the aromatic peaks while the transition around 178 nm remained unchanged. For GR_18 the peak #1 (in green) shifted from 237 nm up to 272 nm. However, in all these cases reported (also for GR_22), the maximum absorption peaks suffer a red shift except for GR_27. The intermediate structures GR_12 and GR_22 lack of complete and closed sp^2 structures meanwhile the sp^2 network in GR_27 has the same arrangement of the atoms in the anthracene. Roughly speaking, a possible explanation can be found by plotting TDOS of GR_10, GR_18, GR_22 and GR_27 (figure 1d). In GR_10, around the Fermi energy, only one peak (named *a* in black) has been detected. This peak corresponds to the convolution of the highest occupied orbitals with similar energy as well as a single peak LUMOs (named a') above the Fermi energy was present. The optical transition is of the type *a -> a'*. By increasing the ratio sp^2/sp^3

of carbons (GR_12), we noted an initial separation of the single peak in two components labeled as *a* and *b* with a slight shift towards higher energy of the peak labeled *b*. In GR_18 two well distinct states *a* and *b* in green were present with b shifted up to higher energy. Finally in GR_27, we noted again a separation between *a* and *b* peaks (in cyan). The peaks *a* in cyan (GR_27) is shifted to lower energy than the peak *a* in green belonging to GR_18. Similar behavior was found for LUMO orbitals. Hence, the adding of adjacent double bond increases the number electronic states with π symmetry around of the Fermi level, producing a bathochromic effect on the optical spectra, reducing the energy gap. For GR_27 and GR_22 we noted a similar behavior that is an enlargement between the peaks *a* and *b* (and the corresponding a' and b') increasing the gap between occupied and virtual electronic states. In particular, for GR_27 the peak labeled as #1 in absorption spectrum suffers an evident blue-shift with respect to the analogue transition in GR_22, spectra reported in figure 1c. If the electronic transition *a* -> *b'* occurs between two states having maximum reduction of the electronic gap then maximum redshift was obtained for the peak labeled #1 in GR_22 (line blue). On the contrary, for GR_27 we noted an opposite behavior that is an enlargement of the gap between of the peaks *a* and *b'*, then this peak suffers a blue-shift with respect to the analogue transition in GR_22, blue-shift recorded for spectrum reported in figure 1c.

GR_57 (figure 2a) is a structure where the carbons with sp^2 hybridization is connected like in a coronene arrangement. GR_38 and GR_50 are intermediate structures with an open sp^2 network whose atomic arrangement is not complete as in an aromatic cyclic structure.

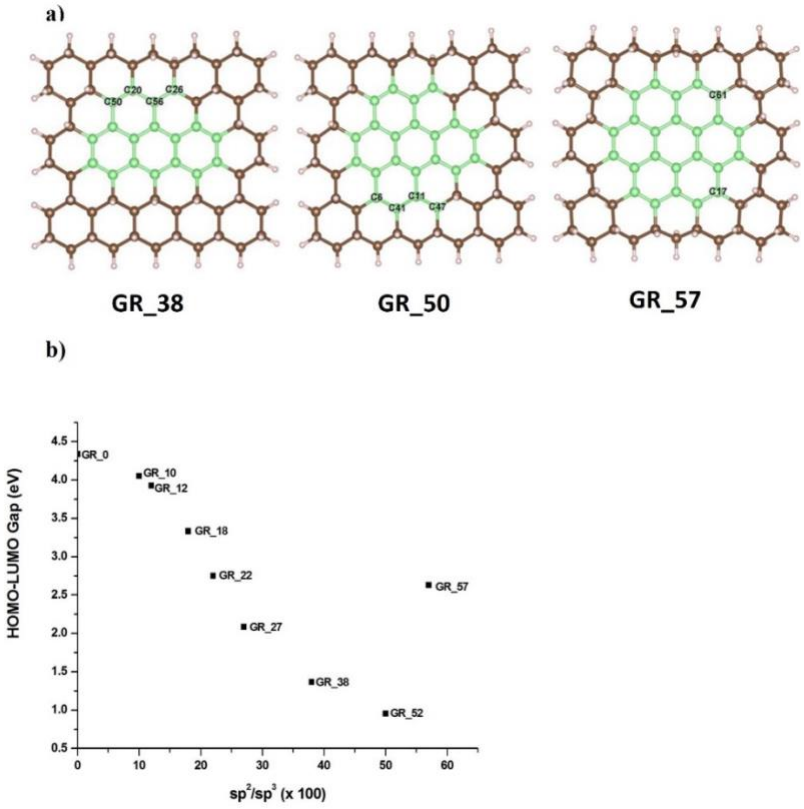


Fig. 2. Figure 2 shows the arrangements of the atoms in the structures GR_38, GR_12, GR_50, GR_57 (a). In (b) the energy gap values were reported in function of the sp^2/sp^3 fraction.

Energy gap in the molecular structure can be used like an empirical index to measure their stability. A sinusoidal decreasing behavior reveals a reduction of the gap typical of polyenes substantially due to the increase of the electronic states with π symmetry around the gap. In this framework the energy gap was calculated between HOMO and LUMO orbitals that in figure 1d are labeled as a and a' for GR_10 and b and b' for the rest of the structures. Hence also for the GR_18 and GR_27 the energy gap decreases. However, it is worth noting that outstanding discontinuity of the curve was detected in GR_57. In this case the effect of the superaromaticity (for example realized through a sextet migration [2]) on the trend in figure 2b and on the stability of these structures is predominant.

Acknowledgments

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [4].
CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see http://www.cresco.enea.it/english for information.

References

[1] M. L. Protopapa, E. Burresi, M. Palmisano, E. Pesce, L. Latterini, N. Taurisano, G.

Quaglia, R. Mazzaro, V. Morandi, Changing the Microstructural and Chemical Properties of

Graphene Oxide Through a Chemical Route, Applied Spectroscopy **76**, pp. 1452-1464, (2022). https://doi.org/10.1177/00037028221127048.

[2] E. Burresi, M.L. Protopapa, Clar's aromaticity in mixed sp^2/sp^3 structures in High Performance Computing on CRESCO Infrastructure: research activity and results 2020, pp 224-227, (2021). ISBN: 978-88-8286-429-3.

[3] P. Giannozzi *et al.* QUANTUM ESPRESSO: a modular and open-source software project for

quantum simulations of materials. J.Phys.:Condens. Matter **21**, 395502, (2009).

[4] F. Iannone *et al.*, "CRESCO ENEA HPC clusters: a working example of a multifabric GPFS

Spectrum Scale layout," *2019 International Conference on High Performance Computing & Simulation (HPCS)*, Dublin, Ireland, pp. 1051-1052, (2019). Doi:10.1109/HPCS48598.2019.9188135.

MACHINE LEARNING FOR THE STUDY OF CATHODE MATERIALS

Marco Catillo[1*], Serena D'Onofrio[2], Francesco Buonocore[1], Simone Giusepponi[1] and Massimo Celino[1]

[1]*Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA)-C.R. Casaccia, Via Anguillarese 301, 00123 Rome, Italy*[6]

[2] *Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA)- C.R. Bologna, Via dei Mille, 21, 40121 Bologna, Italy*

ABSTRACT. In this paper we present an application of the crystal graph neural network GeoCGNN for the prediction of the formation energy per atom and redox potential. These two quantities are fundamental for accelerating the discovery of new cathode materials for batteries, which need to be both stable and efficient in the recharging phase. This is also an important step to researching more environmentally friendly alternatives to lithium-ion batteries.

Introduction

Research into new alternatives to lithium-ion cathode material batteries is becoming extremely important nowadays for enhancing stability, efficiency, and safety. In this direction, one of the main aims of material scientists and chemists is to explore among all material/crystal structures to find cathode materials which allow batteries to meet many different needs [1]. For instance, being able to store large amount of energy; being fast in the recharging phase; lasting longer; being stable; being made by materials abundant on earth and in general with low cost for extraction and manipulation; and most importantly, being environmentally sustainable in the broadest meaning that this word can have. Trying to accomplish all these requirements as well as understanding what weight to give to each of them can be extremely hard to tackle. However, some of them are related to physical quantities that can be straightforwardly computed on HPC systems, which allows computational scientists and theorists to give important hints to experimentalists about which cathode materials are interesting to synthetize. In this regard, we study here the formation energy per atom (E_f), and redox potential (V_{redox}). Specifically, it is desirable to have cathode materials with low E_f and high V_{redox}. The first requirement is a prerequisite for stability while the second makes the battery more efficient and faster in its recharging phase. Up to now, computational methods based on density functional theory (DFT) [2] have been widely employed for studying E_f and V_{redox}. However, when the compositional space of materials is too large, DFT becomes computationally expensive and very time-consuming. In fact, on one side DFT requires several iterative steps to optimize crystal geometry and obtain an accurate description of the material, while on the other side, the computational limits make it difficult to use for large scale analysis. To achieve this task, many machine learning (ML) methods based on graph neural networks (GNNs) have been proposed in the past few years [3-7]. They can learn physical properties from large material databases during the training phase and predict the values of such properties when new material structures are given to them, during the inference phase. The advantage of ML is that once the training is performed, we can predict physical quantities in about a few seconds for thousands of

6 Corresponding author. E-mail: marco.catillo@enea.it.

crystals using GPUs. Conversely, DFT methods can take many core/hours over CPUs, and many attempts to move DFT algorithms on GPUs [8-10] are still far from achieving the same speed of many ML methods.

Methods

In this paper we make use of GeoCGNN as ML method alternative to DFT. GeoCGNN is a geometric enhanced convolutional graph neural network introduced for predicting crystal properties [11]. Its structure is the following. It takes as input a file called CIF, containing information about the atomic positions and cell parameters of the crystal and then it generates a multi-graph considering the information of the topological and geometric structure of crystals, periodicity, lattice vector, and cell volume. The multi-graph consists of nodes representing atoms, and a set of edges representing atomic interactions. This network proceeds by updating the node features with a series of gated convolutional blocks (CGN). Each output of these blocks goes into a pooling layer and finally, a multilayer perceptron (MLP) is applied, which returns the prediction of the physical quantity of interest. Even if it sounds simple, this network has several hyperparameters to be adjusted to obtain accurate predictions. Some of them are the dimension of the feature vectors for the nodes of the multi-graph, the number of CGN blocks and MLP layers, and other important parameters related to the training strategy adopted, such as learning rate and batch size. Once the optimal values of the hyperparameters are found, we can set them for our network and perform training to find the best values of the learnable parameters. Finally, the network is ready to be used in the inference phase for predicting new crystal materials useful for battery applications. However, the tuning of the hyperparameters can be computationally expensive. Indeed, we need to check several values of them and perform training to optimize the learnable parameters. The optimal values of hyperparameters must be the ones which minimize a given cost function. More details are reported in the next subsection.

Hyperparameter tuning and training

For tuning the hyperparameters, we consider the Material Project (MP) database [12]. It contains more than 150 thousand CIFs of crystals associated with the formation energy value and more than 4 thousand couples of discharged and charged cathode materials associated with a redox potential value, both computed via DFT. The first dataset is therefore used for predicting E_f, while the second for V_{redox}. In this latter case, a single instance consists of a couple of CIFs, respectively, for the discharged and charged crystals involved in the so-called, intercalating reaction happening in the battery. The GeoCGNN network for the redox potential prediction is therefore slightly modified to take two CIFs as input. The optimal hyperparameters for the prediction of E_f has been already found in [11], so we do the tuning only for the GeoCGNN applied on V_{redox}, since its network structure has been modified.
For this purpose, we split the cathode material dataset into 20% for testing and 80% for training and validation. This latter is used for tuning. We also employ a 15-fold cross validation technique, which means we do fifteen training optimizations, each time choosing a different validation set. For each training process we use the Adam optimizer and mean squared error (MSE) as the loss function to minimize. Details of this analysis can be found in Ref. [13].

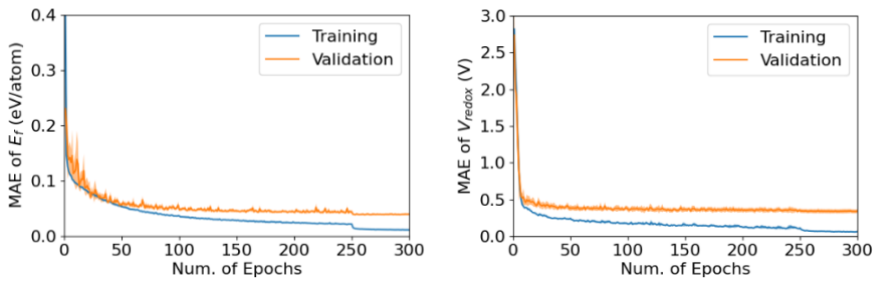


Fig.1: Training for the formation energy per atom (left), and redox potential (right).

On Fig. 1, we show the training for E_f and V_{redox} where the shaded area comes from the error calculated with the 15-fold cross validation. Also, for E_f, we have split the dataset as 80% for training and validation, and 20% for testing. In both cases, the network learns very well from the data with no sign of over- or under-fitting. The small error bars of Fig. 1 indicate a small variability by the training sample used.

Computational details

We tested the GeoCGNN code on the CRESCO ENEA infrastructure, to check its performance over different GPU models available in it: V100 accessible on the cluster CRESCO6; RTX6000 on CRESCOB, A100 and H100 on CRESCO5F (currently dismissed). In Fig. 2 (left), we report the speed-up over different GPUs for the training of E_f, that we define as $s_i = t_{RTX6000}/t_i$, with t_i the average time spent on each training step for the GPU i. We find H100 to be the most performing one.

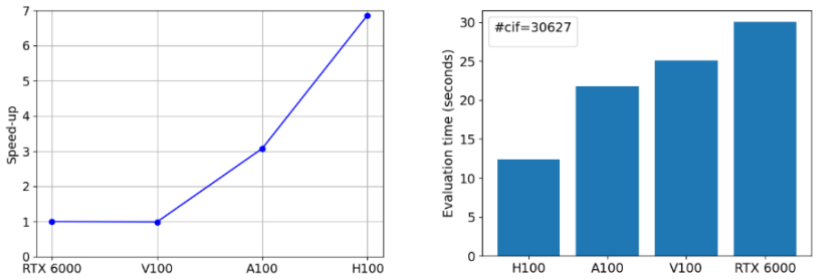


Fig.2: Speed-up on the training steps for formation energy per atom (left). Time spent for evaluating the formation energy per atom of 30627 crystal structures (right).

The right plot of Fig. 2 shows the time spent in the evaluation of E_f for 30627 crystal structures. It is evident that a single crystal structure evaluation takes not more than $10^{-3}s$ in the worst case (RTX6000) and around $4 \times 10^{-4}s$ in the best case (H100). We have obtained similar and consistent results also in the V_{redox} study. Upon this benchmark, we have run our training, finetuning and testing calculations on the cluster XCRESCO, which is the largest cluster of GPUs in CRESCO. It contains sixty nodes with 4 x NVIDIA Volta V100 GPUs/node, each one with Nvlink 2.0 connection and 16GB of memory. The large availability of GPUs of this cluster has been fundamental for launching in parallel many training calculations for the k-fold cross validation and over many different sets of hyperparameters. Regarding V_{redox}, the time taken for training and a single cathode material evaluation has been on average $55\ m$ and $1.5 \times 10^{-3}s$ respectively; while for E_f, we got $7\ h$ and $10^{-3}s$, respectively for training and single crystal evaluation. To give a comparison with DFT, traditionally used for such computations, we report in Fig. 3 the time distribution of 765 crystal structure relaxations with Quantum Espresso [14,15].

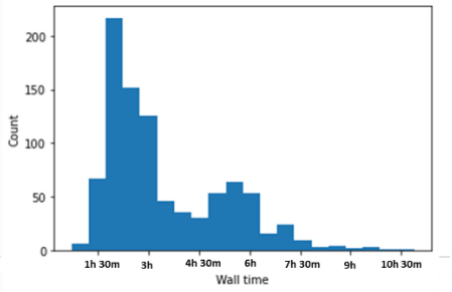


Fig.3: Distribution relaxation times of 765 random crystal structures with Quantum Espresso [14,15].

Except for the training phase, which must be done only once, the predictions of physical quantities take less than a second, while using DFT methods the study of each crystal structure can take several hours. This makes GeoCGNN extremely advantageous compared to DFT.

Testing and conclusion

The remaining 20% of the MP dataset of all crystals and cathode materials is used for testing the prediction capabilities of GeoCGNN of E_f and V_{redox} . In Fig. 4, we show the plot of the predicted values of E_f and V_{redox} with respect to the target quantities given in the MP database, together with the MAE, averaged over 15 training optimizations according to the 15-fold cross validation. The small MAE for both these quantities indicates good agreement with DFT calculations and a good generalization of the network. In conclusion, we have seen how GeoCGNN is fast and accurate in the prediction of physical quantities of interest for cathode material search. The training of this network is therefore a first step through a large screening of many unexplored crystals to find cathode materials which can be good candidates for new energy storage technologies, conceivably better than the current widely used lithium-ion batteries.

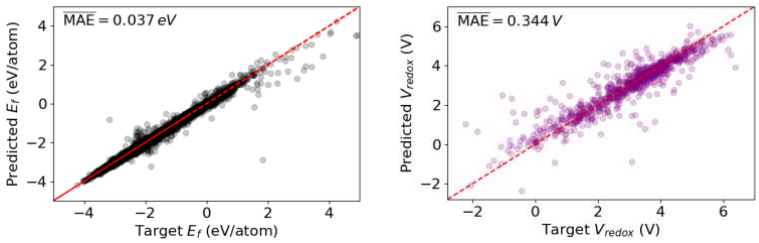


Fig.4: Formation energy per atom (left) and redox potential (right) GeoCGNN prediction compared to DFT calculated values in Materials Project [12]. The dotted red line is the optimal case $y = x$.

Acknowledgments

This work was supported by the European Union through the Next Generation EU funds through the Italian MUR National Recovery and Resilience Plan, Mission 4 Component 2 Investment 1.4, National Center for HPC, Big Data and Quantum Computing, CUP I33C22001270007. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID HPC infrastructure and its staff [16]. CRESCO/ENEAGRID HPC infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programs.

References

[1] M. S. Whittingham. Lithium batteries and cathode materials. *Chemical reviews* **104.10**, pp. 4271-4302, (2004).

[2] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Physical review*, **140(4A)**, pp. A1133, (1965).

[3] T. Xie, and J. C. Grossman. Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties. *Physical review letters*, **120(14)**, pp. 145301, (2018).

[4] C. Chen, W. Ye, Y. Zuo, C. Zheng, and S.P. Ong. Graph networks as a universal machine learning framework for molecules and crystals. *Chemistry of Materials*, **31(9)**, pp. 3564-3572. (2019).

[5] S.Y. Louis, Y. Zhao, A. Nasiri, X. Wang, Y. Song, F. Liu, and J. Hu. Graph convolutional neural networks with global attention for improved materials property prediction. *Physical Chemistry Chemical Physics*, **22(32)**, pp. 18141-18148, (2020).

[6] C.W. Park, and C. Wolverton. Developing an improved crystal graph convolutional neural network framework for accelerated materials discovery. *Physical Review Materials*, **4(6)**, pp. 063801, (2020).

[7] K. Meng, K. Bai, and S. Sun. Artificial intelligence driven design of cathode materials for sodium-ion batteries using graph deep learning method. *Journal of Energy Storage*, **101**, pp. 113809, (2024).
[8] I. Carnimeo, et al. Quantum ESPRESSO: One further step toward the exascale. *Journal of Chemical Theory and Computation* **19.20**, pp. 6992-7006, (2023).
[9] M. Hacene, et al. Accelerating VASP electronic structure calculations using graphic processing units. *Journal of computational chemistry*, **33.32**, pp. 2581-2589 (2012).
[10] D. Sangalli, et al. Many-body perturbation theory calculations using the yambo code. *Journal of Physics: Condensed Matter*, **31.32** pp. 325902 (2019).
[11] J. Cheng, C. Zhang, and L. Dong. A geometric-information-enhanced crystal graph network for predicting properties of materials. *Communications Materials,* **2.1**, pp. 92, (2021).
[12] A. Jain, et al. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL materials* **1.1**, (2013).
[13] M. Catillo, F. Buonocore, S. D'Onofrio, S. Giusepponi, S. Ferlito et al. *In preparation*.
[14] P. Giannozzi et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *Journal of physics: Condensed matter* **21.39**, pp. 395502, (2009).
[15] P. Giannozzi et al. Advanced capabilities for materials modelling with Quantum ESPRESSO. *Journal of physics: Condensed matter*, **29.46**, pp. 465901 (2017).
[16] F. Iannone, et al. CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. In *2019 International Conference on High Performance Computing & Simulation (HPCS),* pp. 1051-1052, IEEE, (2019).

CLIMATE PROJECTIONS OF MEDITERRANEAN CYCLONES AND ASSOCIATED PRECIPITATION WITH A REGIONAL EARTH SYSTEM MODEL UNDER THREE SSP SCENARIOS

Marco Chericoni[123*], Giorgia Fosser[1], Alessandro Anav[23]

[1]*Scuola Universitaria di Studi Superiori IUSS Pavia, 27100, Italy*

[2]*Agenzia nazionale per le nuove tecnologie, l'energia e lo sviluppo economico sostenibile ENEA, 00196, Roma, Italy*

[3]*ICSC Italian Research Center on High-Performance Computing, Big Data and Quantum Computing, Bologna, Italy*

[1] Corresponding author. E-mail: marco.chericoni@iusspavia.it.

ABSTRACT. This In recent years, the Mediterranean region has experienced intense cyclones with heavy precipitation, resulting in severe flooding with multiple fatalities and significant damage to infrastructures. Additionally, the basin is increasingly recognized as a climate change hotspot. For these reasons, it is fundamental to understand how Mediterranean cyclones respond to climate change, identifying the key processes driving these changes.

Our study is the first combining CMIP6 models with a high-resolution atmosphere-ocean coupled regional climate model (ENEA-REG) over the Mediterranean to investigate changes of intense cyclones under three SSP scenarios (SSP5-8.5, SSP2-4.5, SSP1-2.6). Despite a strong reduction in the number of intense cyclones and a subsequent decline in seasonal precipitation, our results underscore a significant increase in cyclone-related extreme precipitation, especially under the SSP5-8.5 scenario.

These results offer valuable information for regional climate impact assessments in the Mediterranean basin and highlight the importance of combing different models within coordinated frameworks to disentangle the influences of large-scale forcings and regional climate processes on the future Mediterranean climate under varying radiative forcing levels.

Introduction

Recent floods in southern Europe, primarily driven by extratropical cyclones and intensified by climate change, have had severe socioeconomic consequences. Notable episodes include the floods in central Italy in 2023 and in Valencia in 2024, both resulted in large economic losses, societal disruption, and fatalities. Additionally, the Mediterranean basin is increasingly recognized as a climate change hotspot [1]. For these reasons, understanding the impact of climate change on the interplay between cyclones and precipitation is more critical than ever.

Cyclogenesis in the Mediterranean region is driven by the intrusion of upper-tropospheric systems, caused by Rossby wave breaking over the Atlantic Ocean [2]. Cyclone development then depends on

both large-scale baroclinic forcing and deep convection near the cyclone centre, which generates potential vorticity critical for its intensification. Maximum intensity, denoted by the lowest sea level pressure (SLP), tend to occur over maritime areas, in the Gulf of Genoa and in the Adriatic, as well as in the Ionian and Aegean Seas. Global warming is expected to increase atmospheric moisture and potential energy available for convection at the surface, fostering moisture convergence, uplift and thus the precipitable water during intense cyclones. However, changes in baroclinic forcing can either amplify or inhibit these processes, highlighting the complex interplay of factors influencing precipitation response to climate change in the Mediterranean region.

Previous studies based on CMIP5 [3] highlight a climate paradox for future Mediterranean precipitation: while mean precipitation is expected to decline, leading to aridification and water scarcity due to reduced cyclonic activity, cyclone-related precipitation is projected to increase in the north-central Mediterranean but decrease in the southeast. However, there is a pronounced inter-model spread regarding these changes, and the underlying mechanisms are not yet fully understood.

This study uses both global (CMIP6, Coupled Model Intercomparison Project Phase 6) and regional (ENEA-REG [4], Table 1) climate projections to evaluate the changes in Mediterranean cyclones characteristics and associated precipitation under three Shared Socioeconomic Pathways (SSP5-8.5, SSP2-4.5, and SSP1-2.6) [5].

Methodology

One high-resolution atmosphere-ocean coupled regional climate model (ENEA-REG [4]) and six CMIP6 global climate models (Table 1) are used in this study. Four different simulation periods are analysed, i.e. one historical (1982-2014) and three future scenarios (2015-2100): SSP5-8.5, SSP2-4.5 and SSP1-2.6.

To identify Mediterranean cyclones, a cyclone tracking algorithm [6] is applied to ENEA-REG, CMIP6 and CMIP6-HR simulations, as well as to ERA5 reanalysis, over the Med-CORDEX domain (Fig. 1). The tracking method uses Mean Sea Level Pressure (MSLP), at 6 hourly intervals, to detect cyclone centres. Only cyclones with a minimum SLP below 1000 hPa and located within the area outlined by solid lines in figure 1 are classified as “intense Mediterranean cyclones” and included in this study. This allows to remove small cyclonic features that typically have limited influence on climate dynamics and extremes of the Mediterranean area.

The same methodology has been already applied in the hindcast (ERA5 driven) simulation of ENEA-REG [7], which highlighted how the model accurately represents the cyclone track statistics, i.e. intensity, lifetime and speed, as well as season cycle and spatial distribution.

Table 1: List of analysed regional and global climate models.

Type of model	Model name	Institution	Atmospheric horizontal resolution
Regional climate model	ENEA-REG	ENEA, Italy	12 km
CMIP6	BCC-CSM2-MR	BCC, China	100 km
CMIP6	CMCC-ESM2	CMCC, Italy	100 km
CMIP6	EC-Earth3	European Consortium	80 km
CMIP6	MPI-ESM1.2-HR	MPI, Germany	80 km
CMIP6	MRI-ESM-2.0	MRI, Japan	100 km
CMIP6	NorESM2-MM	NCC, Norway	100 km

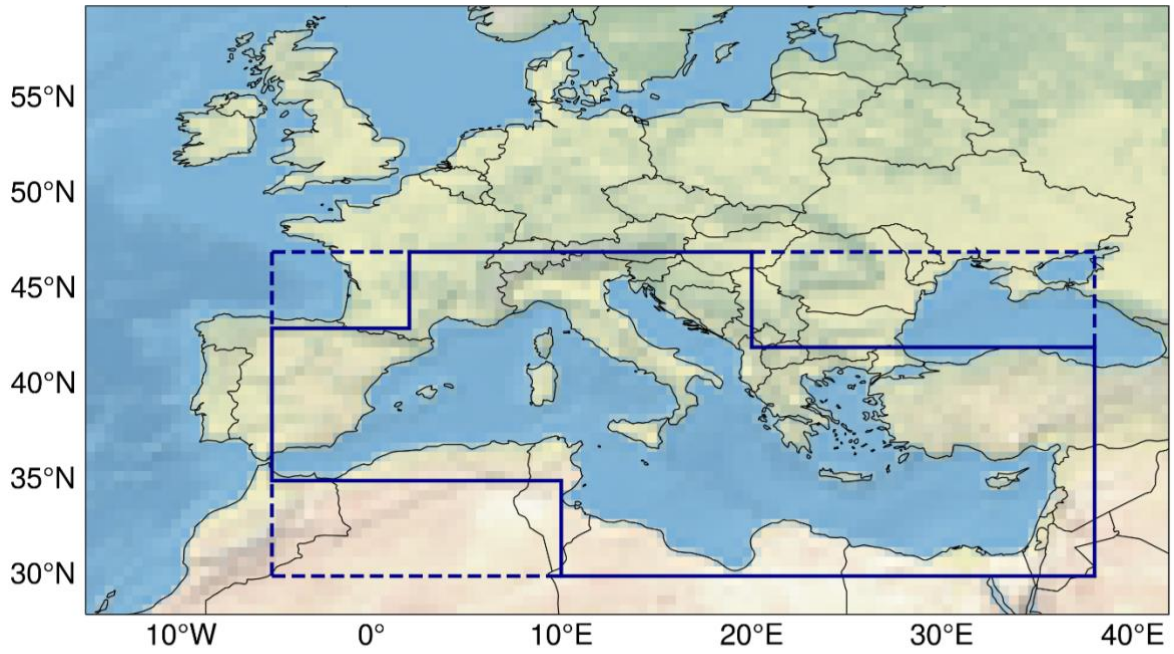


Fig.1: Med-CORDEX domain. Cyclones are retained if their minimum SLP tracking point is present within the area outlined by solid lines.

Future Mediterranean precipitation paradox

During the extended cold season, both ENEA-REG and the ensemble of CMIP6 simulate a significant decrease in the number of intense Mediterranean cyclones (34-39% under SSP5-8.5; Fig. 2a) and in seasonal precipitation (12-14% under SSP5-8.5; Fig. 2b) with increasing greenhouse gas (GHG) emissions across the three analysed SSP scenarios. In clear contrast, extreme (99th percentile) precipitation associated with these cyclones (Fig. 2c) is projected to increase by up to ~19% by the end of the century under the high-emission scenario (SSP5-8.5). This intensification aligns with radiative forcing trends under the various emission scenarios (inset Fig. 2a), as increased energy in the atmosphere fuels these cyclonic events. Notably, ENEA-REG simulates more intense cyclones and stronger extreme precipitation than the CMIP6 ensemble, benefiting from its finer resolution.

Despite the large spread in future projections, reflected by the standard deviation among CMIP6 models (shaded areas in Fig. 2), our analysis highlights the important role of Mediterranean cyclones on the extreme precipitation trends and reinforces the climate paradox for the Mediterranean basin: if GHG emissions continue to rise, the region will experience lower seasonal precipitation due to reduced cyclonic activity, while even more extreme precipitation during intense cyclones. The combination of increased aridity and amplified precipitation extremes is expected to heighten the risk of inland flooding from these storms, posing substantial challenges for the densely populated areas of the Mediterranean.

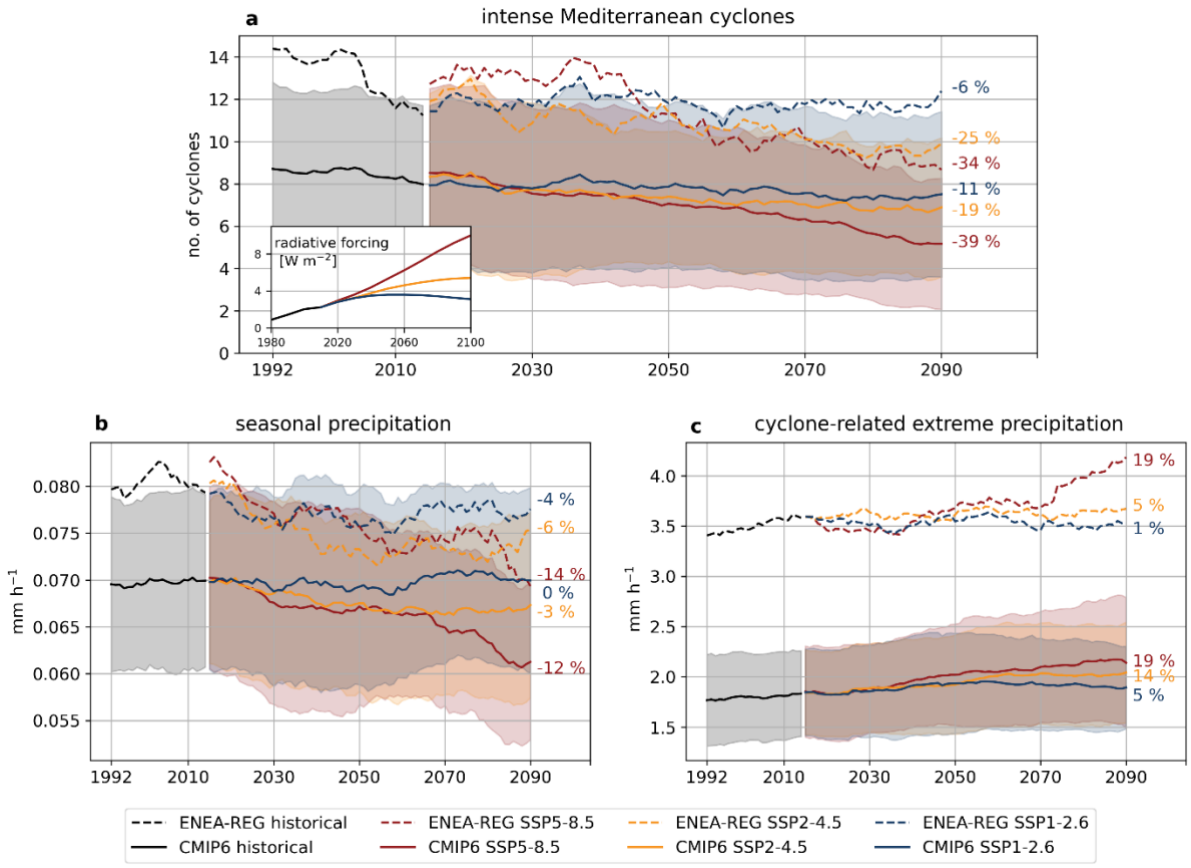


Fig.2: Temporal evolution of intense Mediterranean cyclones and associated extreme precipitation. Time series for (a) the number of intense Mediterranean cyclones, (b) seasonal Mediterranean precipitation (averaged over the Mediterranean domain outlined by dashed lines in Figure S1) and (c) the extreme (99th percentile) precipitation associated with intense Mediterranean cyclones (averaged over the number of cyclones each year). All time series are smoothed using a 20-year moving average. Results are shown for the ONDJFM cold season in ENEA-REG (dashed lines) and the CMIP6 ensemble (solid lines), covering the historical period (black lines) and future scenarios SSP 5-8.5 (red), SSP2-4.5 (orange) and SSP1-2.6 (blue). The colour bands indicate the standard deviation for the CMIP6 ensemble. The percentages at the end of the time series show the changes of the end of the century compared to the historical mean. Panel (a) includes an inset showing the evolution of anthropogenic radiative forcing under different SSP scenarios.

Acknowledgments

We acknowledge the World Climate Research Programme, which, through its Working Group on Coupled Modelling, coordinated and promoted CMIP6. Within this we thank the CMIP6 endorsement of the High-Resolution Model Intercomparison Project (HighResMIP) and Martin Schupfner for providing additional data from the MPI-ESM. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEA-GRID High Performance Computing infrastructure and its staff.

References

[1] Giorgi, F. Climate change hot-spots. *Geophys Res Lett* **33**, (2006).
[2] Flaounas, E. et al. Mediterranean cyclones: current knowledge and open questions on dynamics, prediction, climatology and impacts. *Weather and Climate Dynamics* vol. **3** 173–208 https://doi.org/10.5194/wcd-3-173-2022 (2022).
[3] Reale, M. et al. Future projections of Mediterranean cyclone characteristics using the Med-CORDEX ensemble of coupled regional climate system models. *Clim Dyn* **58**, 2501–2524 (2022).

[4] Anav, A. et al. Dynamical downscaling of CMIP6 scenarios with ENEA-REG: an impact-oriented application for the Med-CORDEX region. *Clim Dyn* doi:10.1007/s00382-023-07064-3 (2024).
[5] Chericoni, M. et al. Unravelling drivers of the future Mediterranean precipitation paradox during cyclones. *npj Clim Atmos Sci* **8**, 260. https://doi.org/10.1038/s41612-025-01121-w (2025).
[6] Flaounas, E. et al. CycloTRACK (v1.0) – tracking winter extratropical cyclones based on relative vorticity: sensitivity to data filtering and other relevant parameters. *Geosci Model Dev* **7**, 1841–1853 (2014).
[7] Chericoni, M. et al. Extreme Mediterranean cyclones and associated variables in an atmosphere-only vs. an ocean-coupled regional model, *Weather Clim. Dynam.*, **6**, 627–643, https://doi.org/10.5194/wcd-6-627-2025, (2025).

NUMERICAL INVESTIGATION OF RICH-MILD-QUENCH-LEAN AMMONIA COMBUSTION IN A DUAL-STAGE BURNER WITH CYCLONIC AND QUAD JICF CONFIGURATION

Matteo Cimini[1*], Donato Cecere[1] and Eugenio Giacomazzi[1]

[1]*Laboratory of Sustainable Combustion and Advanced Thermal and Thermodynamic Cycles, ENEA, 00124 Rome, Italy*

ABSTRACT. This study numerically investigates a dual-stage NH3/H2 burner system that integrates a cyclonic Moderate or Intense Low-oxygen Dilution (MILD) combustor in the primary stage with a Jet-In-Crossflow (JICF) arrangement in the secondary stage. Operating under a Rich-Quench-Lean (RQL) strategy, the burner facilitates ammonia pyrolysis and hydrogen production in the rich primary stage, followed by lean, diffusion combustion in the secondary stage to achieve low NOx emissions while consuming the unburned fuel. The results demonstrate the effectiveness of the proposed design in achieving stable, low-emission operation, providing valuable insights for the development of advanced ammonia-hydrogen combustion systems.

Introduction

The growing need for efficient and clean-burning solutions has driven the evolution of modern combustion systems. Among the various approaches, NH3/H2 blends have been identified as a potential high-efficiency alternative to hydrocarbon fuels for gas turbine engines [1]. Compared to hydrogen, ammonia offers significant advantages [2], and, in many cases, it is feasible to consider generating the necessary H2 for co-firing processes by locally decomposing a part of the available NH3 supply. Recent findings suggest that adopting a two-stage RQL strategy enhances flame stability and lowers emissions in NH3-fired combustors [3,4]. The actual technologies implemented in the first stages use classical premixed jet combustors [4]. An alternative solution is investigated in this research work, where MILD is applied to the rich conditions of a primary-stage cyclonic burner at high pressure (5 bar). Operating under rich MILD conditions, this stage generates incomplete combustion products and NH3 conversion into H2 by maintaining a low temperature of the hot products. These are subsequently burnt in a lean secondary stage that incorporates four distinct JICF arrangements for air dilution. The objectives of this research are twofold. First, with the aid of RANS simulations, we aim to characterize the flow field and combustion behavior in the dual-stage high-pressure burner, with particular emphasis on the transition dynamics between the two stages. Second, we investigate only in the second stage the complex pollutant formation pathways, focusing on NOx by means of DNS.

Computational Setup

In this study, we first perform a set of RANS simulations to investigate different equivalence ratios coupling between the Laboratory Unit Cyclonic (LUCY) burner geometry [5], used as the first stage (ϕ=1.0-1.3), and a square duct with four normal JICF as the second stage (ϕ=0.125-0.3), at 5 bar (see Fig. 1). The LUCY configuration consists of a prismatic chamber of 20×20×5 cm^3. The feeding system is composed of two antisymmetric pairs of inlet jets, while the gas exit is located orthogonally to the inlet

jets and in the centre of the chamber bottom face. The oxidizer (air) and fuel (NH3) injectors consist of cylindrical ducts with internal diameters of 0.8 cm (oxidizer) and 0.15 cm (fuel). The oxidizer flow was preheated to the temperature of 550 K, whereas the fuel stream was fed at 300 K. In the second stage, the square duct (L=2.2 cm) exiting from LUCY is equipped with four slotted rectangular air jets (6×1.1 mm) preheated at the temperature of 750 K. The RANS simulations are carried out using the commercial code ANSYS Fluent. The numerical setup related to the LUCY first stage (not reported here) was validated through comparison with the experimental data available in the literature at one atmosphere [5]. Through this series of RANS simulations, the optimal equivalence ratio coupling between the first and second stages was evaluated to achieve the highest NH3-to-H2 conversion. The equivalence ratios chosen are ϕ=1.025 for the first stage and ϕ=0.125 for the second one.
In the second part of this study, the RANS simulation of the chosen configuration is used to initialize the flow domain for the DNS of the second stage. The final mesh has approximately 1.5B cells. The DNS simulation is conducted through the code HeaRT [3]. A staggered finite-difference scheme is used to solve the compressible, reactive N-S equations, where each chemical species has its individual transport equation. Diffusive fluxes are calculated using a second-order accurate centred scheme, and convective terms are modelled through the AUSM$^+$-up method, coupled with a hybrid third/fifth-order accurate WENO interpolation. The Hirschfelder and Curtiss approximate expression [7] is used to model the molecular transport in the multicomponent mixture. Moreover, the Soret thermo-diffusive effect and pressure gradient diffusion are considered. Kinetic theory is used to compute the thermo-diffusion coefficients and binary diffusion coefficients of the n-th species. Dynamic viscosity and thermal conductivity for individual chemical species are pre-calculated using software libraries [8,9]. Mixture-average properties are estimated using Wilke's formula with Bird's correction for viscosity [10,11], and Mathur's expression for thermal conductivity [10,12]. A sketch of the computational domain used in DNS is shown in Fig. 1.

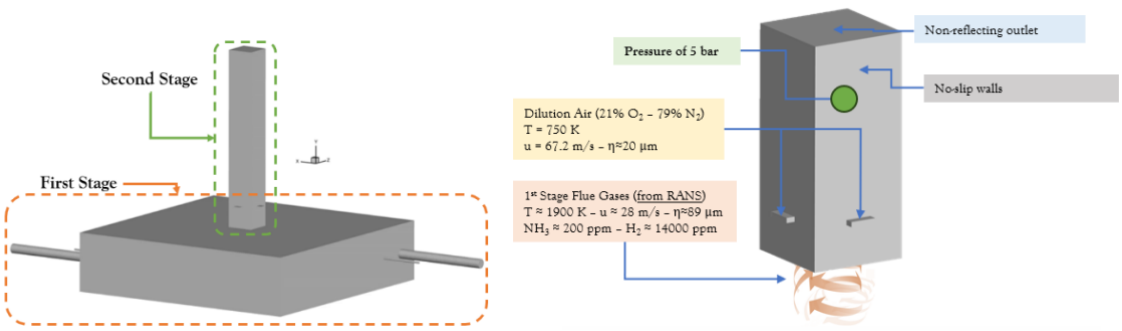


Fig. 1: (Left) Schematic of the computational domain for the RANS simulations with highlights of the first and second stage. (Right) Schematic of the computational domain and boundary conditions for the DNS simulation.

Results

The MILD first-stage cyclonic burner field exhibits an organized flow structure. In Fig. 2, the velocity streamlines illustrate a robust recirculation zone that enhances residence time and stabilizes the combustion process, enabling partial decomposition of the unburnt NH3 into H2. At the exit of the first stage, the hot products contain a mean concentration of unburnt NH3≈200 ppm and H2≈14000 ppm. These high-temperature gases are rapidly diluted with excess air from the second stage to completely oxidize the fuel they contain. Figure 2 also shows an instantaneous three-dimensional visualization of an iso-surface of the Q-criterion coloured by temperature, highlighting the complex turbulent field inside the chamber. As the dilution jets penetrate inside the chamber, Kelvin-Helmholtz hydrodynamic instabilities immediately arise, followed by a rapid breakup of the shear layers that leads to smaller turbulent structures.
Figure 3 depicts the temperature field in the central longitudinal x-y plane. The high-temperature flue gases exiting the first stage undergo a local temperature reduction upon mixing with fresh air; however,

the subsequent combustion of hydrogen results in a temperature rise, with the flow reaching approximately 1950 K at the outlet of the second stage. Figure 3 also shows the NO (dppm) concentration field in the same plane. The NO production is mainly located in the leeward side of each jet, as well as in the windward side and at the centre of the chamber in front of the jets. As the flow develops, the concentration drops, reaching an average value of the order of 200 ppm at 4 cm from the jet injection point (see also Fig. 2).

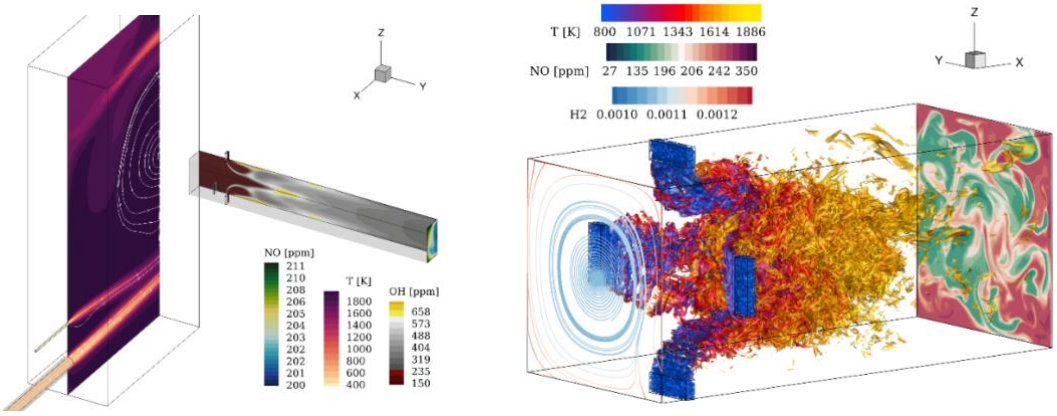


Fig. 2: (Left) RANS results of the burner featuring contours of NO at the outlet section, temperature field at the middle plane of LUCY, and OH at x=0.0 m; (Right) DNS snapshot of Q-criterion iso-surface coloured by temperature, contour of NO in the outlet section, and the stream traces of the swirl-driven flow coming from the first LUCY stage, coloured by Y_{H2}.

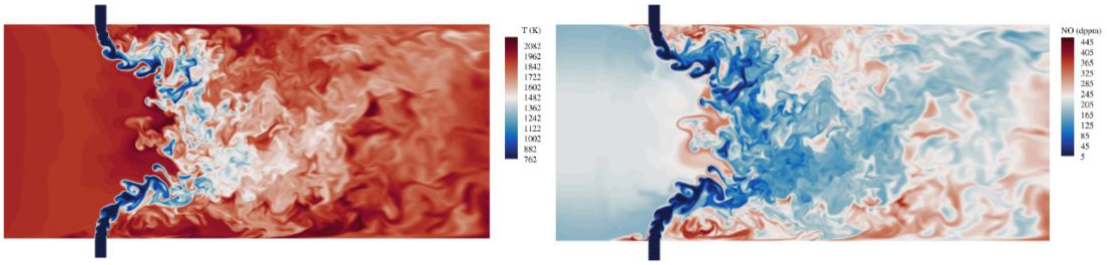


Fig. 3: (Left) Contours of the temperature, and (Right) NO [dppm] in the central longitudinal x–y plane.

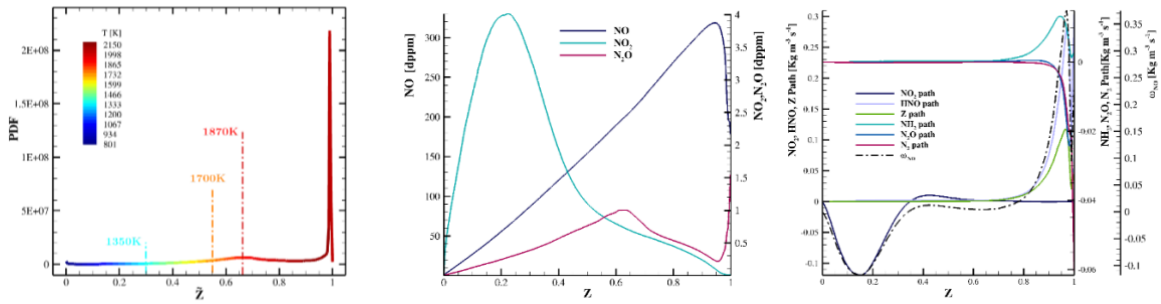


Fig. 4: (Left) Mixture fraction Z PDF coloured by mean temperature. (Centre) Mean NO, NO2, N2O [dry-ppm] as function of the mixture fraction Z, and (Right) NO formation pathways [Kg m^{-3} s^{-1}] as a function of the mixture fraction Z, (dashed-dot line) total NO formation rate.

Figure 4 shows the mixture fraction probability distribution function in the combustion chamber coloured by mean temperature. There are two visible peaks, the more intense is related to the inlet

boundary conditions with a mean inlet temperature of ≈2000 K and a second one at Z=0.67 with a mean temperature of 1870 K. The higher temperatures are located at Z>0.3 so the NO production is in rich combustion regions. The same figure shows the mean NO, NO2, N2O concentration as a function of mixture fraction (a Bilger mixture fraction defined by the elemental mass fractions of H and O is adopted being 0 in the oxidizer and 1 at the inlet fuel mixture coming from the combustor first stage). The NO concentration peak is around Z≈0.9 at rich conditions in contrast to what happens when a fuel jet is injected in air crossflow where the heat release maximum rate is located at low value of the mixture fraction. Figure 4 also shows the total NO formation rate (dash-dot line) and the most relevant NO formation pathways (their sum captures very well the total NO formation rate) as a function of the mixture fraction. NO is mostly formed through HNO and Zeldovich pathways reactions, while it is consumed in lean regions through NO2 pathways. This suggests that it is of paramount importance to rapidly decrease the fuel temperature through mixing with air and thus decrease the effect of Zeldovich mechanism.

Conclusion

The numerical investigations employed in this study, spanning RANS and DNS methodologies, provide critical insights into the fluid dynamics, combustion behavior, and pollutant formation pathways. The first stage produces an average NO value of 150 ppm at the outlet with an average temperature of about 1900 K, but is unable, even under stoichiometric conditions, to burn all the fuel. The second stage burns the residual fuel under globally lean conditions with an average temperature at the outlet of about 2000 K and an average NO value of 200 dppm. Overall, this investigation paves the way for the development of advanced combustion systems that leverage NH3/H2 blends as viable alternatives to conventional hydrocarbon fuels. The dual-stage burner system, with its innovative configuration and operational strategy, offers a promising solution for achieving sustainable and clean energy generation in gas turbine applications. Next steps of this research will focus on the validation of these findings at higher pressures (25 bar) and different operating conditions (equivalence ratios of the first and second stage) to further refine and validate the proposed configurations.

Acknowledgments

The computational resources and associated technical support utilized for this study were made available by the CRESCO/ENEAGRID High-Performance Computing infrastructure and its staff [13]. This study was performed within the Italian project "RICERCA E SVILUPPO DI TECNOLOGIE PER LA FILIERA DELL'IDROGENO POR-H2" ("Research and Development of Technologies for Hydrogen Chain", POR-H2, WP2, LA2.2.5), funded by the European Union (NextGeneraionEU) through the Italian Ministry of Environment and Energy Security, under the National Recovery and Resilience Plan (PNRR, Mission 2, Component 2, Investment 3.5, project code I83C22001170006).

References

[1] N. Yao, et al. In Energy Convers. Manag., 2024.
[2] O. Mathieu and E.L. Petersen. Combust. Flame, 162: 554–570, 2015.
[3] E. Giacomazzi, et al. Energies, 16(23): 7704, 2023.
[4] T. Heggset, et al. Journal of Engineering for Gas Turbines and Power, 146(8), 2024.
[5] G.B. Ariemma, et al. In Proceedings of the combustion Institute, 38(4): 5147-5154, 2021.
[6] Y. Jiang, et al. International Journal of Energy Research, 44(2): 795-810, 2020.
[7] Hirschfelder, J. O. et al. John Wiley & Sons, (1964).
[8] Ern, A. and Giovangigli, V. Springer Science & Business Media, 24 (1994).
[9] Ern, A. and Giovangigli, V. Journal of Computational Physics, 120 (1995), 105-116.
[10] Bird, R. B. Appl. Mech. Rev., 55 (2002), R1-R4.
[11] Wilke, C. R. The Journal of Chemical physics, 18 (1950), 517-519.
[12] Mathur, S., et al. Molecular physics, Taylor & Francis, 12 (1967), 569-579.
[13] Ponti, G. Proc. of the 2014 Int. Conf. on HPC and Simulation, 2014 1030-1033.

NUMERICAL STUDY ON LASER IRRADIATION OF 3D-PRINTED MICRO-STRUCTURES UNDER HIGH-POWER LASER IRRADIATION

M. Cipriani[1], F. Consoli[1], M. Scisciò[1], P. Andreoli[1], G. Cristofari[1], E. Di Ferdinando[1], and G. Di Giorgio[1]

[1]ENEA, Fusion and Technologies for Nuclear Safety and Security Department-C. Frascati, Italy

ABSTRACT Inertial confinement fusion requires a constant search for the most effective materials for improving the efficiency of the compression of the capsule and of the laser-to-target energy transfer. Foams could provide a solution to these problems, but they require further experimental and theoretical investigation. The new 3D-printing technologies, such as the two-photon polymerization, are opening a new era in the production of foams, allowing for the fine control of the material morphology. Detailed studies of their interaction with high-power lasers in regimes relevant for inertial confinement fusion are very few in the literature so far and more investigation is needed. 3D simulations of the laser-target interaction performed with the FLASH code on the CRESCO6 cluster have been used to investigate the hydrodynamics due to laser-matter interaction, obtaining a good agreement with the experimental data collected on the ABC facility.

Introduction

The unique interaction of high-power lasers with micro-structured materials, such as foams, and the evolution of the plasma generated by laser action have proven to be valuable for various applications. These materials have been utilized as efficient sources of X-rays [1], enabling the study of shock waves [2], equations of state [3], and efficient particle acceleration, including neutron and betatron generation with short laser pulses [4]. In particular, their applications in Inertial Confinement Fusion (ICF) are significant due to their ability to enhance laser absorption [5], smooth laser inhomogeneities in the transverse energy profile [6], and enhance the ablation loading on a substrate [7]. Recently, low-density foams have been shown to reduce the SBS instability, which is linked to an increased ratio between the ion and electron temperatures [8, 9]. They are also being suggested to be employed within hohlraums to prevent wall expansion [10, 11] and function as support material for liquid nuclear fuel, known as wetted foams [12, 13]. Recent advancements in manufacturing, simulations, and diagnostics are paving the way for a deeper and more fundamental exploration of the characteristics of the plasma generated by the action of high-power lasers. Until recently, chemical processes were used to manufacture foams of light elements. This involves creating a gel and removing the solvent (see [14] for a review). These foams, called “chemical” foams, are cost-effective and can be produced quickly. However, their structural parameters, such as pore size, solid part thickness, and average density, can only be precisely controlled within specific ranges and may vary from sample to sample due to stochasticity. Additionally, handling these soft and fragile materials during manufacturing and experiments poses challenges. In contrast, Laser Direct Writing (LDW) Multi-Photon 3D Lithography (MP3DL) technology, based on two-photon polymerization, offers a promising solution. This technique enables the production of predictable and reproducible foam samples, ideal for laser-matter experiments. LDW is particularly appealing for rapid prototyping high-resolution complex architectures for experimental research that demands

advanced material engineering. Recent advancements in increasing throughput up to 10^7 voxels/s offer promising results for scaling up production for mass-customization. LDW allows precise sample morphology design, meeting experimental requirements and enhancing performance and diagnostics. Limited high-quality samples still require several hours, resulting in high prices. This technology has a wide margin for new developments to meet the required parameters. Foams with densities of 10 mg/cm^{3}or lower and pore sizes below 1 μm are necessary for
smoothing laser energy deposition or enhancing particle acceleration, as per [4, 6]. In this range, solid parts can only be 10s of nm thick. However, 3D printed filaments can only reach 100 nm, usually set at most at 500 nm for better stability. With this constraint, the average energy density obtainable through 3D printing is approximately 10 mg/cm^3 with a 10 μm separation, as per [15]. Making foams with larger densities above 50 mg/cm^3 and pore sizes of a few microns is less demanding. Various foam morphologies, such as body-centered, octect truss, or Voronoi lattices, can be explored, leading to smoother shock fronts and high-performance implosions. Such high densities, of the order of 100s of mg/cm^3, are required for studying hydrodynamic surrogates of wetted foams [13, 16] or laboratory astrophysics experiments [17].

Simulating the interaction of a high-power laser beam with chemical foams is challenging due to their varying scales. Chemical foams have filament thicknesses from 10s to 100s of nanometers and pore sizes from a few to tens of micrometers, making the structure stochastic. To address this, sub-grid models have been implemented in radiation-hydrodynamic codes, which don't significantly impact computational performance but may lack accuracy. Simulations with a more fundamental description of the laser-matter interaction are preferable. Recent technological advances in target manufacturing, simulation codes, and computational hardware enable this approach. Printed foams can have regular structures that can be replicated in numerical codes, while parallel radiation-hydrodynamics codes and HPC infrastructures allow for more demanding simulations in less time.

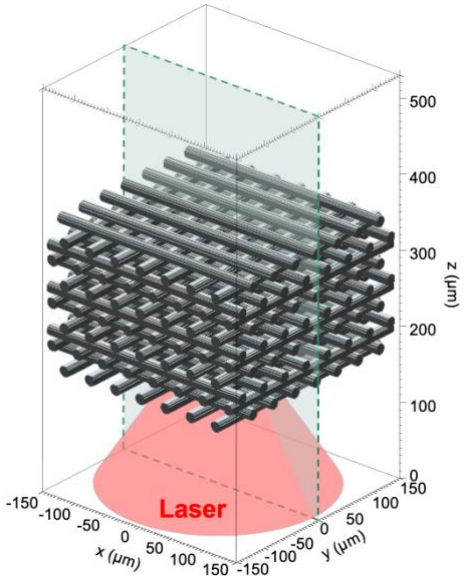


Figure 1: The initial configuration of the simulation. The domain was 300 μm wide along the *x* and *y* axis and 530 μm wide along the *z* axis. The plane along which the domain has been cut to show the results of the simulations presented is shown with a teal fill and a dashed contour.

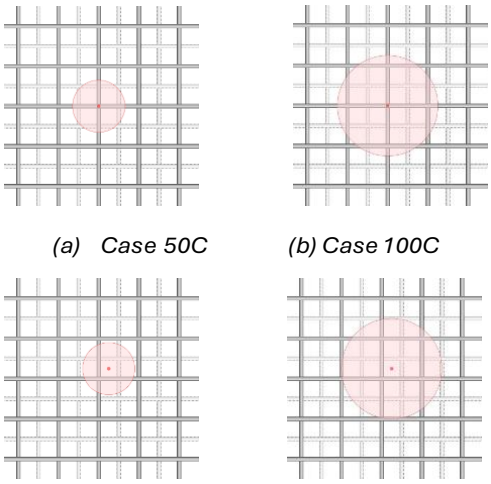


Figure 2: The different cases of focusing. The light red circle indicates the focal spot, while the small red circle indicates its center. The labels correspond to the following cases: 50C, 50 μm diameter spot with the center on the crossing of filaments; 100C, the same as 50C but with a spot diameter of 100 μm; 50H, 50 μm diameter spot with the center on the hole between the filaments of all layers; 100H, the same as 50H but with a spot diameter of 100 μm.

In this work, we present 3D radiation-hydrodynamics simulations with the FLASH code, reproducing the target microstructure using the code's parallelization capabilities on the HPC cluster ENEA-CRESCO6 [18]. The filament thickness and separation in the targets were chosen to reduce computational cost and ease sample production. This lengthened the time to fill gaps, enhancing physical processes like laser scattering due to plasma refraction from filament heating.
The log-pile morphology reduced material fabrication complexity for a better result. Future fusion targets need more random morphologies to reduce asymmetries during compression, especially when the foam constitutes the whole capsule, as in Ref. [13].
Irradiation of regular planar structures is relevant for understanding fundamental processes while keeping manufacturing complexity low and allowing for direct representation of the sample in hydrocodes, as done here.

Simulation results

Before discussing the experimental outcomes, we show the results of the numerical simulations, which clarify the hydrodynamic behavior of the material under irradiation, in particular regarding the degree of homogenization of the plasma. As we will see in Section 2.3, due to the large thickness of the filaments, the plasma does not reach the homogenization. Therefore, we cannot speak about a shock wave being generated in the material and we will refer to the propagation of the front of the region where the micro-structure has been ablated as an *erosion* wave.

Simulation parameters

To simulate the behavior of the material under irradiation we used the FLASH code [19,20]. FLASH is an open-source, modular, multi-physics code in which the plasma fluid equations are solved with finite-volume methods on an Eulerian grid. The code has Adaptive Mesh Refinement (AMR) capabilities, implicit solvers for diffusion, laser ray-tracing, radiation diffusion, and multi-material support with tabulated Equations of State (EoS) and opacities. All the simulations were run on the ENEA CRESCO high performance computing

infrastructure [18]. The large filament thickness and inter-filament spacing allowed to directly reproduce the sample structure in the code as the initial condition and complete the simulation in a reasonable time, therefore avoiding approximations or specific modeling for describing the plasma evolution. The mesh has been chosen to be static rectangular with a cell size of about 1 μm in each spatial direction. Such resolution allowed us to properly resolve the shock dynamics inside the filaments. The 3D ray tracing available in the FLASH code allowed to reproduce the angular incidence of the ABC laser beam on the sample and its scattering on the filaments due to refraction by the plasma at the beginning of the interaction, before the closure of the empty spaces due to the plasma fill. The EoS used for modeling the material under study was the 7592 from the SESAME database [21]. This EoS has been tested in experiments involving TMPTA (trimethylolpropane triacrylate) foam in [16] and a good agreement with the experiments was found. The density of the filaments was set to be 1.2 g/cm^3, in accordance with the density of the printed material.

Laser pointing and laser scattering

The initial setup of the simulation is shown in Figure 1. We used two focal spot diameters. The 50 μm diameter was of the same order of the separation between the filaments in each single layer, while the 100 μm diameter could completely cover a square of four elementary cells. For each focal spot size, we can identify two limit cases in terms of the position of the center of the focal spot, as depicted in Figure 2. Since we used an f /1 lens in the experiments, distinguishing between these two cases has an important relevance, as we will see shortly. We will refer to each of the cases, from here on, with the following labels: 100C,
100 μm diameter spot with the center on the crossing of filaments of the first layer; 50C, the same as 100C but with a spot diameter of 50 μm; 100H, 100 μm diameter spot with the center on the through hole between the filaments of all layers; 50H, the same as 100H but with a spot diameter of 50 μm.

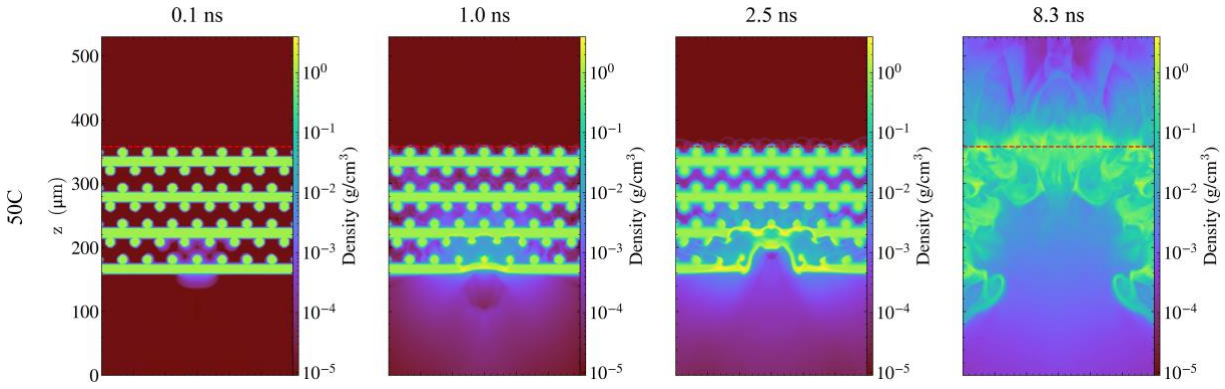


Figure 3: The density plots at different times for the 50C case. The red dashed line indicates the initial position of the back side of the target.

Erosion wave speed

Figure 3 shows the time evolution of the density of the target material and of the laser-generated plasma in the 50C case as an example. We see that after 1 ns the target structure is still almost intact, with shock waves propagating into the filaments of the first two layers. In these case we can also observe that the centers of the filaments of the first two layers are slightly displaced from their original position in correspondence of the laser focal spot. The motion of the heated filaments appears to be symmetric. From the images in Figure 3 we can readily see that during the laser pulse only a partial homogenization of the plasma is achieved.
The breakout time was measured also experimentally, finding a good agreement with the numerical results.

Conclusions and future developments

Understanding the physics underlying the plasma evolution in foams presents a significant challenge, both from a theoretical and experimental perspective. The inherent randomness of the internal structure of chemical foams has hindered our ability to probe the plasma within the material and accurately simulate it for decades.

This work represents one of the few numerical and experimental efforts in the literature [13, 22–24] devoted to the study of the behavior under irradiation of printed foams. The structure of the samples used in this work was conceived to enable 3D simulations with a reasonable computational effort, widening the spatial scales at play. This allowed to explore the evolution of the plasma in detail. As discussed in Section 2, in our conditions no shock wave developed, because the homogenization of the foam was never reached. We rather observed the process of progressive destruction of the material, which we called an erosion wave. The simulations also show that a regular structure with a large separation between the filaments as in our samples leads to a fraction of laser energy being scattered inside the material in the beginning of the interaction. This leads to a peculiar behavior of the plasma, with the formation of zones where the erosion is more efficient, with a less smooth front of the erosion wave. The resolution of the mesh adopted in the simulations and the number of laser rays used in the ray tracing package proved to be adequate for the problem at hand, allowing to properly resolve the main physical processes at play.

Acknowledgements

The software used in this work was developed in part by the DOE NNSA- and DOE Office of Science-supported Flash Center for Computational Science at the University of Chicago and the University of Rochester. The computing resources and the related technical support used for this work have been provided by CRESCO-ENEAGRID High Performance Computing infrastructure and its staff [18]. CRESCO-ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see http://www.cresco.enea.it/english for information.

References

[1] F. P´erez, J. R. Patterson, M. May, J. D. Colvin, M. M. Biener, et al. Bright x-ray sources from laser irradiation of foams with high concentration of Ti. Physics of Plasmas, 21(2):023102, February 2014.

[2] Alessandra Benuzzi, Michel Koenig, Jyothi Krishnan, Bernard Faral, Wigen Nazarov, et al. Dynamics of laser produced shocks in foam–solid targets. Physics of Plasmas, 5(8):2827–2829, August 1998.

[3] Dimitri Batani, Antonio Balducci, Wigen Nazarov, Thorsten L¨ower, Tom Hall, et al. Use of low-density foams as pressure amplifiers in equation-of-state experiments with laser-driven shock waves. Physical Review E, 63(4):046410, March 2001.

[4] O N Rosmej, M Gyrdymov, M M Gu¨nther, N E Andreev, P Tavana, et al. High-current laser-driven beams of relativistic electrons for high energy density research. Plasma Physics and Controlled Fusion, 62(11):115024, November 2020.

[5] M. Cipriani, S. Yu. Gus'kov, F. Consoli, R. De Angelis, A. A. Rupasov, et al. Time-dependent measurement of high-power laser light reflection by low-Z foam plasma. High Power Laser Science and Engineering, 9:e40, 2021.

[6] S. Depierreux, C. Labaune, D. T. Michel, C. Stenz, P. Nicola¨ı, et al. Laser Smoothing and Imprint Reduction with a Foam Layer in the Multikilojoule Regime. Physical Review Letters, 102(19):195005, May 2009.

[7] R. De Angelis, F. Consoli, S. Yu. Gus'kov, A. A. Rupasov, P. Andreoli, et al. Laser-ablated loading of solid target through foams of overcritical density. Physics of Plasmas, 22(7):072701, July 2015.

[8] D. A. Mariscal, O. S. Jones, R. L. Berger, S. Patankar, K. L. Baker, et al. Laser transport and backscatter in low-density SiO2 and Ta2 O5 foams. Physics of Plasmas, 28(1):013106, January 2021.

[9] Lubom´ır Hudec, Jiri Limpouch, Oldrich Renner, Vladimir T Tikhonchuk, Roman Dudzak, et al. Investigation of ion temperature in low-density undercritical foams. Plasma Physics and Controlled Fusion, January 2025.

[10] Sam Iaquinta, Peter Amendt, Jose Milovich, Eduard Dewald, Laurent Divol, et al. Characterization of foam-filled hohlraums for inertial fusion experiments, June 2024.

[11] A. S. Moore, N. B. Meezan, J. Milovich, S. Johnson, R. Heredia, et al. Foam-lined hohlraum, inertial confinement

fusion experiments on the National Ignition Facility. Physical Review E, 102(5):051201, November 2020.
[12] R. E. Olson, M. J. Schmitt, B. M. Haines, G. E. Kemp, C. B. Yeamans, et al. A polar direct drive liquid deuterium–tritium wetted foam target concept for inertial confinement fusion. *Physics of Plasmas*, 28(12):122704, December 2021.
[13] I. V. Igumenshchev, W. Theobald, C. Stoeckl, R. C. Shah, D. T. Bishel, et al. Proof-of-Principle Experiment on the Dynamic Shell Formation for Inertial Confinement Fusion. *Physical Review Letters*, 131(1):015102, July 2023.
[14] Keiji Nagai, Christopher S. A. Musgrave, and Wigen Nazarov. A review of low density porous materials used in laser plasma experiments. *Physics of Plasmas*, 25(3):030501, March 2018.
[15] T. Wiste, O. Maliuk, V. Tikhonchuk, T. Lastovicka, J. Homola, et al. Additive manufactured foam targets for experiments on high-power laser–matter interaction. *Journal of Applied Physics*, 133(4):043101, January 2023.
[16] R. W. Paddock, M. W. von der Leyen, R. Aboushelbaya, P. A. Norreys, D. J. Chapman, et al. Measuring the principal Hugoniot of inertial-confinement-fusion-relevant TMPTA plastic foams. Physical Review E, 107(2):025206, February 2023.
[17] B. Albertazzi, P. Mabey, Th. Michel, G. Rigon, J. R. Marqu`es, et al. Triggering star formation: Experimental compression of a foam ball induced by Taylor–Sedov blast waves. Matter and Radiation at Extremes, 7(3):036902, May 2022.
[18] F. Iannone, F. Ambrosino, G. Bracco, M. De Rosa, Funel, et al. Cresco enea hpc clusters: a working example of a multifabric gpfs spectrum scale layout. In 2019 International Conference on High Performance Computing Simulation (HPCS), pages 1051–1052, 2019.
[19] B. Fryxell, K. Olson, P. Ricker, F. X. Timmes, M. Zingale, et al. FLASH: An Adaptive Mesh Hydrodynamics Code for Modeling Astrophysical Thermonuclear Flashes. The Astrophysical Journal Supplement Series, 131(1):273–334, November 2000.
[20] Anshu Dubey, Katie Antypas, Murali K. Ganapathy, Lynn B. Reid, Katherine Riley, et al. Extensible component-based architecture for FLASH, a massively parallel, multiphysics simulation code. Parallel Computing, 35(10-11):512–522, October 2009.
[21] S.P. Lyon and J.D. Johnson. SESAME 7592, in Los Alamos National Laboratory Report No. LA-UR-92-3407. Technical report, 1995.
[22] O. S. Jones, G. E. Kemp, S. H. Langer, B. J. Winjum, R. L. Berger, et al. Experimental and calculational investigation of laser-heated additive manufactured foams. Physics of Plasmas, 28(2):022709, February 2021.
[23] J. Limpouch, V. Tikhonchuk, O. Renner, Sh. Agarwal, T. Burian, et al. Laser interaction with undercritical foams of different spatial structures. Matter and Radiation at Extremes, 10(1):017402, January 2025.
[24] C. Parisuan˜a, M. P. Valdivia, V. Bouffetier, K. Kurzer-Ogul, G. P´erez-Callejo, et al. Shock propagation in aerogel and TPP foams for inertial fusion energy target design. Physics of Plasmas, 32(8):082707, August 2025.

SELF-ASSEMBLED MONOLAYERS FOR INVERTED PEROVSKITE SOLAR CELLS: A DFT STUDY

Carmen Coppola[1,2,3*], Alessio Dessì[2], Lorenzo Zani[2], and Adalgisa Sinicropi[1,2,3]

[1]*Università degli Studi di Siena, Dipartimento di Biotecnologie, Chimica e Farmacia, R²ES Lab, via Aldo Moro 2, 53100 Siena, Italia.* [7]

[2]*Consiglio Nazionale delle Ricerche - Istituto di Chimica dei Composti OrganoMetallici (CNR-ICCOM), via Madonna del Piano 10, 50019 Sesto Fiorentino, Italia.*

[3] *CSGI, Consorzio per lo Sviluppo dei Sistemi a Grande Interfase, via della Lastruccia 3, 50019, Sesto Fiorentino, Italia.*

ABSTRACT. Over the last decade, interface engineering utilizing self-assembled monolayers (SAMs) has emerged as a crucial approach for enhancing the power conversion efficiency (PCE) of inverted perovskite solar cells (PSCs). Indeed, when used as hole-transport layers (HTLs) anchored on metal oxide surfaces, SAMs could form ordered, thin layers that improve the charge extraction from the perovskite and reduce recombination losses. In this work, we present the *in silico* design of three novel organic SAMs. The molecular and electronic properties of the three molecules have been investigated by means of Density Functional Theory (DFT) and Time Dependent Density Functional Theory (TDDFT) when in solution and when anchored on the ITO surface for a possible application in inverted PSCs. The outcomes of this study will contribute to rationalizing SAMs' adsorption features, as well as structural factors affecting the mechanism of ordered film formation.

Introduction

Perovskite solar cells (PSCs) have garnered significant attention within the photovoltaic community due to their remarkable power conversion efficiency (PCE) [1]. However, their commercialization is still limited by issues of long-term stability. There are two main configurations for PSCs: direct (n-i-p) and inverted (p-i-n). The p-i-n configuration demonstrates enhanced stability and greater compatibility with flexible and tandem devices. It features a transparent conductive oxide (TCO) layer in contact with the hole transport layer (HTL), on which the perovskite is grown (Fig. 1). Thus, the HTL is a vital component of PSCs, as it significantly influences the crystal properties of the perovskite absorbing layer. Additionally, it plays a key role in the extraction and transport of holes from the perovskite to the counter electrode [2]. In this framework, self-assembled monolayers (SAMs) anchored on metal oxides have recently been used as HTLs in p-i-n PSCs, as they are cost-effective, stable, and can be used without additives, boosting the PCE of p-i-n above 26% [3]. SAMs are generally composed of three groups: the anchoring group, usually a carboxylic or phosphoric acid, which should have a good affinity for the metal oxide surface and allows the SAMs to be adsorbed, the terminal group capable of hole extraction that determines the surface and interface properties, and a linker that connects the two groups (Figure 1). In recent years, research efforts have been made to design novel SAM molecules to be employed as HTLs, demonstrating that interface engineering with the use of SAMs could be a valid strategy to improve the charge transfer at the perovskite/HTL [4]. Indeed, SAMs could establish covalent bonds with the metal oxide surface, generally indium tin oxide (ITO), forming a long-range molecularly ordered film which improves the charge extraction from perovskite thanks to better energy levels alignment and

[7] Corresponding author. E-mail: carmen.coppola2@unisi.it

decreases the recombination losses [5]. However, the design of closely packed SAMs is still challenging, and efforts should be made to test alternative molecules with improved terminal groups to tune the interfacial interaction with the perovskite. In this context, this study dissects the *in silico* design of a novel class of organic SAMs with a hole-extracting unit functionalized with various substituents and linked to a phosphoric acid anchoring group. Building on our experience on the design and characterization of novel HTLs for PSCs [2, 6, 7], we fully characterized the novel compounds by *state-of-the-art* Density Functional Theory (DFT) methods to investigate their possible application as SAMs for HTLs in p-i-n PSCs. We uncovered their structural and electronic properties when in solution, but also their interfacial contacts with the ITO surface, exploring the factors influencing the formation of a densely packed self-assembled film. The designed compounds were further synthesized and, in perfect agreement with the computed results, the experimental characterization further confirmed the designed SAMs as valid HTLs in p-i-n PSCs. Ultimately, PSCs containing the designed SAMs were fabricated, yielding outstanding PCE values [8].

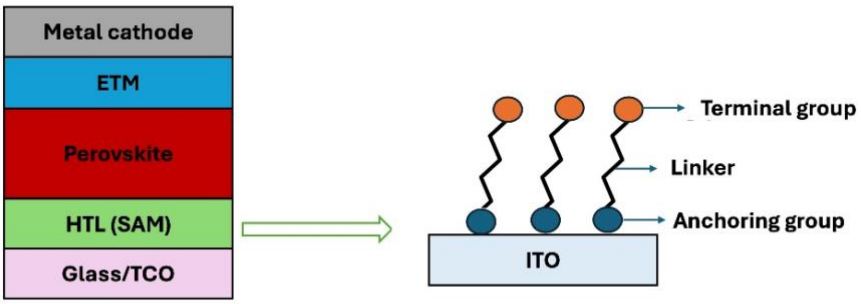


Fig.1: Schematic representation of an inverted PSC and the SAMs' composition.

Computational Methods

All the calculations on isolated molecules were performed by using the Gaussian 16, Revision C.01 suite of programs [9], while the chemisorption of the designed compounds on the ITO surface was carried out with the SIESTA 4.1 program package [10]. The latter was implemented on HPC CRESCO/ENEAGRID infrastructures. We employed the DFT and TDDFT methods, as implemented in Gaussian 16, to calculate the minimum energy structures of the novel molecules and to investigate their molecular dipole moments, energetic alignment with the perovskite, and UV-Vis properties in solution. Stacking models of the three SAMs were also simulated to understand the interactions and arrangements that influence the formation of self-assembled films. Additionally, we used periodic DFT simulations along with Troullier–Martins norm-conserving pseudopotentials on SIESTA 4.1 to characterize the ITO surface and to optimize the most stable SAMs/ITO systems, taking into account dispersion forces by means of the D3BJ damping scheme [11]. A standard calculation of SAMs/ITO interface required 20 nodes, 192GB of RAM per node, and 50GB of storage.

Results and Discussion

The investigation of the molecular and electronic properties of the isolated novel compounds revealed that they have i) correct energetic alignment with the perovskite, allowing for efficient hole extraction, ii) absorption maxima in the UV region of the spectrum, thus not interfering with the perovskite absorption in the visible region, and iii) suitable molecular dipole moments, favouring a strong dipole formation at the SAMs/ITO interfaces. The qualitative comparison of the three designed molecules on the ITO surface helped us to rationalize the adsorption modes. Following the work of Paramonov et al. [12], the simulation of different adsorption geometries allowed us to uncover the most stable binding geometries and calculate their binding energies. The simulations revealed that the anchoring group of the designed compounds can be chemisorbed on the ITO surface with bidentate and tridentate binding

modes. Fig.2 shows a typical scheme of the OH-terminated ITO surface employed in this work and a scheme of possible binding modes on the surface. Additionally, the terminal groups of the three molecules were found to be responsible for the formation of homogeneous and stable SAMs. This finding was further confirmed by simulating stacking models of the three molecules and by advanced experimental characterization. Thus, PSCs containing the designed SAMs were fabricated, yielding outstanding PCE values, highlighting the potential of the novel SAMs in achieving good performance when employed in inverted perovskite solar cells and paving the way for further advancements in the field of photovoltaic technologies [8].

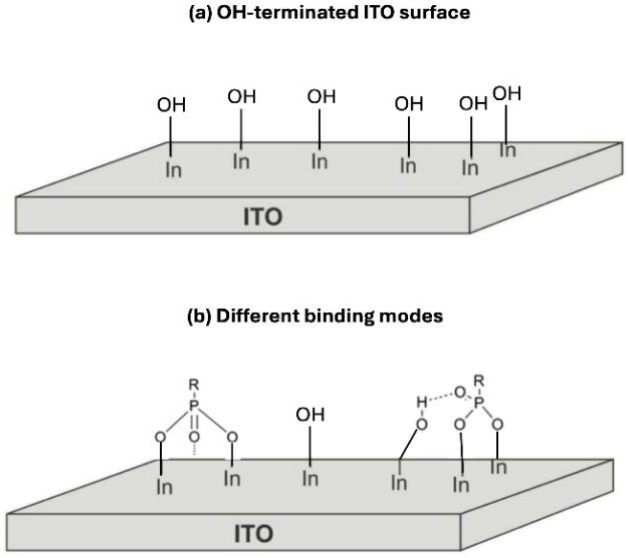


Fig.2: a) scheme of OH-terminated ITO surface; (b) scheme of possible binding modes.

Conclusions

In this study, three novel SAMs were designed and fully characterized to assess their suitability as HTLs in p-i-n PSCs. A complete computational analysis allowed us to rationalize the SAMs' adsorption mechanisms and the factors influencing self-assembly. The results emphasize the importance of interface engineering in enhancing PSC performance and may offer theoretical guidance for creating efficient and stable SAMs. Furthermore, the successful application of these molecules has demonstrated benefits in charge extraction, stability, and efficiency of the fabricated p-i-n PSCs.

Acknowledgments

The computing resources and the related technical support have been provided by the CRESCO/ENEAGRID High Performance Computing Infrastructure and its staff. CRESCO/ENEAGRID High Performance Computing Infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development, and by Italian and European research programmes, see https://www.cresco.enea.it for information. We also acknowledge the hpc@dbcf for providing computational resources.

References

[1] NREL. Best Research-Cell Efficiency Chart, https://www.nrel.gov/pv/assets/pdfs/best-research-cell-efficiencies.pdf

[2] L. A. Castriotta, R. Infantino, L. Vesce, M. Stefanelli, A. Dessì, C. Coppola, M. Calamante, G. Reginato, A. Mordini, A. Sinicropi, A. Di Carlo and L. Zani. Stable Methylammonium-Free p-i-n Perovskite

Solar Cells and Mini-Modules with Phenothiazine Dimers as Hole-Transporting Materials. *Energy & Environmental Materials* **6**, e12455, (2023).
[3] S Liu, J. Li, W. Xiao, R. Chen, Z. Sun, Y. Zhang, X. Lei, S. Hu, M. Kober-Czerny, J. Wang, F. Ren, Q. Zhou, H. Raza, Y. Gao, Y, Ji, S. Li, H. Li, L. Qiu, W. Huang, Y. Zhao, B. Xu,Z. Liu, H.J. Snaith, N.-G. Park, W. Chen. Buried Interface Molecular Hybrid for Inverted Perovskite Solar Cells. *Nature* **632**, 536–542, (2024).
[4] E. Aktas, R. Pudi, N. Phung, R. Wenisch, L. Gregori, D. Meggiolaro, M. A. Flatken, F. De Angelis, I. Lauermann, A. Abate, and E. Palomares. Role of Terminal Group Position in Triphenylamine-Based Self-Assembled Hole-Selective Molecules in Perovskite Solar Cells. *ACS Appl. Mater. Interfaces* **14**, 17461−17469, (2022).

[5] Q. Liao, Y. Wang, Z. Zhang, K. Yang, Y. Shi, K. Feng, B. Li, J. Huang, P. Gao, and X. Guo. Self-assembled donor-acceptor hole contacts for inverted perovskite solar cells with an efficiency approaching 22%: The impact of anchoring groups. *Journal of Energy Chemistry* **68**, 87–95 (2022).

[6] C. Coppola, R. Infantino, A. Dessì, L. Zani, M.L. Parisi, A. Mordini, G. Reginato, R. Basosi and A Sinicropi. DFT and TDDFT investigation of four triphenylamine/phenothiazine-based molecules as potential novel organic hole transport materials for perovskite solar cells. *Materials Chemistry and Physics* **278**, 125603, (2022).
[7] C. Coppola, A. Pecoraro, A.B. Muñoz-García, R. Infantino, A. Dessì, G. Reginato, R. Basosi, A. Sinicropi and M. Pavone. Electronic structure and interfacial features of triphenylamine- and phenothiazine-based hole transport materials for methylammonium lead iodide perovskite solar cells *Phys. Chem. Chem. Phys*. **24**, 14993-15002, (2022).
[8] D. Takhellambam, L. A. Castriotta, M. Salvi, A. Dessì, C. Coppola, L. Poggini, D. Franchi, M. Calamante, A. Sinicropi, L. Zani, G. Reginato, A. Mordini, A. Di Carlo*, manuscript in preparation*.
[9] Gaussian 16, Revision C.01. M. J. Frisch, et al. D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.
[10] J. M. Soler, E. Artacho, J. D. Gale, A. García, J. Junquera, P. Ordejón and D. Sánchez-Portal. The SIESTA method for *ab initio* order-*N* materials simulation. *J. Phys.: Condens. Matter* **14**, 2745–2779 (2002).
[11] S. Grimme, J. Antony, S. Ehrlich and S. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys*. **132**, 154104, (2010).
[12] P.B. Paramonov, S. A. Paniagua, P. J. Hotchkiss, S. C. Jones, N. R. Armstrong, S. R. Marder, and J.-L. Brédas. Theoretical Characterization of the Indium Tin Oxide Surface and of Its Binding Sites for Adsorption of Phosphonic Acid Monolayers. *Chem. Mater*. **20**, 5131–5133 (2008).

FIRST-PRINCIPLES INSIGHT ON FUNCTIONAL MATERIALS: FROM TRIBOLOGICAL TO ELECTROCHEMICAL PERFORMANCE

Domenico Corona[1*], Stefanos Giaremis[2], Margherita Marsili[2], Paola Gori[3], Maria Clelia Righi[2] and Olivia Pulci[1]

[8]*Department of Physics, University of Rome Tor Vergata and INFN, Via della Ricerca Scientifica 1, 00133 Rome, Italy*

[2]*Department of Physics and Astronomy, University of Bologna, Viale Carlo Berti Pichat 6/2, Bologna, 40127, Italy*

[3]*Department of Industrial, Electronic and Mechanical Engineering, Roma Tre University, Via della Vasca Navale 79, 00146, Rome, Italy*

ABSTRACT. First-principles methods are particularly valuable due to their ability to predict material behaviour with high accuracy, their applicability to a wide range of systems, and the consistency of the outcomes they produce. Nonetheless, their application to large or complex atomic systems can be computationally intensive, frequently requiring the use of high-performance computing (HPC) resources. We report *ab initio* results obtained for two separate types of functional materials: tribological properties of diamond surfaces and electrochemical properties of mixed-valence transition-metal oxides. The results shown here have been obtained, either partially or entirely, by calculations on CRESCO/ENEAGRID HPC infrastructure in the last year.

Introduction

The advent of HPC has made it feasible to investigate, entirely *in silico*, the mechanical, electronic, thermodynamic and optical properties of novel materials. This capability is of significant technological relevance, enabling the design and optimization of advanced functional materials for tribology applications, energy conversion and storage, sensing, optoelectronic devices, and emerging quantum technologies. Concerning the mechanical and tribological properties, we report on two relevant studies. One work involves the use of DFT calculations for evaluating the role of B dopants to the initialization and evolution of wear in diamond surfaces at the atomistic level. As diamond is the hardest known material and exhibits remarkable mechanical properties, it is commonly used in protective coatings in numerous technological applications [1]. Doping with B is explored in many applications for improving adhesion to substrate, wear resistance, biocompatibility and reducing friction, but mixed results have been reported in the literature [2, 3, 4]. Therefore, the scope of this study is to assess the

8 Corresponding author. E-mail: domenico.corona@roma2.infn.it.

impact of B dopants on the behavior of wear in diamond surfaces at the atomistic level. In the other work, DFT is being used to probe and analyze the way in which the electronic charge of crystalline solids (Cu, Al and diamond) responds to uniaxial strain affecting their mechanical properties. In this report we present some preliminary results concerning stress-strain curves and Poisson's ratio variation with applied strain of the three materials. Concerning the electronic and electrochemical properties, we report on a relevant study. The work involves mixed transition-metal ($Na_xTi_{0.17}Ni_{0.33}Mn_{0.5}O_2$ and $Na_xFe_{0.17}Ni_{0.33}Mn_{0.5}O_2$) layered sodium manganese oxides (x = 0, 1) [5, 6]. Included in the DFT are the on-site U Hubbard interactions to accurately model the properties arising from the strongly correlated nature of transition-metal *3d* states in these lamellar systems. The on-site U Hubbard parameters are computed fully from first-principles using linear response theory. For these mixed-valence cathode materials, we provide calculated oxidation states and intercalation voltages that show strong agreement with experimental data. In addition, sodium-ion migration pathways are analyzed in detail, revealing activation energies, underlying diffusion mechanisms, and how these can be modulated through substitutional doping.

Methods

Wear mechanisms of B-doped diamond

The B-doped C(001), the C(110), and Pandey reconstructed C(111) diamond surface orientations were used, modelled as slabs in orthorhombic periodic cells with a side dimension of at least 8.75 Å, thickness of at least 6 atomic layers and 15 Å of vacuum. In each configuration, a single C atom was substituted with B on the surface ([B] = 0.5-1.0 at. %), as B dopants in diamond have been shown to occupy substitutional sites [7, 8], preferably towards the surface [9]. Atomistic pulling simulations were performed as follows. One surface atom for each model was displaced by 0.25 Å perpendicular to the surface and then the model was relaxed by the z-coordinate of the displaced atom fixed. Then, starting from the previous positions, the selected atom was further pulled by 0.25 Å and this procedure was repeated up to a total displacement of 2.5 Å. In all cases, the middle-most atomic layer of the slabs was kept fixed in all directions. The results for the respective undoped diamond surfaces were obtained by [10]. For the determination of the ground state properties of our system, i.e. its geometry, total energy, single-particle Kohn-Sham orbitals, etc., we employed DFT by using the electronic structure code *pw.x* of the Quantum Espresso suite [11], using of plane-wave basis set and ultrasoft pseudopotentials [12]. Quantum Espresso is an open-source suite, released under the GPL license, it makes use of LAPACK and BLAS libraries, and presents a hybrid MPI and OpenMP parallelism.

Mechanical properties under uniaxial load

DFT calculations are performed using the Quantum Espresso (QE) package [11] using ultrasoft pseudopotential [12] and plane waves expansion. For the exchange-correlation functional, the Perdew-Burke-Ernzerhof (PBE) [13] was chosen, adopting the parameterization by Rabe-Rappe-Kaxiras-Joannopoulos [14]. All materials are simulated as cubic cells with the xy plane parallel to the crystallographic (001) direction, replicated six times along the z-direction, two atoms belonging to each atomic plane. 18×18×3, 12×12×2, and 16×16×3 Monkhorst-Pack grids were used to sample the Brillouin zone for Al, Cu, and diamond, respectively. The kinetic energy cut-off of 40 Ry (90 Ry), for the wavefunctions, and 320 Ry (720 Ry), for the charge density were employed for aluminum and diamond (copper). Forces were minimized by employing variable cell calculations, optimizing at the same time both atomic positions and cell edges.

Doping-engineered electrochemical properties

DFT calculations have been performed employing the Quantum ESPRESSO suite [11], using ultrasoft pseudopotentials (USPP) generated with scalar-relativistic calculations [12] and an exchange-correlation (XC) energy scheme based on the General Gradient Approximation (GGA) by Perdew-Burke-

Ernzerhof (PBE) XC functional [13]. Grimme's D3 dispersion correction [15] has been inserted to account for van der Waals (vdW) interactions notoriously playing an important role in layered structures. Given the proportions of the supercell, the integration over the Brillouin zone required an 8 ×8×4 Monkhorst-Pack grid of k points Due to the presence of magnetic transition-metal atoms in the various systems, spin-polarization has been accounted on the magnetic atoms. The on-site Hubbard interaction parameters are computed self-consistently in the linear response approach with respect to both the structural and electronic properties of the stoichiometry under investigation.

Results

Wear mechanisms of B-doped diamond

The resulting relaxed models after the pulling sequence are shown in Fig.1. While pulling a C surface atom from the undoped diamond surface was shown to lead to the formation of an atom chain [10], introducing a B dopant alters this behavior, as pulling the B dopant leads to the extraction of only this atom and the formation of a vacancy in its place. The B-C bond is known to be longer than the C-C one, so it appears that it is also weaker and cannot support the further formation of a chain when B is pulled. However, when a C atom in the vicinity of B is pulled, a chain was formed in most cases. The exception was in C(001), in which pulling either component of the surface dimer in the B-doped system did not lead to the formation of a chain, but only to the removal of the single atom that was pulled.

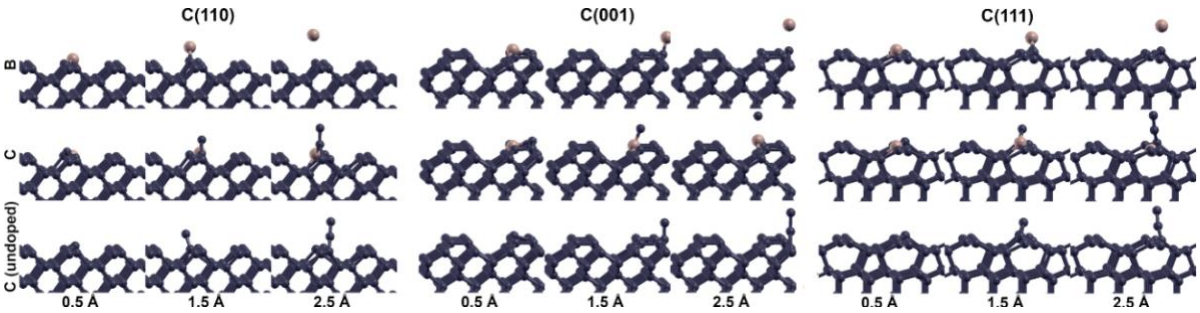


Fig.1: Relaxed configurations for the three considered diamond orientations. Results for the undoped systems are from [8].

Mechanical properties under uniaxial load

Pressure is applied to the systems reducing the bulk cells dimension along the z-direction and allowing the cell to deform within the xy plane, and all the atoms to fully relax. The stress-strain curves of the three materials differ qualitatively: while, within diamond, stress grows linearly in the whole strain range, both Cu and Al presents points, labelled by capital letters in the top panels of Fig. 2, where the stress-strain curves change their slope. At the same strain values also the Poisson's ratio curve presents deviations from a linear behavior.

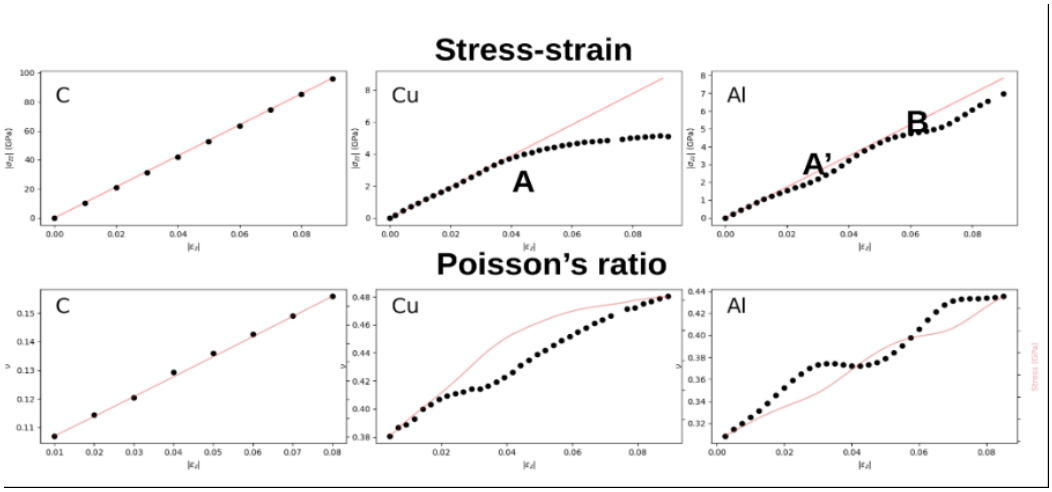


Fig.2: Stress-strain curves (top panels) and Poisson's ratio (bottom panels) of diamond, copper and aluminium.

Doping-engineered electrochemical properties

In the case of the $Na_xFe_{0.17}Ni_{0.33}Mn_{0.5}O_2$, in addition to Mn^{3+}/Mn^{4+} and Ni^{2+}/Ni^{4+}, also the Fe sublattice has a redox active role, with the oxidation of Fe^{3+} to Fe^{5+} during sodium removal, this is consistent with the experimental observations presented in Ref. [5]; while for $Na_xTi_{0.17}Ni_{0.33}Mn_{0.5}O_2$ the Ti adopts a quadrivalent oxidation state, remaining Ti^{4+} upon charging with the exploitation of Mn^{3+}/Mn^{4+} and Ni^{2+}/Ni^{4+} as electroactive redox couples. We confirm the empirical observation that substitutional doping creates a shift in the redox potentials: the voltage of the $Mn^{3+/4+}$ couple is increased by 1.1 V and 1.4 V going from the pristine Na_xMnO_2, to the doped manganites $Na_xFe_{0.17}Ni_{0.33}Mn_{0.5}O_2$, $Na_xTi_{0.17}Ni_{0.33}Mn_{0.5}O_2$, as well as a decrease of the sodium migration barrier to 0.19 eV for both the structures shown in Fig. 3, resulting from a simultaneous enhancement of the structural stability.

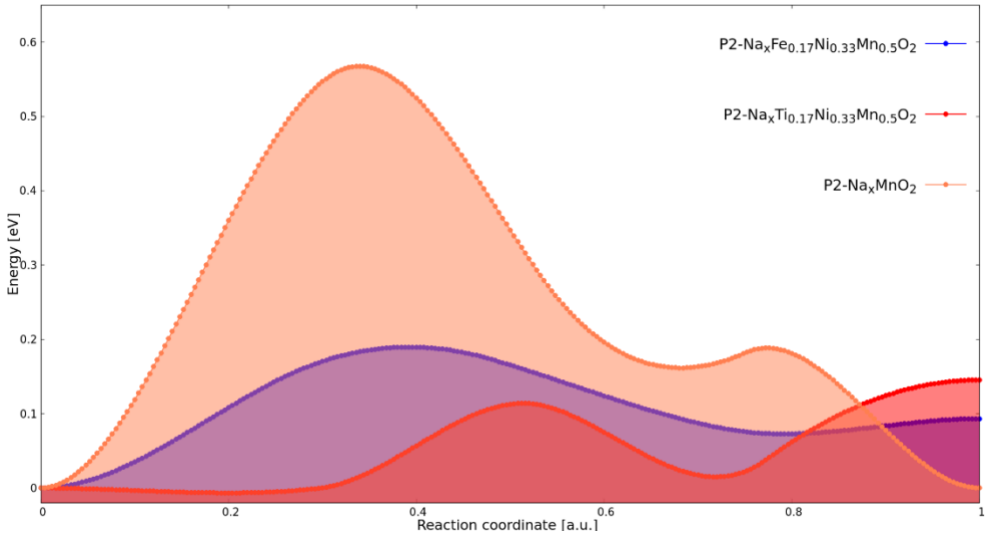


Fig.3: Energy profiles as a function of the reaction coordinate for sodium-ion diffusion from the edge-share site to the face-share site in Na_xMnO_2, $Na_xFe_{0.17}Ni_{0.33}Mn_{0.5}O_2$ and $Na_xTi_{0.17}Ni_{0.33}Mn_{0.5}O_2$.

References

[1] Liu, Y., Ali Erdemir, and E. I. Meletis. *Surface and Coatings Technology* **82.1-2** (1996): 48-56.
[2] Liang, Qi, Andrei Stanishevsky, and Yogesh K. Vohra. *Thin Solid Films* **517.2** (2008): 800-804.
[3] Lu, Ming, et al. *Journal of Manufacturing Processes* **88** (2023): 145-156.
[4] Buijnsters, Josephus G., et al. *ACS Applied Materials & Interfaces* **8.39** (2016): 26381-26391.
[5] Cui, Jing, et al. *Electrochimica Acta* (2025): 146397.
[6] Huang, Rui, et al. *Journal of Alloys and Compounds* **1005** (2024): 176168.
[7] Mukuda, H., et al. *Physical Review B* **75.3** (2007): 033301.
[8] Oguchi, Tamio. *Science and Technology of Advanced Materials* **7.S1** (2006): S67.
[9] Giaremis, Stefanos, and Maria Clelia Righi. *Surfaces and Interfaces* **46** (2024): 104105.

[10] Ta, Huong T.T., Nam V. Tran, and Maria Clelia Righi. *Langmuir* **39.40** (2023): 14396-14403.
[11] Giannozzi, Paolo, et al. *Journal of Physics: Condensed Matter* **29.46** (2017): 465901.
[12] Laasonen, Kari, et al. *Physical Review B* **43.8** (1991): 6796.
[13] Perdew, John P., Kieron Burke, and Matthias Ernzerhof. *Physical Review Letters* **77.18** (1996): 3865.
[14] Rappe, Andrew M., et al. *Physical Review B* **41.2** (1990): 1227.
[15] Grimme, Stefan, et al. *The Journal of Chemical Physics* 132.15 (2010).

LICENSES DETECTION IN THE SOFTWARE HERITAGE DATABASE

Serena D'Onofrio[1*], Lorenzo De Biase[2], Marco Catillo[3] and Angelo Mariano[4]

[1]*Italian National Agency for New Technologies, Energy Technologies and Renewable Sources Department, ICT Division, ENEA, Bologna Research Center, Bologna, Italy*[9]

[2]*Italian National Agency for New Technologies, Energy Technologies and Renewable Sources Department, ICER Division, ENEA, Bologna Research Center, Bologna, Italy*

[3]*Italian National Agency for New Technologies, Energy Technologies and Renewable Sources Department, ICT Division, Casaccia Research Centre, Rome, Italy*

[4]*Italian National Agency for New Technologies, Energy Technologies and Renewable Sources Department, ICT Division, ENEA, Bari Territorial Office, Bari, Italy*

ABSTRACT. In this work we present a methodology for studying the Free Open-Source Software Licenses database extracted from Software Heritage, the universal archive of source code. Such approach involves a mix of Machine Learning and customised clustering parallelised algorithms, to detect and identify some classes of software licenses.

Introduction

Software Heritage [3] is an initiative that provides a uniform, universal archive of source code. The rise of Free Open-Source Software (FOSS) over the last decades has helped to promote source code accessibility. Software Heritage goal is to collect, organise, preserve and make easily available all the software source code. Maintaining such an archive is easier if there are more copies of it around the world, and ENEA owns the first operative copy of Software Heritage, a mirror. One of the aspects of building an archive like Software Heritage is to provide a place where all information about software, their public source code and their development history are made available in a uniform data model.

From this large amount of data, it is possible to develop new analyses by applying recent Machine Learning (ML) and Deep Learning (DL) approaches.

The present work concerns a subset of data from the Software Heritage archive, specifically the dataset presented in [5], where the corpus is a collection of 6.9 million unique software license files with associated metadata. Starting from this dataset, our goal is to develop an algorithm capable of automatically detecting and identifying different software license texts.

The study of software licenses started in the 1990s, when the first licenses appeared, and the only technique used in the beginning was manual inspection by the legal team. In the last decade, with the advent of ML and DL, there has been a new way of approaching the task of detecting and identifying software licenses. One of the pioneering tools in this field is FOSSology [6], the first detection method implemented with ML algorithms. Subsequently, ScanCode was released, a tool derived from combining Natural Language Processing (NLP) techniques and multiple matching methods. ScanCode

[9] Corresponding author e-mail: serena.donofrio@enea.it

was used in the dataset presented in [9] to identify the licenses name on the basis of its high performance.

In the present work, we introduce a method based on an ML approach combined with a clustering algorithm. This method identifies 11 different types of software licenses, exploiting text embedding based on a Deep Learning architecture of Transformers, and performs customised clustering that identifies and classifies unlabelled license texts. This technique proves effective when analysing texts that have a structure based on a standard template.

We implement the code on the CRESCO/ENEAGRID High Performance Computing infrastructure, on the CRESCO7 cluster [7, 8].

Method

For the creation of a custom clustering algorithm capable of identifying and classifying 11 different types of FOSS licences, we use an embedding based on Bidirectional Encoder Representation from Transformers (BERT). This allows each sentence to be represented in a multidimensional space. We then construct 11 cluster centres based on the model texts of the selected licence types in order to define the centroid of the classes to be detected. Finally, we use ground truth to calculate the accuracy and compare our model with the results obtained with ScanCode. In the following sections, we describe the dataset and details of the model.

Dataset description

The dataset under consideration, described in the [5] article, is a collection of 6859189 unique license files obtained from Software Heritage by querying all file names that are likely to contain license information. All license files in the collection are associated with a list of metadata, including: file names, length measures, detected MIME type, contained FOSS licenses detected using ScanCode.

In addition, there is a subset of 8102 annotated documents randomly selected from the entire dataset. These files were manually checked to see whether they were single licenses (4601), multiple licenses (713), simple license notices (579) or no licenses at all (or even plain text). This manual check is essential to verify whether the licenses noted with ScanCode are correct.

For our purposes, we only consider the annotated data set for the development of the algorithm.

BERT Embedding

The first step in our approach involves embedding, i.e. the vector representation of texts. The method we use for embedding is BERT, a Bidirectional Encoder Representation from Transformers (introduced in [1] and first released on Github [4]).

BERT is a pre-trained language model that uses bidirectional context to improve performance in natural language processing tasks. It exploits the attention model to gain a deeper understanding of linguistic context. The architecture of BERT consists of a stack of transformer encoder blocks. The sentences of the input text are separated into tokens, two special tokens are added to determine the beginning and end of the sentences and each sentence is transformed into a vector as the output of the process. There are two training strategies implemented in BERT to improve text comprehension: Masked LM and Next Sentence Prediction. The first consists of randomly masking tokens to improve their prediction from context. The second separates consecutive sentences to see if there is a link between them. In this way,

the model learns an inner representation of the language, which can be used to extract useful features for subsequent tasks.

In the present work we use a pretrained BERT base uncased model with an architecture consisting of 12 layers, 768 hidden nodes, 12 attention heads, and 110M parameters.

The output of this BERT model has a dimension (number of layers×number of batches×number of tokens×number of features), where the number of layers is 13 because we have 12 layers plus the input embedding layer, in our case the number of batches is 1 because we process a single sentence at a time, the number of tokens is the number of different tokens identified in the sentence by the model and the number of features is 768.

To feed BERT with our files, pre-processing is necessary. In fact, the dataset consists not only of text licenses, but also of html files, images, code, and many other different things. Therefore, we have to divide files containing plain text into sentences. When syntactically natural splitting is not possible, we divide text every 512 characters, which is the maximum length supported by the embedding process. Following this, there are several possibilities in the pooling process, i.e. the process of obtaining a single vector of size 1×number of features for each sentence. For our purposes, we consider the concatenation of the last four layers of the embedding model in order to obtain an embedding vector of size 3072 for each sentence of the license text.

Clustering algorithm

Once the embedding is defined and applied to the entire dataset, each license file is associated with a list of vectors representing the different sentences or parts of the text. Similarly, we apply the BERT embedding to the templates of the 11 selected open-source software licenses, the archetypes, which are the classes we detect. The selected licenses that define the classes are the most frequent in the annotated dataset, namely MIT, Apache, GPL, BSD, ISC, LGPL, AGPL, Artistic, CC0, CCBY, WTFPL. To these classes, we added “Unknown” for all other files. With this procedure, we define the centre of gravity of our clusters.

The comparison between a selected license and an archetype is based on the distances between the embedding vectors x and y of both texts. The idea is to check how many embedding vectors have a Euclidean distance less than a parameter ε i.e.:

$$d(x,y) = \left(\sum_{i=0}^{n} |x_i - y_i|^2\right)^{\frac{1}{2}} < \varepsilon.$$

A score is then associated with each class. Based on this score, one has the probability of belonging or not belonging to a certain class.

Results and Conclusions

In this section, we report preliminary results on the accuracy of the model and compare it with the performance of ScanCode. For further details we refer to the paper [2] in preparation.

After postprocessing that homogenises all predicted labels with the ground truth from the dataset, we check the accuracy, i.e. the ratio of the number of correct predictions to the total number of input samples. Our model reaches an accuracy of 0.9463 on single licenses texts.

We see that the predictions made by the clustering method are comparable to the detection made by ScanCode. The confusion matrices in Fig. 1 show that the detection made by our clustering based on the incorporation of BERT and that made by ScanCode are similar.

The class in which our approach performs better than ScanCode is the “Apache” class, because many of these files are in html format and ScanCode does not detect them as licenses at all.

To perform the BERT embedding on 8100 different files, we ran the code on a Cresco7 cluster, implementing a parallelisation with the MPI standard, on 1 node with 48 processes (average time is 1668 seconds per process).

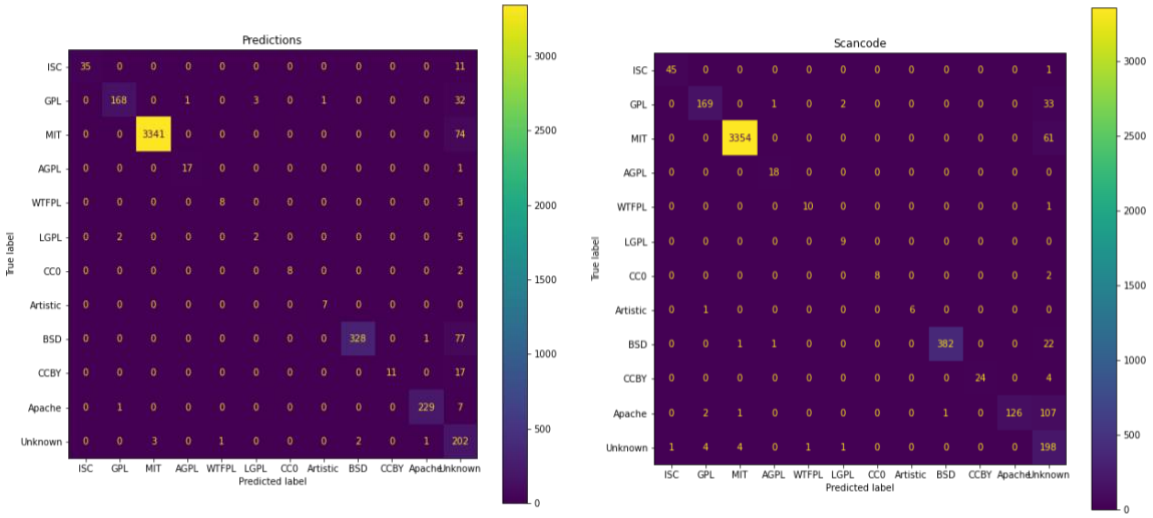


Fig.1: Confusion matrices of the prediction of the model (left) and the prediction of ScanCode (right) on single license texts.

Acknowledgments

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID HPC infrastructure and its staff [7, 8]. CRESCO/ ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes.

References

[1] J. Devlin, M.W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. *In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies*, **1**, pp. 4171-4186. (2019).

[2] S. D’Onofrio, L. De Biase and A.Mariano. Machine Learning approach for software Licences detection and identification in the Software Heritage database. *In preparation*.

[3] R. Di Cosmo, S. Zacchiroli. Software heritage: Why and how to preserve software source code. *In iPRES 2017*. (2017).

[4] GitHub repository of BERT. *https://github.com/google-research/bert*

[5] J.M. Gonzalez-Barahona, S. Montes-Leon, G. Robles, and S. Zacchiroli. The software heritage license dataset (2022 edition). *Empirical Software Engineering*, **28(6)**, 147, (2023).

[6] FOSSology project. *https://www.fossology.org/*

[7] F. Iannone, et al. NEW HPC ENEA CRESCO DEPLOYMENT. *High Performance Computing on CRESCO infrastructure: research activities and results 2023*, **7**. (2024).

[8] F. Iannone, et al. CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. *International Conference on High Performance Computing & Simulation (HPCS),* pp. 1051-1052, IEEE, (2019).

[9] ScanCode toolkit. *https://scancode-toolkit.readthedocs.io/en/latest/index.html*

FIRST PRINCIPLES INVESTIGATION OF 2D MIXED SB/BI- PEROVSKITE-INSPIRED MATERIALS (PIM)

Francesca Fasulo[1*], Michele Pavone[2], Ana B. Muñoz-García[1]

[1]*University of Naples Federico II, Department of Physics "E. Pancini", Complesso Universitario Monte S. Angelo, via Cintia 26, 80126 Naples Ital*[10]

[2]*University of Naples Federico II, Department of Chemistry, Complesso Universitario Monte S. Angelo, via Cintia 26, 80126 Naples Italy*[11]

ABSTRACT. We investigated the structural and electronic features of the new synthetized 2D mixed Sb/Bi-Perovskite Inspired Materials (PIMs) via state-of-the-art Density functional theory (DFT) calculations. Our DFT results predict that the mixed $Cs_{2.4}MA_{0.5}FA_{0.1}Sb_{1.8}Bi_{0.2}I_{8.5}Cl_{0.5}$ forms stable 2D-ASbX layered structure and possesses the necessary features of the nearly direct band gap for a high indoor power conversion efficiency.

Structural and electronic features of 2D Sb/Bi PIM

Perovskite-inspired materials (PIMs) are attracting growing interest as emerging photovoltaic absorbers, thanks to their inherent air stability and low toxicity. In this study, we investigate a novel and promising PIM composition, CsMAFA-Sb/Bi, developed through antimony:bismuth co-alloying of a triple-cation, vacancy-ordered, antimony-based PIM [1].

We investigated the structural and electronic features of the new synthetized 2D mixed Sb/Bi-PIMs $Cs_{2.4}MA_{0.5}FA_{0.1}Sb_{1.8}Bi_{0.2}I_{8.5}Cl_{0.5}$ via state-of-the-art Density functional theory (DFT) calculations, similarly to our previous study on 2D-$A_3(Sb)_2X_9$ PIMs [2]. In particular, we focused our DFT analysis on parent full inorganic 2D-$Cs_3Sb_2I_9$ and $Cs_{2.4}MA_{0.5}FA_{0.1}Sb_2I_{8.5}Cl_{0.5}$ composition. The mixed $A_3(Sb/Bi)_2X_9$ PIM were built up by $A_3(Sb)_2X_9$ PIM reported in a previous work [2] (Figure 1a). We performed DFT calculations with periodic boundary conditions (PBC) employing the light–tier1 basis set of numerical atom-centered orbitals (NAO) for each atom [3], as implemented in the Fritz Haber Institute ab initio molecular simulations (FHI-aims) code [4,5]. Within the FHI-aims framework, the electrons were described by the zero–order regular approximation (atomic ZORA). The Perdew-Burke-Erzenhof (PBE) [6] exchange correlation functional was used for all geometry optimizations including the Tkatchenko–Scheffler (TS) correction [7,8] accounting for van der Waals dispersion forces. Our relaxed structures present maximum forces acting on each atom below 0.02 eV/Å. The Γ-centered (2×2×1) k-points sampling mesh was used; these values ensure converged energies within 3 meV/f.u. Since the PBE functional usually underestimates the band energy gap, we employed the HSE06-TS hybrid functional to calculate the projected density of states (pDOS) and the band energy structures of the PBE-TS optimized geometries.

According to our calculations, the mixed A(Sb/Bi)X PIM exhibits the typical 2D layered phase [9-14], consistently with our previous findings [2], and the introduction of Bi atoms has only a minor effect on bulk structure, i.e. both lattice parameters and the stacking between SbI_8 octahedra (d1 and d2 in Figure 1b) are preserved with respect to the $Cs_{2.4}MA_{0.5}FA_{0.1}Sb_2I_{8.5}Cl_{0.5}$ composition (Table 1).

10 Corresponding author. E-mail: francesca.fasulo@unina.it

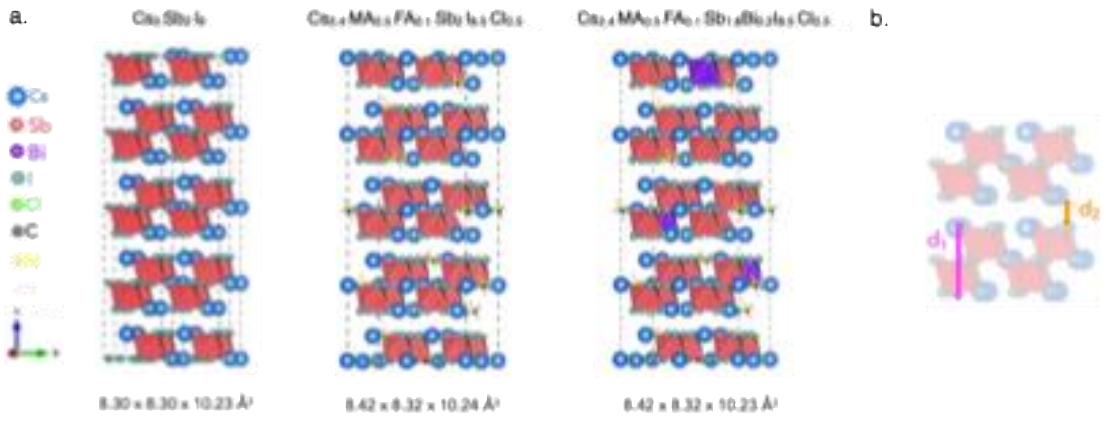


Figure 1: (a) Bulk structure and lattice parameters of the 2D-$A_3(Sb/Bi)_2X_9$ PIMs with A=Cs, FA, MA and X=I, Cl. Color code: Cs (blue), Sb (red), Bi (violet), I (teal), Cl (green), C (black), N (yellow), H (light pink). **(b)** Stacking between layers: d_1 (magenta) and d_2 (orange).

Table 1: Lattice parameters, stacking between layers (d_1 and d_2 in Figure S14) at the PBE-TS level of theory and bandgap (indirect/direct) at HSE06-TS level of theory of 2D-$A_3(Sb/Bi)_2X_9$ PIMs relaxed structures.

2D-$A_3Sb_2X_9$	Lattice parameters/f.u. (Å)			Stacking (Å)		Octahedral distortion		BandGap (eV)	
	a	b	c	d_1	d_2	$\%D_{oct}$	σ^2	Indirect	Direct
$Cs_3Sb_2I_9$	8.30	8.30	10.23	6.95	3.28	3%	14.4	1.72	1.77
$A_3Sb_2X_9$	8.42	8.32	10.24	6.94	3.29	4%	17.6	1.81	1.85
$A_3(Sb/Bi)_2X_9$	8.42	8.32	10.23	6.95	3.29	4%	18.2	1.82	1.87

Additionally, we assessed the average structural distortion for each PIMs structure by evaluating the octahedral distortion factors: the bond length distortion index (D_{oct}) and the bond angle variance (σ^2), calculated according to:

$$D_{oct} = \frac{1}{n}\sum_{i}^{n}\frac{|l_i - l_{av}|}{l_{av}} \qquad \sigma^2 = \frac{1}{m-1}\sum_{i=1}^{n}(\phi_i - \phi_0)^2$$

where l_i and l_{av} are the bond length (B-X) and the average bond length, respectively, in BX_6 octahedra, while m is the number of bond angles (i.e. 12 for octahedra), ϕ_i and ϕ_0 are the bond angle (X-B-X) and the ideal bond angle for a regular octahedron (vertex-core-vertex angles in a perfect octahedron is 90°), respectively [15,16]. D_{oct} and σ^2 represent the distortion in bond length from the centre of the BX_6 octahedron and the degree of angular distortion from a perfect octahedron, respectively. We observed a decrease in B(Sb/Bi)-X(I/Cl) bond length that correlated to an increase in the bond length distortion index and a high bond angle variance in the CsMAFA-Sb and CsMAFA-Sb/Bi structures (Table 1). These findings highlight a slight octahedral distortion, with a more pronounced effect on the bond angle variance attributed to the incorporation of bismuth.

The electronic features of the investigated Bi-doped PIMs derivative were analyzed via atom- and angular-momentum projected Density of States (pDOS) and band structures (Figure 2). The direct and indirect band gap values reported in Table 1 highlight the quasi-direct behavior [2,9-14] is preserved even when Bi atoms are present. As for $Cs_{2.4}MA_{0.5}FA_{0.1}Sb_2I_{8.5}Cl_{0.5}$ PIMs [2], the nature of parental $Cs_3Sb_2I_9$ band gap is retained, with an increase of both direct and indirect band gap of ~ 0.1 eV (~5%). The slight widening of the bandgap and increased octahedral distortion into SbI_6 may indicate a slight reduction in the electronic dimensionality of CsMAFA-Sb and CsMAFA-Sb:Bi compared to the parental $Cs_3Sb_2I_9$. The high electron and hole effective mass (Table 2) is consistent with reduced electronic dimensionality of the 2D mixed PIMs. However, the incorporation of Bi atoms gets a minor effect in the electronic dimensionality of the 2D-PIM, which can be preferred for high-performance photovoltaic materials [17].

Table 2: Effective hole and electron masses calculated with effmass package [18] via a finite difference fitting of valence and conduction bands calculated at HSE06-TS level of theory of 2D-$A_3(Sb/Bi)_2X_9$ PIMs relaxed structures. Masses are expressed as multiple of the electron rest mass m_0.

	$Cs_3Sb_2I_9$	$A_3Sb_2X_9$	$A_3(Sb/Bi)_2X_9$
$m_h(m_0)$	0.53	0.93	0.83
$m_e(m_0)$	0.25	0.31	0.31

Furthermore, the investigated mixed Sb/Bi PIM-composition exhibits electronic features very similar to $Cs_3Sb_2I_9$ and $A_3(Sb)_2X_9$ [1,9,10], as shown in Figure 2. The MA/FA/Cs cations do not contribute to the valence band maximum (VBM) and conduction band maximum (CBM) directly. The valence band is given by the typical s–p interaction from the hybridization of Sb(5s) atomic orbitals and I(5p) atomic orbitals, with a minor contribution from Bi(6s) atomic orbitals, while the conduction band is dominated by the strong p–p interactions resulting from the overlap of I(5p)–Sb(5p) and significative I(5p)–Bi(6p). Overall, our DFT results predict that the mixed $Cs_{2.4}MA_{0.5}FA_{0.1}Sb_{1.8}Bi_{0.2}I_{8.5}Cl_{0.5}$ forms stable 2D-ASbX layered structure and possesses the necessary features of the nearly direct band gap for a high indoor power conversion efficiency.

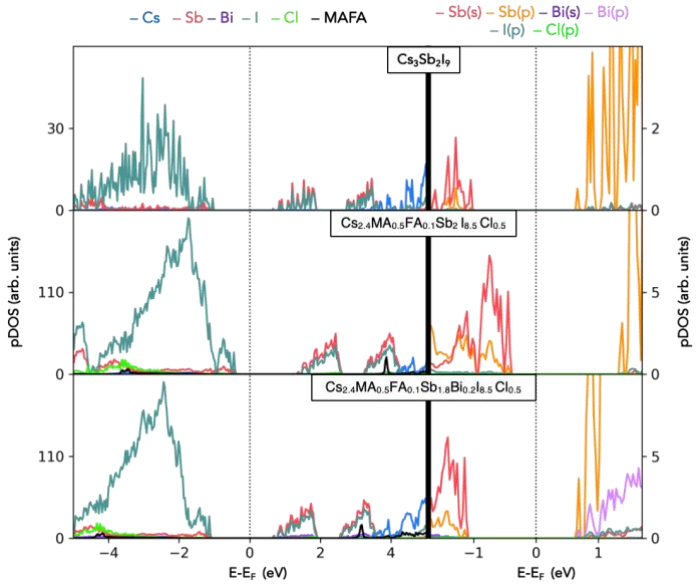


Figure 2: Projected density of states (pDOS) of the 2D-$A_3(Sb/Bi)_2X_9$ PIMs with A=Cs, FA, MA and X=I, Cl. Color code left panel: Cs (blue), Sb (red), Bi (violet), I (teal), Cl (green), MA/FA (black) (on left). Color code right panel: s-states Sb (red), p-states Sb (orange), s-states Bi (violet), p-states Bi (lilac), p-states I (teal), p-states Cl (green).

Citing Related Work

[1] N. Lamminen, J. Karlsson, R. Kumar, N. S. M. Viswanath, S. Lal, F. Fasulo, M. Righetto, M. Krishnaiah, K. Lahtonen, A. Tewari, A. Katerski, J. Lahtinen, I. Oja Acik, E. M. J. Johansson, A. B. Muñoz-García, M. Pavone, L. M. Herz, G. K. Grandhi, P. Vivo. *The Promise of Operational Stability in Pnictogen-Based Perovskite-Inspired Solar Cells*. EES Sol. **1**(2), 139–156 (2025).

References

[1] N. Lamminen, J. Karlsson, R. Kumar, N. S. M. Viswanath, S. Lal, F. Fasulo, M. Righetto, M. Krishnaiah, K. Lahtonen, A. Tewari, A. Katerski, J. Lahtinen, I. Oja Acik, E. M. J. Johansson, A. B. Muñoz-García, M. Pavone, L. M. Herz, G. K. Grandhi and P. Vivo. *The Promise of Operational Stability in Pnictogen-Based Perovskite-Inspired Solar Cells*. *EES Sol.* **1**(2), pp. 139–156, (2025).

[2] N. Lamminen, G. K. Grandhi, F. Fasulo, A. Hiltunen, H. Pasanen, M. Liu, B. Al-Anesi, A. Efimov, H. Ali-Löytty, K. Lahtonen, P. Mäkinen, A. Matuhina, A. B. Muñoz-García, M. Pavone and P. Vivo. Triple A-Site Cation Mixing in 2D Perovskite-Inspired Antimony Halide Absorbers for Efficient Indoor Photovoltaics. *Advanced Energy Materials* **13** (2022).

[3] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler. *Ab initio molecular simulations with numeric atom-centered orbitals*. *Comput. Phys. Commun.* **180**, pp. 2175–2196 (2009).

[4] V. Havu, V. Blum, P. Havu and M. Scheffler. *Efficient integration for all-electron electronic structure calculation using numeric basis functions.J. Comput. Phys.* **228**, pp.8367–8379 (2009).

[5] E. Van Lenthe, J. G. Snijders and E. J. Baerends. *The zero-order regular approximation for relativistic effects: The effect of spin–orbit coupling in closed shell molecules. J Chem. Phys.* **105**, pp. 6505 (1998).

[6] J. P. Perdew, K. Burke and M. Ernzerhof. *Generalized Gradient Approximation Made Simple. Phys. Rev. Lett.* **77**, 3865 (1996).

[7] A. Tkatchenko and M. Scheffler. *Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data. Phys. Rev. Lett.* **102**, 073005 (2009).

[8] A. Tkatchenko, R. A. DiStasio, R. Car and M. Scheffler. *Accurate and efficient method for many-body van der Waals interactions. Phys. Rev. Lett.* **108**, 236402 (2012).

[9] Y. L. Liu, C. L. Yang, M. S. Wang, X. G. Ma and Y. G. Yi. *Theoretical insight into the optoelectronic properties of lead-free perovskite derivatives of $Cs_3Sb_2X_9$ (X = Cl, Br, I). J. Mater. Sci.* **54**, pp. 4732–4741 (2019).

[10] T. Geng, Z. Ma, Y. Chen, Y. Cao, P. Lv, N. Li and G. Xiao. *Bandgap engineering in two-dimensional halide perovskite $Cs_3Sb_2I_9$ nanocrystals under pressure. Nanoscale* **12**, pp. 1425–1431 (2020).

[11] S. Berri. *Theoretical analysis of the structural, electronic, optical and thermodynamic properties of trigonal and hexagonal $Cs_3Sb_2I_9$* compound. *Eur. Phys. J. B* **93**, 191 (2020).

[12] J. P. Correa-Baena, L. Nienhaus, R. C. Kurchin, S. S. Shin, S. Wieghold, N. T. P. Hartono, M. Layurova, N. D. Klein, J. R. Poindexter, A. Polizzotti, S. Sun, M. G. Bawendi and T. Buonassisi. *A-Site Cation in Inorganic $A_3Sb_2I_9$ Perovskite Influences Structural Dimensionality, Exciton Binding Energy, and Solar Cell Performance. Chem. Mater.* **30**(11), pp. 3734–3742 (2018).

[13] A. Koliogiorgos, S. Baskoutas and I. Galanakis. *Electronic and gap properties of Sb and Bi based halide perovskites: An ab-initio study. Comput. Condens. Matter* **14**, pp. 161–166 (2018).

[14] B. Saparov, F. Hong, J.-P. Sun, H.-S. Duan, W. Meng, S. Cameron, I. G. Hill, Y. Yan and D. B. Mitzi. *Thin-Film Preparation and Characterization of $Cs_3Sb_2I_9$: A Lead-Free Layered Perovskite Semiconductor*. *Chem*. *Mater.* **27**(16), pp. 5622–5632 (2015).

[15] J. H. Van Vleck. *The Jahn-Teller Effect and Crystalline Stark Splitting for Clusters of the Form XY_6*. *J. Chem*. *Phys*. **7**(1), pp. 72–84 (1939).

[16] W. H. Baur. *The Geometry of Polyhedral Distortions: Predictive Relationships for the Phosphate Group*. *Acta Crystallogr. B* **30**(5), pp. 1195–1215 (1974).

[17] Z. Xiao, W. Meng, J. Wang, D. B. Mitzi and Y. Yan. *Searching for Promising New Perovskite-Based Photovoltaic Absorbers: The Importance of Electronic Dimensionality*. *Mater. Horiz*. **4**(2), pp. 206–216 (2017)

[18] L. Whalley. *effmass: An Effective Mass Package*. *J. Open Source Softw.* **3**(28), pp. 797 (2018).

KURAMOTO OSCILLATORS IN RANDOM NETWORKS

Agostino Funel[1*]

[1]*ENEA, TERIN-ICT-HPC Department, P.le. E. Fermi 1, Portici (Naples), Italy*[12]

ABSTRACT. By means of numerical analysis conducted with the aid of the computer, the collective synchronization of coupled phase oscillators in the Kuramoto model in the connected regime of random networks of various sizes is studied. The oscillators synchronize and achieve phase coherence, and this process is not significantly affected by the level of connectivity of the network. If the probability that two oscillators are coupled is around the network connectivity threshold synchronization still occurs, although in a more attenuated way. If the size of the network is sufficiently large the oscillators have a phase transition.

Introduction

The phenomenon of collective synchronization of an ensemble of interacting entities can be observed in different areas. For example in biology, in processes involving the regulation of the cell cycle *[1, 2]*; in physics, in the study of coupled lasers*[3, 4, 5]* and Josephson junctions*[6]*; but also in human *[7]* and ecological networks *[8]*. One of the first important works on the subject was written by A. Winfree *[9]* in the 1960s who, studying a population of interacting oscillators, discovered that if the natural frequencies of the oscillators were very close, there existed a threshold condition for the coupling constant that, if exceeded, caused a phase transition in which some oscillators suddenly synchronized. Winfree's ideas were taken up by Kuramoto *[10]* who proposed a model whose analytical solution could be studied in the mean-field approximation. The equations of dynamics, for a collection of oscillators, are:

$$\dot{\theta}_\iota = \omega_i + \frac{K}{N}\sum_{j=1}^{N} sin\left(\theta_j - \theta_i\right) (1)$$

The temporal evolution of the phase of each oscillator $\theta_i$depends on its natural frequency ω_i , and on the phases of the other oscillators with which it interacts. If $\omega_i = \omega_j \forall i, j$ the model is called homogeneous, otherwise inhomogeneous. The magnitude of the interaction is given by the coupling constant $K \geq 0$, which is the same for all oscillators. In order to have a geometric representation of the dynamics, it is convenient to imagine that the motion occurs on a circle of unit radius so that the phase of an oscillator is represented by the azimuthal angle defined by its position, see Fig. 1.

[12] Corresponding author. E-mail: first.author@first.institution.

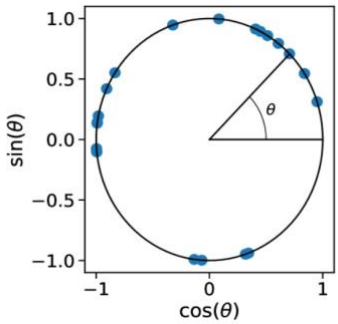


Fig.1: Geometric representation of oscillator dynamics in the Kuramoto model. Motion occurs on a circle of unit radius. The azimuthal angle represents the phase of the oscillator.

Collective motion can be studied by introducing the order parameter

$$Re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} \ (2)$$

The modulus $0 \leq R \leq 1$ measures phase coherence. If the oscillators are scattered along the entire circle then $R \approx 0$ while if they are in phase coherence $\theta_i = \theta_j, \forall i,j$ then $R \approx 1$. The angle ψ is the average phase. If $K = 0$ the oscillators decouple and the phase of each of them evolves according to the equation $\theta_i(t) = \omega_i t + \theta_i(0)$ and synchronization, that is $\dot{\theta}_\iota = \dot{\theta}_j = \dot{\theta}, \forall i,j$ can only occur if $\omega_i = \omega_j \forall i,j$. Note that by summing the members of the

Eq.(1) we obtain $\sum_{i=1}^{N} \dot{\theta}_\iota = \sum_{i=1}^{N} \omega_i$ which implies that if all the oscillators synchronize they will have a frequency equal to the average of the natural frequencies: $\dot{\theta} = \langle\omega\rangle = N^{(-1)} \sum_{i=1}^{N} \omega_i$. In terms of the order parameter Eq.(1) can be written as:

$$\dot{\theta}_\iota = \omega_i + KRsin(\psi - \theta_i)(3)$$

where it is evident how the single oscillator can be thought as interacting with a mean-field with effective coupling constant equal to KR and phase ψ. The numerical analysis of the Kuramoto model shows the existence of a threshold value of the coupling constant below which the oscillators evolve while remaining in a state of phase incoherence; gradually increasing K beyond this threshold the incoherent state becomes unstable, a small nucleus of oscillators is formed which synchronize at a frequency equal to $\langle\omega\rangle$ while the remaining part moves non-uniformly around the unit circle; finally, if K reaches a critical value K_c, collective synchronization begins. Kuramoto proved that $K_c = 2/\pi\, g(0)$ in the limit $N \to \infty$, where $g(\omega)$ is the distribution of natural frequencies, see *[11]* for details. The Kuramoto model refers to a fully (all-to-all) connected network. However, it is interesting to study collective synchronization also in the case of heterogeneous networks, because of this kind are those of particular importance in nature, social interactions and technological infrastructures. Considering the network topology, Eq.(1) becomes:

$$\dot{\theta}_i = \omega_i + (K/k_i)\sum_{j=1}^{N} A_{ij}\, sin(\theta_j - \theta_i) (4)$$

where A_{ij} is the adjacency matrix ($A_{ij} = 1$ if θ_i and θ_j are coupled, 0 otherwise) and k_i is the number of connections of oscillator θ_i. In this paper, only connected and symmetric networks ($A_{ij} = A_{ji}$) are considered. The critical threshold value K_c for the onset of synchronization has been theoretically calculated in the context of heterogeneous degree mean-field and the quenched mean-field theories. The former predicts that K_c is proportional to the ratio between the first two moments of the degree distribution, while the latter that it scales with the inverse of the largest eigenvalue of the adjacency matrix *[12]*. In *[13]* these predictions are compared for scale-free networks. Many other works also investigated the finite-size scaling of the order parameter *[14, 15, 16, 17, 18]*. In this paper we focus on the study of the Kuramoto model for random networks. This type of networks are extensively studied both for their theoretical interest and for their statistical properties that make them a basic model in different application fields. The novelty presented in this paper is the accurate numerical analysis of the synchronization process in the connectivity regime of random networks in the inhomogeneous Kuramoto model. In particular, the numerical results support the conjecture according to which the critical threshold of the coupling probability of oscillators above which they almost certainly synchronize is that of the network connectivity. Furthermore, to the best of our knowledge, this analysis has never been conducted for networks of the size of the largest network analyzed in this paper.

Random networks

The structure of a random network $G(N, p)$ of N nodes, also known as Erdős-Rényi, is determined by the probability $0 \leq p \leq 1$ that a link exists between two randomly chosen nodes. The degree distribution of a random network, which is the probability that a randomly chosen node has degree k, is a binomial $P(k) = \big((N-1)!/(k!\,(N-1-k)!)\big)p^k(1-p)^{(N-1-k)}$ and the average value of the degree is $\langle k\rangle = p(N-1)$. For $p < 1/N$ (subcritical region) the network is mainly composed of isolated nodes, tiny clusters and $\langle k\rangle < 1$. The largest cluster is a tree whose size is $N_G \sim lnN$. For $p = 1/N$ (critical point) $\langle k\rangle = 1$, the network is composed of many small disconnected components, mainly trees, and $N_G \sim N^{(2/3)}$. For $1/N < p \leq ln\,N/N = p_c$ supercritical region) $\langle k\rangle > 1$ and a giant component is formed which contains a finite part of the network and grows rapidly as p increases. However, the network is not yet connected in this region. The probability p_c is called connectivity threshold because the network is supposed to became connected for p just above p_c. For $p > p_c$ one has $\langle k\rangle > lnN$ and $N_G \sim N$. At the end of the connected region $p = 1$and the network becomes fully (all-to-all) connected. See*[20]* for details.

Method

In this work, networks of size $N = 50{,}100{,}500$ and 1000 are considered. For each type of network analyzed 10 simulations are performed and for each of them the distribution of natural frequencies is a Gaussian with $\langle\omega\rangle = 0$ and $\sigma = 0.1$. Since isolated oscillators evolve independently, their dynamics do not affect the collective synchronization of those connected to the network. Therefore, we consider the region with $p > p_c$. To study the limit $p \to p_c$, we set $p = p_c$ as parameter in the numerical library used to randomly generate the networks *[21]*, exclude disconnected configurations and accept only connected ones that correspond to the case in which the giant component absorbs all the nodes of the network and has the minimum number of links. The range of values of the coupling constant examined

in this work is $0 \leq K \leq 1$ and we choose to divide it into 100 steps. For each value of K the system of Eq.(4) is numerically integrated up to 5000 time iterations of amplitude 0.1 s. The order parameter R is calculated by averaging the last 1000 iterations in order to exclude values due to initial transient phase fluctuations. The value of K_c was estimated by finite-size scaling analysis.

Results

To study how the density of links affects synchronization we consider the probability increment $\Delta p = (1 - p_c)/10$ and calculate the order parameter $R(K, p)$ for $p = p_c + n\Delta p$ for $n = 0, \ldots, 10$. The results are shown in Fig. 2.

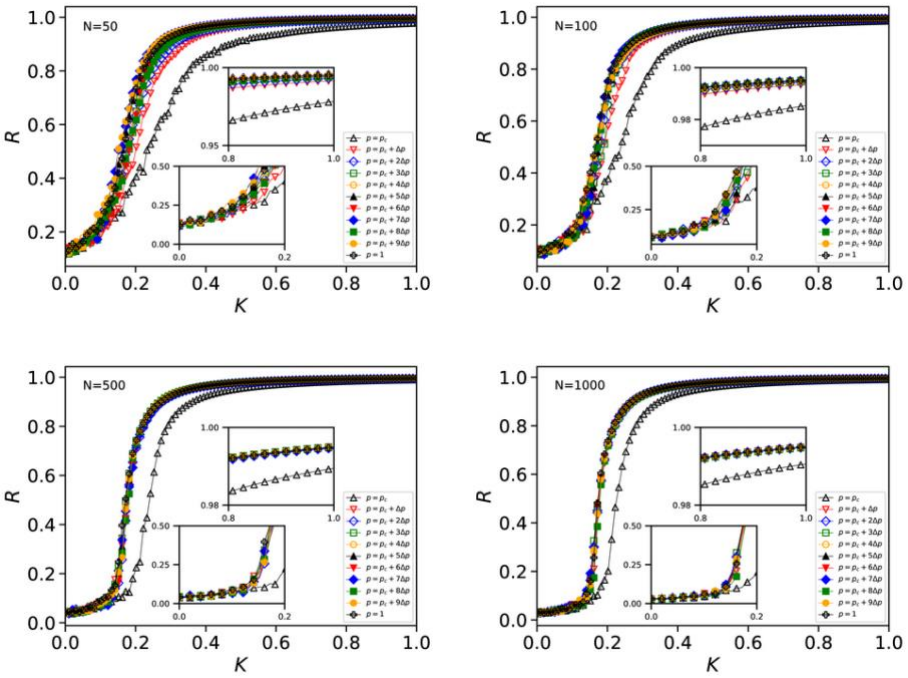


Fig.2: Order parameter R as a function of the coupling constant K for random networks of size $N = 50,100,500,1000$. For each value of K the system of Eq.(4) is numerically integrated up to 5000 time iterations of amplitude 0.1 s. R is calculated by averaging the last 1000 iterations. The figure shows $R(K)$ for different values of the probability p that determines the structure of the network in the connected regime. For $p = p_c \simeq ln\, N/N$ the largest connected component absorbs all the nodes and has the minimum number of links. This is the configuration for which the network has the minimum level of connectivity. For $p = 1$ the maximum level of connectivity is achieved and every node is connected to all the others. The function $R(K)$ behaves roughly the same for $p > p_c$ while for $p = p_c$ it grows more slowly as K increases. The figure shows in more detail the regions where the onset of synchronisation occurs and those where the oscillators are in phase coherence. For $p > p_c$ the value of the critical threshold of the coupling constant for which synchronization occurs is $K_c \sim 0.16 - 0.19$, while for $p = p_c$ it is $K_c \sim 0.21$.

What is observed is that the oscillators synchronize and reach the state of phase coherence for all values of p examined in the connected region. The behavior of $R(K)$ is very similar for all values of $p > p_c$. For $p = p_c$, the system reaches phase coherence, although more slowly. As the size of the network

grows, for $K > K_c$ the transition from state $R \sim 0$ to $R \sim 1$ is more and more sudden, as the shape of $R(K)$ shows, which means that the system has a phase transition. This also happens when $p = p_c$, as shown in more detail in Fig. 3, where the derivative dR/dK is also shown. It can be seen that as N increases the maximum value of dR/dK increases, while its variation around the maximum shrinks.

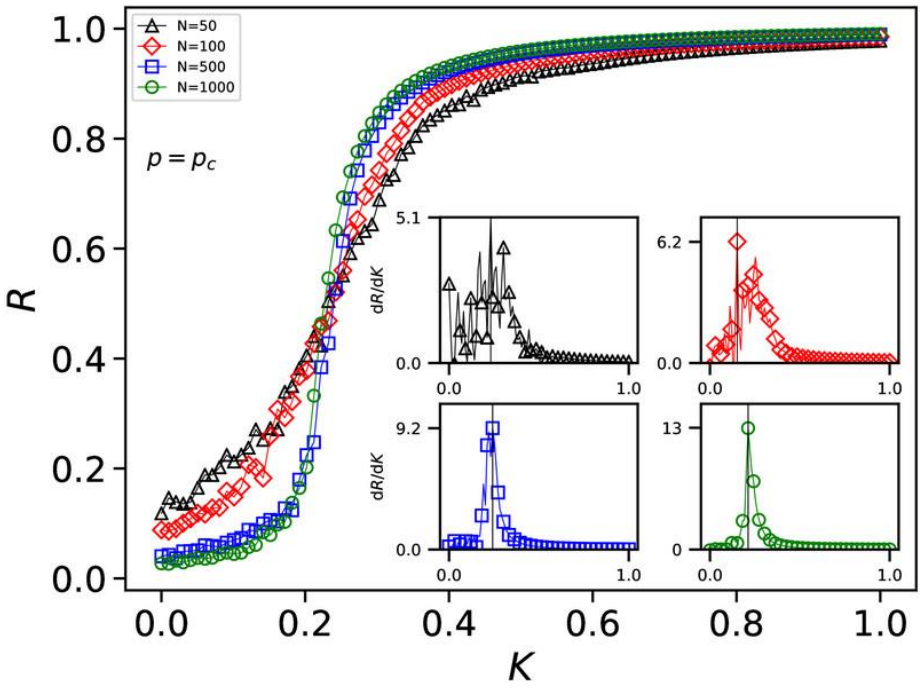


Fig.3: The figure shows $R(K)$ when $p = p_c = ln\, N/N$ for all analyzed networks. As the size N of the network increases the transition from $R \sim 0$ to $R \sim 1$ becomes abrupt and the oscillators have a phase transition. This is also evidenced by the inner figures which show that as N increases the derivative dR/dK has a larger maximum and the variation around it becomes smaller.

To estimate K_c we consider that for $K < K_c$ the oscillators do not synchronize and R decays around a value of $O\left(1/\sqrt{N}\right)$. For $K > K_c$ R increases and saturates at a value $R_\infty < 1$, although with fluctuations of $O\left(1/\sqrt{N}\right)$. Therefore K_c can be estimated numerically considering that at the onset of synchronization $R(N, K_c)$ has the smallest variation with N. Fig. 4 shows $R(N, K)$ for all values of p examined in this work. For $p > p_c$ we have that $K_c \sim 0.16 - 0.19$. This suggests that the onset of synchronization does not depend too much on the level of connectivity of the network itself. The theoretical value predicted by Kuramoto is $K_c \sim 0.16$. For $p = p_c$ the best estimate is $K_c \sim 0.21$, which confirms that for this network configuration the onset of synchronization is slower.

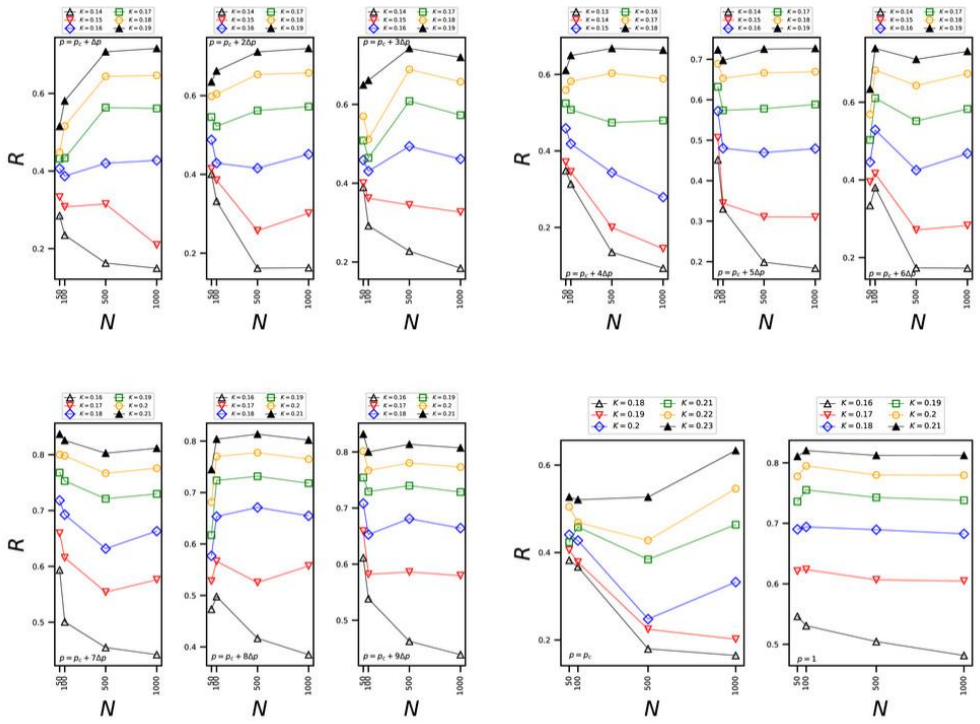


Fig.4: The figure shows $R(N,K)$ for all values of p. For $K < K_c$, R decays towards a residual of $O\left(1/\sqrt{N}\right)$. For $K > K_c$, R saturates towards a value $R_\infty < 1$ with fluctuations of $O\left(1/\sqrt{N}\right)$. For $K = K_c$ the variation $R(N,K_c)$ is minimal. The best estimate is $K_c \sim 0.16 - 0.19$ for $p > p_c$, and $K_c \sim 0.21$ for $p = p_c$.

Discussion

The conjecture that the probability threshold above which synchronization is highly probable is $p_c = ln\,N/N$ has been considered by several authors. In *[22]*, in the context of optimization of nonconvex cost functions, authors proved, for the homogeneous Kuramoto model, that if the oscillators synchronize with high probability for $N \to \infty$. In *[23]*, with arguments from spectral graph theory, authors proved that the same occurs if $p \gg (lnN)^2/N$. The issue has also been addressed in the context of graphon theory *[24, 25, 26, 27, 28]*. In *[29]* authors proved, still for the case of the homogeneous model and for $N \to \infty$, that the oscillators reach phase synchronization with high probability if $p > p_c$. They performed numerical simulations for networks of size $N \leq 100$. In this work the above conjecture has been proved numerically even in the more general case of inhomogeneous Kuramoto model and for larger network sizes.

Conclusion

We have numerically studied the collective synchronization of coupled phase oscillators in the Kuramoto model in the connected regime of random networks of sizes $N = 50,100,500,1000$. The oscillators synchronize and reach phase coherence for all connectivity levels of the networks. The behavior of the order parameter R as a function of the coupling constant K does not show significant discrepancies between different levels of link density. For the limiting case in which oscillators are coupled with probability around the connectivity threshold of the network, synchronization still occurs, although in a more attenuated way. As the size of the network increases, the transition from the incoherent to the coherent phase state occurs more and more abruptly and the oscillators have a phase transition. The critical threshold value for the onset of synchronization varies in the range $K_c \sim 0.16 -$

0.19 for all the connectivity levels of the networks. In the limiting case when the network goes from disconnection to connection $K_c \sim 0.21$. A random network of a certain size can be obtained by perturbing a complete network of the same size by randomly eliminating links. Therefore, from this analysis it can be deduced that a random perturbation of a complete network of oscillators coupled according to the Kuramoto model does not affect the synchronization as long as the perturbation does not disconnect the network.

Acknowledgments

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff *[30]*. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see http://www.cresco.enea.it/english for information.

References

[1] Banfalvi G 2011 Methods Mol Biol. 761 1

[2] Manukyan A, Abraham L, Dungrawala H and Schneider B L 2011 Methods Mol Biol. 761 173

[3] Ziping J and McCall M 1993 J. Opt. Soc. Am. 10 155

[4] Kourtchatov S Y, Likhanskii V V, Napartovich A P, Arecchi F T and Lapucci A 1995 Phys. Rev. A 52 4089

[5] Takemura N, Takata K, Takiguchi M and Notomi M 2021 Sci. Rep. 11 8587

[6] Wiesenfeld K, Colet P and Strogatz S H 1998 Phys. Rev. E 57 15631

[7] Shahal S, Wurzberg A, Sibony I, Duadi H, Shniderman E, Weymouth D, Davidson N and Fridman M 2020 Nat. Commun. 11 3854

[8] Vandermeer J, Hajian-Forooshani Z, Medina N and Perfecto I 2021 R. Soc. Open Sci. 8 210122

[9] Winfree A T 1967 J. Theoret. Biol. 16 15

[10] Kuramoto Y 1984 Chemical Oscillations, Waves and Turbulence vol 19 (Berlin: Springer)

[11] Strogatz S H 2000 Physica D: Nonlinear Phenomena 143 1

[12] Rodrigues F A, Peron T, Ji P and Kurths J 2015 Physics Reports 610 1

[13] Peron T, Messias B, Mata A S, Rodrigues F A and Moreno Y 2019 Phys. Rev E 100 042302

[14] Moreno Y and Pacheco A F 2004 Europhysics Letters 68 603

[15] Lee D-S 2005 Phys. Rev E 72 026208

[16] Hong H, Park H and Tang L-H 2007 Phys. Rev E 76 066104

[17] Hong H, Um J and Park H 2013 Phys. Rev E 87 042105

[18] Um J, Hong H and Park H 2014 Phys. Rev E 89 012810

[19] English L Q 2007 European Journal of Physics 29 143

[20] Barabási A 2016 Network science (Cambridge University Press)

[21] Hagberg A A, Schult D A and Swart P J 2008 Proc. of the 7th Python in Science Conference (SciPy2008) (Pasadena, California) p 11 doi: 10.25080/PFVC8793

[22] Ling S, Xu R and Bandeira A S 2019 SIAM J. Optim. 29 1879

[23] Kassabov M, Strogatz S H and Townsend A 2022 Chaos: An Interdisciplinary Journal of Nonlinear Science 32 093119

[24] Lovász L 2012 American Mathematical Soc. 60

[25] Borgs C, Chayes J, Cohn H and Zhao Y 2019 Transactions of the American Mathematical Society 372 3019

[26] Medvedev G S 2014 SIAM Journal on Mathematical Analysis 46 2743

[27] Medvedev G S 2014 Archive for Rational Mechanics and Analysis 212 781

[28] Medvedev G S 2019 Communications in Mathematical Sciences 17 883

[29] Nagpal S V, Nair G G, Strogatz S H and Parise F 2024 Synchronization in random networks of identical phase oscillators: A graphon approach arXiv:2403.13998 [math.DS]

[30] Iannone F et al2019 Proc. Int. Conf. on High Performance Computing Simulation (HPCS) (Dublin, Ireland. Publisher: IEEE) p 1051–1052 doi: 10.1109/HPCS48598.2019.9188135

HIGH-FIDELITY NUMERICAL SIMULATIONS OF HYDROGEN AND AMMONIA-HYDROGEN FLAMES

Davide Laera, Francesco Gabriele Schiavone, Raffaele Intranuovo, Nicola Misceo, Sergio Mario Camporeale, and Marco Torresi[13]

[1]Department of Mechanics, Mathematics and Management, Polytechnic University of Bari, Via Orabona 4, Bari 70125, Italy

Abstract. Numerical simulations have been performed on CRESCO 7 using the explicit massively parallel compressible Navier-Stokes solver AVBP 7.15 to investigate hydrogen and ammonia-hydrogen flames. The first paragraph treats flame-turbulence interaction in partially cracked ammonia flames under different levels of hydrogen content and turbulence intensities. The second one deals with the investigation of noise emission in laminar hydrogen flames via fully-resolved numerical simulations, to investigate flame-wall and flame-flame interaction.

Impact of thermodiffusive effects on the turbulent flame speed in partially cracked ammonia flames

Ammonia (NH_3) combustion is considered a promising solution to reduce anthropogenic carbon emissions, but it poses several technological issues due to the low reactivity of NH_3 [1]. To solve these issues, two strategies are commonly used: (1) doping ammonia with other more reactive fuels and (2) partial cracking of ammonia molecules (PCA). This second solution, which is the focus of this study, is particularly interesting, being easier to implement it in practical combustors. High-fidelity numerical simulations are required to complement experiments and gain an understanding of such complex phenomena. Therefore, this work aims to improve the knowledge in ammonia combustion for both low and high Karlovitz number (*Ka*) combustion regimes, focusing on flame-turbulence interaction, thermodiffusive (TD) effects and their impact on turbulent flame properties. Specific attention is given to turbulent flame speed to gain an understanding of the mechanism behind the peculiar "bending effect" found in [2], i.e., the decrease of turbulent flame speed at high level of hydrogen content in the fuel mixture.

A set of DNS of planar ammonia-air flames stabilized in a three-dimensional (3D) box-shaped numerical domain submitted to Homogeneous Isotropic Turbulence (HIT) is performed with the AVBP code (https://cerfacs.fr/avbp7x). The HIT field is forced within the computational domain following the linear forcing model provided by Bassenne et al. [3]. The grid resolution is defined to discretize the flame front with at least 10 points, and ensure the solution of the smallest turbulent scales by applying the Pope criterion. An Analytically Reduced Chemistry mechanism [4] is adopted to solve the ammonia-hydrogen-air chemical reactions. Based on the experimental work by Zitouni et al. [2], simulations are performed considering three levels of ammonia cracking γ=5%, 20% and 40%, one level of turbulent intensity u'=1.13 m/s, and one value of equivalence ratio ϕ=0.8. These operating conditions are chosen to investigate multiple flame-turbulence interaction regimes.

Figure 1 shows the flame fronts obtained for the three levels of ammonia cracking. It is evident that the effects of turbulence-flame interaction and flame dynamics impact the flame surface, with a stronger wrinkling observed for a higher hydrogen content, due to the impact of thermodiffusive (TD) effects. This has an influence also on the turbulent flame speed S_T.

[13] Corresponding author. E-mail: marco.torresi@poliba.it

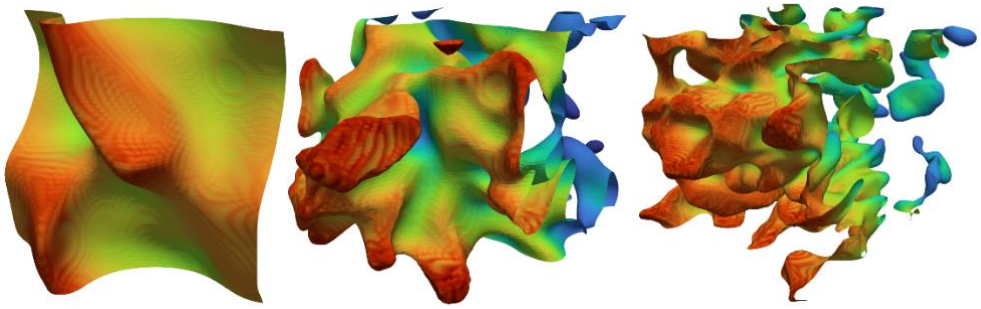


Fig.1: Isosurfaces of progress variable *c* at the value of maximum heat release rate in the corresponding unstretched laminar 1D flame for the flame at γ=5% (left), 20% (centre) and 40% (right).

In fact, as illustrated in Fig. 2, it is evident that for γ=5% the S_T increases predominantly because of turbulence, which tends to augment the flame surface area A_T, rather than the TD effects, which vary the local flame structure (e.g., local reactivity) [5]. These are represented by the stretch factor I_0, which remains close to unity in this case. In contrast, for γ=20%, I_0 is greater than this value. Thus, the combined effect of enhanced H_2 and low *Ka* is such that TD instabilities are more pronounced than in the high *Ka* case, resulting in a more wrinkled flame surface and in a more evident deviation from the 1D flamelet behavior, as can be seen in Fig. 3. Consequently, in the latter case, turbulence and TD effects both impact S_T. These phenomena are even more evident in γ=40% case. In fact, here, most of the flame burns at values of φ that are leaner than the target (Fig. 3): the flame burns locally at slower conditions, resulting in the bending effect on S_T.

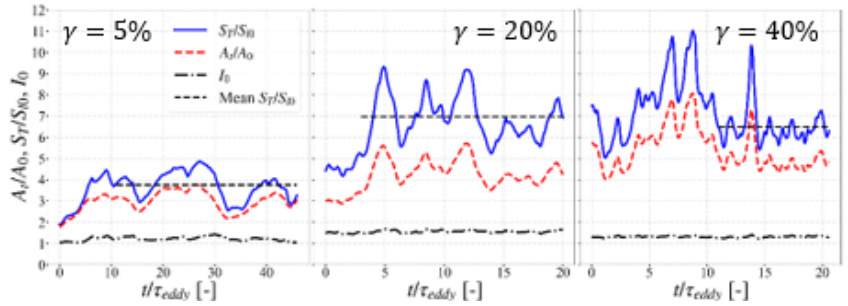


Fig. 2: Temporal evaluation of the stretch factor I0, normalized turbulent flame speed ST/Sl0 and flame surface area AT/A0 ($\tau_{eddy} = 2.3 \cdot 10^{-3} s$).

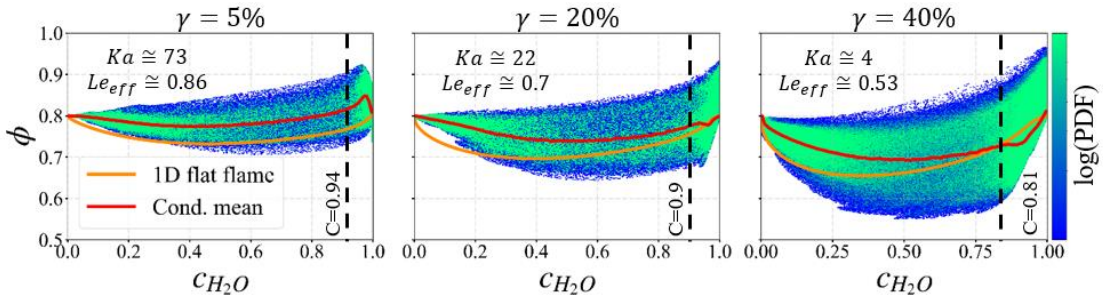


Fig. 3: Scatter plots of ϕ and progress variable *c*. Vertical lines represent the value of *c* in which Heat Release Rate is maximum in the 1D flat flame.

This work has been published in: "R. Intranuovo, F.G. Schiavone, D. Laera, "*DNS study of the impact of thermodiffusive effects on the turbulent flame speed in partially cracked ammonia flames*". In Proceedings of 47th Meeting of the Italian Section of the Combustion Institute, ISBN 978-88-88104-30-0."

Investigation of flame quenching and acoustic emission of premixed laminar methane-air and hydrogen-air flames

Combustion noise is a significant issue for engineering combustion systems, which are subject to stringent noise reduction goals [6]. Moreover, the interaction between pressure waves generated by the flame and the acoustics of the surrounding confinement can lead to thermoacoustic instabilities, a well-known issue that can cause severe mechanical failures [7]. Given hydrogen (H_2) distinct physicochemical properties, including enhanced reactivity with respect to conventional fuels, preferential diffusion and thermodiffusive effects [5], a more in-depth investigation into the sound generation mechanisms in H_2 flames is required. In the literature [6], flame-generated noise has been linked to unsteady heat release rate (HRR) processes, involving creation or destruction of flame fronts. This can occur due to different mechanisms. This work focuses on flame-wall interaction and flame-flame interaction. One-dimensional (1D) simulations are carried out in the AVBP code using a canonical Head-On Quenching (HOQ) configuration, consisting of a flame traveling towards a rigid wall. The grid size is set to 2 μm in the near-wall region, and then smoothly stretched towards the outlet, considered as non-reflecting and modelled with the NSCBC formalism. Pressure signals are extracted from a far-field probe (P_{ff}). A parametric study is performed to explore the effects of fuel (H_2 and CH_4), equivalence ratio, and wall thermal condition (adiabatic and isothermal). The chemical mechanisms employed include the Analytically Reduced Chemistry scheme CH4_15_256_9_AP [8] for CH4 flames, and the San Diego mechanism [9] for H_2 flames. The 'HEAT_LOSS' wall boundary condition is set with a thermal resistance value (R_w = 1.0e-4 m^2K/W), effectively making the wall isothermal, as confirmed by additional verification. This condition is representative of a flame-wall interaction, while the adiabatic wall is equivalent to a symmetry condition and thus represents a flame-flame interaction.

Figure 4 presents the time evolution of the domain-averaged heat release rate variation (solid line) and the acoustic pressure (dashed line) recorded at the far-field probe for CH_4 (left) and H_2 (right) flames featuring the same laminar flame speed (S_L). For each fuel, simulations with adiabatic (blue curves) and heat-loss (yellow curves) wall boundary conditions are shown. In methane flames, heat release rate (HRR) variation differs markedly between adiabatic and heat-loss cases, both in waveform and fluctuation intensity – sharper in the former, smoother in the latter – due to distinct quenching mechanisms: reactant depletion in the adiabatic case and thermal dissipation in the heat-loss case, occurring before wall contact [10]. The adiabatic case shows a positive HRR peak, indicating superheating of the unburnt mixture and increased reactivity before extinction, whereas in the heat-loss case, thermal exchange governs a slower HRR decay, with a time derivative at least one order of magnitude lower. Differently, in lean hydrogen flame (with $Le < 1$), HRR evolution for adiabatic and heat-loss cases are nearly superimposable, indicating a similar dominant extinction mechanism governed by thermo-diffusive effects regardless of the wall thermal conditions: the unequal diffusion rates of temperature and the deficient reactant (fuel) cause fuel to keep diffusing toward the flame front as it nears the wall. This leads to combined depletion by reaction and diffusion, resulting in gradual extinction before wall contact. This third mechanism dominates the others, rendering quenching of lean H_2 flames independent of wall heat losses. For same laminar flame speed, CH_4 flames exhibit higher thermal power and generate stronger SPL upon quenching. The peak sound pressure level (SPL) generated by flame quenching is compared with the predictive model proposed by Talei et al. [10]:

$$|p'| = \rho_u c_b \left(1 - \frac{T_u}{T_b}\right) S_L$$

The Talei model accurately predicts the peak SPL in CH4 cases with adiabatic walls (thus also for FFI), but it performs less accurately for H_2 flames and in heat-loss configurations.

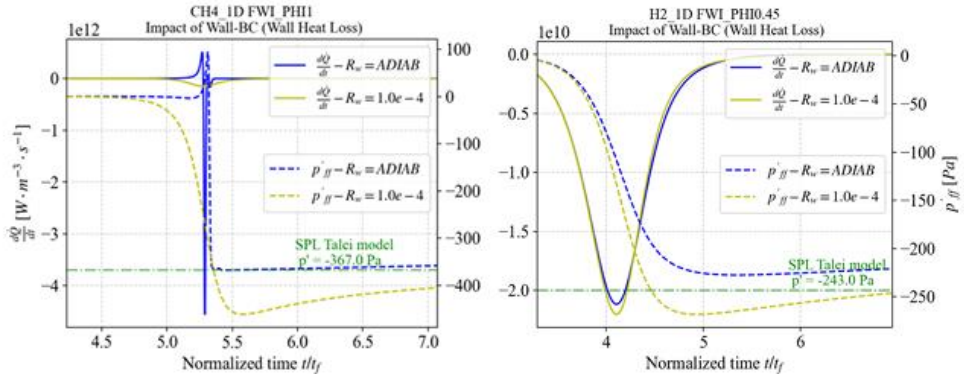


Fig. 4: HRR variation (solid lines) and pressure wave generated (dashed lines) for CH_4 (left) and H_2 (right) flames employing ARC chemistry, with same laminar flame speed (S_L). Adiabatic (blue curves) and heat-loss (yellow curves) wall boundary conditions are presented for each fuel.

Figure 5 summarizes the impact of equivalence ratio and thermal wall conditions for the two fuels. At same equivalence ratio, the hydrogen flames have a stronger sound pressure peak, due to the higher heat release rate and laminar flame speed. In the presence of an isothermal wall, the emitted noise is higher than in the case of adiabatic wall, despite its slower dynamics. Furthermore, for lean hydrogen flames, as observed before, the differences between adiabatic and isothermal wall conditions are smaller than for stoichiometric and rich flames.

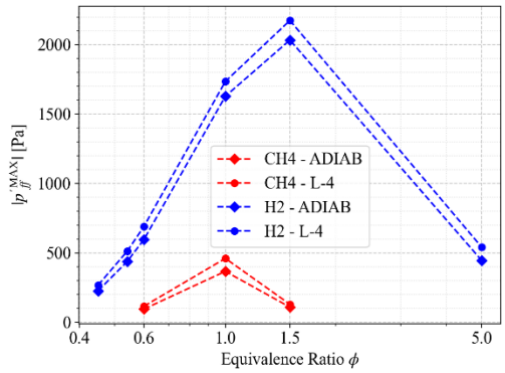


Fig. 5: Sound pressure peaks for hydrogen and methane flames at different equivalence ratios under adiabatic (ADIAB) and isothermal (L-4) wall conditions.

This work has been published in: "N. Misceo, F.G. Schiavone, S. M. Camporeale, M. Torresi, D. Laera, "*flame quenching mechanisms and noise generation in premixed methane and hydrogen flames*". In Proceedings of 47th Meeting of the Italian Section of the Combustion Institute, ISBN 978-88-88104-30-0."

The study is being extended to two-dimensional flames, in order to investigate the aforementioned flame-quenching mechanisms in the presence of thermodiffusive instabilities, which cannot be observed in a 1D configuration. Preliminary results are reported in Fig. 6, which reports the instantaneous pressure field resulting from the flame-wall interaction of a thermodiffusively unstable lean hydrogen-air flame. It can be observed the radiation of acoustic wave from the interaction between flame sheets in the presence of cellular structures.

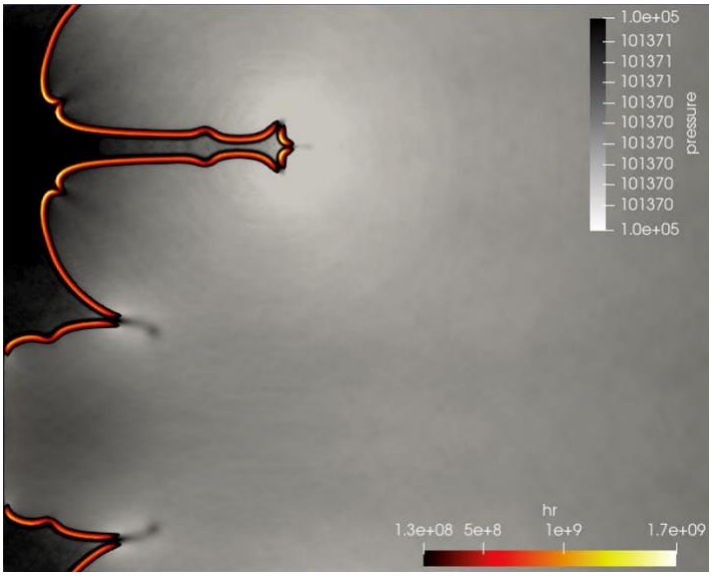


Fig. 6: Two-dimensional simulation of a lean hydrogen-air flame impacting a wall (on the left): acoustic field and heat release rate.

References

[1] Valera-Medina, A., Amer-Hatem, F., Azad, A.K., Dedoussi, I.C., De Joannon, M., Fernandes, R.X., Glarborg, P., Hashemi, H., He, X., Mashruk, S. and McGowan, J., "Review on ammonia as a potential fuel: from synthesis to economics", *Energy & Fuels*, *35*(9), pp.6964-7029, (2021).

[2] Zitouni, S., Brequigny, P. and Mounaïm-Rousselle, C., "Turbulent partially cracked ammonia/air premixed spherical flames" *Fuel Communications*, 20, p. 100126, (2024).

[3] Bassenne, M., Urzay, J., Park, G.I. and Moin, P., "Constant-energetics physical-space forcing methods for improved convergence to homogeneous-isotropic turbulence with application to particle-laden flows", *Physics of Fluids*, *28*(3), (2016).

[4] Ruiz, H.V., Laera, D., Lartigue, G., Mashruk, S., Valera-Medina, A. and Gicquel, L., "Extension of the dynamic Thickened Flame model for partially-premixed multi-fuel multi-injection combustion and application to an ammonia–hydrogen swirled flame" *Combustion and Flame*, 274, p. 113992, (2025).

[5] Tingas, E.A. (Ed.), "Hydrogen for Future Thermal Engines", *Springer Nature*, pp. 93-265, (2023).

[6] Dowling, A. P., Mahmoudi, Y. "Combustion noise", *Proceedings of the Combustion Institute* 35, pp. 65-100, (2015).

[7] Schuller, T., Poinsot, T., & Candel, S., "Dynamics and control of premixed combustion systems based on flame transfer and describing functions". *Journal of Fluid Mechanics* 894, p. P1, (2020).

[8] Pestre, A., PhD Thesis (2023).

[9] Saxena, P., Williams, F. A. "Testing a small detailed chemical-kinetic mechanism for the combustion of hydrogen and carbon monoxide", *Combustion and Flame* 145, pp. 316-323, (2006).

[10] Talei, M., Brear, M. J., Hawkes, E. R., "Sound generation by laminar premixed flame annihilation". *Journal of Fluid Mechanics*, 679, pp. 194-218, (2011).

THE BOUND STATES OF NANOLUC COMPLEXES WITH THE KETO- AND ENOLIC- FORM OF COELENTERAZINE AND FURIMAZINE VIA HAMILTONIAN REPLICA EXCHANGE SIMULATIONS

Marina Macchiagodena[1], Piero Procacci[1*]

[1]*University of Florence, Chemistry Dept. U. Schiff, Via della Lastruccia 3, 50019 Sesto Fiorentino, Italy*

ABSTRACT. The recent engineering of the luminescent NanoLuc (Nluc) protein from Oplophorus luciferase has revolutionized biological research due to its small size, extreme brightness, and stability. Nluc is an invaluable tool for high-throughput drug discovery, particularly in NanoBRET assays, which efficiently assess drug-promoted protein-protein interactions via Bioluminescence Resonance Energy Transfer (BRET). The bioluminescence mechanism relies on the oxidative decarboxylation of the luciferin substrate (either coelenterazine or furimazine). Although the reactants and products of this reaction are well characterized, the structure of the short-lived enzyme-substrate complex remains unknown due to the rapid oxidation that occurs upon binding. In this study, we performed extensive Hamiltonian Replica Exchange molecular dynamics simulations on the CRESCO6 HPC platform to provide a thorough structural characterization of the Nluc-substrate complex.

Background

NanoLuc (Nluc) is a novel luminescent protein engineered via directed evolution from the luciferase of the deep-sea opossum shrimp (Oplophorus gracilirostris) by a team at Promega Corporation led by Mark Hall.[1] It has brought exquisite sensitivity to a wide range of biochemical applications. The NanoLuc platform has gained enormous popularity for bioluminescence assays due to its enhanced brightness, approximately 150-fold higher than that of standard firefly luciferase.[2]
Unlike most known luciferases, Nluc is characterized by its small size (19 kDa), remarkable thermal and pH stability, unbiased distribution within cells, and independence from ATP for activation.
Nluc luminescence is produced through the oxidation of a luciferin substrate bound within the protein's internal pocket, followed by light emission. It produces significantly brighter luminescence with a specific synthetic substrate, furimazine (FMZ), compared to the natural and less efficient luciferin, coelenterazine (CTZ), common in many marine organisms. The structures of the Nluc complexes with either coelenterazine or furimazine have never been resolved by X-ray crystallography, cryo-EM, or NMR. This is because, upon entering the binding pocket, the substrate undergoes an immediate oxidative reaction, producing coelenteramide or furimamide, which are rapidly released into the solvent.
Recently, a study in Nature Communications hypothesized an activation-deactivation mechanism regulated by the mutually exclusive binding of FMZ or CTZ to either an intrabarrel catalytic site or an allosteric site on the protein surface.[3] This work suggested a cardinal role for the intrabarrel residue ARG162, which interacts strongly with the substrate in the catalytic site. However, these speculations were based primarily on static data from crystal structures and molecular docking, supplemented by "epochs" of standard MD simulations using an adaptive sampling algorithm[4] driven by root-mean-square deviation (RMSD) metrics relative to crystal structures.
In this report, we further characterize the salient structural features of the Nluc-FMZ and Nluc-CTZ complexes. Our approach utilizes molecular docking followed by extensive and *unbiased* Hamiltonian Replica Exchange (HREX) simulations of the enolic and ketonic tautomers of FMZ and CTZ, initially positioned within the intra-barrel catalytic site of Nluc in both its open and closed conformations. The

HREX protocol is defined by a "hot-zone" encompassing the ligand and all residues lining the catalytic site. As demonstrated in Ref. [5], the conformational sampling of these complexes benefits from a combination of MD replicates and HREX sampling, whereby multiple concurrent HREX simulations are launched within a single parallel job on the CRESCO6 HPC platform[6] engaging hundreds of cores.
The analysis from these simulations is ongoing and far from complete. However, to ensure transparency and enable independent scrutiny and validation, we have uploaded the PDB trajectories of the HREX-sampled target states for all eight simulated systems (encompassing both protein conformations and both tautomers of each ligand) to a public Zenodo repository.

Methods

Structural Preparation and Molecular Docking: The open (active) structure of taken from the Protein

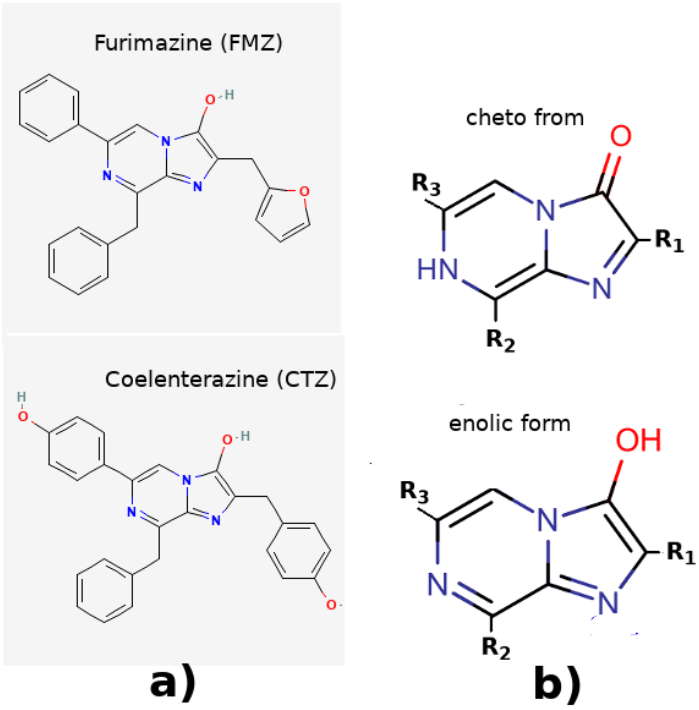


Fig. 1: a) Chemical structure of furimazine and coelenterazine; b) keto-enolic tautomerism of the central imidazo-pyrazine core

data bank (PDB 5b0u) corresponding to the *apo* form of the molecule. The closed (inactive) structure was taken from the PDB entry 8aq6, removing the bound furimamide on the allosteric site on surface. The 3D structures of the FMZ and CTZ ligands were obtained by converting to PDB format the SMILES codes of the ketonic and enolic forms using the OpenBabel program. The two ligands in their enolic form are shown in Figure 1a. Keto-enolic tautomerism in the central and reactive imidazo[1,2-a]pyrazine moiety is illustrated in Figure 1b. Both tautomers can exist at physiologiocal pH, based on pKa ranges obtained using the Chemicalize online platform by ChemAxon. The ground state energies of the two tautomers of the imidazo[1,2-a]pyrazine core common to FMZ and CTZ, computed with the Gaussian program at the B3LYP/cc-pVTZ level of theory, differ by only 0.4 kcal/mol, implying that the two species can easily interconvert in physiological conditions. While the ketonic specie is generally believed to be the active specie[7] the oxidative decarboxylation can occur also for the enolic form. Due to the lack of a firm experimental validation for the structure of the complexes, both forms were considered in this study.
Preliminary molecular docking calculations was performed using the Autodock4 and Autodock Vina software. For the open (5b0u) and closed (8aq6) forms, docking was performed by centering a cubic

docking box approximately at the intra-barrell catalytic sites of the A chain. The coordinates of the box center are set to at x=23Å , y=29 Å, z=41 Å for 5b0u, and at x=15 Å, y=-40 Å, z=15 Å for 8aq6. The sidelenght of the box was set to 22 Å, thus affording the screening of ligand poses in the catalytic site as well as and on the protein surface.
For the open form, the side chains of residue Leu22, Tyr109, Phe110, Tyr114, Ile137, Asp139, Phe151, Arg162, Tyr94, Val15, all known to be implied in the binding[3], were considered as flexible. In the closed conformation, the exposed Tyr94 is kept rigid; its role is replaced by the flexible side chaincores of His93, which is oriented towards the interior of the catalytic pocket.[3]
HREX simulations: Molecular dynamics (MD) simulations were initiated from the best binding poses identified in the docking stage for each of the eight systems (2 protein conformations × 2 ligands × 2 tautomers). Each Nluc complex was solvated in a cubic water box with a side length of 60 Å (open form) or 63 Å (closed form) using TIP3P water molecules at standard density. The AMBER99SB-ILDN force field was used for the protein. Potential parameters for the ketonic and enolic tautomers of the ligands (Fig. 1) were derived using the PrimaDORAC web application.[8] Long-range electrostatic interactions were treated using the Particle-Mesh Ewald (PME) method. The equations of motion were integrated using a multiple time-step r-RESPA propagator optimized for biological systems.[9] Each system underwent a 100 ps preliminary equilibration in the NPT ensemble (P=1 atm, T=300 K) on a local workstation, employing a Parrinello-Rahman barostat and a Nosé-Hoover thermostat chain.PME) method.
Production simulations were conducted using Hamiltonian Replica Exchange (HREX) molecular dynamics on the CRESCO6 cluster for all eight complexes, maintaining the potential parameters, PME, and thermostat/barostat settings described above. The HREX protocol implemented a solute-tempering approach. A "solute" hot-zone was defined to include the ligand and all residues lining the intra-barrel catalytic pocket. Across the replica progression, only the intramolecular potential of this solute was scaled, with a minimum scaling factor of 0.15 for the torsional potential (equivalent to a reference temperature of 3000 K) and of 0.3 (equivalent to a temperature of 1500 K) for the intrasolute nonbonded interactions. This approach enables enhanced conformational sampling of the ligand and its binding pocket residues while the remainder of the protein and solvent environment remains at 300 K.
For each complex, a single job was launched on CRESCO6, comprising 7 concurrent, independent HREX simulations. Each HREX simulation employed 16 replicas with a solute temperature progression defined by the standard scaling for harmonic oscillators.[5] This setup engaged 672 cores in total. Each independent HREX simulation was run for 4 ns, producing 28 ns of sampled data per complex in the target state (~0.5 microseconds in the extended ensemble) with a wall-clock time of ~20 hours. All simulations were done on the CRESCO6 cluster[6] managed by ENEA in Portici using the program ORAC.[10]

Results and discussion

Docking results

The reliability of AutoDock Vina for pose prediction was assessed using the known X-ray structure of the NanoLuc-D9R/H57A/K89R mutant in complex with azacoelenterazine (a-CTZ) bound at the intra-barrel catalytic site (PDB: 8BO9).[7] The a-CTZ compound is a strict (non reactive) analog of native coelenterazine (CTZ), wherein the carbon atom adjacent to the carbonyl moiety on the imidazole ring is replaced by nitrogen.

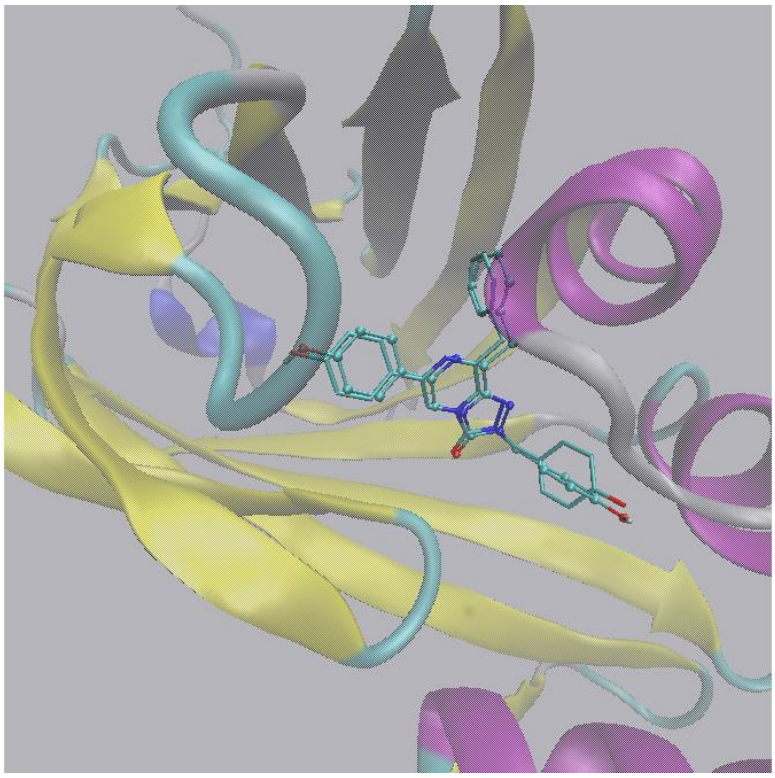


Fig. 2*:* Predicted (in CPK representation) and experimental (bond representation) binding pose of a-CTZ in the intra-barrel pocket of D9R/H57A/K89R Nluc.

Table 1: Autodock-Vina binding affinities (in Kcal/mole) of the FMZ/CTZ Nluc complexes

Nluc form	**Open (active)**		**Closed (inactive)**	
Tautomer	keto	enolic	keto	enolic
FMZ	-9.86	-9.98	-9.03	-9.19
CTZ	-10.63	-9.72	-10.29	-9.13

As shown in Figure 2, the pose predicted by AutoDock Vina is in close agreement with the experimental conformation. The predicted pose is nearly superimposed onto the crystallographic orientation, differing only by a tilt of one of the phenolic rings.

In Table 1, we report the binding affinities obtained with AutoDock-Vina for the eight Nluc complexes examined. While the CTZ ketonic form exhibits a stronger binding affinity than the enolic form in both the open and closed structures of Nluc, in the case of FMZ the enolic form binds slighly stronger than the ketonic tautomer. Furthermore, the binding affinity for the two tautomers is slightly stronger in the open structure than in the closed one. Docking results show that Nluc, whether in its open or closed form, can indeed hold the luciferins in both the ketonic and enolic state. Moreover, allowing side chain flexibility for the intra-barrel residues in the apo 5B0U structure revealed multiple, diverse ligand docking poses with similar binding affinities across all eight cases. These results imply that the conformational landscape of Nluc complexes in solution is more intricate than typically considered.

HREX results

Prior to analyzing the structure of the eight Nluc complexes, we conducted two standard molecular dynamics simulations of 150 ns on the unligated *apo* closed and open structures. These simulations were performed to assess the shape and volume modulation of the internal and external pockets of the monomers in aqueous solution under standard conditions.

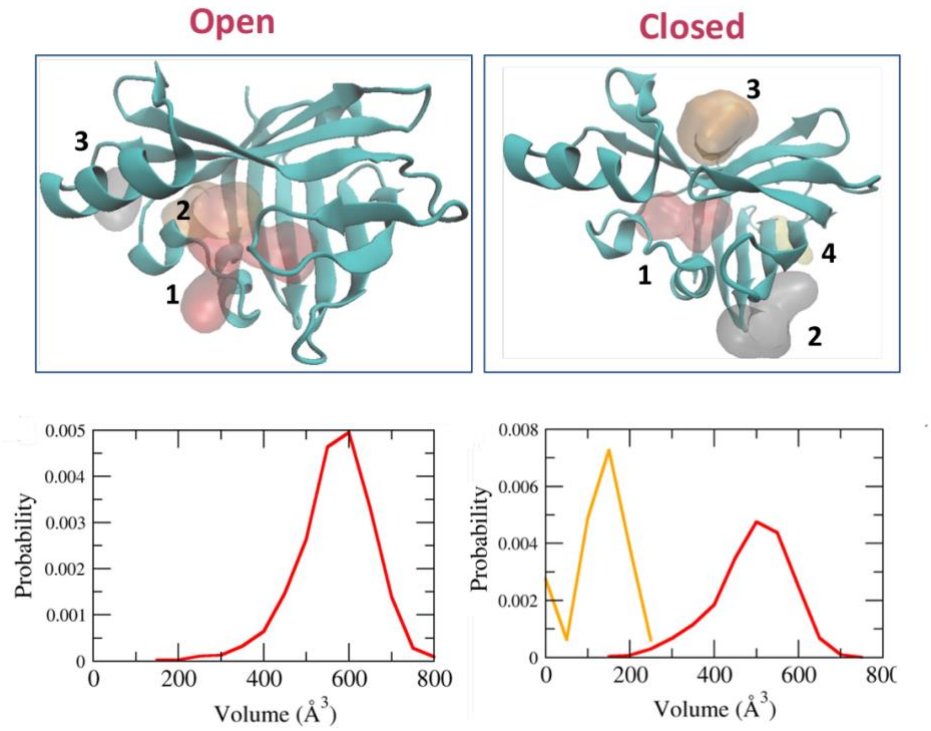


Fig. 3*:* Fpocket results for the open (left) and closed structure of Nluc (right)

The resulting trajectories were analyzed using the Fpocket software [LeGuilloux2009], with the results presented in Figure 3. In the open structure, the intra-barrel cavity (numbered 1, light red) is prominent. In the closed structure, the analogous intra-barrel cavity (also numbered 1, light red) is significantly shrunk, while an allosteric site (numbered 3, light orange) appears at the surface location previously identified in Ref. [3]. The volume distributions for cavity 1 in the open structure and cavities 1 and 3 in the closed form were evaluated using 200 snapshots taken at regular intervals from the 150 ns trajectories. For the open form, the volume probability distribution of the intra-barrel cavity 1 is markedly left-skewed, peaking at approximately 600 Å^3 and extending up to 800 Å^3. In the closed form, the intra-barrel cavity (cavity 1) exhibits a similar volume range but its distribution is shifted, now peaking at ~500 Å^3. This represents a 17% decrease in the average cavity volume compared to the open form. Notably, in both conformations, the size of the intra-barrel cavity remains sufficiently large to easily accommodate CTZ and FMZ, whose molecular volumes are on the order of 230–260 Å^3. In contrast, the surface allosteric site (cavity 3) in the closed form displays a narrower volume distribution (peaked below 200 Å^3) suggesting a relative rigidity of its constituent residues.

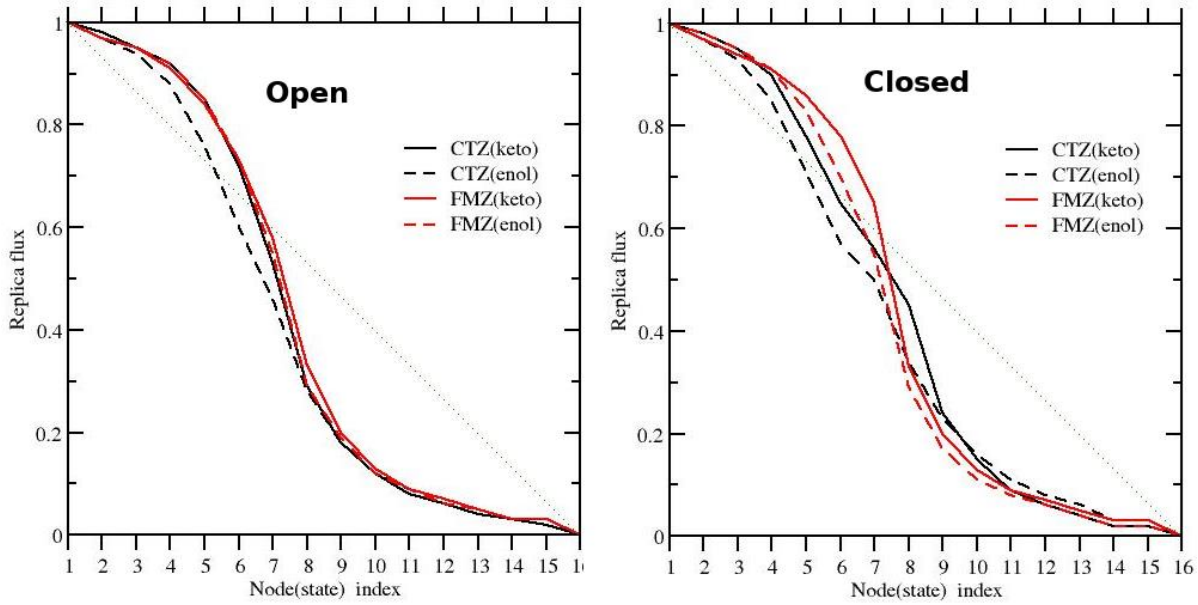


Fig. 4*:* Replica flux in the HREX simulations across the 16 states in the eight simulated complexes for the open (left) and closed Nluc structure. The greem dotted line represents the ideal flux.

In Figure 4, we show the replica flux across the 16-nodes (states) during the simulation as evaluated on all the seven concurrent HREX simulations for all eight analyzed complex. Details on the calculation of the flux based on the mean free passage time can be found in Ref. [5]. The flux, althought far from the ideal linear behaviour, is decent irrespective of the ligand, its tautomeric state, or the open/closed form of Nluc, with mean round trip times of the order of 1.3 ± 0.8 ns and average exchange rate of 38.1% ± 11.0%, indicating a good sampling efficiency, especially considering the size of the hot-zone.

Prior to presenting the results from the Hamiltonian replica exchange (HREX) simulations of the complexes, it is important to emphasize two key points. First, the simulated state represents a single ligand-Nluc complex in solution. Second, a well-converged HREX simulation, with a properly defined hot region, should be independent of the initial configuration (which, for each complex, was taken from docking calculations). In this context, it must also be noted that the binding pattern characterized by HREX for an Nluc monomer in solution may not align with that observed in X-ray crystal structures. The crystalline environment can induce structural features, such as the tight homotetrameric association observed in all available closed-form Nluc structures,[3] which are absent under physiological conditions, whereby Nanol functions as a monomeric enzyme. Consequently, the information derived from our HREX simulations provides valuable, complementary insights into the potential binding mode of the Nluc-substrate association in a near-physiological, monomeric state.

Figure 5 depicts the distribution of the distance between the centers of mass (COM) of the ligand and the protein for the open and closed conformations of Nluc, as obtained from the 8 HREX simulations of the complexes. This metric serves as a proxy for ligand mobility within the intra-barrel pocket.

In the open form, the peak positions for the enolic tautomers are shifted to higher values compared to their ketonic counterparts for both CTZ and FMZ luciferins, indicating a distinct binding mode. The ketonic form of FMZ exhibits the greatest mobility within the cavity of open Nluc, with the COM-COM distance sampling the 4–7 Å range with significant probability. Similarly, the CTZ ligand appears more mobile in its enolic form.

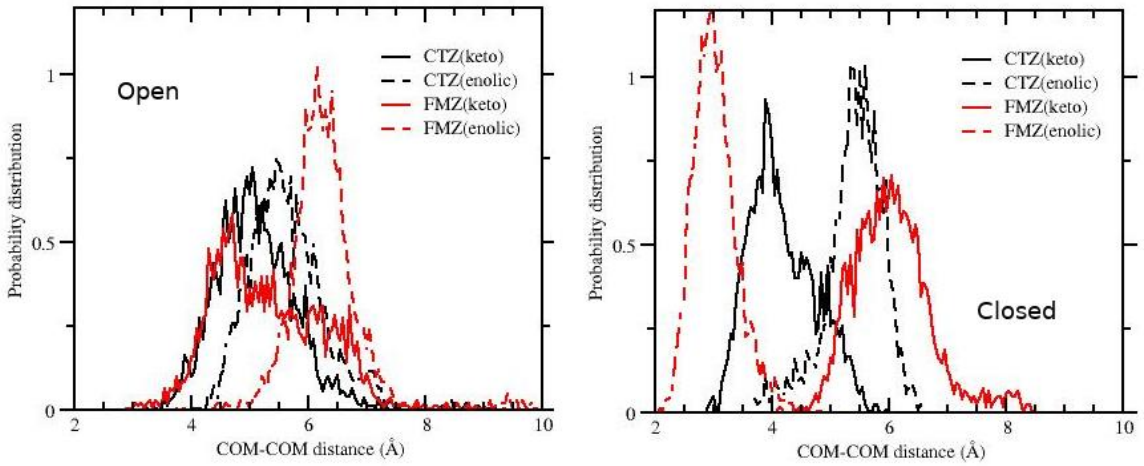


Fig. 5: Probability distributions of the ligand-Nluc COM-COM distance for the eight complexes with Nluc in the odpn (left) and closed (right) form.

The COM-COM distribution for the closed conformation of Nluc reveals a more complex behavior. For CTZ, the enolic form still shows an upward shift in its distribution relative to the ketonic tautomer. In contrast, the situation is reversed for FMZ, where a pronounced downward shift is observed for the enolic tautomer. As in the open form, the distributions for the enolic tautomers in the closed form are sharper than those of the ketonic forms, suggesting that the enolic states are more tightly bound within the intra-barrel cavity. It is important to note that this observation does not necessarily imply a stronger binding affinity for the enolic state. In fact, greater ligand mobility can enhance binding through a more favorable entropic contribution.

Table 2 reports the strength of contacts between the ligand and the predefined hot-zone residues (see Methods section) in the eight simulated complexes, as quantified by the Hamiltonian replica exchange (HREX) simulations. The target-state configurations were analyzed using ancillary software tools from the Zenodo distribution's "bin" directory. A contact was defined as occurring if any ligand atom was within 4.5 Å of any residue atom. Contact probabilities (p) derived from the HREX trajectories were classified as strong (S; $p > 0.75$), medium (M; $0.35 < p < 0.75$), or weak (W; $p < 0.35$). In the open form, most hot-zone residues exhibit persistent contact with both FMZ and CTZ ligands, regardless of their tautomeric state. However, ARG162, a residue implicated in binding to the open Nluc enzyme,[1] shows only weak interaction with the enolic FMZ tautomer. This suggests a distinct binding mode for this tautomer compared to its ketonic counterpart. Furthermore, Tyr114 displays weak binding in three of the four complexes, with only ketonic CTZ forming a strong interaction. In the closed form, His93 points into the binding pocket and replaces Tyr94 in the hot-zone (see Methods section). This residue interacts strongly with all ligands except for the ketonic FMZ tautomer, for which the interaction is surprisingly weak. Tyr144 in the closed form follows a pattern similar to that observed in the open form, showing weak interactions with all ligands except ketonic CTZ.

Table 2: Ligand-receptor contact analysis involving the residues of intra-barrel catalytic sice for the eight Nluc complex (see text for details)

					OPEN					
	L22	Y94	Y109	F110	Y114	I137	D139	F151	V153	R162
CTZ(k)	S	S	S	S	S	S	S	S	S	M
CTZ(e)	S	S	S	S	W	S	S	M	S	M
FMZ(k)	S	S	S	S	W	S	S	S	S	S
FMZ(e)	S	S	S	S	W	M	M	M	S	W
					CLOSED					
	L22	H93	Y109	F110	Y114	I137	D139	F151	V153	R162
CTZ(k)	S	S	S	S	S	S	S	S	S	S
CTZ(e)	S	S	S	S	W	S	S	S	S	S
FMZ(k)	S	W	S	S	W	W	S	S	S	S
FMZ(e)	S	S	S	S	W	S	S	S	S	S

Conclusion

We have presented the results of extensive Hamiltonian replica exchange (HREX) simulations for eight Nluc complexes, involving the ligands FMZ and CTZ in both their ketonic and enolic tautomeric states, with the protein in either the open or closed conformation. The analysis of ligand mobility within the intra-barrel pocket and the ligand-Nluc contact patterns only partially confirms the binding modality speculated from available experimental data, which is based on complexes with non-reactive analogs or unligated structures.

Our simulations highlight notable differences in binding modalities that depend on the tautomeric state of the ligand and the conformational state (open or closed) of the protein. These insights could inform the design of more effective substrates based on the reactive imidazo-pyrazine core.

This study represents a preliminary foray into this structural space, and work is ongoing to further elucidate the binding mechanisms. The HREX trajectories and docking calculations have been deposited in the public Zenodo repository (https://zenodo.org/uploads/17177177) and are available with no restrictions for independent analysis by researchers interested in this subject.

Acknowledgments

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes (see www.cresco.enea.it for information).

References

[1] Hall, M. P.; Unch, J.; Binkowski, B. F.; Valley, M. P.; Butler, B. L.; Wood, M. G.;Otto, P.; Zimmerman, K.; Vidugiris, G.; Machleidt, T.; Robers, M. B.; Benink, H. A.; Eggers, C. T.; Slater, M. R.; Meisenheimer, P. L.; Klaubert, D. H.; Fan, F.; Encell, L. P.; Wood, K. V. Engineered Luciferase Reporter from a Deep Sea Shrimp Utilizing a Novel Imidazopyrazinone Substrate. ACS Chem. Biol. 2012, **7**, 1848–1857.

[2] Green, A. A.; McElroy, W. Crystalline firefly luciferase. Biochim. Biophys. Acta 1956, **20**, 170–176.

[3] Nemergut, M.; Pluskal, D.; Horackova, J.; Sustrova, T.; Tulis, J.; Barta, T.; Baatallah, R.; Gagnot, G.; Novakova, V.; Majerova, M.; Sedlackova, K.; Marques, S. M.; Toul, M.; Damborsky, J.; Prokop, Z.; Bednar, D.; Janin, Y. L.; Marek, M. Illuminating the mechanism and allosteric behavior of NanoLuc luciferase. Nature Communications 2023, **14**, 7864.

[4] Doerr, S.; Harvey, M. J.; Noé, F.; De Fabritiis, G. HTMD: High-Throughput Molecular Dynamics for Molecular Discovery. J. Chem. Theory Comput. 2016, **12**, 1845–1852.
[5] Procacci, P. Dealing with Induced Fit, Conformational Selection, and Secondary Poses in Molecular Dynamics Simulations for Reliable Free Energy Predictions. J. Chem. Theory Comput. 2023, **19**, 8942–8954.
[6] Iannone, F.; Ambrosino, F.; Bracco, G.; De Rosa, M.; Funel, A.; Guarnieri, G.; Migliori, S.; Palombi, F.; Ponti, G.; Santomauro, G.; Procacci, P. CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. 2019 International Conference on High Performance Computing Simulation (HPCS). 2019; pp 1051–1052.
[7] Takatsu, K.; Kobayashi, N.; Wu, N.; Janin, Y. L.; Yamazaki, T.; Kuroda, Y. Biophysical analysis of Gaussia luciferase bioluminescence mechanisms using a non-oxidizable coelenterazine. BBA Adv. 2023, **3**, 100068.
[8] Procacci, P. PrimaDORAC: A Free Web Interface for the Assignment of Partial Charges, Chemical Topology, and Bonded Parameters in Organic or Drug Molecules. J. Chem. Inf. Model. 2017, **57**, 1240–1245.
[9] Marchi, M.; Procacci, P. Coordinates scaling and multiple time step algorithms for simulation of solvated proteins in the NPT ensemble. J. Chem. Phys. 1998, **109**, 5194–5202.
[10] Procacci, P. Hybrid MPI/OpenMP Implementation of the ORAC Molecular Dynamics Program for Generalized Ensemble and Fast Switching Alchemical Simulations. J. Chem. Inf. Model. 2016, **56,** 1117–1121.

INSIGHTS FROM MD SIMULATIONS ON COMPLEX MOLECULAR SYSTEMS: THE CASES OF THE ALDH2 AND THE HUMAN CNT3 TRANSPORTER

Tiziana Marino*, Giada Ciardullo, Alessia Dodaro, Mario Prejanò, and Nino Russo

Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, Via P. Bucci, 87036 Rende, Italy.

ABSTRACT. MD simulations were employed to analyze the dynamic behaviour of two different systems: Aldehyde dehydrogenase 2 (ALDH2) and human Concentrative Nucleoside Transporter 3 (hCNT3). In the case of ALDH2 MD simulations included the NAD+ cofactor and in that of hCNT3 the lipid mitochondrial membrane. Using molecular docking and molecular dynamics (MD) simulations, the structural and energetic impacts of each ligand on ALDH2 and hCNT3 were deeply analyzed.

Computational point of view of the inhibition of ALDH2 by quercetin and Miquelianin: comparative molecular insights

Alcohol is metabolized in the liver to acetate in a two-step process: alcohol (ethanol) is converted to acetaldehyde by the enzyme alcohol dehydrogenase (ADH) followed by conversion of acetaldehyde to acetate by the aldehyde dehydrogenase (ALDH, Figure 1A).

When the body consumes large amounts of alcohol in a short period of time, it leads to the accumulation of acetaldehyde, which in turn causes headaches, nausea, and dizziness. ALDH2 is the main functioning among the most relevant enzymes in the family of ALDHs for alcohol metabolism, located at mitochondrial cell region. [1,2]

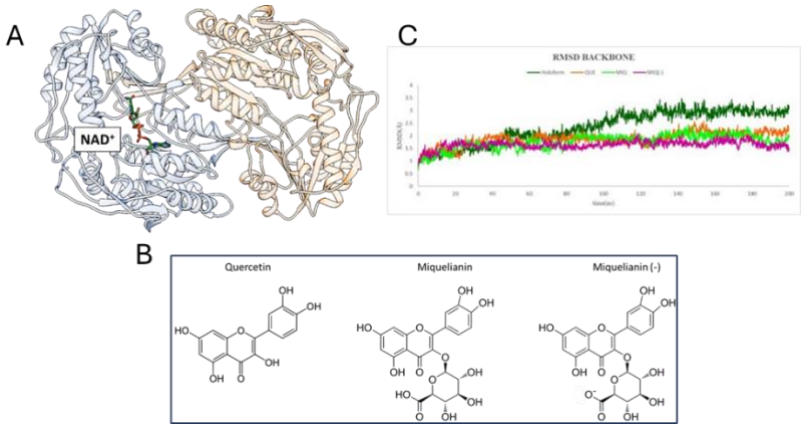


Fig. 1 A) Representation of the A-chain (in corn flower blue) and B-chain (in orange) of ALDH2 together with the cofactor (NAD^+); B) The chemical structure of the three antioxidants (Quercetin, Miquelianin and its anionic species (Miquelianin (-)) examined; C) RMSD of backbone atoms calculated for the ALDH2:NAD^+:Que; ALDH2:NAD^+:MIQ and ALDH2:NAD^+:MIQ(-) complexes compared to the holoform ALDH2.

The reduction of enzymatic activity of ALDH2 results in aldehyde accumulation, which generates signs of alcohol usage, like headaches, flushing, tachycardia, nausea, and vomiting after consuming alcohol. [3,4] Some people who can drink other alcoholic beverages without experiencing any headaches, experience headaches when they consume red wine. The cause for this effect has been attributed to the high level of phenolics in red wine. Recently it was observed that headaches can be attributed to a dysfunctional ALDH2 variant, allowing it to accumulate.[5] In the same work the inhibition screening assay proposed the quercetin-3-glucuronide (miquelianin) as a strong inhibitor for additional research compared to quercetin. In order to gain a deeper understanding of the triggers of headache by red wine (RWH) and to explain the inhibitory effects of phenolics on ALDH2, the present study explored by molecular dynamics the European isoform of the ALDH2 (in holo state with nicotinamide adenine dinucleotide (NAD+)) in complex with two flavonoids, Miquelianin in both neutral (MIQ) and negatively charged [$MIQ^{(-)}$] forms and quercetin (QUE). Their structure is depicted in Figure 1B. The present study investigates the conformational behaviour of the ternary complexes formation (ALDH2:NAD^+:Que; ALDH2:NAD^+:MIQ and ALDH2:NAD^+:MIQ(-)) to provide atomistic insights helpful to better clarify the molecular mechanisms of their formation.

Molecular docking and MD simulations were employed to assess the structural and energetic influences of these ligands on ALDH2, with a focus on binding affinities, protein flexibility, and enzymatic stability. The crystal structure of ALDH2 (PDB ID: 8DR9) [6] was utilized as a model. AutoDock4 was used for docking calculations, targeting residues near the active site (Cys302).

200 ns of MD simulations were conducted to evaluate structural stability and ligand-induced flexibility. All the simulations have been carried out with GROMACS20. [7,8] A very detailed post-MD analysis (Root-mean-square deviation (RMSD), root mean square fluctuation (RMSF), hydrogen (H)-bond analysis, RMSD-based clustering of MD trajectories, analysis of the distances) was carried out to assess the conformational behaviour during the molecular dynamics simulations changes. MD trajectories have been evaluated by calculation of RMSD of the three complexes in comparison with the holo-form shown in Figure 1C. It is possible to observe as the presence of the ligand, in all the three complexes, induces the structural stability of the holoform ALDH2, indicating a good performance of the adopted computational protocol. Despite the presence of the sugar ring making bulkier MIQ and $MIQ^{(-)}$, MIQ binds in a similar region of QUE as observed from the most populated clustered conformation (Figure 2). Instead, the flavonoid rings are oriented in very comparable fashion to that of QUE keeping contact with the same amino acid residues as reported in Figure 2.

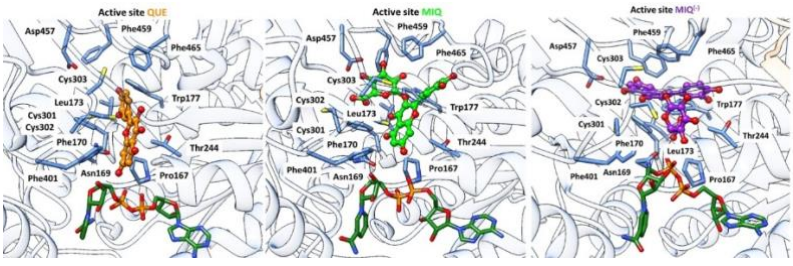


Fig. 2 The most representative cluster from MD simulations of the ALDH2:NAD^+:Que; ALDH2:NAD^+:MIQ and ALDH2:NAD^+:MIQ(-) complexes.

Furthermore, by MM-PBSA free energy calculations reported in Figure 3, miquelianin, especially in its neutral form, exhibited higher binding affinity to ALDH2 compared to quercetin. Such results are in agreement with the experimental observations. . The MM-P/GBSA calculation was carried out using the MMPBSA.py [9] module of AMBER16. [10]

QUE		
Energy Component	Average	Std. Err. Of Mean
$E_{VDWAALS}$	-36,8937	0,208
E_{EL}	-32,4635	0,3317
E_{PB}	59,1729	0,3422
E_{NPOLAR}	-3,2269	0,0051
E_{DISPER}	0	0
ΔG gas	-69,3572	0,3271
ΔG solv	55,946	0,3408
Δ TOTAL	-13,4112	0,2874

MIQ		
Energy Component	Avarage	Std. Err. Of Mean
$E_{VDWAALS}$	-44,5843	0,305
E_{EL}	-97,8099	0,5731
E_{PB}	130,3668	0,5175
E_{NPOLAR}	-4,6839	0,0089
E_{DISPER}	0	0
ΔG gas	-142,3942	0,5322
ΔG solv	125,6829	0,5171
Δ TOTAL	-16,7113	0,4074

$MIQ^{(-)}$		
Energy Component	Average	Std. Err. Of Mean
$E_{VDWAALS}$	-39,7778	0,3608
E_{EL}	-119,7481	1,659
E_{PB}	144,5112	1,3316
E_{NPOLAR}	-4,5269	0,0107
E_{DISPER}	0	0
ΔG gas	-159,526	1,4939
ΔG solv	139,9844	1,3385
Δ TOTAL	-19,5416	0,57

Fig. 3 Average of binding free energy (kcal/mol) of the three examined complexes

From the comparative analysis of the MD simulations of the ALDH2:NAD^+:Que; ALDH2:NAD^+:MIQ and ALDH2:NAD^+:MIQ(-) complexes, important differences in terms of interactions and structural aspects emerged contributing to rationalize the experimental evidence.[4] Deeper analysis of the investigation will be available in the manuscript in preparation for submission.

Molecular Dynamics Insights into the conformational behaviour of the Platinum-Modified guanines in the human CNT3 Transporter in the Plasma Membrane

A transporter as a membrane protein facilitates the movement of ions, molecules, or other substances across a biological membrane. These proteins are essential for maintaining cellular homeostasis by regulating the import and export of nutrients, waste products, and signalling molecules.

Nucleosides, which function as signalling molecules and precursors to nucleic acids, enter cells by concentrative nucleoside transporters (CNTs). Additionally, CNTs are essential for the uptake of nucleoside-derived drugs showing antiviral and anticancer activities. Gaining insight into how CNTs identify and import their substrates may result helpful in the development of nucleoside-derived drugs that can be recognized by their targets more effectively as well as can improve our understanding of biological processes related to nucleosides. Computational methodologies hold a distinct advantage in offering mechanistic insights that may be beyond the reach of experimental approaches but are essential for guiding the rational design of drugs.

In this work, we studied model complexes based on guanosines N7-coordinated to a platinum center, working as modified nucleosides with unaltered sugar moieties, guided by experimental data that tested the possibility that N7-platinated purines could be transported into cells through HELA cells plasma membrane.[11] So, we focused on the structural behavior observed within the site cleft of CTN3 (a nucleosides transporter) and its neighboring area when interacting with the considered 1–3 platinated nucleosides shown in Figure 4A. To better understand the specificity of platinated nucleoside uptake by this trans-membrane transporter and its potential access to specific cell types and intracellular compartments, also the canonical dGuo was studied for comparison.

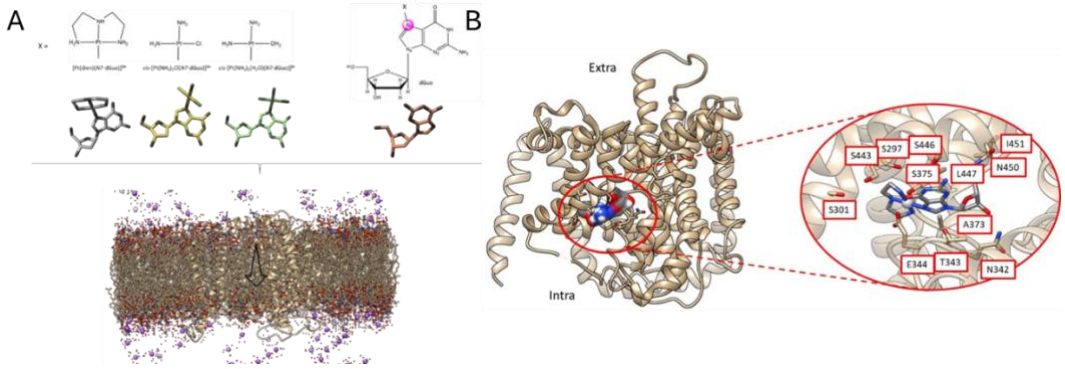


Fig. 4 A) Schematic representation of complexes: [Pt(dien)(N7-dGuo)] (1), cis-[Pt(NH_3)$_2$Cl(N7-dGuo)] (2), and cis-[Pt(NH_3)$_2$(H_2O)(N7-dGuo)] (3); dGuo = 5'-(2'-deoxy)-guanosine that bind the transporter inside the membrane cell; B) Nucleoside binding region of the hCNT3 considered in the docking of all the four nucleosides. The example reported in the figure is related to Pt(dien)(N7-dGuo)].

The structures of all the complexes are based on the human Cryo-EM structure of the (PDB 6KSW, resolution of 3.6 Å).[12] The structure of the enzyme consists of 632 amino acid residues and includes three chains. Next, the four nucleosides were docked using Autodock4 [13]. on human CTN3 at the site of the transporter where the molecules bind with more affinity corresponding to the nucleoside binding region of the transporter. (Figure 4B)

To obtain the model of the transporter systems inside the membrane cell CHARMM-GUI, [14] a web-based tool has been used. By Membrane Builder option, one can create the sequence of CHARMM inputs required to construct a protein/membrane complex to run molecular dynamics simulations has been created. All the simulations were run using GROMACS20 package [7,8].

In the MD simulations, all the nucleosides reside in the nucleoside pocket of the protein, in particular [Pt(dien)(N7-dGuo)]$^{2+}$, interacts by the nucleobase portion with T342 and E343 as the dGuo which in turn is in contact also with Q341, accordingly to the residues indicated for hCNT3.[12]

The number of hydrogen bonds formed between ligands and the membrane protein was determined for each system during the 200 ns of simulation. A donor-acceptor cutoff distance within 3.0 Å and a maximum donor-hydrogen-acceptor angle of 20° were used in the calculation. Figure 5 displays the results of the donor-acceptor H-bond analysis between the transporter and the nucleosides present in the four systems under investigation. A thorough examination was conducted on the key diagonal barrier scaffold domain TM9 (residues 432 to 475 of hCNT3), which plays a role in the translocation of the bound nucleoside across the membrane into the cell.

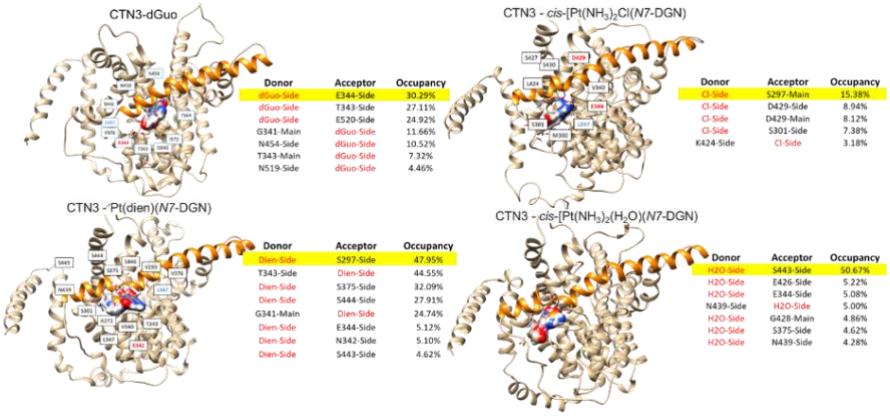


Fig. 5. Hydrogen bonding between the ligand and the transporter encountered during the simulation time of 200 ps.

The TM9 helix's S297, S375, S443, and S444 residues are the most commonly involved in hydrogen bonding with the [Pt(dien)(N7-dGuo)] moiety (Figure 5). This implies that the interaction with the positively charged platinated molecule depends critically on these polar residues. In order to promote the transportability process of the three platinated nucleosides on the transporter, serine residues were discovered to be very significant.

More detailed analysis of the investigation will be found in the manuscript submitted.

Computer code

Simulations and analysis of trajectories were carried out using GROMACS 2020 (gmx_mpi module)

Acknowledgments

Financial support from Dipartimento di Chimica e Tecnologie Chimiche, Università della Calabria, is acknowledged. Some of the computer simulations were carried out adopting local resources (Amber for MM-PBSA).
The computing re- sources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High-Performance Computing infrastructure and its staff. CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programs, see http://www.cresco.enea.it/english for information.[15]

References

[1] A. A. Klyosov; L. G. Rashkovetsky; M. K. Tahir; W.-M. Keung, Possible Role of Liver Cytosolic and Mitochondrial Aldehyde Dehydrogenases in Acetaldehyde Metabolism. *Biochemistry* **35(14)**, 4445–4456, (1996).

[2] Y. Zhang, Y. Qiu, H. Zhang, Computational Investigation of Structural Basis for Enhanced Binding of Isoflavone Analogues with Mitochondrial 2c00032.Aldehyde Dehydrogenase. *ACS Omega*, **7 (9)**, 8115–8127, (2022).

[3] D. R., Greenberg, H. P., Bhambhvani, S. S., Basran, B. P., Salazar, L. C., Rios, S. J., Li, C. H., Chen, Mochly-Rosen, D., & M. L. Eisenberg, ALDH2 expression, alcohol intake and semen parameters among East Asian Men. *Journal of Urology,* **208(2)**, pp. 406–413, (2022).

[4] C. L., Lai, C. T., Yao, G. Y., Chau, L. F., Yang, T. Y., Kuo, C. P., Chiang, & S. J. Yin, Dominance of the inactive Asian variant over activity and protein contents of mitochondrial aldehyde dehydrogenase 2 in human liver. *Clinical and Experimental Research,* **38(1)**, 44–50, (2014).

[5] Devi, A., Levin, M. & Waterhouse, A.L. Inhibition of ALDH2 by quercetin glucuronide suggests a new hypothesis to explain red wine headaches. *Sci Rep* **13**, 19503, (2023).

[6] S. Y. Xu, J. K. Weng Crystal Structure of Human ALDH2 in Complex with NAD^+ and PEG MME 550. August 30, 2023. https://doi.org/10.2210/pdb8dr9/pdb.

[7] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, SoftwareX 1 (2), 19–25 (2015).

[8] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, H. J. Berendsen, GROMACS: fast, flexible, and free. *J. Comp. Chem.* 26(16), 1701-1718 (2005).

[9] B.R. Miller III, D. T, Jr. McGee, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA.py: an efficient program for end-state free energy calculations, *J. Chem. Theory Comput*. 8 (9), 3314(2012)

[10] D.A. Case, H.M. Aktulga, K. Belfon, I.Y. Ben-Shalom, J.T. Berryman, S.R. Brozell, D. S. Cerutti, T.E. Cheatham III, G.A. Cisneros, V.W.D. Cruzeiro, T.A. Darden, R. E. Duke, G. Giambasu, M.K. Gilson, H. Gohlke, A.W. Goetz, R. Harris, S. Izadi, S. A. Izmailov, K. Kasavajhala, M.C. Kaymak, E. King, A. Kovalenko, T. Kurtzman, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, M. Machado, V. Man, M. Manathunga, K.M. Merz, Y. Miao, O. Mikhailovskii, G. Monard, H. Nguyen, K. A. O'Hearn, A. Onufriev, F. Pan, S. Pantano, R. Qi, A. Rahnamoun, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, A. Shajan, J. Shen, C.L. Simmerling, N. R. Skrynnikov, J. Smith, J. Swails, R.C. Walker, J. Wang, J. Wang, H. Wei, R. M. Wolf, X. Wu, Y. Xiong, Y. Xue, D.M. York, S. Zhao, P.A. Kollman, Amber, University of California, 2016. San Francisco.

[11] F. De Castro, E. De Luca, C.R. Girelli, A. Barca, A. Romano, D. Migoni, T, Verri, M, Benedetti, F.P. Fanizzi First evidence for N7-Platinated Guanosine derivatives cell uptake mediated by plasma membrane transport processes. *J Inorg Biochem*. Jan;226:111660 (2022).

[12] Y., Zhou, L., Liao, C., Wang, J., Li, P., Chi, Q., Xiao, D. Deng, Cryo-EM structure of the human concentrative nucleoside transporter CNT3. *PLoS Biology*, 18(8), e3000790, (2020).

[13] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, A.J. Olson AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility, *J. Comp. Chem*., Dec;30(16):2785-91, (2009).

[14] J. Lee, D. S. Patel, J. Ståhle, S.-J. Park, N. R. Kern, S. Kim, J. Lee, X. Cheng, M. A. Valvano, O. Holst, Y. A. Knirel, Y. Qi, S. Jo, J. B. Klauda, G. Widmalm, W. Im, CHARMM-GUI Membrane Builder for Complex Biological Membrane Simulations with Glycolipids and Lipoglycans, *J Chem Theory Comput* 15, 775–786, (2019).

[15] F. Iannone, F. Ambrosino, G. Bracco, M. De Rosa, A. Funel, G. Guarnieri, S. Migliori, F. Palombi, G. Ponti, S. Santomauro, P. Procacci, International Conference on High Performance Computing and Simulation, HPCS), Dublin, Ireland, pp. 1051–1052, (2019).

EVALUATING ENTROPY EFFECTS IN DESIGN OF LAYERED OXIDES AS FUTURE NIB CATHODES: NEW HINTS FROM FIRST PRINCIPLES

Arianna Massaro[1,*], Ana B. Muñoz-García[2], Michele Pavone[1]

[1]*Department of Chemical Sciences, University of Naples Federico II, Via Cintia 21, Naples, Italy*

[2]*Department of Physics "E. Pancini", University of Naples Federico II, Via Cintia 21, Naples, Italy*

ABSTRACT. Using density functional theory, we investigate structural and electronic features of P2-type Na_xMO_2 systems that have been proposed for use in NIB positive electrodes. While simulating the cathode charging process (*i.e.*, desodiation), we correlate the changes in structural and electronic properties to describe the electrochemical behaviour. By comparing three formulations encompassing different elements at M site, we aim at unveiling the role of increasing entropy in regulating the structure-property relationship in this class of materials. Our theoretical insights can provide insights for future design principles towards structurally stable high-entropy layered oxides as efficient NIB cathodes.

Introduction

The research conducted at the University of Naples "Federico II" has focused on the *ab initio* investigation of high-entropy oxides that have been proposed as promising cathode materials for Na-ion batteries (NIBs) [1]. While selecting the suitable TM composition in layered P2-Na_xTMO_2 has emerged as the crucial point to tune and regulate specific capacity and energy density of the cathode [2-5], complex formulations featuring high configurational entropy can be achieved by mixing different metal elements. In principle, different metals can exert different roles, either being directly involved in the redox activity and charge compensation upon Na^+ removal or contributing to structure stabilization. Herein, three compositions have been explored, $Na_{0.69}(Mn_{0.44}Ni_{0.19}Co_{0.25}Ti_{0.09}Mg_{0.03})O_2$, $Na_{0.69}(Mn_{0.44}Ni_{0.19}Co_{0.20}Al_{0.05}Ti_{0.09}Mg_{0.03})O_2$ and $Na_{0.69}(Mn_{0.44}Ni_{0.19}Co_{0.18}Al_{0.05}Fe_{0.02}Ti_{0.09}Mg_{0.03})O_2$ (MEO-1, MEO-2, and HEO, respectively), the labelling criteria following the calculation of configuration entropy according to Eq. 1. Direct comparison with the low-entropy counterpart $Na_{0.69}(Mn_{0.69}Ni_{0.31})O_2$ (LEO) is also provided. The aim is to unveil the specific role of TM mixing and correlate the increasing entropy to potential effects into structural and electronic features upon charge.

Methods and Computational details

Spin-polarized DFT calculations are performed with projector-augmented wave (PAW) potentials and plane-wave (PW) basis set, as implemented in the Vienna Ab-initio Simulation Package (VASP) code. We use the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional with the +U Hubbard-like correction scheme to overcome the large self-interaction error (SIE) that affects DFT when applied to mid-to-late first-row TM oxides with tightly localized d electrons. The on-site correction is applied to the d electrons of Mn, Ni, Co, Fe atoms with a unified average effective U-J parameter, U_{eff}, equal to 4 eV [6]. The D3-BJ dispersion correction is also added to account for interlayer van der Waals (vdW) interactions [7]. A kinetic energy cutoff of 600 eV and a 1×2×2 Γ-centered k-points sampling mesh are required to converge the PW basis set within 10^{-5} eV threshold. Criteria for ionic minimization is set to $3x10^{-2}$ eVÅ^{-1}. Electronic structure has been refined at HSE06 level of theory, including the TS correction to account for vdW dispersion forces, as implemented in the FHI-aims code. Light-tier1 basis sets of NAOs have been employed [8]. Structural models consist of 8x4x1 supercells containing 236 atoms in the highest sodiated state. To accurately consider the configurational entropy at both Na_{Edge}/Na_{Face} and M sites at

each stoichiometry, we employ the special quasi-random structure (SQS) approach, that ensures occupancy of each site consistently respected throughout the analysis [9].

Results and discussion

The investigated formulations, $Na_{0.69}(Mn_{0.69}Ni_{0.31})O_2$, $Na_{0.69}(Mn_{0.44}Ni_{0.19}Co_{0.25}Ti_{0.09}Mg_{0.03})O_2$, $Na_{0.69}(Mn_{0.44}Ni_{0.19}Co_{0.20}Al_{0.05}Ti_{0.09}Mg_{0.03})O_2$ and $Na_{0.69}(Mn_{0.44}Ni_{0.19}Co_{0.18}Al_{0.05}Fe_{0.02}Ti_{0.09}Mg_{0.03})O_2$ feature configurational entropy values of S = 0.62R, 1.34R, 1.47R and 1.54R (according to Eq. 1). As the following ranges are usually considered: S < 1R low-entropy, 1R < S < 1.5R medium-entropy, S > 1.5R high-entropy, we will name the compounds as LEO, MEO-1, MEO-2, and HEO, respectively.

$$S = -R \sum x_i \ln(x_i) \qquad (1)$$

The structural models consist of 8x4x1 supercells within the $P6_3/mmc$ space group. The SQS approach was used to simulate the M disorder in TMO_2-layers. The Na atoms are placed in edge (e) and face (f) sites to achieve an e/f ratio of 2 [10]. To mimic the desodiation process, we model various Na contents (x = 0.69, 0.50, 0.30, 0.12). Fig. 1 illustrates the structural models, the in-plane projections displaying the random metal atoms distributions obtained from SQS, and the 3D-perspective view of the cells.

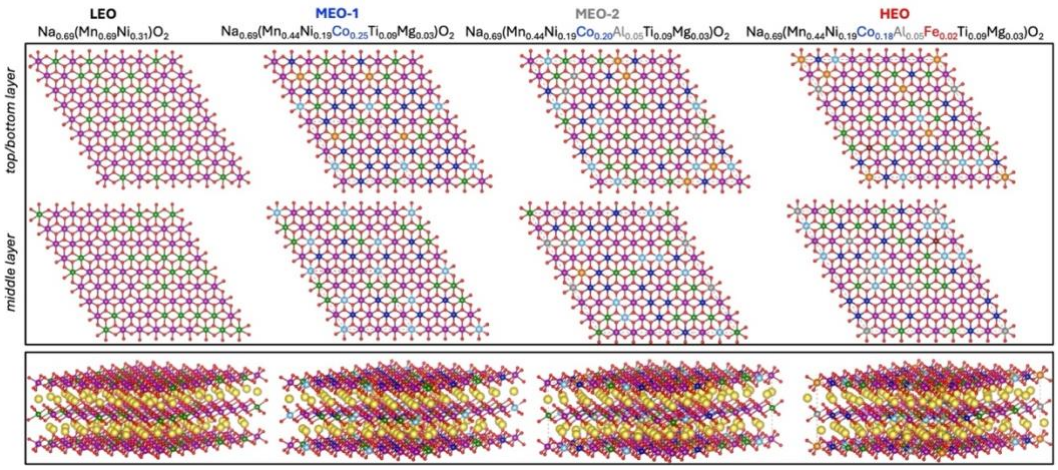


Fig.1: Structural models for the P2-Na_xMO_2 series: LEO, MEO-1, MEO-2 and HEO. **Top:** random distribution of metal atoms within the MO_2-layers. **Bottom:** 3D view of the hexagonal cells in the $P6_3/mmc$ space group.

As listed in Table 1, the minimum-energy structures obtained at PBE+U-D3BJ level of theory features a and c constants in agreement with experimental data (deviation below 1% is observed). Lattice compression is evidenced for all compounds across desodiation, as a result of the stronger attractive interactions associated to the lesser Na cations content. Moreover, cell volume variations are also highlighted along the oxides series as a function of increasing entropy, suggesting that the substitution at M site with different M in different oxidation states does alter the geometry arrangements.

Table 1: Lattice parameters of the minimum-energy structures obtained at the PBE+U-D3BJ level of theory. Deviation on a and c constants from experimental data is reported for the highly sodiated structure.

	LEO		MEO-1		MEO-2		HEO	
x Na	**a (Å)**	**c (Å)**	**a (Å)**	**c (Å)**	**a (Å)**	**c (Å)**	**a (Å)**	**c (Å)**
	dev	**dev**	**dev**	**dev**	**dev**	**dev**	**dev**	**dev**
0.69	23.272	11.233	23.097	11.155	23.138	11.175	23.127	11.169
	0.6%	0.6%	0.4%	0.1%	0.6%	0.3%	0.5%	0.1%
0.50	23.165	11.181	23.007	11.112	23.063	11.139	22.808	11.487
0.30	23.097	11.149	23.031	11.123	22.921	11.070	22.773	11.469
0.12	22.951	11.078	22.927	11.073	22.888	11.054	22.794	11.480

To address the role of increasing entropy in tuning the structural features of the oxides, we dissect the contribution of each M sublattice to the magnitude of distortion (related to bond distance variations) and shear fraction parameters (associated with bond angle variations), as defined by Van Vleck: these figures of merits can quantify the geometrical rearrangements occurring within a M-centered octahedron (MO_6 units in our structures) related to Jahn-Teller effects or to increasing electrostatics owing to oxidated M centers. In Fig.2, r and h are plotted as function of desodiation (top and bottom panels, respectively) for each oxide compound. From the average values, we extract the individual contributions from the specific M atoms (see displayed colour legend). Increasing/decreasing trend of r with decreasing Na content are ascribed to the higher/lower content of Jahn-Teller active centers as a result of M oxidation occurring upon charge. As for h trend, we can generally assess that the highest the shear contribution, the highest the tendency to P-to-O gliding underlying detrimental solid-state phase transition [11]: that being considered, the Ti sublattice seems to significantly mitigate this effects, suggesting an active role in improving the structural retention of the cathode by preventing the P2-O2 phase transition at high voltage (cyan lines in Fig.2).

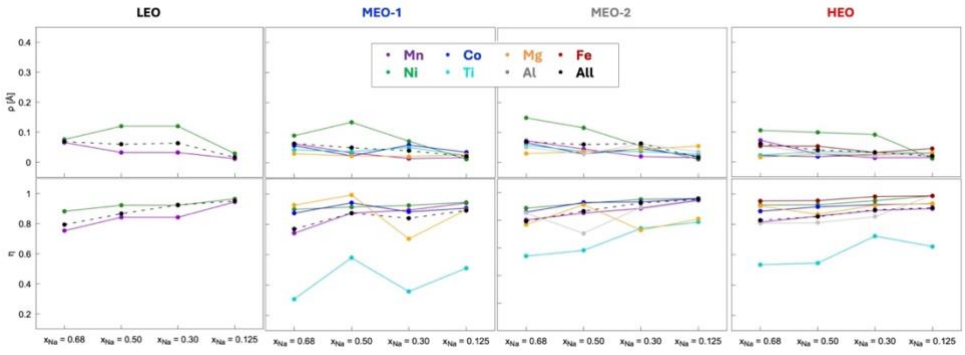


Fig. 2: Magnitude of distortion, r, and shear fraction parameter, h, calculated for (from left to right) LEO, MEO-1, MEO-2, and HEO, with the VanVleckCalculator. Colour legend is displayed.

To further inquire the structural stability upon charge, we focus on the formation of Mn out-of-layer migration, which is preliminary to metal migration commonly observed in this kind of materials during cycling. We model the process as the formation of an anti-site defect, that is Mn occupying a Na site left vacant after desodiation step. Fig.3 shows the corresponding structures obtained after geometry optimization, together with the magnetization values on the Mn migrated site. In all explored cases, the

displaced Mn features a reduced 2+ oxidation state. The formation energy results demonstrate that desodiation facilitates the accommodation of Mn in Na vacancies, and more interestingly that Mn is less prone to migrate in a cell with high configurational entropy. This finding unveils that metal migration, which can be harmful to the cell capacity because it reduces the available sites for reversible Na storage, can be prevented by implementing tailored formulations with high configurational entropy. Further investigation is ongoing, exploring additional M migration (*e.g.*, Ni, Co) and the resulting chemical environment so as to shed light on this mechanism and pursue this direction towards rational design principles.

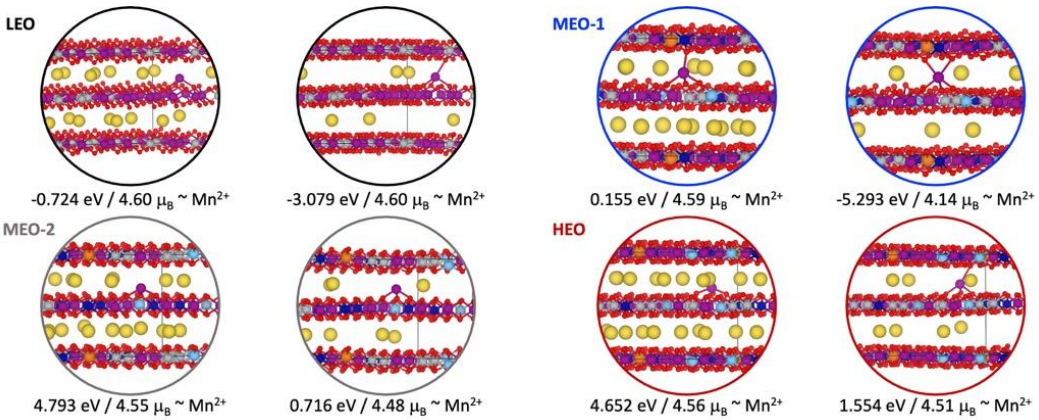


Fig. 3: Mn-out-of-layer migration: structural details, anti-site defects formation energy, Mn magnetic moment.

Acknowledgments

The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff; CRESCO/ ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes. See: http://www.cresco.enea.it/english for information [12].

References

[1] J. Wang, *et al*. P2-type layered high-entropy oxides as sodium-ion cathode materials, *Mater. Futures,* **1**, pp. 035104 (2022)

[2] C. Delmas, *et al.* Structural classification and properties of the layered oxides. *Physica*, **99**, 1, pp. 81–85 (1980)

[3] A. Massaro, *et al.* Unveiling Oxygen Redox Activity in P2-Type $Na_xNi_{0.25}Mn_{0.68}O_2$ High-Energy Cathode for Na-Ion Batteries. *ACS Energy Lett*, **6**, 7, pp. 2470–2480 (2021)

[4] A. Massaro, *et al.* First-principles insights on anion redox activity in $Na_xFe_{1/8}Ni_{1/8}Mn_{3/4}O_2$: Toward efficient high- energy cathodes for Na-ion batteries, *J. Am. Ceram. Soc.*, **1**, 11 (2022)

[5] A. Massaro, *et al.* Ru-Doping of P2-$Na_xMn_{0.75}Ni_{0.25}O_2$-Layered Oxides for High-Energy Na-Ion Battery Cathodes: First-Principles Insights on Activation and Control of Reversible Oxide Redox Chemistry, *ACS Appl. Energy Mater.*, **5**, 9, pp. 10721-10730 (2022)

[6] N. J. Mosey, *et al.* Rotationally invariant ab initio evaluation of Coulomb and exchange parameters for DFT+U calculations. *J. Chem. Phys.*, 129, 1, (2008)

[7] S. Grimme, *et al.* A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*,**132**,15,pp.154104 (2010)

[8] A. Tkatchenko, *et al.* Accurate molecular Van Der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* **102**, pp. 073005 (2009)

[9] H. Wang, *et al.* Special Quasirandom Structures for Gadolinia-Doped Ceria and Related Materials. *Phys. Chem. Chem. Phys.* **14**, 33, pp. 11737−11742 (2012)

[10] R. Fielden, *et al.* Investigation of the $NaNi_xMn_{1-x}O_2$ ($0 \leq x \leq 1$) system for Na-ion battery cathode materials. *J. Electrochem. Soc.*, **162**, 3, A453–A459 (2015)
[11] A. Langella, *et al.* Atomistic Insights into Solid-State Phase Transition Mechanisms of P2-Type Layered Mn Oxides for High-Energy Na-Ion Battery Cathodes, *ACS Energy Lett.*, **10**, pp. 1089–1098 (2025)
[12] G. Ponti, *et al.* The Role of Medium Size Facilities in the HPC Ecosystem: The Case of the New CRESCO4 Cluster Integrated in the ENEAGRID Infrastructure. International Conference on High Performance Computing & Simulation (HPCS), IEEE, pp. 1030–1033 (2014)

SATELLITE DATA ASSIMILATION OF SO_2, CO, HCHO, NO_2 AND O_3 IN ATMOSPHERIC MODELLING SYSTEM MINNI

Mihaela Mircea[1], Andrea Bolignano[1], Mario Adani[1], Gino Briganti[1], Felicita Russo[1], Massimo D'Isidoro[1], Guido Guarnieri[2]

[1]*Models and Measures for Air Quality and Climate Observation Laboratory, Italian National Agency for New Technologies, Energy and Sustainable Economic Development – ENEA, Bologna, 40121, Italy.*

[2]*Scientific and High-Performance Computing Infrastructure Laboratory, Italian National Agency for New Technologies, Energy and Sustainable Economic Development – ENEA, Portici, 80055, Italy*

ABSTRACT. This is the first application of satellite data assimilation in the MINNI Atmospheric Modelling System (MINNI-AMS). The tool was developed in the CAMEO project, in the framework of Copernicus Atmosphere Monitoring Service. Copernicus Sentinel-5 total columns were used to adjust model simulation fields, in order to improve the accuracy of air quality forecasts and analyses. The ENEA CRESCO 6 HPC facility was used.

Introduction

Data measured by the TROPOspheric Monitoring Instrument (TROPOMI) instrument on board the Copernicus Sentinel-5 Precursor satellite, launched in October 2017, provides a comprehensive atmospheric composition dataset at high horizontal resolution (5.5 km - 7 km). The total columns of SO2, CO, HCHO, NO2 and tropospheric O3 L2 products were assimilated in MINNI-AMS, by means of an Optimal Interpolation (OI) algorithm. The data assimilation routine has recently been integrated into the hybrid MPI/OMP FARM [1][2] (v. 6.0) chemical dispersion code and can be activated via a dedicated flag, upon providing a specific namelist. Here, we present some preliminary evaluations of its application to five pollutants, SO2, CO, HCHO, NO_2 and O_3, retrieved by Sentinel 5P satellite [3] for August 2023. The aim of this work was to investigate which is the impact of the covariance matrix definition used by OI and of multipollutant assimilation. The changes in total column content and surface concentrations caused by different assimilation setup are shown.

To the authors' knowledge, this is the first application of a satellite assimilation technique by an Italian research group.

The optimal interpolation algorithm

The Optimal Interpolation (OI) data assimilation scheme is based on Adani and Uboldi [4] and operates in the observation space. Its central objective is to minimize a quadratic cost function, which provides the Best Linear Unbiased Estimate (BLUE) by weighting the differences between observations and the forward model based on their respective error covariance matrices. The observational operator is realized by the CAMS Satellite Operator (CSO, version 1.9, for reference [5][6]), which transforms the 3D model field into the satellite's observation space at each pixel location. This process involves profile interpolation onto the satellite's pressure levels and the subsequent application of the observation's

averaging kernel to modulate the sensitivity and influence of the retrieved profile across different heights, culminating in a total column estimate. Once the total column analysis increment is calculated at the pixel level, a reverse algorithm allocates this assimilated change back to the corresponding 3D model grid points (the analysis increment spreading).

Operationally, assimilation is performed cyclically on each pollutant at every time step where observation availability is confirmed. Since no chemistry calculations are performed between the assimilation of one pollutant and the next, the processing order is non-critical. Model corrections were strictly localized to the areas where observations were available. Furthermore, observations were subject to quality control: negative observed values were excluded, and a minimum quality flag threshold of 0.75 was enforced for all retained data.

Whereas observational errors are directly available in the satellite retrievals (observations) downloaded using CSO package, the model background error covariance matrix was computed with NMC [7] for 48h-0h forecast (D0 and D-2) using logarithmic concentration. Two different approaches for estimating the covariances were used: the first was based on averaged monthly covariance matrix (Bavg) and the second on monthly hourly maximum (Bmax). The latter method assigns a higher weight to the model error, leading to a stronger correction of the model using the observations. The covariance matrix was estimated in the same way for all five pollutants and, for simplicity, was factorized assuming to be proportional to the product of standard deviations at different locations and gaussian shaped, provided the horizontal correlation length of 100 km; the vertical scale was set to 0.18, expressed in pressure natural logarithm units.

The simulations set-up

Five simulations (one base case and four with assimilation) were performed with MINNI-AMS over Europe using the same configuration as in CAMS Regional Air Quality [8]. The setup of the base case simulation is shown in Table 1.

Table 1: Base case setup (no assimilation).

horizontal resolution	0.15 x 0.1 lat/lon
number grid points	468x421
number of vertical levels	17
top of domain	11970 m
meteorological driver	D0 IFS, 1 hour time resolution
boundary conditions	CAMS global
emission inventory	CAMS-REG-V6.1.1_year2022

The four simulations with assimilation used the setup of base case, but they performed assimilation of pollutants as follows: 1) assimilation of five pollutants (5poll) with Bavg; 2) 5poll with Bmax (5polls_BgkCov_max); 3) assimilation of four pollutants (NO2_CO_HCHO_SO2, excluding ozone) with Bavg; 4) O_3 assimilation with Bavg.

Results of OI assimilation

In Fig. 1 the percentage differences on monthly averaged (August 2023) surface concentrations (in $\mu g/m^3$) are shown for SO_2 and NO_2. The assimilation of SO_2 retrievals make just “visible” Etna volcano by

model that did not consider its emission in the middle of August 2023. The assimilation increases column content all over domain both for Bavg and Bmax, with higher values for Bmax (top left graph). As expected, the effect of Bmax is lightly enhancing the differences with respect to basecase, because a larger specified model error increases the weight applied to the observations.

Similar results are observed for NO_2. These increases are observed also for total column content of both pollutants shown in Fig. 2. In Fig. 3 the time series of global biases over all stations are shown for CO, SO_2, NO_2 and O_3. For CO: none of the simulations with satellite data assimilation has significant effect with respect to "basecase" model results. For SO_2 a slight increase and decrease in BIAS and CORR are observed for Bavg and Bmax simulations; again, Bmax has a higher impact with respect to Bavg. Only Bmax produce a visible effect increasing slightly both BIAS and CORR for NO_2.

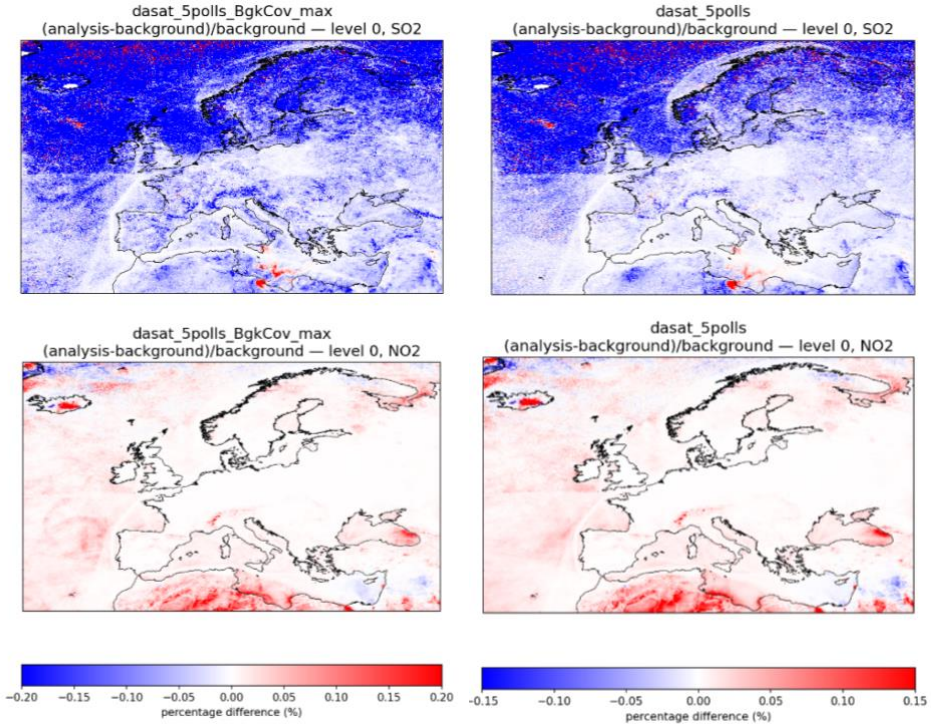


Fig.1: Monthly (August 2023) averaged percentage differences (%) of surface concentrations of NO_2 and SO_2 due to different definition of Bavg and Bmax

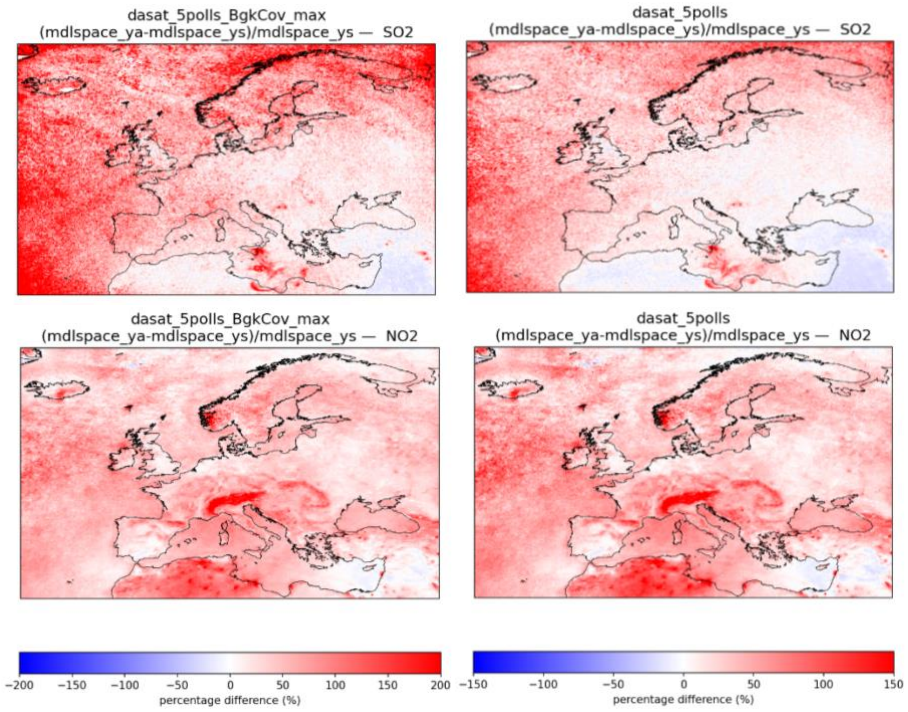


Fig.2: Monthly (August 2023) averaged percentage differences (%) of total column concentrations of NO_2 and SO_2 due to different definition of Bavg and Bmax

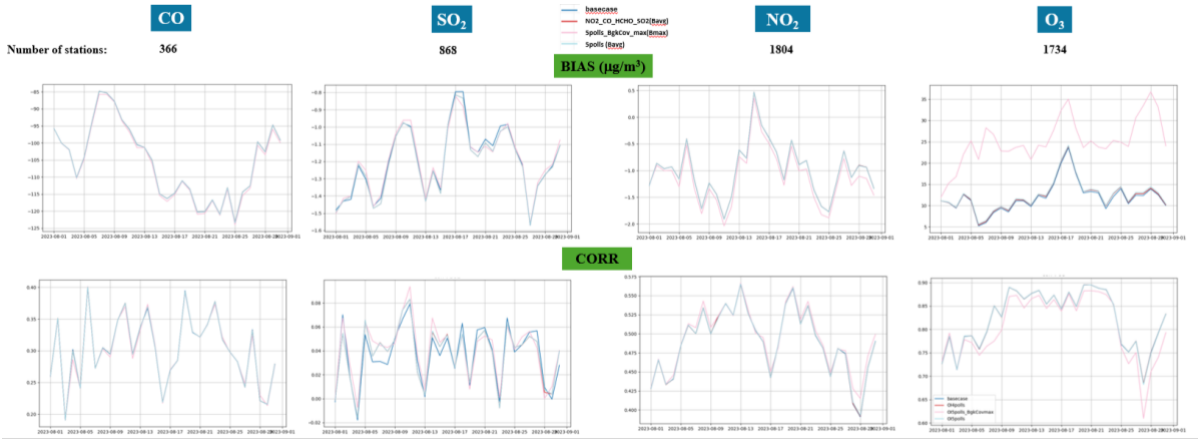


Fig.3: Daily bias (BIAS) and correlation (CORR) for CO, SO_2, NO_2 and O_3.

It worth noting that the effect of Bmax on O_3 is to increase the BIAS up to a factor 3, while CORR is slightly decreased. The assimilation of only O_3 satellite retrievals has indeed a lower impact on surface concentrations than the simultaneous assimilation of HCHO, NO_2, CO, SO_2 and O_3 (Fig. 4). The increase of O_3 concentration over most of the domain has a pattern like the increase of HCHO. However, the pattern of differences in NO_2 concentrations shows little resemblance with that of O_3. This may be due to horizontal spatial resolution of the simulations and gas phase chemistry formulation of HCHO, NO_2, and O_3 interactions as well as due to assimilation approach and quality of satellite retrievals.

Conclusions

The application of OI assimilation method to different pollutants has shown that the effect of assimilation is pollutant dependent, with low impact of satellite data on simulated CO and high impact on model performances in simulating O_3.

The use of different definitions for Background Error Covariance Matrix (B) is important, the choice of average or maximum values leads to different results according to pollutant. By using Bmax, an increase of surface concentration in areas with high emissions as Etna volcano for SO_2 and Po Valley, northern Italy for NO_2, was observed.

The simultaneous assimilation of interacting pollutants such as HCHO and O_3 leads to different results in modelled surface concentration of O_3 with respect to the simulation when only O_3 is assimilated.

More research is needed for finding the best definition of covariance matrix according to pollutant, for finding the best approach for simultaneous assimilation of satellite retrievals for interacting pollutants and for simultaneous assimilation of satellite retrievals and surface measurements.

The results also show that higher knowledge on atmospheric gas phase chemistry involving HCHO and O_3 as well as on biogenic volatile organic carbon (VOC) emissions largely responsible for the variabilities of HCHO is necessary.

Computer code and computational times

A non-official fork version of the 3D Eulerian chemical transport model, FARM, based on release 6.0, was used. The adopted chemical scheme was SAPRC-99 [9], with aerosol AERO3 module [10]. The integration of CSO 1.9 assimilation subroutines in FARM will be introduced into the next official releases. The computation was performed on the CRESCO 6 cluster, whose nodes are equipped with 48 CPU cores. The computation utilized 7 nodes, with each node configured with two sockets, each hosting 24 CPU cores. Elapsed time was about 33.5 minutes per day. Therefore, each of the five simulations completed the entire month of August 2023 in approximately 17 hours.

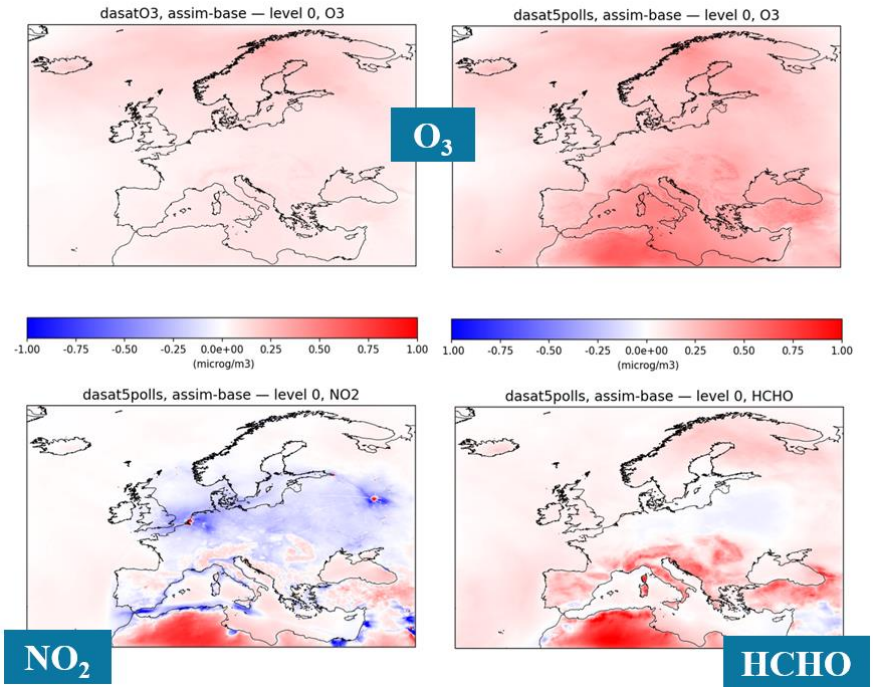


Fig.4: Monthly averaged differences of hourly surface concentrations of O_3, NO_2 and HCHO.

Acknowledgments

This study was developed under the CAMEO project (grant agreement No 101082125) is 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 Commission. Neither the European Union nor the granting authority can be held responsible for them."

References

[1] M. Mircea, L. Ciancarella, G. Briganti, G. Calori, A. Cappelletti, I. Cionni, M. Costa, G. Cremona, D'Isidoro, S. Finardi, et al. Assessment of the AMS-MINNI System Capabilities to Simulate Air Quality over Italy for the Calendar Year 2005. *Atmos. Environ.* **84**, pp. 178-188 (2014).
[2] M. Mircea, G. Grigoras, M. D'Isidoro, G. Righini, M. Adani, G. Briganti, L. Ciancarella, A. Cappelletti, G. Calori, I. Cionni, et al. Impact of Grid Resolution on Aerosol Predictions: A Case Study over Italy. *Aerosol Air Qual. Res.* **16**, pp. 1253-1267, (2016).
[3] https://sentinels.copernicus.eu/copernicus/sentinel-5p, last accessed 12/11/2025.
[4] M. Adani and F. Uboldi. Data Assimilation Experiments over Europe with the Chemical Transport Model FARM. *Atmos. Environ.* **306**, 119806, (2023).
[5] https://cams.ci.tno.nl/cso/, last accessed 12/11/2025.
[6] https://ci.tno.nl/gitlab/cams/cso/, last accessed 12/11/2025.
[7] D.F. Parrish and J.C. Derber. The National Meteorological Center's Spectral Statistical-Interpolation Analysis System. *Monthly Weather Review*, pp. 1747-1763, (1992).
[8] https://atmosphere.copernicus.eu/charts/packages/cams_air_quality/products/europe-air-quality-forecast-regulated, last accessed 12/11/2025.
[9] Carter, W.P.L. Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment. *Final Report to California Air Resources Board*, (2000). Available online: https://intra.engr.ucr.edu/~carter/pubs/s99doc.pdf (last accessed 12/11/2025).
[10] F.S. Binkowski, S.J. Roselle. Models-3 Community Multiscale Air Quality (CMAQ) Model Aerosol Component 1. Model Description. *J. Geophys. Res.*, **108**, D6, (2003).

REVEALING IMPROVED HOLE TRANSFER MECHANISMS IN PEROVSKITE SOLAR CELLS VIA SELF-ASSEMBLED MONOLAYERS AND A [1]BENZOTHIENO[3,2-B][1]BENZOTHIOPHENE INTERLAYER

Adriana Pecoraro[1*], Francesca Fasulo[1], Michele Pavone[2] and Ana Belén Muñoz-García[1]

[1]*Department of Physics "E. Pancini", University of Naples Federico II, 80126 Napoli, Italy*

[2] *Department of Chemical Sciences, University of Naples Federico II, 80126 Napoli, Italy*

ABSTRACT. Perovskite Solar Cells (PSCs) offer exceptional photoconversion efficiencies and are poised to revolutionize photovoltaics. Their performance and stability hinge on the interfaces between the perovskite absorber and charge transport layers. Introducing interlayers has boosted efficiency and durability, but the underlying mechanisms remain unclear. Using density functional theory, we explore the role of a C10-BTBT interlayer between a triple-cation perovskite and MeO-2PACz. Our findings reveal a novel bridge-mediated charge transfer mechanism where C10-BTBT enhances interfacial chemical bonding and accelerates hole injection. This molecular-level insight establishes a versatile framework for optimizing PSC interfaces to achieve improved performance and longevity.

Introduction

Perovskite solar cells (PSCs) have rapidly advanced in renewable energy, achieving efficiencies comparable to silicon cells while offering lower costs and flexible designs.[1] Continued progress relies heavily on interfacial engineering, where inserting interlayers between the perovskite and charge transport layers (CTLs) enhances energy alignment, interface quality, charge extraction, and reduces defects—boosting both efficiency and stability. Self-assembled monolayers (SAMs) are a proven interlayer approach, providing precise surface control. They passivate defects, tune energy levels, stabilize interfaces, and sometimes serve as CTLs themselves. Examples include N719 and N749 dyes, which improve defect passivation and moisture resistance, and 2PACz, which modifies ITO work function to enhance hole extraction. ZnO, SnO_2, and NiOx also benefit from SAMs or dipole molecules to improve charge transport.[2,3] Despite their advantages, SAMs often struggle with thermal stability and adhesion under temperature fluctuations. Self-assembled bilayers (SABs), composed of two distinct SAMs, offer improved stability and functional versatility. For example, the TFMBTA interlayer between MeO-2PACz HTL and perovskite enhances charge transfer, surface morphology, and defect passivation.

Computational studies of perovskite/CTL interfaces vary: some focus on isolated CTLs and band alignment; others simulate full interfaces with polarization and charge transport effects; only a few examine interlayer roles. Recently, Takhellambam et al.[4] introduced a C10-BTBT interlayer between MeO-2PACz and a triple-cation perovskite in a p-i-n PSC. This interlayer preserves perovskite structure while significantly improving charge extraction. Preliminary DFT modeling of the MeO-2PACz/C10-BTBT dyad showed favorable HOMO alignment aiding hole collection. Here, we extend this work with DFT[5] simulations of the complete perovskite/MeO-2PACz interface—with and without C10-BTBT, offering deeper insight into how the interlayer enhances interfacial interactions and charge transfer.

Results and Discussion

Structural model and electronic structure features of TriLHP bulk are depicted by **Figure 1** For the interfaces, we consider the thermodynamically favored (010) perovskite surface with the PbX_2 (with X=I, Br), being the most suitable for photovoltaic applications.according to literature.[6–8] We consider either the MeO-2PACz SAM or the MeO-2PACz/C10-BTBT dyad.

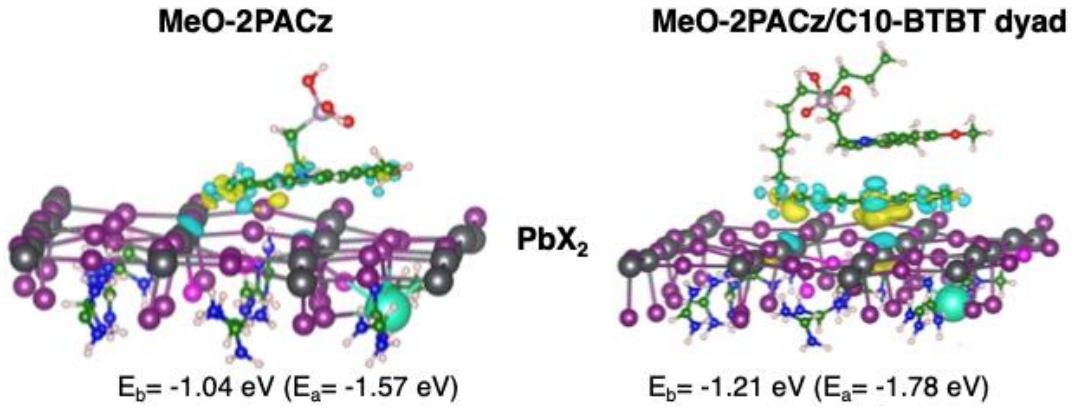


Figure 1. Lateral views of the relaxed TriLHP/Molecule(s) interfaces. Molecule(s) refer to MeO-2PACz or to the C10-BTBT/ MeO-2PACz dyad. Only the two outermost layers of the perovskite are shown for clarity. Electron density difference plots (iso-surface value = 0.005 a.u) are shown; yellow and blue regions denote charge accumulation and depletion zones upon interface formation, respectively. Binding (E_b) and adhesion (E_a) energies calculated at the DFT-HSE06 level of theory according to Eqs.1 and 2, respectively, are shown for each interface. Atomic color code: Cs, turquoise; Pb, dark gray; I, plum; Br, magenta; C, green; N, blue; H, light pink; P, lavender; O, red.

Binding and adhesion energies, which are calculated as:

$$E_b = E_{int} - E_{mol} - E_{surf} \quad (1)$$

$$E_a = E_{int} - E*_{mol} - E*_{surf} \quad (2)$$

Where the binding energy (E_b) is the difference between the total energy of the relaxed interface (E_{int}) and those of the isolated perovskite surface (E_{surf}) and molecule (E_{mol}). The adhesion energetic contribution (E_a) refers instead to the energies of the surface (E^*_{surf}) and the molecules (E^*_{mol}) at the interface minimum-energy geometry. Notably, the introduction of C10-BTBT improves the binding between PbX_2 termination and MeO-2PACz, achieving the most favorable binding energy (-1.21 eV).

A qualitative estimate of the band alignment for the PbX_2 interface model can be derived from density of states (DOS), as shown in **Figure 2**.

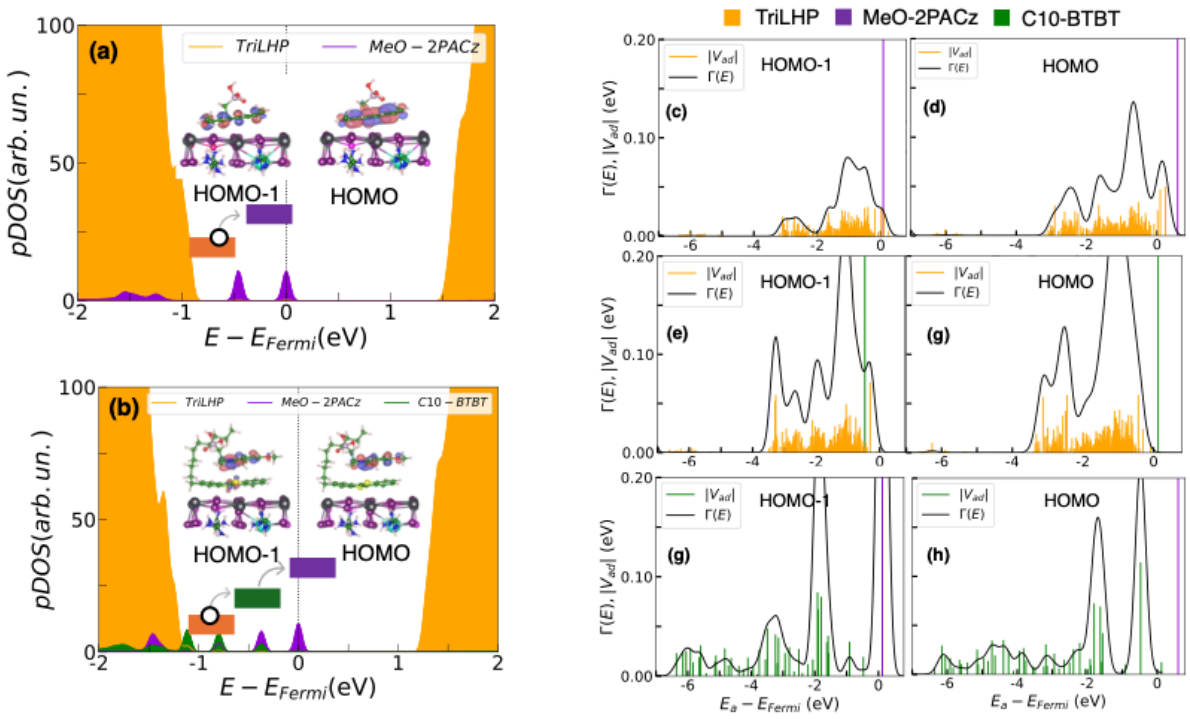


Figure 2. DFT-HSE06 projected density of states for the interfaces with MeO-2PACz (a) and dyad (b). Density of states plots are resolved for the different moieties (TriLHP, MeO-2PACz, and C10-BTBT). HOMO and HOMO-1 are depicted for each interface (isosurface value: 0.09 *a.u.*). DFT-HSE06 computed coupling matrix elements (V_{ad}), considering as Donor/Acceptor the couples MeO-2PACz/TriLHP (c and d), C10-BTBT/TriLHP (e and f), and MeO-2PACz/C10-BTBT (g and h) systems, for both the HOMO and HOMO-1 donor states and a wide range of acceptor states (E_a). The solid black line represents the spectral function G(E) calculated according to Eq. 3. In each plot, MeO-2PACz and C10-BTBT states are represented by violet and green vertical lines, respectively.

Computed DOS show favorable alignment between the molecule's HOMO and the perovskite valence band. At the bare perovskite/HTL interface, MeO-2PACz's HOMO and HOMO-1 lie above the perovskite VBM. Adding the interlayer maintains this alignment and increases the energy gap, enhancing the driving force for hole injection. The molecule's HOMO stays localized on the SAM, while HOMO-1 and lower states delocalize across both SAM and interlayer, as supported by molecular orbitals in Figure 2 (a and b).

Beyond thermodynamics, we examine C10-BTBT's impact on charge transport kinetics using the projection operator diabatization (POD) method by Futera et al.[9] This technique projects the Kohn-Sham Hamiltonian onto an orthogonal basis localized on donor and acceptor atoms, isolating donor-acceptor couplings to extract transition probabilities and times. We calculated these coupling elements at the HSE06 level using the POD implementation in CP2K. **Figure 2** shows the coupling elements VnmV_{nm}Vnm between donor (n) and acceptor (m) states for charge transfer at PbX_2/MeO-2PACz interfaces, with and without the interlayer. Without the interlayer, electron injection occurs directly from the SAM to the perovskite valence band. With the dyad, transfer happens in two steps: from MeO-2PACz to the interlayer's occupied states, then from C10-BTBT to the perovskite. Both HOMO and HOMO-1 orbitals are considered as transfer channels due to their energy levels above the perovskite VBM. Characteristic transfer times are evaluated via the spectral function G(E), defined as:

$$\Gamma_n(E) = 2\pi \sum_m V_{nm}^2 \delta(E - \varepsilon_m) \quad (3)$$

$$\tau = \frac{h}{2\pi \max(\Gamma)} \quad (4)$$

Calculated injection times are summarized in **Table 1**.

Table 1: Charge transfer times for the TriLHP/MeO-2PACz and the full TriLHP/dyad interfaces with PbX_2 termination, calculated at DFT-HSE06 level from $G_n(E)$ (Eq. 3).

Interface	**TriLHP/MeO-2PACz**	**TriLHP/C10-BTBT/MeO-2PACz**	**TriLHP/C10-BTBT/MeO-2PACz**
Donor/Acceptor	**MeO-2PACz/TriLHP**	**MeO-2PACz/C10-BTBT**	**C10-BTBT/TriLHP**
HOMO-1	8.21 fs	1.12 fs	2.18 fs
HOMO	4.84 fs	3.04 fs	2.15 fs

The computed transition times serve as comparative indicators of charge transfer kinetics between models, rather than absolute values for direct experimental comparison. At the bare TriLHP/MeO-2PACz interface, the fastest charge transfer (CT) originates from the molecule's HOMO, with a time of 4.84 fs. Adding C10-BTBT enables even faster two-step CT pathways: from MeO-2PACz HOMO-1 to BTBT, then from BTBT HOMO or HOMO-1 to the perovskite, with total times of 3.27 fs and 3.30 fs, respectively, both faster than the direct transfer without the interlayer.

In conclusion, our DFT modeling reveals that the C10-BTBT interlayer enhances both the stability and charge transfer efficiency at the TriLHP/MeO-2PACz interface. The PbX_2-terminated (010) perovskite surface offers the strongest binding with MeO-2PACz, further stabilized by C10-BTBT. Electronically, the interlayer shifts molecular orbitals to increase the hole injection driving force and accelerates charge transfer, with faster injection times confirming a more efficient, two-step transfer mechanism via C10-BTBT.

Computational Details

All-electron DFT calculations with periodic boundary conditions were performed using numerical atom-centered orbitals (NAOs) in FHI-aims.[10,11] Electrons were treated with atomic ZORA, using a 1×10^{-6} eV energy convergence threshold. Geometry optimizations employed the PBE functional with Tkatchenko–Scheffler van der Waals corrections, relaxing structures until forces were below 0.02 eV/Å.[12] Calculations were run at the Γ-point using the light-tier1 basis. Electronic properties and couplings were refined using the HSE06 functional and computed with CP2K

References

1 J. Han, K. Park, S. Tan, Y. Vaynzof, J. Xue, E. W.-G. Diau, M. G. Bawendi, J.-W. Lee and I. Jeon, Perovskite solar cells, *Nat Rev Methods Primers*, 2025, **5**, 1–27.

2 F. H. Isikgor, S. Zhumagali, L. V. T. Merino, M. De Bastiani, I. McCulloch and S. De Wolf, Molecular engineering of contact interfaces for high-performance perovskite solar cells, *Nat Rev Mater*, 2023, **8**, 89–108.

3 B. Dong, M. Wei, Y. Li, Y. Yang, W. Ma, Y. Zhang, Y. Ran, M. Cui, Z. Su, Q. Fan, Z. Bi, T. Edvinsson, Z. Ding, H. Ju, S. You, S. M. Zakeeruddin, X. Li, A. Hagfeldt, M. Grätzel and Y. Liu, Self-assembled bilayer for perovskite solar cells with improved tolerance against thermal stresses, *Nat Energy*, 2025, 1–12.

4 D. Takhellambam, L. A. Castriotta, G. Zanotti, L. Mancini, V. Raglione, G. Mattioli, B. Paci, A. Generosi, M. Guaragno, V. Campanari, G. Ammirati, F. Martelli, E. Calabrò, A. Cricenti, M. Luce, N. Yaghoobi Nia, F. Di Giacomo and A. Di Carlo, Enhancing Hole Transfer in Perovskite Solar Cell with Self-Assembled Monolayer by Introducing [1]Benzothieno [3,2-b][1]Benzothiophene Interlayer, *Solar RRL*, 2023, **7**, 2300658.
5 W. Kohn and L. J. Sham, Self-Consistent Equations Including Exchange and Correlation Effects, *Phys. Rev.*, 1965, **140**, A1133–A1138.
6 D. H. Cao, C. C. Stoumpos, C. D. Malliakas, M. J. Katz, O. K. Farha, J. T. Hupp and M. G. Kanatzidis, Remnant PbI2, an unforeseen necessity in high-efficiency hybrid perovskite-based solar cells?a), *APL Materials*, 2014, **2**, 091101.
7 C. Roldán-Carmona, P. Gratia, I. Zimmermann, G. Grancini, P. Gao, M. Graetzel and M. K. Nazeeruddin, High efficiency methylammonium lead triiodide perovskite solar cells: the relevance of non-stoichiometric precursors, *Energy Environ. Sci.*, 2015, **8**, 3550–3556.
8 A. Pecoraro, F. Fasulo, M. Pavone and A. B. Muñoz-García, First-principles study of interfacial features and charge dynamics between spiro-MeOTAD and photoactive lead halide perovskites, *Chem. Commun.*, 2023, **59**, 5055–5058.
9 Z. Futera and J. Blumberger, Electronic Couplings for Charge Transfer across Molecule/Metal and Molecule/Semiconductor Interfaces: Performance of the Projector Operator-Based Diabatization Approach, *J. Phys. Chem. C*, 2017, **121**, 19677–19689.
10 V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler, Ab initio molecular simulations with numeric atom-centered orbitals, *Computer Physics Communications*, 2009, **180**, 2175–2196.
11 V. Havu, V. Blum, P. Havu and M. Scheffler, Efficient O(N) integration for all-electron electronic structure calculation using numeric basis functions, *Journal of Computational Physics*, 2009, **228**, 8367–8379.
12 A. Tkatchenko and M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.*, 2009, **102**, 073005.

APROTIC SOLVATION THROUGH MOLECULAR DYNAMICS

Vanessa Piacentini, Adriano Pierini, Enrico Bodo, Sergio Brutti

Department of Chemistry, University of Rome La Sapienza, P.le Aldo Moro 5, 00185 Roma, Italy

ABSTRACT. The increasing demand for enhanced energy storage solutions has catalyzed substantial investigation into alternative electrolytes for lithium-ion batteries (LIBs). Significant findings reveal that appropriate solvent selections and their respective dielectric properties play crucial roles in the efficiency of these systems. For instance, studies have demonstrated that dimethyl sulfoxide (DMSO) and tetraethylene glycol dimethyl ether (TEGDME) exhibit contrasting solvation characteristics, wherein the latter's higher dielectric constant can lead to improved ion transport mechanisms when employed in binary mixes with organic carbonates. Similarly, the application of glyoxal acetals has emerged as a promising direction in maximizing battery performance, demonstrating effective accommodation of volumetric changes in silicon-based anodes and enabling high capacity operations at elevated temperatures. The incorporation of polarizable force fields in molecular dynamics simulations enhances the accuracy of modeling electrolyte properties, thereby fostering a deeper understanding of ion solvation and dynamic behaviors. Collectively, these advancements underscore the profound significance of solvent chemistry on the mechanisms governing lithium transport and overall battery efficacy, paving the way for the development of next-generation sustainable electrochemical energy storage systems..

State of the art and rationale of the research activity

The research activity embedded within the current state of lithium-ion battery (LIB) technologies emphasizes the necessity to explore novel electrolytes that can enhance the performance and safety of these energy storage systems. Traditional carbonate-based electrolytes, while widely employed, show limitations in terms of thermal stability, safety, and environmental sustainability. Consequently, there is a pivot towards understanding and employing alternative solvent systems to optimize ion transport properties and enhance battery performance. The exploration of glyoxal acetals as potential electrolyte components signifies a notable advancement in this domain. These compounds exhibit film-forming capabilities which are essential for forming stable solid-electrolyte interphases (SEIs) on silicon and graphite anodes. This capability to maintain structural integrity during significant volume changes during charge and discharge cycles could lead to improved cycling stability, thereby realizing higher capacity outcomes in LIB applications. Studies indicate that the incorporation of glyoxal acetals can result in impressive electrochemical performance metrics, facilitating stable operation under various conditions. Simultaneously, the comparative investigation of dimethyl sulfoxide (DMSO) and tetraethylene glycol dimethyl ether (TEGDME) reveals substantial insights into solvation dynamics and lithium-ion transport properties. The marked differences in solvation behavior between these solvents underscore the necessity for a nuanced understanding of their interactions at the molecular level, informing the design and efficiency of battery electrolytes. Notably, DMSO's lithium coordination features significant implications for the ionic conductivity of the electrolyte, suggesting that solvent choice is fundamentally linked to the efficacy of lithium transport mechanisms. In tandem with experimental approaches, molecular dynamics simulations have emerged as a powerful tool for investigating electrolyte properties. The integration of polarizable force fields in simulations can capture the dynamic and interacting behaviors of ionic species more accurately than traditional fixed

charge models. These advancements enable researchers to predict the impacts of different solvent compositions and to fine-tune the physicochemical attributes of novel electrolytes. Findings have demonstrated that polarizable force fields can yield enhanced descriptions of the structural and dynamical properties critical to electrolyte performance. Furthermore, the identification and optimization of novel electrolyte formulations often rely on qualitative comparisons across a broad spectrum of solvent systems. This methodological framework facilitates the exploration of various solvate configurations, enabling a systematic understanding of the coupling between ionic mobility and the electric field within the electrolyte medium. By articulating these relationships through both experimental and computational lenses, researchers can derive more coherent insights into how compound structures impact ionic conduction and overall battery performance. The trend toward diversifying solvent options extends the landscape for innovative battery chemistries, incorporating ethers, ionic liquids, and other organic solvents into the mix. This exploration not only serves to alleviate constraints posed by conventional carbonate-based systems but also aligns with global pushes toward greener and more efficient energy storage solutions. As the field advances, the collective understanding cultivated from empirical findings and theoretical models will support the development of sophisticated electrolytes tailored for improved battery designs. In conclusion, the ongoing research activity in exploring alternative electrolytes is critical for surmounting the challenges posed by traditional lithium-ion battery technologies. The dual focus on glyoxal acetals and the comparative study of solvation effects provides a promising avenue for enhancing the structural stability, safety, and efficiency of next-generation batteries. These investigations ultimately contribute to broader goals of sustainability in energy storage technology and the quest for more efficient electrochemical systems (figure 1)..

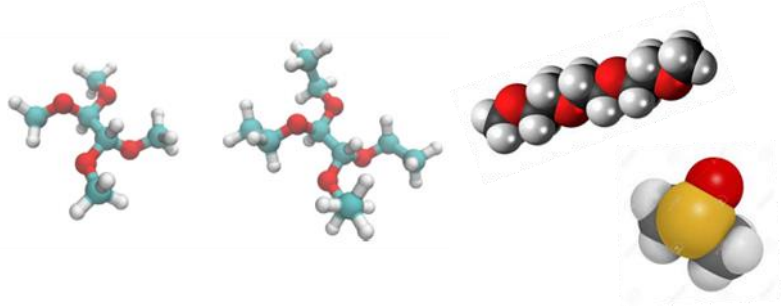


Figure 1. Chemical systems studied within the research activity

Results

Advances in Electrolyte Choices for Lithium-Ion Batteries. The current landscape of lithium-ion battery (LIB) technology increasingly underscores the importance of alternative electrolyte systems, particularly in the context of growing energy demands and sustainability concerns. Traditional carbonate-based electrolytes, such as LP30, have demonstrated several limitations relating to thermal stability, safety, and toxicity. These deficiencies become especially critical when paired with high-capacity anode materials like silicon, which undergo significant volumetric changes during lithiation and delithiation. Recent studies have focused on the formulation of new electrolyte components that enhance performance and mitigate issues attributed to conventional solvents. The exploration of glyoxal acetals represents a significant shift in this regard. These compounds have been shown to form stable solid-electrolyte interphases (SEIs) on silicon-based anodes. By effectively accommodating the volume changes during operation, glyoxal acetals facilitate improved cycling stability and longevity of the battery. Their high ionic conductivity further contributes to the overall efficiency of the LIB system, enabling rapid charge-discharge cycles without the typical issues associated with SEI degradation. Moreover, the adoption of polarizable force fields in molecular dynamics simulations has revolutionized the research approach. Unlike fixed point-charge models, which often lack precision in predicting the behavior of highly mobile charge carriers in liquid electrolytes, polarizable force fields provide deeper insights into the structural and dynamical properties. This refinement in modeling allows researchers to

accurately gauge the influence of various solvent compositions on ionic transport, thus fostering the design of next-generation electrolytes that can meet the evolving demands of energy storage technologies.

Solvation Dynamics and Ionic Transport Mechanisms. An essential aspect of enhancing electrolyte performance lies in understanding solvation dynamics and the mechanisms of lithium-ion transport. Evaluations of the solvents dimethyl sulfoxide (DMSO) and tetraethylene glycol dimethyl ether (TEGDME) reveal significant variances in solvation behavior that impact lithium mobility. Experimental analyses, coupled with molecular dynamics simulations, underscore the predominant role of these solvents in governing the fluid dynamics within the electrolyte system.

DMSO has emerged as a favored solvent due to its superior solvation properties, which enhance ionic mobility compared to TEGDME. Temperature-dependent studies demonstrate that as the operating temperature increases, DMSO-based electrolytes exhibit decreased bulk resistance, allowing for enhanced lithium transport. Conversely, TEGDME displays less optimal performance with observable disparities in diffusion coefficients among the different ionic species. The nuanced interactions between the solvent and lithium ions not only affect conductivity but also determine the structural characteristics of the electrolyte, such as its viscosity and thermal stability.

The disparities in solvent behavior emphasize the importance of meticulous selection of solvent mixtures designed to promote optimal ion transport. Through focused empirical and computational approaches, researchers can tailor solvent compositions to address specific performance goals, aligning with the operational requirements of next-generation batteries.

Implications and Future Directions in Battery Technology. The implications of these research findings extend well beyond immediate performance improvements in LIBs. The advancements in exploring alternative electrolytes and solvation strategies contribute to a broader narrative focusing on sustainability in battery technology. As the demand for cleaner, safer, and more efficient energy storage solutions escalates, the integration of less toxic and environmentally friendly solvents becomes paramount.

Glyoxal acetals, for instance, represent a forward-thinking response to concerns surrounding the safety of conventional carbonate-based electrolytes. Their intrinsically lower toxicity and enhanced thermal stability renders them attractive candidates for developing eco-friendly battery systems. Moreover, the new insights gained from advanced simulations facilitate the optimization of these components for diverse applications, including renewable energy storage systems and electric vehicles.

Looking ahead, it is essential to continue bridging the gap between experimental findings and theoretical advancements to facilitate practical applications of novel electrolytes. Emerging research efforts should also aim to explore the potential of synthetic and biological alternatives in solvent formulation, considering the intersection of chemistry and environmental impact.

In conclusion, the ongoing exploration of alternative electrolytes, particularly the promising glyoxal acetal formulations, alongside enhanced solvation dynamics in traditional solvents like DMSO and TEGDME, signals a transformative era for lithium-ion battery technology. The collective advancements catalyzed by this research hold immense potential to enhance the performance, safety, and sustainability of energy storage solutions, ultimately aligning with global initiatives toward cleaner energy systems and efficient resource utilization. As researchers continue to refine and develop

innovative electrolyte systems, the future of battery technologies promises to be more resilient, adaptable, and environmentally sustainable.

Acknowledgments

The authors want to acknowledge the financial support of the European Union´s Horizon Europe transport program under the project SiGNE (grant agreement No 101069738).

References

[1] doi.org/10.1002/cssc.202301962

[2] doi.org/10.1002/open.202400134.

IN SILICO ANALYSIS OF THE NUTRIEPIGENOMIC POTENTIAL OF SAFFRON BIOACTIVE MOLECULES

Paolo Roberto Saraceni[1*], Caterina Arcangeli[1], Barbara Benassi[1]

[1]*Division of Biotechnologies, Department for Sustainability, Italian National Agency for New Technologies, Energy and Sustainable Economic Development (ENEA), Via Anguillarese 301, 00123 Rome, Italy*[14]

ABSTRACT. This report describes the interaction dynamics between bioactive compounds from Saffron (*Crocus sativus* L.) and two specific DNA methyltransferases. Molecular docking and molecular dynamics (MD) simulations were performed using the high-performance parallel computing resources of CRESCO6 and the GROMACS 2021 suite. The integrated in silico approach provides novel insights into the nutri-epigenomic potential of saffron, supporting its classification as a promising functional food.

Introduction

Recent studies have increasingly elucidated the complex interplay between organisms and their environment, with epigenetic regulation playing a central role in linking genotype to phenotype. Nutri-epigenomics explores how functional foods modulate these epigenetic mechanisms, influencing health outcomes [1,2].

Saffron is recognized for its protective effects against chronic diseases. While its antioxidant and anti-inflammatory properties are well documented, its nutri-epigenomic potential remains underexplored. Emerging evidence indicates that saffron bioactive compounds interact with key epigenetic enzymes, affecting histone modifications and DNA methylation [3,4]. The detailed molecular mechanisms underlying these interactions, however, require further investigation.

This study applies molecular docking and molecular dynamics simulations to characterize the binding of saffron constituents to epigenetic targets. This study suggests that specific bioactive molecules of saffron can form stable interactions with epigenetic enzymes, potentially modulating their activity. These findings support the role of saffron as a functional food with nutri-epigenomic properties [5].

Computational methods

The ligands and protein structures identified through molecular screening were prepared for the docking procedure using AutoDock Tools 1.5.6 [6], where Gasteiger charges were assigned to saffron bioactive molecules (BM) and Kollmann charges to the epigenetic targets (ET). The highest-scoring ligand poses within the enzyme binding sites, obtained from docking with VINA [7], served as the initial configurations for subsequent molecular dynamics (MD) simulations. MD simulations were conducted using GROMACS 2021. Table 1 lists all complexes analyzed through these in silico experiments. The epigenetic enzymes were parameterized with the AMBER 99sb-ildn force field [8], while ligand

[14] Corresponding author. E-mail: paolo.saraceni@enea.it

parameters were generated via ACPYPE (AnteChamber PYthon Parser interfacE) [9] utilizing the AnteChamber tool [10]. Each system was solvated using TIP3P water models [11] and neutralized with 150 mM NaCl. Bond constraints involving hydrogen atoms were applied using the LINCS algorithm [12], and initial velocities were assigned according to a Maxwell-Boltzmann distribution. Unrestrained MD simulations were performed in the NPT ensemble for 200 nanoseconds. Key simulation details are summarized in Table 2. Trajectories of MD simulated complexes were visualized using VMD and PyMOL software, widely recognized tools for their effectiveness in graphically representing complex molecular systems.

Table 1: List of simulated systems.

Ligand	Enzyme ET1[e]	Enzyme ET2[f]
BM1[a]	ET1:BM1	ET2:BM1
BM2[b]	ET1:BM2	ET2:BM2
BM3[c]	ET1:BM3	ET2:BM3
BM4[d]	ET1:BM4	ET1:BM4

a) Crocin, PubChem ID: CID 10368299
b) Crocetin, PubChem ID: CID 5281232
c) Picrocrocin, PubChem ID: CID 130796
d) Safranal, PubChem ID: CID 61041
e) DNA methyltransferase 1, PDB ID: 3swr.pdb
f) DNA methyltransferase 3a, PDB ID: 5yx2.pdb

Table 2: MD simulation parameters.

Simulations	All-atom based
Energy Minimization	Steepest descent
Cut-off scheme	Verlet
Long-range electrostatic force	PME method [13]
Equilibration	300ps
Thermostat	Bussi-Donadio-Parrinello [14]
Temperature	300K
Barostat	Parrinello-Rahman [15]
Pressure	1bar
Dielectric constant	1
Integration step	2fs
Length of simulations	200ns

Results and Discussion

To assess the stability of the complexes and the dynamics of individual residues, the RMSF (Root Mean Square Fluctuation) analysis was performed. This analysis measures the average fluctuation of the Cα atoms of the enzyme and all the atoms of the ligand relative to their average positions over the simulation time, providing insights into the flexibility and conformational variability of the complexes.

The analysis highlighted notable differences in residue mobility between the ET1 and ET2 systems.

Within the ET1 systems, the ET1:BM4 complex exhibits fluctuations primarily concentrated in the amino acid region spanning residues 700 to 1100, with a notable peak at residue 1125, reflecting an increase in conformational flexibility in this region that is not observed in the other complexes. Conversely, the ET1:BM1 complex shows the lowest RMSF values, revealing greater structural rigidity and suggesting it represents the most stable binding conformation.

Regarding the ET2–ligand complexes, all complexes exhibit similar RMSF fluctuations across the binding interface, with ET2:BM1 showing the lowest RMSF values. In contrast, residues 700–750 of ET2:BM4 display increased fluctuations that are not present in the other complexes, indicating reduced local stability and greater conformational flexibility and greater conformational flexibility in this region.

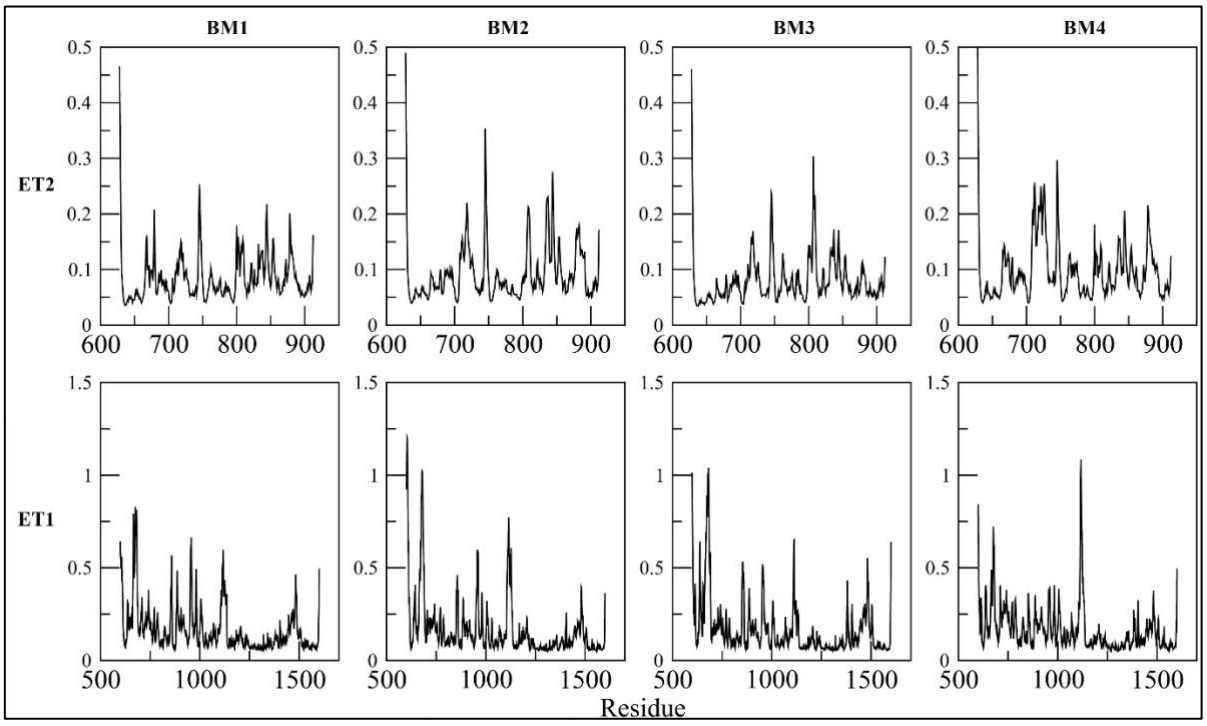


Figure 1: RMSF values, as a function of simulation time, of the systems in which ET1 and ET2 enzymes are in complex with BM1, BM2, BM3 and BM4.

Conclusion

Emerging evidence suggests that bioactive compounds from dietary sources can modulate epigenetic enzyme activity, thereby influencing gene regulation and potentially impacting human health and disease. This study applied a structure-based in silico approach, combining molecular docking and molecular dynamics simulations, to explore the interaction between four saffron-derived molecules and two key epigenetic enzymes. The majority of the simulated complexes showed stable ligand-target associations, indicating a possible epigenetic modulatory role for these compounds. Given the limited data currently available on the epigenetic actions of saffron, these computational findings highlight the need for further simulation studies and experimental validation, both to assess their therapeutic relevance and to support their classification as functional food components. The results presented in this report constitute a part of a broader study entitled ‘Epigenomics Nutritional Insights of Crocus

sativus L.: Computational Analysis of Bioactive Molecules Targeting DNA Methyltransferases and Histone Deacetylases', which has been submitted to the International Journal of Molecular Sciences.

Acknowledgments

The computational analyses were conducted on the CRESCO (Computational RESearch Centre on COmplex Systems) cluster [16], part of ENEAGRID — the infrastructure granting access to the high-performance computing resources of the ICT (Information and Communication Technologies) division, established by ENEA (Italian National Agency for New Technologies, Energy and Sustainable Economic Development).

References

1] L. Pérez et al. Nutritional epigenetics: The role of diet in gene expression regulation. Nutrients, 15(5), pp. 1172, (2023).
[2] D. J. McClements, F. Li, and H. Xiao. The nutraceutical bioavailability classification scheme: Classifying nutraceuticals according to factors limiting their oral bioavailability. Annual Review of Food Science and Technology, 14, pp. 299–327, (2023).
[3] S. Ashrafi et al. Saffron, its active components, and their association with DNA and histone modification: A narrative review of current knowledge. Antioxidants, 12(8), pp. 1619, (2023).
[4] M. De la Fuente Muñoz et al. Effects of supplementation with standardized saffron extract (affron®) on the kynurenine pathway and melatonin synthesis in rats. Antioxidants, 12(8), pp. 1619, (2023).
[5] G. Zhao and L. N. Rusche. Sirtuins in epigenetic silencing and control of gene expression in fungi. Annual Review of Microbiology, 76, pp. 157–178, (2022).
[6] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson. AutoDock Tools 1.5.6: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16), pp. 2785–2791, (2009).
[7] O. Trott and A. J. Olson. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2), pp. 455–461, (2010).
[8] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins*, 78(8), pp. 1950–1958, (2010).
[9] J. C. Sousa da Silva and W. F. Vranken. ACPYPE - AnteChamber PYthon Parser interfacE. *BMC Research Notes*, 5, 367, (2012).
[10] J. Wang, W. Wang, P. A. Kollman, and D. A. Case. Antechamber: an accessory software package for molecular mechanical calculations. *Journal of the American Chemical Society*, 222, (2001).
[11] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2), pp. 926–935, (1983).
[12] B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12), pp. 1463–1472, (1997).
[13] Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H. and Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.*, 103, pp. 8577–8592, (1995).
[14] Bussi, G., Donadio, D. and Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.*, 126, (2006).
[15] Parrinello, M. and Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.*, 52, 7182–7190, (1981).
[16] Iannone, F., Ambrosino, F., Bracco, G., De Rosa, M., Funel, A., Guarnieri, G., Migliori, S., Palombi, F., Ponti, G., Santomauro, G. and Procacci, P. CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum Scale layout. 2019 International Conference on High Performance Computing Simulation (HPCS), pp. 1051–1052, (2019).

NEUTRONIC ANALYSES FOR THE FEASIBILITY STUDY OF ACCELERATOR-DRIVEN SYSTEMS

Massimo Sarotto[1], Roberto Pergreffi[2] and Francesco Lodi[2]

[1]*ENEA NUC-ENER-PRO, Strada per Crescentino 41, 13040 Saluggia, Italy*

[2]*ENEA NUC-ENER-PRO, Via dei Mille 21, 40121 Bologna, Italy*

ABSTRACT. The Accelerator-Driven System (ADS) concept represents one of the possible strategies to enhance the sustainability of nuclear energy through the minimization of nuclear waste with very long-term radio-toxicity. It consists essentially in a sub-critical core coupled with a proton accelerator that - through the spallation process in a liquid metal target - supplies additional neutrons to the core which can be loaded - and transmute - significant amounts of transuranic nuclides without posing safety issues. Since the 2000s, the laboratory for Design and Analysis of Nuclear Systems has been involved in numerous - national and international - research activities related to the lead-cooled ADSs: as briefly described, accurate neutronic analyses - relying on both deterministic and Monte Carlo codes - are needed for a complete feasibility study of these innovative concepts.

Introduction

Since the 2000s, with an initial impetus provided by Carlo Rubbia (President of ENEA in 1999-2005), the ENEA laboratory for the Design and Analysis of Nuclear Systems (NUC-ENER-PRO) has been engaged in research studies related to the Accelerator-Driven System (ADS) concept, aiming at the closure of the fuel cycle to enhance the sustainability of nuclear energy. Besides the energy production, the main ADS objective is represented by the minimization of nuclear waste, as transuranic (TRU) nuclides having very long-term radio-toxicity. The ADSs investigated in the last two decades - within Euratom projects, national programmes and international collaborations - consisted essentially in a proton accelerator coupled with a sub-critical core that - without posing safety issues - can be loaded with significant amounts of TRU nuclides and minimise their amount by transmutation. The “additional” neutrons for the operability of the fast-spectrum system are provided through the spallation process of the impinging protons in a target located in the centre of the core. Certainly, being a sophisticate and very innovative system, the cost/benefit ratio will have to be accurately verified, and a challenging R&D program would be necessary for its development and effective realization.

In this research field, NUC-ENER-PRO has gained a valuable expertise in the study of the lead-cooled ADSs, having the liquid lead (or lead-bismuth eutectic) the double role of coolant for the sub-critical core and target for the protons’ spallation, as represented in the conceptual scheme shown in the left part of Fig. 1. In the ADS feasibility studies, both deterministic and Monte Carlo (MC) codes are required for a complete neutronic characterisation. Indeed, while deterministic codes can “easily” perform fuel cycle analyses with the typical upper boundary at 20 MeV energy, the MC codes are unable to perform depletion calculations for sub-critical systems with external proton source. At the same time, the deterministic codes need an “external source term” to consider only neutrons with energy lower than 20 MeV deriving from protons’ spallation, that can be computed only by MC codes.

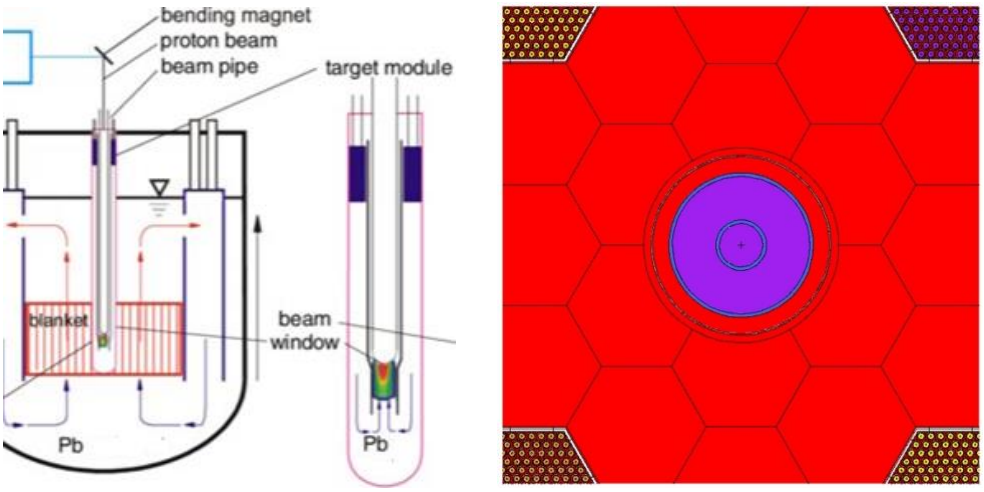


Fig.1: Conceptual scheme of a lead-cooled ADS (left, vertical); 2D horizontal view of the MCNP model of the target region at core mid-plane (right).

Methodology and neutronic codes

In the NUC-ENER-PRO laboratory, the ADS conceptual design and neutronic analyses are usually carried out with the deterministic code ERANOS ver. 2.2 [1], nowadays coupled with the most up-to-date nuclear data library ENDF/B-VIII.0 prepared in a special format readable by ERANOS [2]. The MC code mainly adopted is MCNP ver. 6.1 [3] with the same ENDF/B-VIII.0 nuclear data [4]. The MC simulations start from the impinging protons - typically with an energy ~600÷800 MeV- in the liquid metal target positioned at the centre of the core. The nuclear interaction options used in the MCNP simulations are as follows [3]:

- for interactions between neutrons and all the isotopes, the ENDF/B-VIII.0 neutron library when energy is lower than 20 MeV and the MCNP physics models for upper energies;
- for interactions between protons and target nucleons, the LA150h proton library when energy is lower than 150 MeV and the MCNP physics models for upper energies.

Table 1 indicates the physics models used in MCNP as reference for the different stages of the complex spallation phenomena such as: intranuclear cascade (INC) with a pre-equilibrium stage, equilibrium evaporation and fission of the residual nuclei.

Table 1: Physics models used in MCNP as reference [3].

INC model	Pre-equilibrium model	Evaporation model	Residual nuclei fission model
Bertini	Multistage	Dresner	RAL (Rutherford Appleton Laboratory)

As mentioned, the external source term for deterministic codes as ERANOS is provided by means of MC codes and represents the neutrons below 20 MeV energy appearing in the ADS after protons' spallation and successive neutronic reactions (*e.g.*, scattering and fission). This term includes the average number of neutrons produced by each spallation proton, a representative spectrum (*e.g.*, at 51 energy groups) and a normalized neutron density in each region of the core. To obtain this spatial neutron density, mainly distributed in the central target and surrounding fissile core, each neutron produced by spallation is "followed" in the MC simulations until it reaches energies below 20 MeV.

The ADS multiplicity is represented by the k_S parameter, defined similarly to the standard k_{eff} by considering the neutron multiplication "*M*" (due to fissions) per spallation neutron:

$$k_S = \frac{\iint \phi \upsilon \Sigma_f dEdV}{\iint \phi \upsilon \Sigma_f dEdV + 1} \quad (1)$$

where the integral in equation (1) represents the reaction rate (s^{-1}) of the neutrons production by fission "$\upsilon\Sigma_f$" in the core, calculated by ERANOS through the flux values (ϕ) in the mesh volumes (*V*) of the core regions at all energies (*E*<20 MeV). The correct estimation of the k_S parameter is fundamental for the correct evaluation of the proton current required to sustain the nominal core power in operation.

An example of application

Recently, under commission of the Transmutex Swiss company, NUC-ENER-PRO carried out the feasibility study of a ~300 MW_{th} lead-cooled ADS working as a "dirty" plutonium (Pu) burner of a spent mixed oxide fuel coming from pressurized water reactors. In some details:

- the ADS operates in an open fuel cycle made of 6 batches of 1 year each;
- the sub-critical core is loaded with hexagonal assemblies arranged around the accelerator and spallation modules occupying three rings in the core lattice, as shown in right part of Fig. 1;
- as schematised in the left part of Fig. 2 representing the central part of the ERANOS (axial) core model, the fissile region is divided in two zones (FINN and FOUT having a different dirty Pu enrichment) for radial power flattening, whereas the central two rings for accelerator and spallation modules are divided in three axial regions (beam, window and target);
- the external source term for ERANOS – made of neutrons below 20 MeV energy deriving from spallation – was calculated by Transmutex adopting the GEANT-4 code [5].

As a benchmark, additional analyses were performed in ENEA by means of the MCNP code in "sub-critical mode" (*i.e.*, starting the simulation from the 800 MeV protons impinging on the target), that permitted to retrieve accurate results for the start-up core at the beginning of the 1-year irradiation sub-cycle: the right part of Figs. 1 and 2 point out the level of detail of the MCNP model.

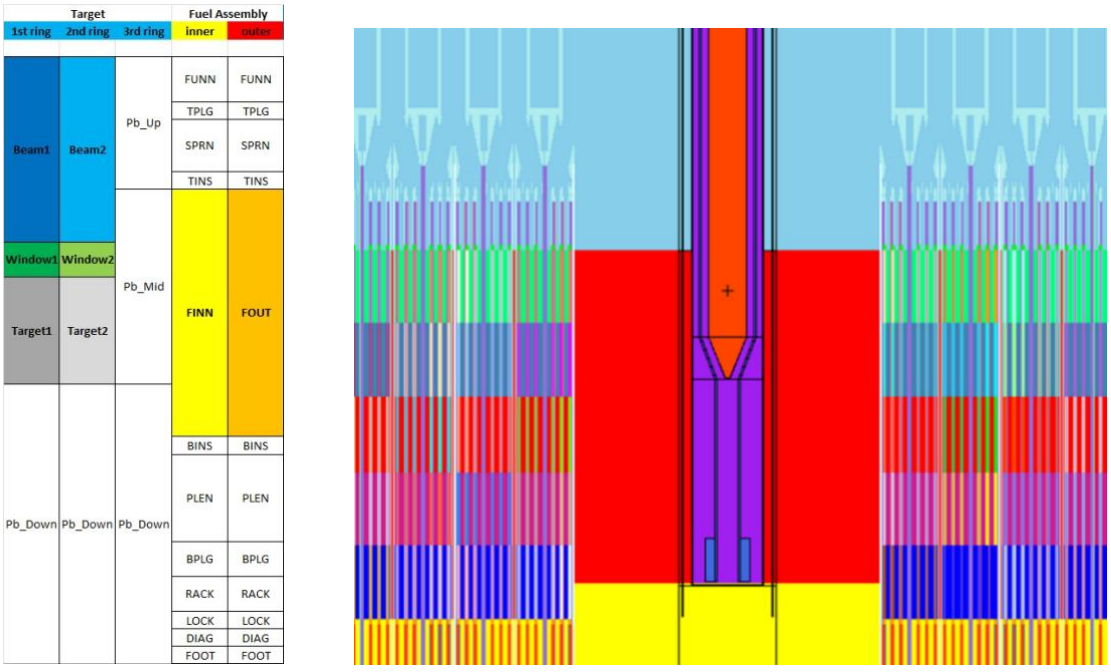


Fig.2: 2D axial views of the ADS: ERANOS (left) and MCNP (right) models.

The ADS "key" parameters (as k_S and resulting proton current) were evaluated in the 6-year fuel cycle with the ERANOS code, by adopting a 3D hexagonal model of the core with a 2p/3 radial symmetry. Fig. 3 shows the obtained power distributions with the values of the three peaking factors: (i) Fuel Assembly Distribution Factor (FADF) in the sub-critical core; (ii) Fuel Pin Distribution Factor (FPDF) among the pins in the hot FAs of the inner and outer fuel zones (circled in blue); (iii) Fuel Pin Axial Factor (FPAF) in the hottest pin of the hot FAs. The results (shown graphically by a python procedure) refer to the end of the first 1-year irradiation and were used as input data for the thermal-hydraulic and thermo-mechanical analyses. Both ERANOS and MCNP (serial and parallel) simulations - performed for the ADS conceptual design and neutronic analyses - were run on the ENEA CRESCO HPC systems.

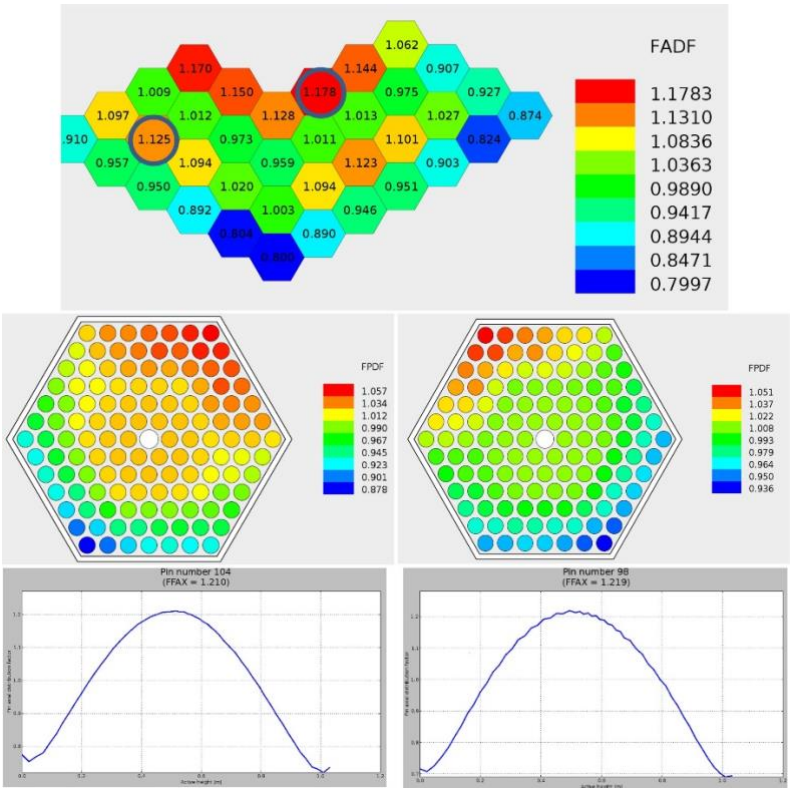


Fig.3: ADS power peaking factors: FADF (2p/3 symmetry, top), FPDF (middle) and FPAF (bottom).

References

[1] G. Rimpault et al. The ERANOS code and data system for fast reactor neutronics analyses. Proc. of *PHYSOR 2002*: Int. Conf. on New Frontiers of Nuclear Technology: Reactor Physics, Safety and High-Performance Computing, Seoul, Korea, October 7-10 (2002).

[2] D.M. Castelluccio et al. Realization of an adjusted nuclear data library based on ENDF/B-VIII.0 nuclear data evaluations for the ALFRED core, Proc. of *FR22*: Int. Conf. on Fast Reactors and related fuel cycles: sustainable clean energy for the future, Vienna, Austria, April 19-22, CN291-422 (2022).

[3] B. Pelowitz, MCNP6 user's manual, Tech. Rep. LA-CP-13-00634 Rev. 0, Los Alamos National Laboratory (2013).

[4] D.A. Brown et al. ENDF/B-VIII.0: The 8th Major Release of the Nuclear Reaction Data Library with CIELO-project Cross Sections, New Standards and Thermal Scattering Data. *Nucl. Data Sheets* **148**, pp. 1-142 (2018).

[5] S. Agostinelli et al. GEANT4 – a simulation toolkit, *Nucl. Instrum. Meth. A*, **506**(3), pp. 250-303 (2003).

MHD MIXED CONVECTION AROUND CURVED PIPES IN WATER-COOLED BREEDING BLANKETS

Simone Siriano*, Lorenzo Melchiorri, Alessandro Tassone

Nuclear Engineering Research Group - DIAEE, Sapienza University of Rome, Corso Vittorio Emanuele II 244, 00186 Rome, Italy

Corresponding author: simone.siriano@uniroma1.it

ABSTRACT. This contribution presents computational fluid dynamics simulation simulations of the magnetohydrodynamics (MHD) mixed convection regime in a variant of the Water-Cooled Lead-Lithium blanket, relevant to nuclear fusion reactors. The study employs Ansys CFX 2021R1 on CRESCO6 to solve the MHD equations for an upward lead-lithium (PbLi) flow, cooled by nested U-pipes, under fusion-relevant magnetic fields and volumetric heating. Simulations cover equatorial, apical, and near-divertor cells. Results show strong suppression of convection by the magnetic field, with heat transfer dominated by conduction across all positions.

Introduction

The Water-Cooled Lead-Lithium (WCLL) Breeding Blanket (BB) is a promising concept for DEMO fusion reactors, designed to optimize tritium breeding and heat extraction [1]. The lead-lithium (PbLi) breeder circulates at reduced velocities to limit MHD effects, while pressurized water cools the system via double-walled tubes (DWT) immersed in PbLi and square channels drilled into the First Wall (FW). The interplay of intense non-uniform volumetric heating, generated by the neutronic power deposition, a strong magnetic field and intricate geometry leads to substantial temperature gradients and to a mixed-convection MHD flow regime where buoyancy and electromagnetic forces interact with the imposed forced convection flow [2]. Accurate characterization of these regimes is essential for blanket thermal-hydraulic performance [3]. This study numerically investigates the magneto-convection regime in a WCLL BB variant, focusing on three elementary cells with distinct gravity orientations: equatorial (EC), apical (AC), and near-divertor (DC).

Problem formulation

In this WCLL configuration, the Breeding Zone (BZ) is composed by rectangular channels that extend throughout the blanket's height where PbLi flows. To cooling the BZ, DWT are radially inserted from the black plate through apertures drilled stiffening plates [4], therefore the fundamental cooling components are two nested U-pipes in the PbLi, and two square channels drilled in the FW, as shown in Figure 3. Ansys CFX code employs a MHD model based on the inductionless approximation, considering the electric potential ϕ as the main electrical variable. Considering an incompressible flow and the Boussinesq approximation for buoyancy, the dimensionless governing equations are [2]:

$$\nabla \cdot \mathbf{u} = \mathbf{0} \quad (1)$$

$$\frac{Gr}{Ha^4}\left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\mathbf{u}\right] = -\nabla p + \frac{1}{\mathrm{Ha}^2}\nabla^2 \mathbf{u} + \mathbf{J} \times \mathbf{B} - \mathbf{g} \quad (2)$$

$$Pe[(\mathbf{u} \cdot \nabla)T] = \nabla^2 T + \mathrm{Q} \quad (3)$$

$$\nabla^2 \phi = \nabla \cdot (\mathbf{u} \times \mathbf{B}) \quad (4)$$

$$\mathbf{J} = -\nabla \phi + \mathbf{u} \times \mathbf{B} \quad (5)$$

The symbols $\mathbf{u}$, p, $\boldsymbol{J}$, $\boldsymbol{B}$, $\boldsymbol{g}$ and T represents, respectively, the velocity, the dynamic pressure, the electric current density, the imposed uniform and constant magnetic field oriented along the Y direction (ref. Figure 3), the gravitational acceleration and the temperature field. The symbol Q is the volumetric power source term that decreases exponentially moving away from the FW, following the law presented in ref. [5].The dimensionless parameters are the Hartmann number $\mathrm{Ha} = \mathrm{L_{Ha}} B\sqrt{\sigma\mu}$, which represents the ratio between the electromagnetic to the viscous forces and governs the thickness of the MHD boundary layers; the Grashof number, $\mathrm{Gr} = \mathrm{g}\rho^2\beta L_{Gr}^3 \Delta T_{Gr}/\mu^2$, which represents the ratio of buoyancy to viscous forces; the Peclet number, $\mathrm{Pe} = \mathrm{PrGr/Ha}^2$, which represents the ratio of advection to conduction heat transfer through the Prandtl number $\mathrm{Pr} = \mathrm{c_p}\mu/k$, the ratio of momentum diffusivity to thermal diffusivity. All the parameters used to define the dimensionless number are collected in

Table **2** with the respective values for all the materials. The structural material has been considered as perfectly electrical conductivity materials which is a conservative assumption considering the braking Lorentz force action on the fluid and the heat transfer [6]. The system operates under a magnetic field of $4.4\,T$, which corresponds to $Ha = 8550$ ($\mathrm{L_{Ha}} = 82\,mm$, half-toroidal length, ref. Figure 3). The Grashof number results 7.08×10^7, where the reference temperature difference $\Delta T_{Gr} = \hat{Q} L_{Gr}^2/\kappa = 185.95\,K$ is estimated from the average volumetric heat source in the region between the pipes ($\hat{Q} = 8.82\,MW/m^3$, where $\mathrm{L_{Gr}} = 18.25\,mm$ half-distance between the pipes in the X direction).

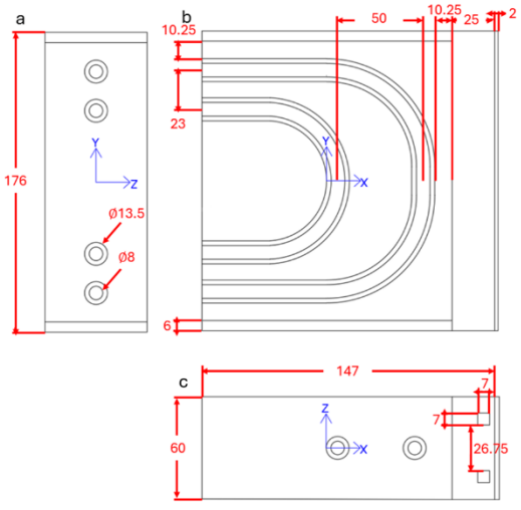


Figure 3. Elementary cell geometry on a toroidal (Y)-poloidal (Z) (a), radial (X)-toroidal (b) and radial-poloidal plane (c). The measurements are in millimeters.

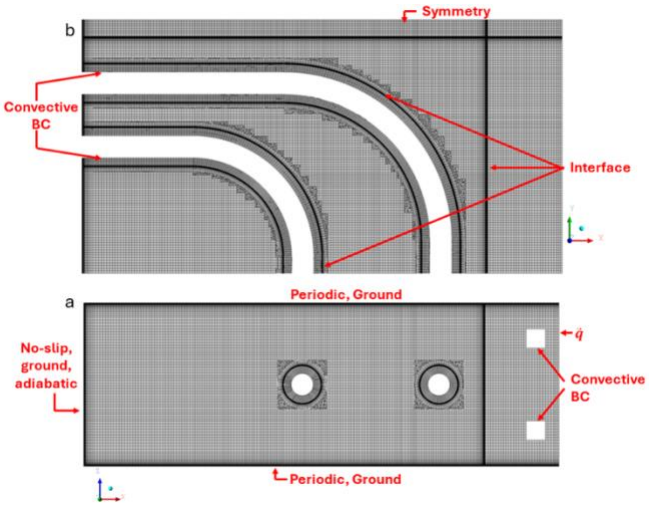


Figure 4. Computational grid and BCs employed on a radial-toroidal plane (a) and radial-poloidal plane (b).

Table 2. Physical properties of PbLi [7], Eurofer (E) [8] and tungsten (W) [9]. T is the temperature expressed in K. The T independent ones are considered at $T_{ref} = 710\ K$.

Property	PbLi value	E value	W value	Unit
Density (ρ)	9675.21	7617.99	19171.91	Kg/m^3
Thermal exp. coef. (β)	1.23×10^{-4}	–	–	$1/K$
Specific heat (c_p)	188.49	607.46	147.98	$J/kg\ K$
Dynamic viscosity (μ)	$1.87 \exp(1400/\mathrm{T}) \times 10^{-4}$	–	–	$Pa\ s$
Thermal conduct. (κ)	$1.95 + 1.95\mathrm{T} \times 10^{-2}$	25.56	124.69	$W/m\ K$
Electrical conduct. (σ)	$(102.3 + 0.0426\mathrm{T}) \times 10^6$	–	–	S/m

Numerical model

The MHD governing equations are solved with High Resolution pseudo-transient scheme with Auto Timescale control for the virtual time step and all the simulations reach the convergence when the RMS residuals are below 10^{-5} and the monitored temperature and velocity local quantities become constant. The computational grid employs hexahedral structured meshes in almost all the domain, with higher resolution near the walls which ensure almost 7 nodes in the MHD layers, and an unstructured part around the pipes which guarantee the minimum use of unstructured elements. The thickness of the walls is discretized with 12 nodes. An example of the mesh with the Bundary Conditions (BCs) are shown in Figure 4. On the back face are applied the no-slip velocity ($\mathbf{u} = \mathbf{0}$) BC, the electric potential "ground" BC ($\phi = 0$) and the temperature adiabatic one ($\partial T/\partial n = 0$) was applied to the front walls. Whereas the top and bottom faces are modelled with a translational periodic BC which model the continuity of the COB with an imposed average velocity $\mathrm{v}_0 = 1.825\ \mathrm{mm/s}$. On the front face a surface heat load of $\ddot{q} = 0.320\ MW/m^2$ model the plasma impact on the tungsten armor [5], while the lateral Eurofer walls are modelled with the symmetry BC which model the left and right cells. The interface between different materials is modelled in CFX ensuring the continuity of the field and the respective fluxes. Regarding the internal patch of the pipes and FW channels, the water heat sink has been modelled using a convective BC with a fix heat transfer coefficient and wall temperature of, respectively, $39287\ W/m^2K$ and $T_{H2O} = 584.65\ K$ for the pipes [6], and $42110\ W/m^2K$ and $587.66K$ for the channels [1]. The MHD model of ANSYS CFX has been validated in the past for pressure-driven and buoyancy-driven benchmarks in rectangular ducts [10].

Results and discussion

Figure 5, Figure 6 and Figure 7 show the velocity and temperature distribution in different planes for, respectively, the EC, AC and DC. Considering that the AC and ND cells form, respectively, an angle of $27.64°$ and $144.03°$ with respect of the horizontal plane [11], we have the following gravity vectors: $\boldsymbol{g_{EC}} = (0{,}0{,}-9.81) m/s^2$, $\boldsymbol{g_{AC}} = (8.69{,}0{,}-4.55) m/s^2$ and $\boldsymbol{g_{ND}} = (-7.94{,}0{,}-5.76) m/s^2$. Velocity and temperature are scaled, respectively, as $u' = u/u_0$ and $T' = (T - T_{ref})/\Delta T_{Gr}$. Regardless of the position, as expected since $Gr/(Ha^4) \sim 10^{-8}$ and $Pe \sim 10^{-2}$, the heat transfer is dampened due to the magnetic field action and the prevalent mechanism is conduction. The convective cells seen in previous hydrodynamic analyses [6] are suppressed and substituted by an almost slug flow in the back part of the cell and the generation of velocity jet near the pipes and the FW. As can be observed by comparing Figures 3, 4, 5, subfigure c, the fundamental component of velocity undergoes a curvature due to the gravity vector, which changes the inclination with which the flow strikes the pipes, specifically in the area immediately below them. For the EC, the streamlines are predominantly vertical, while for the AC they are inclined to the left and for the DC to the right. It is noteworthy the persistence at almost the same vertical position of the small vortex that forms between the tube and the FW, which are not totally suppressed by the field having their axis aligned with it. From the distribution of the

velocity components (see subfigures a and b), the global displacement of the flow is clearly visible. In the AC, it is predominantly concentrated in the anterior zone, where the negative u_z' zone is present only in proximity to the FW due to buoyancy forces, just as u_x' is predominantly positive, being emphasized by the concordant gravity. The opposite behavior occurs in the DC, where the distribution of u_z' returns to be more like the equatorial one but u_x' is negative in most of the domain, slowed down by the g_x component. These changes, however, do not substantially impact the temperature distribution, being due almost exclusively to thermal conduction as mentioned above. The independence of the temperature field as the position of the blanket varies due to the action of suppressing convection due to the magnetic field is clearly visible by comparing the maximum temperature values and the Nusselt numbers in proximity to the FW and the pipes collected in Table 3, which are all closer to the unity that is the pure conduction value. The last ones are defined as $Nu = UL/\kappa$, where the thermal transmittance is $U = \hat{q}/(\hat{T} - T_r)$. For the pipes L is the external diameter which define the surface where the average heat flux $\hat{q}$ is computed, $\hat{T}$ is the average temperature of the toroidal-poloidal plane passing at the half distance between the pipes and $T_r = T_{H2O}$. For the FW, L is the minimum half-length between the outer tube and the FW which also define the toroidal-poloidal plane in which T_r is computed, $\hat{T}$ is the average temperature of the FW/PbLi interface wall where is also computed the average heat flux $\hat{q}$.

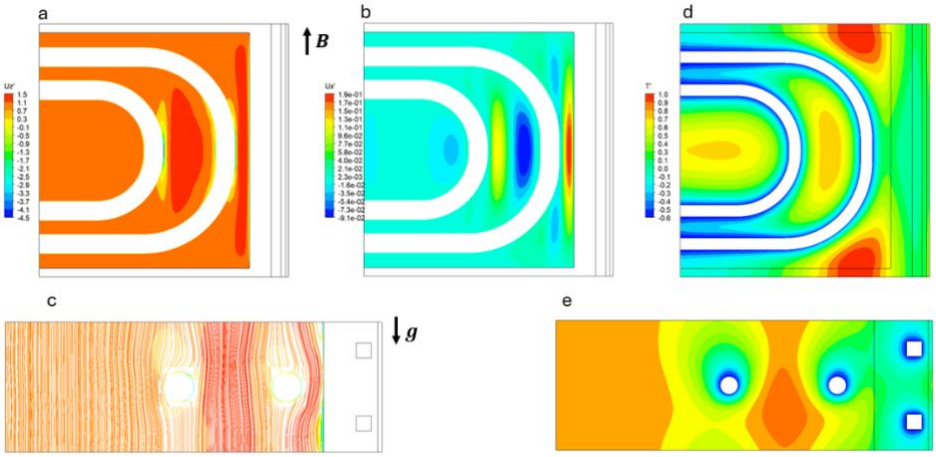


Figure 5. Scaled velocity distribution on the radial-toroidal plane for the Z (a) and X (b) component, velocity streamlines on the radial-poloidal plane (c) and scaled temperature distribution on the radial-toroidal plane (d) and on the radial-poloidal one (e) for the EC.

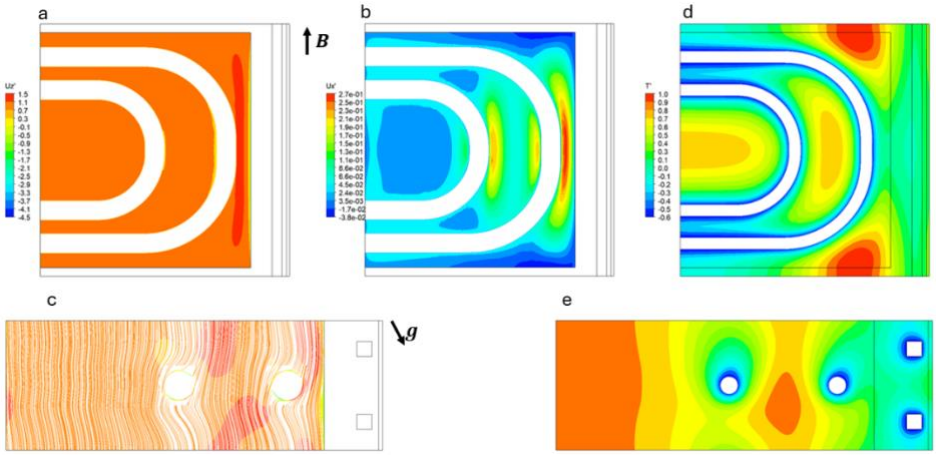


Figure 6. Scaled velocity distribution on the radial-toroidal plane for the Z (a) and X (b) component, velocity streamlines on the radial-poloidal plane (c) and scaled temperature distribution on the radial-toroidal plane (d) and on the radial-poloidal one (e) for the AC.

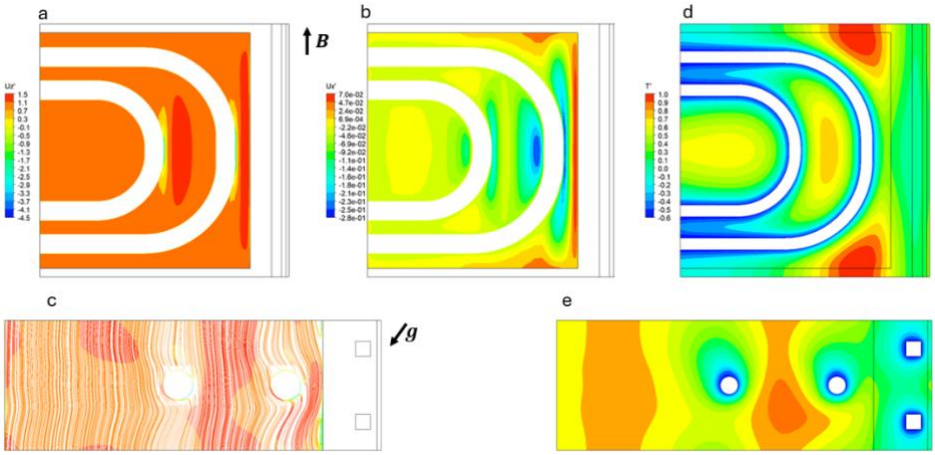


Figure 7. Scaled velocity distribution on the radial-toroidal plane for the Z (a) and X (b) component, velocity streamlines on the radial-poloidal plane (c) and scaled temperature distribution on the radial-toroidal plane (d) and on the radial-poloidal one (e) for the DC.

Table 3. Maximum temperature in the PbLi (T^M_{PbLi}), Eurofer (T^M_E) and Nusselt numbers related to the FW (Nu_{FW}), to the outer pipes ($Nu_{p,OUT}$, near the FW) and to the inner pipes ($Nu_{p,IN}$) for all the cases.

Case	T^M_{PbLi} [K]	T^M_E [K]	Nu_{FW}	$Nu_{p,OUT}$	$Nu_{p,IN}$
EC	904.9	903.4	1.367	1.952	1.752
AC	902.5	901.0	1.309	1.967	1.817
DC	905.9	904.3	1.379	1.963	1.725

Conclusions

This study has analyzed the magneto-convection regime in a variant of the WCLL BB using the Ansys CFX code. This variant anticipates a poloidal-upward flow of the breeder, cooled by nested DWT U-pipes inserted horizontally from the back plate. The study considered fusion-relevant magnetic field intensity and volumetric power source, resulting in a weak inertial magneto-convection flow regime. Simulations were conducted considering the equatorial, the apical, and the near-divertor elementary cell. Regardless of the gravity orientations, the magneto-convection regime was only slightly altered, consistently demonstrating the suppression of convection in favour of conduction.

Acknowledgments

This work has been carried out within the framework of the EUROfusion Consortium, funded by the European Union via the Euratom Research and Training Programme (Grant Agreement No 101052200 — EUROfusion). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be held responsible for them.

References

[1] P. Arena *et al.*, “Design and Integration of the EU-DEMO Water-Cooled Lead Lithium Breeding Blanket,” *Energies*, vol. 16, no. 4, Art. no. 4, Jan. 2023, doi: 10.3390/en16042069.

[2] O. Zikanov, I. Belyaev, Y. Listratov, P. Frick, N. Razuvanov, and V. Sviridov, “Mixed Convection in Pipe and Duct Flows With Strong Magnetic Fields,” *Applied Mechanics Reviews*, vol. 73, no. 1, p. 010801, Jan. 2021, doi: 10.1115/1.4049833.

[3] C. Mistrangelo *et al.*, “MHD R&D Activities for Liquid Metal Blankets,” *Energies*, vol. 14, no. 20, p. 6640, Oct. 2021, doi: 10.3390/en14206640.

[4] R. Mozzillo, A. Del Nevo, E. Martelli, and G. Di Gironimo, “Alternative design of DEMO Water Cooled Lithium Lead internal structure,” *Fusion Engineering and Design*, vol. 146, pp. 1056–1059, Sep. 2019, doi: 10.1016/j.fusengdes.2019.02.001.
[5] E. Martelli, G. Caruso, F. Giannetti, and A. Del Nevo, “Thermo-hydraulic analysis of EU DEMO WCLL breeding blanket,” *Fusion Engineering and Design*, vol. 130, pp. 48–55, May 2018, doi: 10.1016/j.fusengdes.2018.03.030.
[6] A. Tassone, G. Caruso, F. Giannetti, and A. Del Nevo, “MHD mixed convection flow in the WCLL: Heat transfer analysis and cooling system optimization,” *Fusion Engineering and Design*, vol. 146, pp. 809–813, Sep. 2019, doi: 10.1016/j.fusengdes.2019.01.087.
[7] U. Jauch, V. Karcher, B. Schulz, and G. Haase, “Thermophysical properties in the system Li-Pb,” Stuttgart, Germany, Technical Report SWB-057695733, 1986.
[8] K. Mergia and N. Boukos, “Structural, thermal, electrical and magnetic properties of Eurofer 97 steel,” *Journal of Nuclear Materials*, vol. 373, no. 1–3, pp. 1–8, Feb. 2008, doi: 10.1016/j.jnucmat.2007.03.267.
[9] G. Ermile, Schoofs, Frank, G. Aiello, and G. Pintsuk, “DEMO-DEF-1-CD1 - Materials Properties Handbook - Tungsten,” EUROfusion, Technical Report, 2020. Accessed: Jun. 24, 2025. [Online]. Available: https://idm.euro-fusion.org/default.aspx?uid=2P3SPL
[10] A. Tassone, G. Caruso, A. Del Nevo, and I. Di Piazza, “CFD simulation of the magnetohydrodynamic flow inside the WCLL breeding blanket module,” *Fusion Engineering and Design*, vol. 124, pp. 705–709, Nov. 2017, doi: 10.1016/j.fusengdes.2017.05.098.
[11] A. Tassone, “On the magnetic field distribution in the TBM set and a blanket based on the DEMO2017 baseline,” 2022.

MONTE CARLO SIMULATION SUPPORTING EXPERIMENTAL ACTIVITIES FOR MEDICAL TERBIUM-161 PRODUCTION AT ENEA TRIGA RC-1

Lucrezia Spagnuolo[1], Luigi Lepore[1*], Marco Capogni[2]

[1]*ENEA, NUC-IRAD-CRGR, Nuclear Material Characterization Laboratory and Nuclear Waste Management*

ENEA Casaccia Research Center Via Anguillarese, 301, 00123 Roma, RM, Italy[15]

[2]*ENEA, NUC-INMRI, National Institute of Ionizing Radiation Metrology,*

ENEA Casaccia Research Centre, Via Anguillarese 301,00123 Rome, Italy

ABSTRACT. TRIGA (Training, Research, and Isotope production General Atomics) research reactors are widely used worldwide for neutron activation analysis and several neutron-related experiments and applications. Neutron fluence rates in TRIGA reactors can also be used for radionuclide production, e.g. to generate radioactive sources or medical isotope supply. The ENEA TRIGA RC-1 is involved in the EU-funded SECURE Project, aiming at studying the feasibility of Italian local production of radionuclides for medical applications. In particular, the ENEA team has performed some experimental activities to produce terbium-161 (^{161}Tb), a promising isotope for targeted radiation therapy with potentially improved efficacy over lutetium-177 (^{177}Lu) currently used in cancer therapy, through neutron activation of gadolinium target highly enriched in gadolinium-160 (^{160}Gd) exploiting the reaction channel ^{160}Gd(n,γ)à ^{161}Gd(β^-)à^{161}Tb. Monte Carlo simulation for radiation transport has been used for reactor calculations to: i) validate the experimental results performed in 2024, ii) evaluate correction factors needed in case of irradiation of significative amount of ^{160}Gd raw material, iii) evaluate the compliance of such amount of material with the reactor prescriptive body, ensuring appropriate reactivity coefficient values during the irradiation.

Introduction

TRIGA reactors have proven to be valuable assets for the production of medical radioisotopes, despite offering lower neutron fluence rates compared to dedicated production reactors. These facilities can help address the growing demand for medical isotopes by serving as reliable, local sources of supply. Today, the contribution of all types of nuclear reactors to the radioisotope supply chain should not be overlooked.

In this framework, the ENEA TRIGA RC-1 reactor has joined the EU-funded SECURE Project (Strengthening the European Chain of sUpply for next generation medical RadionuclidEs). The project aims to significantly contribute to the sustainability of medical radioisotope production and supply, ensuring their continued and stable use in diagnostics and therapy within Europe. It focuses on promising developments in novel radionuclides, irradiation target design, and innovative production methods for both existing and next-generation medical isotopes.

One of the key radionuclides under investigation is ^{161}Tb, a particularly promising isotope for cancer treatment. It shares similar decay characteristics and chemical behavior with the clinically established ^{177}Lu, but with potentially enhanced therapeutic effects due to its higher emission of conversion and

*Corresponding author. E-mail: luigi.lepore@enea.it

Auger electrons. These emissions may be more effective in targeting micro-metastases and even individual cancer cells. [1,2].

The ENEA Casaccia Research Centre is contributing to the project by focusing on the production of ^{161}Tb, which is generated through neutron activation of a highly enriched ^{160}Gd target. The relevant nuclear reaction is ^{160}Gd(n,γ)à^{161}Gd, followed by beta decay to ^{161}Tb — often abbreviated as ^{160}Gd(n,γ)à^{161}Tb — and takes place at the TRIGA RC-1 reactor.

Once the irradiation cycle is complete, the target is extracted and chemically processed to achieve two main outcomes: i) production of ^{161}Tb in salt form, which serves as a precursor for radiopharmaceutical development, and ii) recovery of the remaining ^{160}Gd to fabricate new targets.

This recovery process enables the reuse of the costly enriched gadolinium material in subsequent production cycles, thereby maximizing resource efficiency and reducing overall costs.

Methodology

Reactor calculation software plays a vital role in predicting neutron fluence rate distributions and radiation fields within nuclear reactors, also performing criticality evaluations. These simulations are crucial for assessing radionuclide production through neutron activation — a key process in nuclear medicine, industrial applications, scientific research — and in the management and safety of reactor operations.

Most reactor calculation codes are based on **Monte Carlo methods**, which model neutron behavior as a series of random interactions with matter, reflecting their uncharged and stochastic nature. When the physical characteristics of the reactor and its environment are accurately replicated in the simulation, these tools can yield highly precise and dependable results.

The primary function of such software is to provide neutron fluence rate **distributions**, which serve as essential input data for other computational tools. These may include codes for radionuclide activation and decay analysis, transmutation studies, or radiation shielding design.

Moreover, criticality assessments and reactivity calculations are essential for planning and coordinating experiments involving the reactor core. These evaluations help ensure that any experimental setup does not introduce excessive perturbations to the neutron multiplication behavior within the nuclear fuel, thus maintaining safe and controlled reactor conditions.

Monte Carlo simulation for radiation transport

At present, Monte Carlo simulation represents the most effective method for modeling radiation transport in complex and large-scale geometries—characterized by numerous volumes ("cells"), multiple radioactive sources, and diverse materials, ranging from a few centimeters to several hundred meters.

This technique involves simulating a large number of particle "histories" to capture all possible paths the radiation may follow within a virtual environment. The simulation process is governed by random sampling from the probability density functions that describe the underlying physical interactions.

When all key elements influencing radiation transport are accurately represented in the model, highly detailed geometries can be constructed, leading to precise and reliable results. However, achieving statistically meaningful outcomes also requires sufficient computational resources to support the simulation of a large number of particle histories.

For the purpose of this study, simulations were performed using MCNPX (Monte Carlo N-Particle eXtended), version 2.5.0 [3].

Simulations have exploited the usage of multi-processing on CRESCO High Parallel Computing Resources extensively. The computing resources and the related technical support used for this work have been provided by CRESCO/ENEAGRID High Performance Computing infrastructure and its staff [4].

CRESCO/ENEAGRID High Performance Computing infrastructure is funded by ENEA, the Italian National Agency for New Technologies, Energy and Sustainable Economic Development and by Italian and European research programmes, see http://www.cresco.enea.it/english for information. In particular, the whole Monte Carlo work has requested 300,000 hours*core.

Irradiation at TRIGA RC-1 inside Central Thimble

The prior determination of the neutron fluence rate is essential for initiating activation calculations using specialized codes, such as ISPACT-II, which enables automated evaluations of all possible activation channels for a given raw material to be irradiated.
From a technical standpoint, the reaction channel $^{160}Gd(n,\gamma)^{161}Tb$ does not exhibit a particularly favorable microscopic cross section. As a result, irradiating a given quantity of ^{160}Gd in a high-flux neutron source would be significantly more efficient compared to lower-flux facilities, such as a TRIGA reactor. Furthermore, the irradiation time should be selected to correspond to at least 1 to 2 mean-life of ^{161}Tb (i.e., approximately 10 to 20 days) to ensure sufficient production.
Under ideal conditions, the saturation activity — the maximum achievable activity for a fixed amount of target material — would be reached after roughly 30 days of continuous irradiation. However, achieving such extended irradiation durations is not feasible at the ENEA TRIGA RC-1 reactor for two main reasons:
a) the reactor core may lack the necessary excess reactivity to sustain such a prolonged irradiation cycle, and b) there are operational constraints, including a maximum irradiation time of 6 hours per day, which results in suboptimal utilization of the enriched gadolinium target material.
The expected irradiation schedule, considering the previous mentioned limits, involves the placement of the sample at the Central Thimble position for 12 days, with 6-7 hours of irradiation per day, excluding Saturdays and Sundays. In total, the sample could be irradiated for approximately 72-84 hours.
On September 2024 a suitable 8.7 mg Gd_2O_3 (98.2% ^{160}Gd enriched Gd) sample was irradiated at the TRIGA RC-1 Central Thimble bottom position (core centre *i.e.* the maximum irradiation position available), 1 MW thermal power with the maximum magnitude neutron spectrum available at the ENEA TRIGA RC-1 reactor i.e. inside the Central Thimble with a neutron fluence rate of $5 \cdot 10^{13}\,cm^{-2} \cdot s^{-1}$[6].
This irradiation condition allows to reach a final activity of about 110 MBq of ^{161}Tb in the irradiated sample, so that ~ 15 $GBq_{Tb\text{-}161}/g_{Gd\text{-}160}$.

The MCNP TRIGA RC-1 calculation neutron Self Shielding Correction Factors at the Central Thimble

The ENEA TRIGA RC-1 MCNPX model is based on detailed material composition retrieved from plant documentation and schemes [5].

The MCNP input available at the ENEA Research Nuclear Reactor Laboratory can be used for reactor calculations to determine neutron self-shielding corrections factors essential when irradiation is performed with a large amount of raw material in the sample. Simulation of neutron self-shielding factors is also useful to understand the behaviour of the sample in the reactor configuration while sample mass is increased over and over. Increasing the target mass at an inherently low-magnitude reactor facility is a key point to increasing the final activity of the expected radionuclide to be produced. Unfortunately, such radionuclide production is not proportional to the mass of the target because of such neutron self-shielding. A compromise in target configuration should be found to maximize the ratio of the radionuclide activity production vs. the initial target mass. In evaluating neutron activation at the ENEA TRIGA RC-1 reactor with the ASTM E262 standard, these correction factors were calculated using the MCNP code: this calculation tool allows us to obtain more realistic results compared to the ones reached with analytical methods that may neglect neutron self-shielding.

To obtain the correction factors, the reaction rates were evaluated in two different conditions: 1) with the sample defined at its actual density, and 2) with the sample defined at a density equal to one-

thousandth of its actual density. The energy bins were constructed up to 0.55 eV (for the G_{th} evaluation) and up to 20 MeV (for the G_{res} evaluation). Finally:

$$G_{th} = \frac{R_\rho|_{1e-11\ eV}^{0.55\ eV}}{R_{\frac{\rho}{1000}}|_{1e-11\ eV}^{0.55\ eV}}$$

$$G_{res} = \frac{R_\rho|_{0.55\ eV}^{20\ MeV}}{R_{\frac{\rho}{1000}}|_{0.55\ eV}^{20\ MeV}}$$

The sample was modelled as a cylinder with a height of 14.4 cm (almost the whole capsule height Fig.1) and three different radius configurations (0.1 cm, 0.546 cm, 1.092 cm) to assess the trends in self-shielding neutron absorption factors and the distribution of the neutron fluence rate inside the sample, from the centre to the maximum radius, while the mass is increasing.

As expected, the value of the G_{th} decreases, while G_{res} is less decreasing. This behaviour is attributed to the fact that neutron self-shielding is more sensitive in the thermal region compared to the epithermal region, due to the significantly higher neutron cross-section.

Consequently, the inner shell of the sample becomes underutilized due to the reduction in neutron fluence rate, which in turn affects the values of the spectrum-averaged microscopic cross-section also.

As expected, in the case of massive samples, a compromise in raw material target configuration should be found to maximize the ratio of the radionuclide activity production vs. the initial target mass. Monte Carlo simulation could be particularly useful to address this topic.

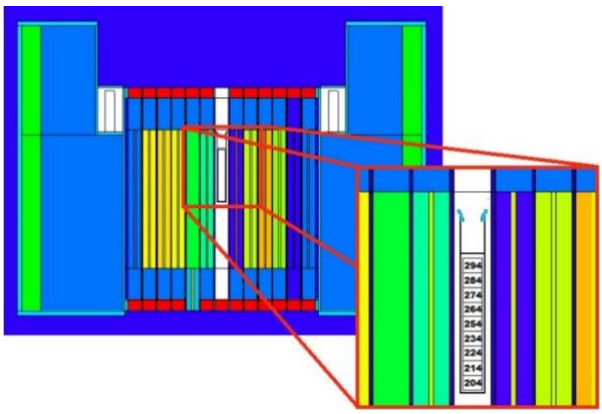


Fig.1 ENEA TRIGA RC-1 in MCNP: section of the reactor core and particular of the Central Thimble irradiation position.

TRIGA RC-1 core criticality and reactivity calculations

Increasing the mass of the sample to be irradiated, when introduced in a geometrical position within the reactor core, means the sample to interact with the neutron multiplication features of the reactor core, sometimes increasing neutrons population, sometimes depleting neutron population.

Monte Carlo simulation has been used to reactor effective k-factor calculations, allowing to demonstrate that even for massive samples (125 g of Gd_2O_3, the maximum sample dimension studied

within SECURE at TRIGA RC-1) the reactivity insertion due to the sample is less than 500 p.c.m. (~0.71 $) as requested by the prescriptive body of TRIGA RC-1.

Results and Conclusions

In recent years, there has been a growing interest in emerging short-lived radionuclides with potential applications in cancer therapy and medical imaging. This increased attention has led the scientific community to investigate the physical and chemical properties of these radionuclides, as well as to explore various production methods, including conventional pathways.
Today, diversifying production methodologies is essential to ensure the continuous and reliable supply of radiopharmaceuticals to healthcare facilities, thereby minimizing the risk of disruptions that could compromise timely and effective patient treatments.
While radionuclide production has traditionally relied on nuclear power reactors, research reactors and accelerator-based facilities are increasingly recognized for their potential to supplement the supply chain—particularly in response to temporary or permanent shutdowns of major production reactors.
In this context, as part of the EU-funded SECURE Project (Strengthening the European Chain of sUpply for next generation medical RadionuclidEs), the production potential of 161Tb at the ENEA TRIGA RC-1 research reactor has been investigated.
From the calculations made, it emerges that it is possible to obtain an activity of 15 GBq of ^{161}Tb per gram of ^{160}Gd, demonstrated with an irradiation of 8.7 mg of Gd_2O_3, enriched at 98.2% in ^{160}Gd, in September 2024, with 77 hours of irradiation time inside the Central Thimble, the maximum neutron fluence rate irradiation position [7].
Since ^{161}Tb is a radionuclide still in experimental phase for medical applications, in order to quantify the activity necessary for a therapy, we can refer to therapies carried out with ^{177}Lu (a radionuclide with similar chemical and physical characteristics). A typical therapy with ^{177}Lu consists in 6 injections of 7 GBq each (at 6-week distance).

TRIGA RC-1 reactor can only fulfill local requests for this radiopharmaceutical: if 10 g of ^{160}Gd were irradiated, it would theoretically be possible to obtain 150 GBq (EOI), i.e. 70 GBq after a week of chemical processing, sufficient to treat 10 patients. ENEA could be also potentially engage in production aimed at experimental studies for early clinical trials.

As expected, in the case of massive samples, a compromise in raw material target configuration should be found. For this reason, the possibility of irradiating larger quantities of gadolinium into the reactor is being considered and studied thorough Monte Carlo simulation.

Acknowledgement

The publication was created within the project SECURE funded by the European Union under grant agreement No. 101061230.

References

[1] N. Gracheva, C. Müller, Z. Talip e e. al., «Production and characterization of no-carrieradded ^{161}Tb as an alternative to theclinically-applied ^{177}Lu for radionuclidetherapy» EJNMMI Radiopharmacy and Chemistry, vol. 4, n. 12, 2019.

[2] S. McNeil, M. Van de Voord, C. Zhang e e. al., «A simple and automated method for 161Tb purification and ICP-MS analysis of 161Tb» EJNMMI Radiopharmacy and Chemistry, vol. 7, n. 31, 2022.

[3] J. F. Briesmeister, Ed., “MCNP: A General Monte Carlo N-Particle Transport Code,” 2000.

[4] F. Iannone et al., “CRESCO ENEA HPC clusters: a working example of a multifabric GPFS Spectrum cale layout,” in 2019 International Conference on High Performance Computing Simulation (HPCS), 2019, pp. 1051–1052.

[5] L. Falconi, N. Burgio, M. Palomba, E. Santoro e M. Carta, «A MCNPX TRIGA RC1 Experimental channels model for the design of a new neutronic diffraction facility» in Proceedings of the Meeting Research Reactors Fuel Management, Berlin, 2016.

[6] D. Chiesa, M. Carta, V. Fabrizio, L. Falconi, A. Grossi e M. Nastasi, «Characterization of TRIGA RC-1 neutron irradiation facilities for radiation damage testing» The European Physical Journal Plus, vol. 135, n. 349, 2020.

[7] L. Lepore, L. Spagnuolo, L. Cozzella, T. Guarcini, F. Limosani, S. Placidi, A. Pagano, L. Falconi, V. Fabrizio, D. Formenton, M. Lammardo, A. Roberti, M. Capogni «ENEA TRIGA RC-1 REACTOR TO MEDICAL RADIONUCLIDES PRODUCTION WITHIN THE EU SECURE PROJECT», RRFM-2025, 2025

UNRAVELLING MAGNETIC EXCHANGE EVOLUTION IN HETERONUCLEAR VANADYL-COPPER PORPHYRIN QUBITS: A PERIODIC DFT EXPLORATION FROM GAS TO BULK PHASES

Federico Totti[1*] and Prem Prakash Sahu[1]

[1]*Department of Chemistry "U. Schiff", University of Florence & UdR INSTM Firenze, 50019, Sesto Fiorentino, Italy*

ABSTRACT. Magnetic exchange interactions in molecular spin qubits are highly sensitive to their environment, with significant modulation arising from condensed-phase packing. This work investigates the heterometallic complex [VO-Cu(DPP)$_2$] by analyzing three computational models: an isolated molecule relaxed in vacuum, a fully periodic condensed bulk phase based on crystallographic data, and a dimeric fragment preserving a key intermolecular VO---phenyl contact. Using periodic density functional theory augmented by self-consistent Hubbard U corrections, we observe pronounced exchange enhancement in the bulk phase, primarily driven by packing and specific intermolecular contacts. The dimeric fragment reproduces this behaviour, confirming the structural origin of the exchange increase. Highly converged broken symmetry calculations ensured precise determination of exchange couplings.

Introduction

Molecular spin qubits based on transition metal complexes have gained significant attention for quantum information processing due to their chemical tunability and inherently long spin coherence times [1, 2]. Porphyrin-based bimetallic complexes offer well-defined localized spins embedded in extended **π**-conjugation systems [3, 4]. The heterobimetallic [VO-Cu(DPP)$_2$] complex, featuring VO^{2+} and Cu^{2+} centres, exemplifies a promising platform for two-qubit gate implementation [5].

In frozen solution, experimental electron paramagnetic resonance (EPR) data indicate minimal magnetic exchange coupling between these spin centres, essential for preserving quantum coherence. However, in the condensed bulk phase, molecular packing and intermolecular contacts can introduce additional exchange pathways absent in isolated molecules. For example, the vanadyl oxygen atom often engages in a close (~3 Å) contact with the phenyl ring of a neighbouring molecule, potentially mediating spin interactions through space.

Standard gas-phase calculations cannot capture such packing effects, necessitating periodic density functional theory (DFT) approaches with electron localization corrections. The self-consistent Hubbard U method, combined with tightly converged broken symmetry calculations, enables accurate resolution of subtle exchange phenomena in these large periodic systems.

This study employs a hierarchical modelling approach: isolated molecule, fully periodic bulk crystal, and a dimeric fragment retaining key contacts, to dissect intrinsic and environment-driven contributions to magnetic exchange in [VO-Cu(DPP)$_2$]. The ENEA CRESCO high-performance computing platform provided the computational resources critical for this investigation.

Structural Models

To elucidate the influence of environment on magnetic exchange, three structural representations were utilized (see Figure 1 below):

a) Condensed Bulk Phase: Based on experimental crystallographic data with lattice parameters a=b=19.276 A˚; c=13.084 A˚, the bulk model contains four molecular pairs (452 atoms). Initial disorder and solvent molecules were removed, and the structure was fully relaxed under periodic boundary conditions while preserving lattice parameters, faithfully representing crystal packing.

b) Isolated Molecule: A single [VO-Cu(DPP)$_2$] molecule was placed in a 24 Å cubic vacuum cell to avoid artificial periodic interactions and fully optimized. This isolated model serves as the baseline for intrinsic magnetic properties.

c) Dimer-Pair Cluster: Extracted from the relaxed bulk, the dimer maintains the critical ~3 Å VO---phenyl intermolecular contact. Relaxation in a 28 Å cubic vacuum cell ensured elimination of spurious periodic image effects while preserving packing-induced interaction motifs.

This multi-scale framework isolates molecular and packing effects, enabling detailed analysis of their roles in magnetic exchange modulation.

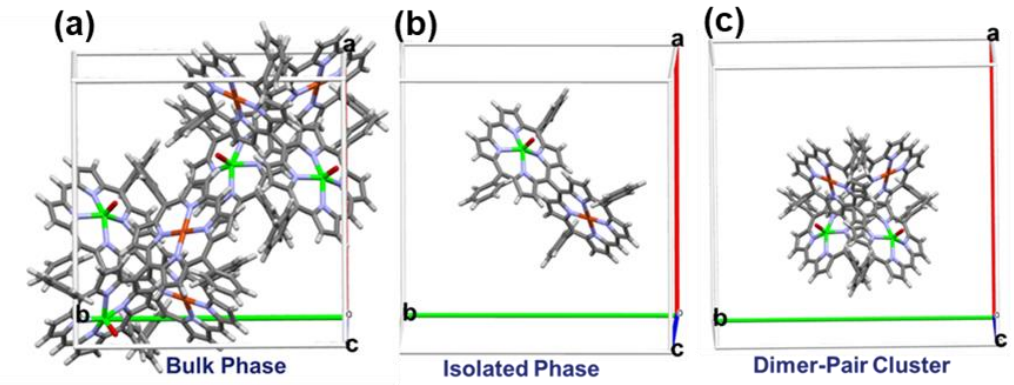


Fig.1: Structural models used to analyze the influence of molecular surroundings on magnetic exchange: (a) Bulk Phase illustrating molecular packing within the crystallographic unit cell; (b) Isolated Molecule oriented with respect to the crystallographic cell axes; (c) Dimer-Pair Cluster preserving the relative molecular arrangement from the bulk.

Computational Methodology

Calculations were performed using Quantum ESPRESSO v7.0.0 [6, 7] with the revPBE functional [8] and Grimme D3BJ dispersion corrections [9], essential for accurately modelling intermolecular forces in the bulk. Ultrasoft pseudopotentials from SSSP Efficiency v1.1 [10, 11] and PSLibrary v0.3.1 [12] were employed for all elements.

1. **Geometry Optimization:** The bulk structure was relaxed via the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [13], with force and energy convergence thresholds set at 10^{-4} Ry/Bohr and 10^{-6} Ry, respectively. Relaxations preserved lattice parameters closely matching

experimental values. The isolated molecule and dimer fragment underwent equivalent relaxations in their vacuum cells.

2. **Hubbard U Parameterization**: Recognizing that standard DFT delocalizes d electrons, compromising magnetic exchange accuracy, self-consistent Hubbard U parameters for V 3d, Cu 3d, and vanadyl O 2p orbitals were computed via linear response in the HP module [14]. These values corrected on-site Coulomb interactions, enhancing spin localization and exchange precision. A subsequent relaxation incorporating U corrections ensured structural and electronic consistency.

3. **Magnetic Exchange Computation:** Exchange couplings were extracted using the broken symmetry method [15-17] by comparing total energies of high-spin and antiferromagnetic states, within the full projected formula. SCF convergence was tightened to 10^{-9} Ry to resolve the subtle energy differences intrinsic to weak exchange. This protocol was consistently applied across all three models.

Results and Discussion

Isolated Molecule

The optimized isolated molecule maintained its structural integrity with minimal distortion. Broken symmetry calculations yielded magnetic exchange values in excellent agreement with available EPR data and previous theoretical studies, confirming that VO-Cu spin coupling in isolation is negligible, consistent with long coherence times observed experimentally.

Condensed Bulk Phase

The relaxed bulk crystal structure faithfully reproduced experimental lattice constants within typical DFT accuracy margins. Incorporation of Hubbard U corrections shifted the relative energies of V and Cu d orbitals, enabling more accurate spin localization. This resulted in a pronounced enhancement of magnetic exchange coupling mediated by crystal packing effects. Notably, the total exchange interaction arises not only from the four intramolecular VO-Cu pairs but also from at least two significant intermolecular VO···phenyl contacts, highlighting the crucial role of packing-induced through-space interactions in the condensed phase.

Dimer-Pair Cluster

The dimer model, containing one intramolecular and one key intermolecular exchange pathway, successfully reproduced the qualitative magnitude of the bulk magnetic exchange coupling. Relaxation in an expanded vacuum cell prevented spurious periodic interactions, validating that the VO···phenyl contact acts as the principal intermolecular exchange conduit. This fragment thus serves as an efficient and physically meaningful reduced model for studying packing effects on spin interactions.

Conclusions

This comprehensive multiscale computational study elucidates the mechanisms by which condensed-phase packing and specific intermolecular contacts modulate magnetic exchange interactions in the [VO-Cu(DPP)$_2$] molecular qubit system. Our findings demonstrate that the planewaves-based p-DFT calculations augmented with self-consistent Hubbard U corrections and stringent broken symmetry convergence are essential for accurately capturing the weak but non-negligible exchange couplings arising from crystal packing. The intermediate dimer fragment model, preserving critical packing motifs, provides a computationally accessible proxy for investigating environment-driven magnetic phenomena relevant to quantum device design. These insights pave the way for rational engineering of molecular q4ubits with tailored spin interactions. The ENEA CRESCO HPC facility was indispensable in enabling this rigorous investigation, offering the computational power required for such large-scale periodic simulations.

Acknowledgments

This work was supported by the National Recovery and Resilience Plan, Mission 4, Component 2 - Investment 1.4, through the National Center for HPC, Big Data, and Quantum Computing, funded by the European Union under NextGenerationEU (CUP: B83C22002830001). Computational resources and technical support were provided by the CRESCO ENEAGRID High Performance Computing infrastructure and its dedicated staff, whose assistance is gratefully acknowledged. We also thank the ERC-Synergy project CASTLE (proj. n. 101071533) and the support of MUR through Progetto Dipartimenti di Eccellenza 2018–2022 (CUP B96C1700020008) and 2023–2027 (CUP B97G22000740001 – DICUS 2.0).

References

[1] M. Atzori and R. Sessoli, The second quantum revolution: role and challenges of molecular chemistry, *J. Am. Chem. Soc.*, **141**, pp. 11339-11352, (2019).

[2] M.R. Wasielewski, M.D. Forbes, N.L. Frank, K. Kowalski, G.D. Scholes, J. Yuen-Zhou, M.A. Baldo, D.E. Freedman, R.H. Goldsmith and T. Goodson III, Exploiting chemistry and molecular systems for quantum information science, *Nat. Rev. Chem.*, **4**, pp. 490-504, (2020).

[3] D. Ranieri, F. Santanni, A. Privitera, A. Albino, E. Salvadori, M. Chiesa, F. Totti, L. Sorace and R. Sessoli, An exchange coupled meso-meso linked vanadyl porphyrin dimer for quantum information processing, *Chem. Sci.*, **14**, pp. 61-69, (2023).

[4] F. Santanni and A. Privitera, Metalloporphyrins as building blocks for quantum information science, *Adv. Opt. Mater.*, **12**, pp. 2303036, (2024).

[5] D. Ranieri, A. Privitera, F. Santanni, K. Urbanska, G. J. Strachan, B. Twamley, E. Salvadori, Y.-K. Liao, M. Chiesa, M. O. Senge, F. Totti, L. Sorace and R. Sessoli, *Angew. Chem. Int. Ed.*, **62**, e202312936, (2023).

[6] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G.L. Chiarotti, M. Cococcioni and I. Dabo, QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials, *J. Phys.: Condens. Matter*, **21**, pp. 395502, (2009).

[7] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M.B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli and M. Cococcioni, Advanced capabilities for materials modelling with Quantum ESPRESSO, *J. Phys.: Condens. Matter*, **29**, pp. 465901, (2017).

[8] J.P. Perdew, K. Burke and M. Ernzerhof, Generalized gradient approximation made simple, *Phys. Rev. Lett.*, **77**, pp. 3865, (1996).

[9] S. Grimme, J. Antony, S. Ehrlich and H. Krieg, A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu, *J. Chem. Phys.*, **132**, pp. 154104, (2010).

[10] G. Prandini, A. Marrazzo, I.E. Castelli, N. Mounet and N. Marzari, Precision and efficiency in solid-state pseudopotential calculations, *npj Comput. Mater.*, **4**, pp. 72, (2018).

[11] The SSSP library of the Materials Cloud:
https://www.materialscloud.org/discover/sssp/table/efficiency.
[12] A. Dal Corso, Pseudopotentials periodic table: From H to Pu, *Comput. Mater. Sci.*, **95**, pp. 337-350, (2014).
[13] D.F. Shanno, Conjugate gradient methods with inexact searches, *Math. Oper. Res.*, **3**, pp. 244-256, (1978).
[14] I. Timrov, N. Marzari and M. Cococcioni, HP- A code for the calculation of Hubbard parameters using density-functional perturbation theory, *Comput. Phys. Commun.*, **279**, pp. 108455, (2022).
[15] L. Noodleman and J.G. Norman Jr, The X α valence bond theory of weak electronic coupling. Application to the low-lying states of $Mo_2Cl_8^{4-}$, *J. Chem. Phys.*, **70**, pp. 4903-4906, (1979).
[16] A. Bencini and F. Totti, A few comments on the application of density functional theory to the calculation of the magnetic structure of oligo-nuclear transition metal clusters, *J. Chem. Theory Comput.*, **5**, pp. 144-154, (2009).
[17] C. Adamo, V. Barone, A. Bencini, F. Totti and I. Ciofini, On the calculation and modeling of magnetic exchange interactions in weakly bonded systems: the case of the ferromagnetic copper (II) μ2-azido bridged complexes, *Inorg. Chem.*, **38**, pp. 1996-2004, (1999).

DESIGNING PEROVSKITE-INSPIRED MATERIALS: STRUCTURAL AND TRANSPORT MODULATION VIA HALIDE TUNING

Michael Zambrano-Angulo[1], Adriana Pecoraro[1*], Michele Pavone[2] and Ana Belén Muñoz-García[1]

[1]*Department of Physics "E. Pancini", University of Naples Federico II, 80126 Napoli, Italy*

[2] *Department of Chemical Sciences, University of Naples Federico II, 80126 Napoli, Italy*

ABSTRACT. This study employs Density Functional Theory (DFT) to investigate the structural and electronic properties of $Cs_3Bi_2I_9$ and $Cs_3Bi_2Br_9$, lead-free perovskite-inspired materials with potential for photocatalytic applications. By studying iodine/bromine mixing, we predict a phase transition near 42% Br content, with distinct stable phases below 39% and above 50% Br. The bandgap increases with Br content, except during the transition range, and structural analysis shows minimized octahedral distortion at 50% Br. Electron mobility exceeds that of holes, which become increasingly localized with more Br. These findings highlight the pivotal role of halide composition in tuning material performance.

Introduction

Bismuth-based perovskite-inspired materials have emerged as promising candidates in this search for safer and more stable materials. Bismuth, with its trivalent oxidation state and similar electronic configuration to lead, forms defect-tolerant perovskite derivatives such as $Cs_3Bi_2X_9$ (X = Br, I), which feature layered or zero-dimensional structures with favorable optoelectronic characteristics[1]. These compounds exhibit improved stability and lower toxicity, making them viable for photocatalytic applications such as dye degradation, hydrogen evolution, and CO_2 photoreduction. Notably, $Cs_3Bi_2I_9$ has been widely studied for CO_2 photoreduction, producing CO and CH_4 as primary products. Despite its promising activity, the low dimensionality of the $Cs_3Bi_2I_9$ structure limits charge transport and light-harvesting efficiency, hindering its overall performance[2]. Conversely, $Cs_3Bi_2Br_9$ offers superior stability but suffers from a large band gap and high exciton binding energy, which restrict its absorption in the visible spectrum[3].

A promising approach to overcoming these limitations involves halide mixing (e.g., I/Br), which can tune the electronic structure, reduce the band gap, and enhance charge transport by increasing the structural dimensionality of the material. However, while theoretical studies have extensively characterized the pure halide compounds or with specific amounts of bromine content[4], there remains a significant knowledge gap regarding the structural and electronic evolution of mixed-halide systems such as $Cs_3Bi_2I_{9-X}Br_X$ across the entire composition range.

This study proposes a systematic theoretical investigation of $Cs_3Bi_2I_9$, $Cs_3Bi_2Br_9$, and their intermediate compositions to elucidate the effect of bromine substitution on structural integrity, electronic properties, and charge transport behavior. Special attention will be given to phase transitions and distortion mechanisms across the compositional range. The insights gained will contribute to the rational design of lead-free perovskite-inspired materials for high-performance energy conversion applications, including photovoltaics and solar-driven CO_2 reduction.

Results and Discussion

We began with unit cells of $Cs_3Bi_2I_9$ and $Cs_3Bi_2Br_9$ from literature: the hexagonal phase ($P6_3/mmc$) for and the trigonal phase (P-3m1) respectively, both stable at room temperature. **Figure 1a-b** compares lattice parameters optimized via PBE-TS with experimental values, showing good agreement and validating our computational approach. To study the phase transition upon bromine doping and octahedral distortions, 56-atom supercells were constructed for both phases, simulating I/Br doping from 0% to 100%. The energy difference between phases at matching I/Br ratios (**Figure 1c**) reveals that up to ~39% Br, the hexagonal phase remains stable. A significant destabilization occurs near 50% Br, indicating a phase transition, consistent with experimental reports[5]. Intermediate compositions (42%, 44%, 47% Br) pinpoint the onset around 42% Br. Cell volume decreases with increasing Br content (**Figure 1d**), reflecting smaller Br ionic radius and shorter Bi–halogen bond distances, which contract from 3.07 Å (pure iodine) to 2.87 Å (pure bromine) (**Figure 1e**).

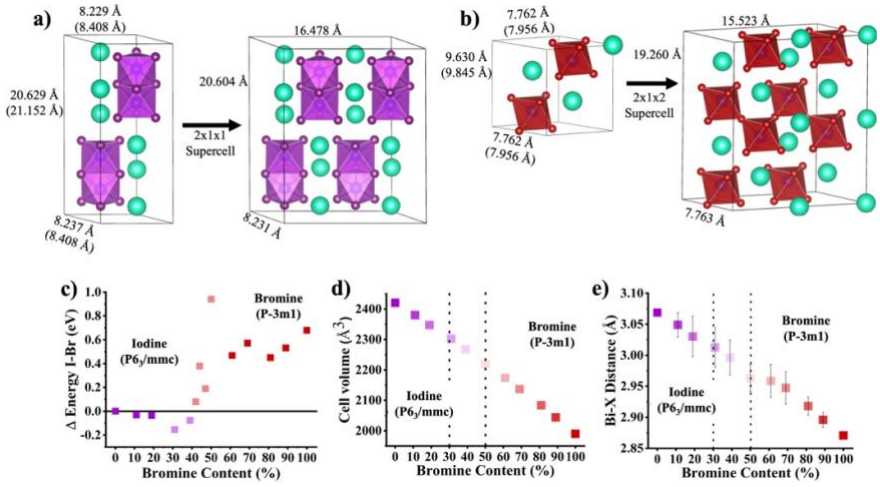


Figure 1: Lattice constants calculated at PBE-TS level of theory for: **a)** $Cs_3Bi_2I_9$ and **b)** $Cs_3Bi_2Br_9$. The corresponding experimental values are reported in parenthesis. **c)** Total energy differences between $Cs_3Bi_2I_9$ in the $P6_3/mmc$ phase at various Br doping percentages and the corresponding $Cs_3Bi_2Br_9$ P-3m1 phase with the same I/Br ratio. Energies are calculated at PBE on top of PBE-TS optimized structures, **d)** Cell volume and **e)** Distance Bi-halogen with error bars calculated through standard deviation.

Focusing on $Cs_3Bi_2I_9$ ($P6_3/mmc$) doped with Br, we analyzed structural parameters driving the phase transition: Bi–X bond distances, cell volume, and octahedral distortions measured by Octadist software[6]—specifically stretching (ζ), angular (Σ), and torsional (Θ) distortions. These quantify bond length variability, cis-angle deviation from 90°, and trigonal distortion from octahedral geometry, respectively. As shown in **Figure 2a-c**, all distortion parameters increase with Br doping up to ~70%, reaching maximum octahedral distortion. A distinct inflection occurs around 42–44% Br, coinciding with the phase transition identified energetically. Afterward, distortions decline toward the bromide-rich phase. Minimum distortion at ~50% Br suggests phase transition ending, highlighting compositional control over structural dynamics that underpin phase stability and physical properties.

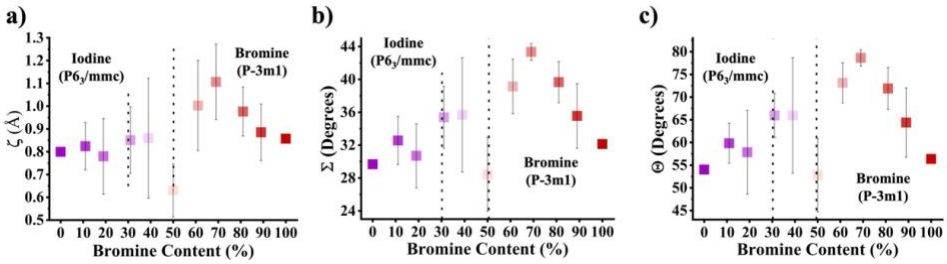


Figure 2. Variation of the structural parameters evaluated on the $Cs_3Bi_2I_9$ in the $P6_3/mmc$ phase at different Br contents with error bars calculated through standard deviation for the **a)** stretching (z), **b)** the angular (S) and **c)** the torsional (q) distortions.

Electronic properties were characterized through band structures and density of states (DOS), exploring multiple functionals and basis sets to optimize accuracy. Inclusion of spin-orbit coupling (SOC) was essential due to heavy Bi atoms. **Table 1** summarizes bandgap predictions: PBE underestimates gaps, while larger basis sets and HSE06 overestimate slightly. The best agreement with experiment comes from HSE06 with intermediate basis and SOC.

Table 1. Band gap values calculated with different basis sets/functional choices for the pure phases of the single halogen compounds. In all calculations a 12x12x6 and a 12x12x12 k-point grids are used for I- and Br-based compounds, respectively. Experimental values are reported in the last row for comparison.

Bandgaps (eV)		
Basis set/Functional	**$Cs_3Bi_2I_9$**	**$Cs_3Bi_2Br_9$**
Light/PBE	1.80	2.30
Intermediate/PBE	2.31	2.95
Light/HSE06	2.32	2.95
Intermediate/HSE06+SOC	2.15	2.48
Experimental	2.00[5]	2.64[5]

For $Cs_3Bi_2I_9$ (**Figure 3a**), the band structure shows an indirect bandgap of 2.14 eV (C → Γ), aligning with literature[7]. $Cs_3Bi_2Br_9$ (**Figure 3b**) also has an indirect bandgap of 2.48 eV between points on the K–Γ segment and A, consistent with previous reports[8]. Projected DOS indicates halogen orbitals dominate occupied bands, while unoccupied bands near the Fermi level arise mainly from Bi and halogen hybridization. This strong interaction contributes to bonding stability and affects optical and electronic properties relevant for applications.

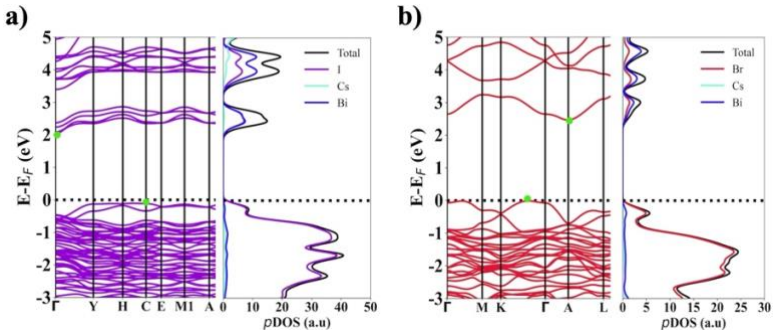


Figure 3. Band structure calculated at HSE06+SOC level of theory for **a)** $Cs_3Bi_2I_9$ and **b)** $Cs_3Bi_2Br_9$. Green circles connect points in which the VBM to CBM transition is observed.

We calculated band gaps (**Figure 4a**) and effective masses for holes and electrons across Br doping levels (**Figure 4b and 4c**). Bandgap increases with Br content, except for a plateau between 40–50% Br, coinciding with the phase transition and decreased octahedral distortion. This behavior matches experimental observations by Ghosh et al[9]. Effective masses, derived via a parabolic band approximation, reveal electrons have lower effective masses (~0.30 m_e for I, ~0.54 m_e for Br) compared to holes (~0.84 m_h for I, ~1.07 m_h for Br), indicating higher electron mobility. These values align with literature[10,11]. Electron effective masses remain roughly constant across compositions (**Figure 4c**), reflecting conduction band insensitivity to halogen substitution. Hole effective masses (**Figure 4b**) increase sharply above 50% Br, linked to a shift in valence band character from iodine to bromine orbitals. This suggests stronger hole self-trapping by Br, hindering hole mobility, as previously noted[12].

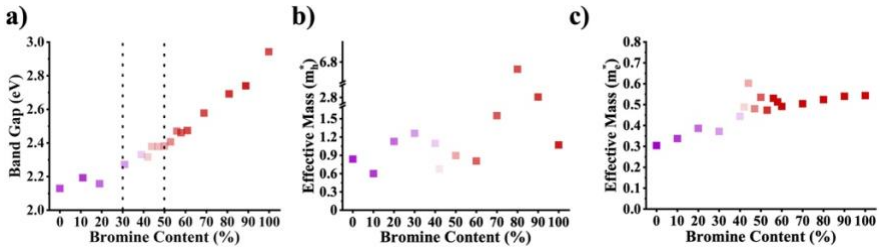


Figure 4: a) Bandgap values calculated at HSE06+SOC level of theory in all the mixed halide compositions investigated. Hole **(b)** and electron **(c)** effective masses calculated from band structures of the mixed I/Br systems, within the parabolic band approximation. Masses are referred to the rest mass of the electron.

Two compositions stand out for optimal charge transport: 11% and 42% Br content show the lowest hole effective masses, echoing findings from Sun et al[13]. and Li et al.[14], who identified similar compositions with superior electron mobility. Notably, the phase transition near 42% Br aligns with Sun et al.'s observations and other experimental data. Structurally, at ~42% Br, bromine atoms bridge adjacent octahedra, enhancing dimensionality and electronic connectivity. This balances localization and delocalization, optimizing charge transport. Higher Br contents introduce localized states acting as traps, raising hole effective masses. At very high Br content (89–100%), effective masses decrease again, reflecting major structural changes as the system completes the transition to the bromide-rich phase with strong Bi–Br hybridization dominating. These shifts significantly influence electronic properties and charge transport.

In conclusion our DFT study of $Cs_3Bi_2I_9$ and $Cs_3Bi_2Br_9$ shows a halogen-driven phase transition near 42% Br, linked to structural changes like Bi–X bond shortening and octahedral distortion. Bandgap increases with Br except during the transition, where distortion alters the trend. Electron mobility is higher than hole mobility, with hole effective masses rising beyond 50% Br from self-trapping. These findings highlight halogen composition as key to tailoring structural, electronic, and transport properties for optoelectronic uses.

Computational Details

Density functional theory (DFT) calculations were performed using FHI-aims with periodic boundary conditions[15]. Geometry optimizations use the PBE functional[16] with Tkatchenko–Scheffler dispersion correction[17] and NAO[18] basis sets (light-tier1), achieving force convergence below 0.02 eV/Å and a 4x4x2 and 4x4x4 k-point grids for iodine and bromine systems respectively. Electronic properties, including band structure and density of states (DOS), are calculated using the hybrid HSE06 functional[19] with

spin–orbit coupling (SOC), refined k-point grid (12x12x6 and 12x12x12) and tier2 basis for Cs atoms, ensuring high accuracy. To capture realistic halide mixing, supercells of 56 atoms are constructed using the Special Quasi-random Structure (SQS) approach[20,21], simulating a broad range of I/Br compositions (0–100% substitution). Each supercell is fully relaxed, and its electronic features—including charge carrier effective masses—are analyzed using the effmass tool[22].

References

1 A. E. Maughan, A. M. Ganose, A. M. Candia, J. T. Granger, D. O. Scanlon and J. R. Neilson, Anharmonicity and Octahedral Tilting in Hybrid Vacancy-Ordered Double Perovskites, *Chem. Mater.*, 2018, **30**, 472–483.

2 M. K. Hossain, G. F. Ishraque Toki, D. P. Samajdar, M. H. K. Rubel, M. Mushtaq, Md. R. Islam, Md. F. Rahman, S. Bhattarai, H. Bencherif, M. K. A. Mohammed, R. Pandey and J. Madan, Photovoltaic Performance Investigation of $Cs_3 Bi_2 I_9$ -Based Perovskite Solar Cells with Various Charge Transport Channels Using DFT and SCAPS-1D Frameworks, *Energy Fuels*, 2023, **37**, 7380–7400.

3 S. Dai, X. Gan, K. Li, Q. Huang, L. Guo and H. Liu, Bandgap lowering in mixed alloys of $Cs_3 Bi_{2-x} Sb_x Br_9$ perovskite powders, *Phys. Chem. Chem. Phys.*, 2023, **25**, 30993–31002.

4 B. Ghosh, S. Chakraborty, H. Wei, C. Guet, S. Li, S. Mhaisalkar and N. Mathews, Poor Photovoltaic Performance of $Cs_3 Bi_2 I_9$: An Insight through First-Principles Calculations, *J. Phys. Chem. C*, 2017, **121**, 17062–17067.

5 S. Bonomi, P. Galinetto, M. Patrini, L. Romani and L. Malavasi, Optical and Structural Property Tuning in Physical Vapor Deposited Bismuth Halides $Cs_3 Bi_2 (I_{1-x} Br_x)_9$ ($0 \le x \le 1$), *Inorg. Chem.*, 2021, **60**, 14142–14150.

6 R. Ketkaew, Y. Tantirungrotechai, P. Harding, G. Chastanet, P. Guionneau, M. Marchivie and D. J. Harding, OctaDist: a tool for calculating distortion parameters in spin crossover and coordination complexes, *Dalton Trans.*, 2021, **50**, 1086–1096.

7 L. Zhang, C. Liu, L. Wang, C. Liu, K. Wang and B. Zou, Pressure-Induced Emission Enhancement, Band-Gap Narrowing, and Metallization of Halide Perovskite $Cs_3Bi_2I_9$, *Angew Chem Int Ed*, 2018, **57**, 11213–11217.

8 K. K. Bass, L. Estergreen, C. N. Savory, J. Buckeridge, D. O. Scanlon, P. I. Djurovich, S. E. Bradforth, M. E. Thompson and B. C. Melot, Vibronic Structure in Room Temperature Photoluminescence of the Halide Perovskite $Cs_3 Bi_2 Br_9$, *Inorg. Chem.*, 2017, **56**, 42–45.

9 S. Ghosh, S. Mukhopadhyay, S. Paul, B. Pradhan and S. K. De, Control Synthesis and Alloying of Ambient Stable Pb-Free $Cs_3 Bi_2 Br_{9(1-x)} I_{9x}$ ($0 \le x \le 1$) Perovskite Nanocrystals for Photodetector Application, *ACS Appl. Nano Mater.*, 2020, **3**, 11107–11117.

10 M. Li, S. Xu, L. Wu, H. Tang, B. Zhou, J. Xu, Q. Yang, T. Zhou, Y. Qiu, G. Chen, G. I. N. Waterhouse and K. Yan, Perovskite $Cs_3 Bi_2 I_9$ Hexagonal Prisms with Ordered Geometry for Enhanced Photocatalytic Hydrogen Evolution, *ACS Energy Lett.*, 2022, **7**, 3370–3377.

11 B. Liu, X. Feng, M. Long, M.-Q. Cai and J. Yang, Designing Surface-Functionalized $Ti_3C_2T_2$-$Cs_3Bi_2Br_9$ (T = O, Cl, OH, or F) Heterostructures for Perovskite Optoelectronic Applications, *Phys. Rev. Applied*, 2022, **18**, 054036.

12 C. Liu, Y. Wang, H. Geng, T. Zhu, E. Ertekin, D. Gosztola, S. Yang, J. Huang, B. Yang, K. Han, S. E. Canton, Q. Kong, K. Zheng and X. Zhang, Asynchronous Photoexcited Electronic and Structural Relaxation in Lead-Free Perovskites, *J. Am. Chem. Soc.*, 2019, **141**, 13074–13080.

13 Q. Sun, B. Xiao, L. Ji, D. Zhao, J. Liu, W. Zhang, M. Zhu, W. Jie, B.-B. Zhang and Y. Xu, Effect of dimensional expansion on carrier transport behaviors of the hexagonal Bi-based perovskite crystals, *Journal of Energy Chemistry*, 2022, **66**, 459–465.

14 X. Li, G. Zhang, Y. Hua, F. Cui, X. Sun, J. Liu, H. Liu, Y. Bi, Z. Yue, Z. Zhai, H. Xia and X. Tao, Dimensional and Optoelectronic Tuning of Lead-free Perovskite $Cs_3 Bi_2 I_{9-n} Br_n$ Single Crystals for Enhanced Hard X-ray Detection, *Angew Chem Int Ed*, 2023, **62**, e202315817.

15 V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter and M. Scheffler, Ab initio molecular simulations with numeric atom-centered orbitals, *Computer Physics Communications*, 2009, **180**, 2175–2196.

16 J. P. Perdew, K. Burke and Y. Wang, Generalized gradient approximation for the exchange-correlation hole of a many-electron system, *Phys. Rev. B*, 1996, **54**, 16533–16539.

17 A. Tkatchenko and M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.*, 2009, **102**, 073005.

18 V. Havu, V. Blum, P. Havu and M. Scheffler, Efficient integration for all-electron electronic structure calculation using numeric basis functions, *Journal of Computational Physics*, 2009, **228**, 8367–8379.

19 J. Heyd, G. E. Scuseria and M. Ernzerhof, Hybrid functionals based on a screened Coulomb potential, *The Journal of Chemical Physics*, 2003, **118**, 8207–8215.
20 A. Van De Walle, R. Sun, Q.-J. Hong and S. Kadkhodaei, Software tools for high-throughput CALPHAD from first-principles data, *Calphad*, 2017, **58**, 70–81.
21 A. Van De Walle, M. Asta and G. Ceder, The alloy theoretic automated toolkit: A user guide, *Calphad*, 2002, **26**, 539–553.
22 L. D. Whalley, effmass: An effective mass package, *JOSS*, 2018, **3**, 797.

MODELLING OF COMPLEX STRUCTURES: NANOCRYSTALS, HETEROSTRUCTURES, AND HYBRID MATERIALS

Juliette Zito[1], Meysoun Jabrane[1], Iyyappa R. Panneerselvam[1] and Gabriele Saleh[1,16]

[1]*Nanochemistry, Istituto Italiano di Tecnologia, Via Morego 30, 16163 Genova, Italy*

ABSTRACT. This document discusses the work on materials modelling performed in our group. Most of this computational work concerns atomistic simulations of complex structures such as nanocrystals, heterostructures, and organic-inorganic materials and of dynamic phenomena such as ionic exchange. Additionally, we illustrate the software we are currently developing for the systematic prediction of epitaxial interfaces and heterostructures.

Introduction

The research in materials science often aims at tuning the properties of materials for a specific application. There are multiple ways to achieve control over the properties of materials, the most obvious being tuning their composition and doping. Other strategies involve developing complex structures with intricate atomic arrangements such as nanocrystals, heterostructures, and hybrid organic-inorganic materials. These systems are characterized by structural heterogeneity, reduced dimensionality, and often, emergent phenomena not present in bulk materials. Nanocrystals exhibit size-dependent quantum effects; heterostructures combine different materials at the nanoscale to yield novel electronic or optical behaviors; and chiral crystals exhibit asymmetry that can influence interactions with polarized light or chiral molecules. The performance of these materials in applications—from catalysis and sensing to photonics and quantum technologies—depends critically on their atomic-scale structure and dynamics.

Understanding such systems requires more than experimental characterization alone. Atomistic simulations are indispensable tools for probing the fundamental mechanisms that govern the stability, reactivity, and functional properties of these complex materials. For example, these simulations can grant access to structural details not easily accessible from experimental probes, or explain the electronic mechanism that leads to the observed optical properties.

In this report, we illustrate our works on atomistic simulations on the above mentioned complex structures, highlighting the interplay with the experiments performed at our institute.

Improving the stability of colloidal CsPbBr3 nanocrystals with an alkylphosphonium bromide as surface ligand pair

Experimentalists in our department synthesized a phosphonium-based ligand, trimethyl(tetradecyl)phosphonium bromide (TTP-Br), and used it to treat Cs-oleate-capped $CsPbBr_3$ nanocrystals (NCs). This surface exchange increased photoluminescence quantum yield (PLQY) from ~60% to over 90% and significantly improved air stability compared to benchmark DDA-Br-treated NCs. To understand the molecular basis for this enhanced optical performance and stability, we carried out density functional theory (DFT) simulations.

Using CP2K[i] with the PBE functional[ii] and a double-ζ basis set with polarization functions[iii], we modeled a 2.4 nm cubic $CsPbBr_3$ NC and replaced one CsBr unit on the surface with either a TTP-Br or DDA-Br ion pair. Optimized structures showed similar ligand orientations, with binding energies of 42.7 kcal/mol (TTP-Br) and 45.2 kcal/mol (DDA-Br), consistent with reported values for other ligand types[iv]

16 Corresponding author. E-mail: gabriele.saleh@iit.it

(Fig. 1a,b). These results support the idea that electrostatic interactions dominate ligand binding to the NC surface.

To explore full surface passivation, we substituted CsBr units on all six facets with simplified phosphonium ligands (ethyl tail), maintaining a surface density of 1.27 ligands/nm^2 in line with experimental conditions. The resulting electronic structure was free of midgap states and showed a delocalized valence band maximum (Fig. 1c). In contrast, prior calculations showed that Cs-oleate-capped NCs exhibit trap states due to oxygen-containing anchoring groups[v]. This supports the experimental observation that phosphonium ligands effectively passivate surface defects and enhance NC optical performance.

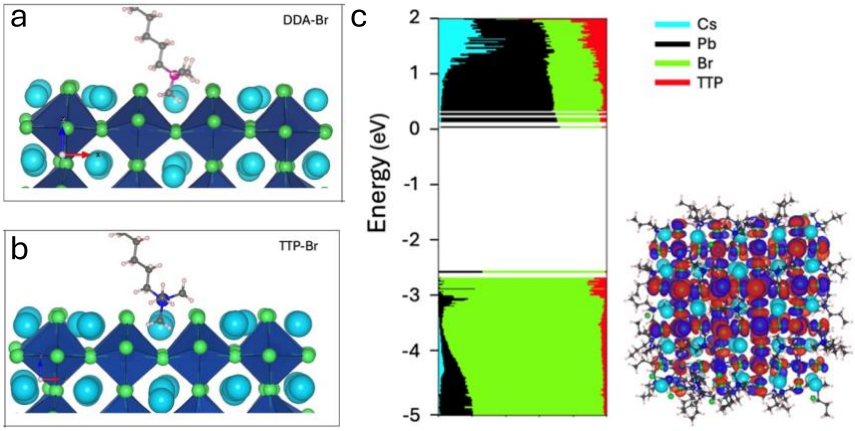


Figure 1. Binding configuration of (a) DDA$^+$ and (b) TTP$^+$ ligands in the A-site of the $CsPbBr_3$ NC surface after DFT/PBE structural relaxation. Both quaternary ammonium and phosphonium ligands adopt similar orientations, with one N–CH_3 or P–CH_3 bond nearly perpendicular to the surface. (c) Electronic structure of a ~2.4 nm $CsPbBr_3$ NC passivated with TTP-Br at 1.27 ligands/nm^2, computed at the DFT/PBE level. Atom-type contributions to molecular orbitals are color-coded; C, H, and P from TTP-Br are grouped. The valence band edge orbital is shown (isosurface: 0.02 e/Bohr^3).

Perovskite-Chalcohalide Nanocrystal Heterostructures as a Platform for the Synthesis and Investigation of CsPbCl3-CsPbI3 Epitaxial Interface

Halide exchange in lead halide perovskites is well understood for Br/Cl and Br/I systems, but Cl/I mixing is typically hindered by large ionic size mismatch and miscibility gaps. To overcome this, our experimental collaborators used $CsPbCl_3$–$Pb_4S_3Cl_2$ heterostructures as a platform to study Cl→I exchange, enabling the formation of intermediate structures not accessible in pure $CsPbCl_3$ systems. These experiments revealed segmented $CsPbI_3$–$CsPbCl_3$–$Pb_4S_3Cl_2$ architectures with partial I-alloying and interface reconstruction.

To rationalize the formation and stability of these intermediates, we performed DFT calculations using the PBE functional[ii] and a double-ζ polarized basis set in CP2K[i]. The Cl→I exchange was modeled using the following reaction:

$HS(x{\cdot}Cl) + x\ PbI_2 \rightarrow HS(x{\cdot}I) + x\ PbCl_2$

with reaction energy:

$\Delta E\_exchange = [(E\_HS(x{\cdot}I) + E\_PbCl_2) - (E\_HS(x{\cdot}Cl) + E\_PbI_2)] / x$.

We began with a published $CsPbCl_3$–$Pb_4S_3Cl_2$ model[vi] and simulated I-for-Cl substitution in five perovskite layers, either from the heterointerface or from the opposite end. The results showed that substitution starting from the bottom of the perovskite domain is slightly favored, supporting the experimental observation of segmented structures.

Next, we modeled I incorporation directly at the interface by fully replacing Cl^- with I^- in a single perovskite monolayer. We then simulated progressive I substitution in the first two chalcohalide layers (up to 100% I). It turned out that the initial 0–20% exchange yields the greatest energetic gain, likely due to interfacial strain relaxation. Further substitution shows nearly constant reaction energies. Entropy,

though not included, is expected to favor intermediate I/Cl ratios (~40–60%), aligning with halide distributions seen in STEM.

We also examined partial I-alloying within the $CsPbCl_3$ domain by substituting ~20% of Cl^- with I^-. Four I^- distributions were tested: random, CsI layers perpendicular and parallel to the interface, and a PbI_2 layer. All were energetically favored, with ordered configurations being 5–10 kcal/mol more stable than random ones. The energy gain scales with system size. Structural analysis shows that I ordering reduces strain by expanding the $CsPbCl_3$ lattice parallel to the interface—explaining the experimentally observed I-rich layering.

Finally, we analyzed the electronic structure of three heterostructures: $CsPbCl_3$–$Pb_4S_3Cl_2$, partially exchanged $CsPbI_3$–$CsPbCl_3$–$Pb_4S_3Cl_2$, and fully exchanged $CsPbI_3$–$Pb_4S_3Cl_2$ (Fig. 2). In all cases, the band edges are shaped by chalcohalide-localized trap states. In the Cl-based system, mixed VB and CB states promote nonradiative recombination, consistent with the lack of emission. In contrast, the fully exchanged heterostructure shows VB states localized on $CsPbI_3$ and reduced electron transfer to trap states, allowing radiative recombination and explaining the emission observed experimentally. In the partially exchanged system, the remaining $CsPbCl_3$ layer acts as a barrier, further enhancing radiative recombination in the I-based domain. These trends align well with steady-state and time-resolved optical data.

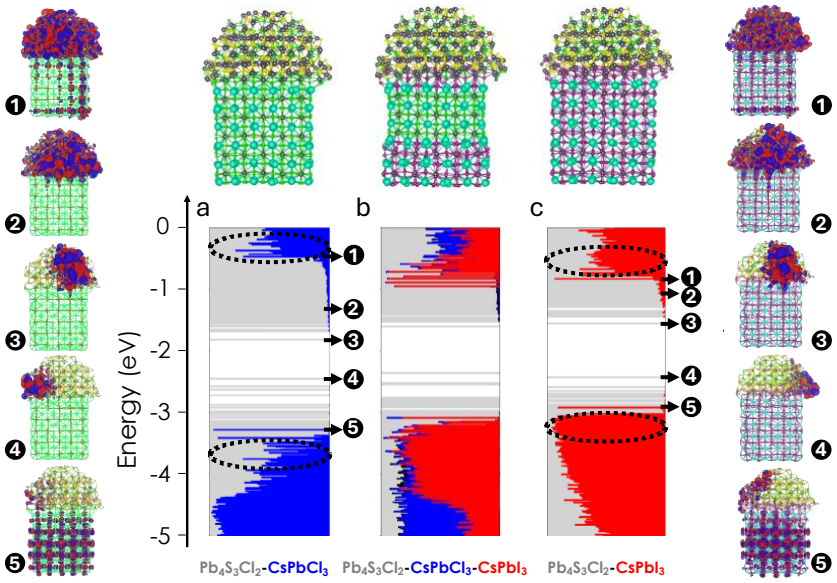


Figure 2. The electronic structure of (a) the $CsPbCl_3$-$Pb_4S_3Cl_2$, (b) the $CsPbI_3$-$CsPbCl_3$-$Pb_4S_3Cl_2$, and (c) the $CsPbI_3$-$Pb_4S_3Cl_2$ heterostructures computed at the DFT/PBE level of theory. The color code indicates the contribution of each domain to each molecular orbital. Representation of molecular orbitals corresponding to band edge delocalized states and to band edge trap states are additionally reported for the parent and fully exchanged heterostructures.

Scanning Materials Combinations for Matching Heterostructures

When two different materials form a hybrid interface, they can display unique physical properties not found in their bulk forms. These carefully designed combinations, known as heterostructures, are engineered by interfacing materials with distinct electronic or optical characteristics to unlock new functionalities such as tailored bandgaps, or quantum confinement effects[vii]. To take advantage of these effects and improve device performance, our project aims to develop a computer program to make predictions about epitaxial interfaces in nanocrystals. The goal is to predict interface formation between materials with different lattice parameters using only the most basic input: the crystal structure of compounds. To this aim, we exploit the libraries of OGRE, an open-source tool that predicts inorganic epitaxial interfaces through lattice and surface matching[viii]. By ensuring commensurability in domain-matched interfaces, OGRE avoids the need for costly DFT simulations.

We integrated the capability to run OGRE package on large databases like Crystallography Open Database (COD)[ix], enabling the identification of chemically plausible and geometrically compatible

epitaxial interfaces across a wide range of material combinations. As a case study, we focused on $CsPbBr_3$ and CdSe, matching them against a post-processed COD dataset containing over 10,000 materials. The screening process involves evaluating all possible facets within defined area and mismatch limits.
In the first step of lattice matching, the program processes several sub-datasets totaling 839,208 Miller matches. The second step involves generating all feasible interfaces from the resulting terminations, as shown in the figure below. These calculations are being performed on Cresco6 using 4 - 20 cores, with runtimes ranging from minutes to days depending on the dataset size and computation stage. So far, 90% of the results have been extracted for $CsPbBr_3$ x COD, and approximately 20% for CdSe x COD. In the final stage, data analysis will be applied to identify trends in interfacial matches. Spotting these patterns can help guide future screening and reveal useful rules for designing better material interfaces.

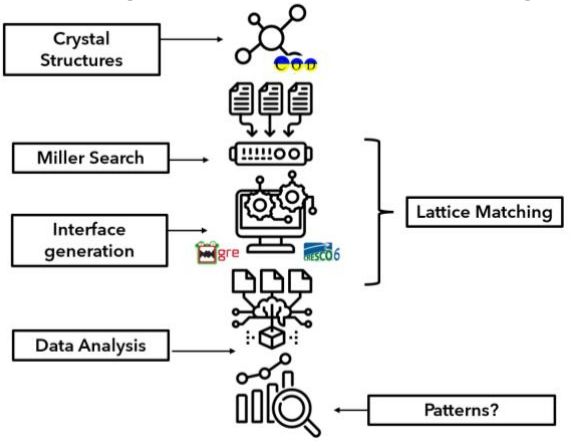


Figure 3. Workflow of Lattice matching for heterostructure prediction

Pushing the boundaries of organic-inorganic hybrid perovskites

In our department, synthetic chemists managed to synthesize $EAPbI_3$ ($EA=CH_3CH_2NH_3^+$) nanocrystals, a result previously thought to be impossible due to the large size of EA. We adopted DFT simulations (through the VASP software[x]) and performed simulated annealing molecular dynamics to explore the structural and electronic properties of $EAPbI_3$[xi]. Eleven low-energy configurations with varying EA cation positions were found to have small energy differences, suggesting that multiple configurations coexist at room temperature (Fig. 4). This structural variability leads to significant band gap fluctuations (up to 0.44 eV), contributing to broadened excitonic features. The electronic structure remains consistent with typical lead halide perovskites, involving Pb–I antibonding orbitals in both valence and conduction bands.

Heping experimental probes to locate dopants in complex nanocrystals

In an ongoing project, we performed simulations on Zn-doped InAs quantum dots to locate the most favourable positions on dopants. The information obtained from simulations are combined with experimental results. In particular, we performed DFT simulations[i,ii] placing Zn in various locations on the quantum dot and seeking the most stable configurations, *i.e.* those in which the energy reaches a minimum.

Acknowledgments

We acknowledge funding from the Project IEMAP (Italian Energy Materials Acceleration Platform) within the Italian Research Program ENEA-MASE (Ministero dell'Ambiente e della Sicurezza Energetica) 2021–2024 "Mission Innovation" (agreement 21A033302 GU No. 133/5-6-2021).

References

[i] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele. cp2k: atomistic simulations of condensed matter systems. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, **4(1)**, pp. 15–25, (2013).
[ii] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, **77(18)**, pp. 3865–3868, (1996).
[iii] J. VandeVondele, and J. Hutter. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *The Journal of Chemical Physics*, **127(11)**, 114105, (2007).
[iv] M. Imran, P. Ijaz, L. Goldoni, D. Maggioni, U. Petralanda, M. Prato, G. Almeida, I. Infante, and L. Manna. Simultaneous Cationic and Anionic Ligand Exchange For Colloidally Stable $CsPbBr_3$ Nanocrystals. *ACS Energy Letters*, **4(4)**, pp. 819–824, (2019).
[v] G. Almeida, O. J. Ashton, L. Goldoni, D. Maggioni, U. Petralanda, N. Mishra, Q. A. Akkerman, I. Infante, H. J. Snaith, and L. Manna. The Phosphine Oxide Route toward Lead Halide Perovskite Nanocrystals. *Journal of the American Chemical Society*, **140(44)**, pp. 14878–14886, (2018).

[vi] S. Toso, M. Imran, E. Mugnaioli, A. Moliterni, R. Caliandro, N. Schrenker, A. Pianetti, J. Zito, F. Zaccaria, Y. Wu, M. Gemmi, C. Giannini, S. Brovelli, I. Infante, S. Bals, and L. Manna. Halide perovskites as disposable epitaxial templates for the phase-selective synthesis of lead sulfochloride nanocrystals. *Nature Communications*, **13(1)**, (2022).
[vii] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*. Les Ulis: Éditions de Physique, 1990.
[viii] S. Toso, D. Dardzinski, L. Manna, and N. Marom, "Structure prediction of ionic epitaxial interfaces with OGRE demonstrated for colloidal heterostructures of lead halide perovskites," *ACS Nano*, vol. 19, no. 5, pp. 5326–5341, (2025).
[ix] S. Gražulis, D. Chateigner, R. T. Downs, A. F. Tupytsin, E. Quirós, L. Lutterotti, M. Kairys, and B. Le Bail, "Crystallography Open Database (COD): An open-access collection of crystal structures," *J. Appl. Crystallogr.*, vol. 42, pp. 726–729, (2009).

[x] G. Kresse, and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational materials science*, **6(1)**, pp. 15-50, (1996).
[xi] C. M. Guvenc, S. Toso, Y. P. Ivanov, G. Saleh, S. Balci, G. Divitini, and L. Manna. Breaking the Boundaries of the Goldschmidt Tolerance Factor with Ethylammonium Lead Iodide Perovskite Nanocrystals. *ACS nano*, **19(1)**, 1557-1565, (2024).