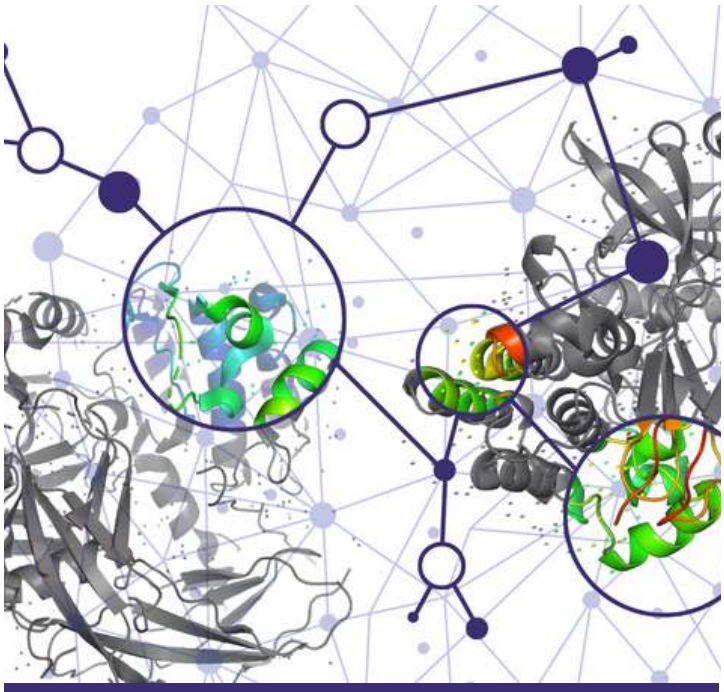
# NEWSLETTER

OCTOBER 2023

ML4NGP COST ACTION | ED. 1

By Rita Vilaça | Science Communication Coordinator



## THE KICK-OFF MEETING

The Action kick-off was held on 25 and 26 October 2022 during the first MC meeting and marked the beginning of our COST Action. This inaugural meeting gathered the main proposer, members of the management committee and the COST officers to discuss the implementation of the Action.

During this event, the members of the core group, including the Chair and Vice-Chair, were elected by the management committee and the main project's goals and objectives were outlined. It was an exciting start that set the stage for a very successful year!

## COST ACTION CA21160

The COST Action CA21160 ML4NGP - Non-globular proteins in the era of Machine Learning - aims to establish an multidisciplinary pan-European network to describe the structural properties of Non-Globular Proteins (NGPs) by combining experimental data and novel machine learning (ML) approaches. Understanding the structure and function of NGPs is crucial as they represent 30-40% of human proteome and are implicated in many biological processes and in several diseases. Our main goal is to improve the design of experimental frameworks to provide information to computational methods, and develop new computational methods that are trained and benchmarked with experimental data. Ultimately, this knowledge can be used for developing new therapies and novel biomaterials.

*Members of the core group together at the 1st ML4NGP annual meeting in Bratislava, Slovakia.*

## OUR COMMITMENT

ALEXANDER MONZON & ZUZANA BEDNARIKOVA, ACTION CHAIRS

We recognize the significance of our mission to unravel the complexity of the structure and function of NGPs through the fusion of experimental data and cutting-edge ML-based methodologies. The scientific community has been biased towards globular proteins for many years, and it is time we shed light on the diverse and intricate world of NGPs. Together with the members of this Action, we shall embark on this journey to unravel the complexities of NGPs and usher in a new era of understanding in the scientific community. Through collective efforts, dedication, and innovation, we shall achieve the goals set forth in this Action, fostering a brighter future for NGP research.

# FIRST YEAR IN REVIEW

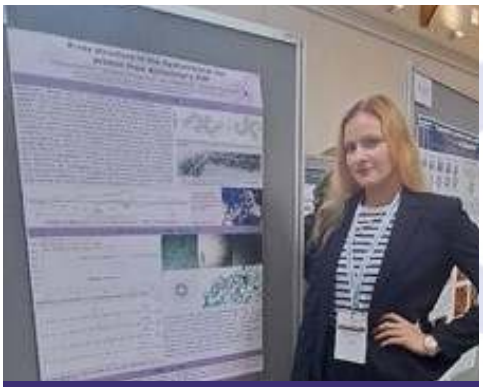
## TRAINING SCHOOL

PORTO, PORTUGAL 2023

The 1st ML4NGP Training School was our first in person meeting and it was a success! The event was co-organized together with PhasAGE Twinning project and IDPfun MSCA-RISE project from 19 to 21 April in Porto, Portugal. The course covered the latest advances in computational resources and experimental methods for studying non-globular protein dynamics, such as aggregation and phase separation. The 22 selected participants (11 countries) learned about the development of machine learning-based tools for protein structure prediction and characterization from internationally recognized speakers, including ML4NGP Working Group leaders, through lectures and hands-on sessions. [More...]

*Participants and speakers during the theoretical lectures and hands-on practical sessions at the 1st ML4NGP Training School in Porto, Portugal.*

> *The organization of the training school was excellent. The speakers and trainers were really interesting. I enjoyed very much listening to the presenters. They covered interesting topics, most of them related to my study area.*



### TRAINING AND MOBILITY OF YOUNG RESEARCHERS

During the first grant period we awarded four researchers from inclusiveness target countries with conference grants to promote dissemination of their research work at high level international conferences.
We also granted funding for three short term scientific missions for mobility and training of researchers from different countries and institutions.
With these funding opportunities, we aim to support scientific research and foster collaboration among Action's members.

## WORKSHOPS

The 1st ML4NGP Workshop "Structural Ensembles of Intrinsically Disordered Proteins" took place at Vila Lanna from 7 to 8 May, in the beautiful city of Prague, Czech Republic. The workshop was attended by 18 participants to discuss new developments of IDPs experimental determination through Nuclear Magnetic Resonance, and generation, comparison and validation of structural ensembles. At the end, the participants express their interest in starting to draft a community paper, addressing the current and future challenges on the topic. [More...]

The 2nd ML4NGP Workshop "Computational Methods for Tandem Repeat Proteins" took place in Montpellier, France from 26 to 28 September. The workshop gathered 30 participants to discuss the state-of-the-art and latest advances on machine learning computational methods, including AlphaFold, for classification of Tandem Repeat Proteins (TRPs). The next steps include preparing a community paper about Class I and II of TRPs and write a biocuration protocol for RepeatsDB.

# FIRST YEAR IN REVIEW

CONFERENCE

## 1ST ML4NGP MEETING ON MACHINE LEARNING AND NON-GLOBULAR PROTEINS

BRATISLAVA, SLOVAKIA 2023

The first meeting of the ML4NGP COST Action took place in Bratislava, Slovakia, from July 5 to 7, 2023. The meeting brought together 108 computational and experimental researchers from 29 COST Countries and 1 International Partner Country interested in adapting machine learning based computational tools for the advancement of research on non-globular proteins (NGPs). The program included 5 invited talks from members of the Scientific Advisory Board and featured 20 selected short talks and 37 posters. Most of the presentations showcased the recent advances in computational and experimental methods for classification and characterization of NGPs and how machine learning can be used for NGPs research, including protein-protein interactions, protein engineering, and protein function prediction. It was also discussed how aggregation of NGPs are associated with human diseases and how their characterization is of major relevance for health and biotech innovation and development. Six presenters were awarded for their outstanding contributions, with two receiving recognition for the best poster presentation, two for the best short talk, and two for the best flash talk. The success of this meeting was acknowledged by the conference chair, "ML4NGP serves as a remarkable and inclusive platform, bringing together interdisciplinary researchers from diverse career stages to collectively explore the vast potential of machine learning in comprehending non-globular proteins", said Alexander Monzon.

*Photos of the invited speakers: Peter Tompa, Sonia Longhi, Andrey Kajava, Miguel Andrade and Salvador Ventura.*

FEEDBACK

*Great program/speakers, adequate time for discussions/networking. The presentations were very interesting and very helpful for my research in the future. Good balance between experimental/computational talks and the scientific networking.*

*Group photo with the participants and speakers in front of Clarion Congress Hotel during the first annual ML4NGP meeting held in Bratislava, Slovakia.*

This meeting also featured specific sessions with presentations from selected members of each working group as a starting point for future engagement and collaborations. Moreover, members of the management committee gathered to discuss on the agenda of future actions and to elect the co-leaders of the four scientific working groups (WG1-WG4) of the ML4NGP COST Action. [More...]

# FIRST YEAR IN REVIEW

## MAIN OUTPUTS

## DELIVERABLES

- ML4NGP Handbook of Operations v1
- ML4NGP Science Communication, Dissemination and Exploitation Plan v1

## PUBLICATIONS


Escobedo, N., et al., "Combining Protein Conformational Diversity and Phylogenetic Information Using CoDNaS and CoDNaS-Q." Curr Protoc, 2023. 3(5): p. e764

Del Conte, A., et al., "CAID prediction portal: a comprehensive service for predicting intrinsic disorder and binding regions in proteins." Nucleic Acids Res, 2023. 51(W1): p. W62-W69

Arrias, P.N., et al., "The repetitive structure of DNA clamps: An overlooked protein tandem repeat." J Struct Biol, 2023. 215(3): p. 108001

Del Conte, A., et al., "Critical assessment of protein intrinsic disorder prediction (CAID) – Results of round 2." Proteins, 2023

Monzon, A.M., et al., "A STRP-ed definition of Structured Tandem Repeats in Proteins." J Struct Biol, 2023: p. 108023

Garcia-Pardo, J., et al. "A3DyDB: exploring structural aggregation propensities in the yeast proteome." Microb Cell Fact, 2023. 22, p186

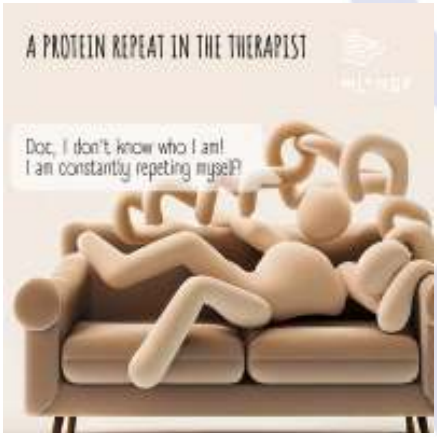
Aspromonte, M.C., et al. "DisProt in 2024: improving function annotation of intrinsically disordered proteins" Nucleic Acids Res, 2023.

Ghafouri, H., et al. "PED in 2024: improving the community deposition of structural ensembles for intrinsically disordered proteins" Nucleic Acids Res, 2023.


## OUR ACTION IN NUMBERS

| | | | |
|---|---|---|---|
| 229 Action Members | 43 Countries | 4 Training Events | 183 Participants in events |
| 4 Short Term Missions | 3 ITC Conference Grants | 5 Publications | 2 Deliverables |

## THANK YOU ALL FOR AN EXCELLENT YEAR!



*This newsletter is based upon work from COST Action ML4NGP, CA21160 supported by COST (European Cooperation in Science and Technology.*

*COST (European Cooperation in Science and Technology) is a funding agency for research and innovation networks. Our Actions help connect research initiatives across Europe and enable scientists to grow their ideas by sharing them with their peers. This boosts their research, career an innovation. https://www.cost.eu/*

Funded by the European Union