

ML⁴NGP

MACHINE LEARNING FOR NON GLOBULAR PROTEINS

2nd MEETING on MACHINE LEARNING AND NON- GLOBULAR PROTEINS

14 – 17 May 2024

Onoma Hotel, Thessaloniki, Greece

2nd MEETING ON MACHINE LEARNING AND NON-GLOBULAR PROTEINS

13 – 17 May 2024, Thessaloniki, GREECE

Welcome message

With great pleasure and on behalf of Non-globular proteins in the era of Machine Learning COST Action CA21160 (ML4NGP) and Centre for Research and Technology Hellas in Thessaloniki (Greece), we cordially welcome you in Thessaloniki for the 2nd ML4NGP Meeting. We are delighted that more than 100 scientists from 30 countries will meet in Thessaloniki in the upcoming days.

This meeting is co-organized by the H2020-funded MSCA project REFRACT ("Repeat Protein Function, Refinement, Annotation and Classification of Topologies", GA 823886).

The program integrates keynote speakers, invited talks by leading experts, short talks, flash presentations, and more than 50 poster presentations by scientists in the field from within and beyond the organizing consortia. During the conference, we invite all participants to join the working group meetings to discuss recent advances in protein structure prediction tools and further contribute to developments in this field.

We hope that your participation at the 2nd ML4NGP Meeting in Thessaloniki will be scientifically rewarding and provide a unique opportunity to experience our city's wonderful setting and culture.

Organizing Committee

Fotis Psomopoulos (Centre for Research and Technology Hellas, Thessaloniki, Greece)

Elena Bakoglidou (Centre for Research and Technology Hellas, Thessaloniki, Greece)

Anastasia Chasapi (Vidavo S.A, Thessaloniki, Greece)

Core Group of ML4NGP COST Action

Keynote speakers

Christine Orengo (University College London, UK)

Michele Vendruscolo (University of Cambridge, UK)

Andrei N. Lupas (Max Planck Institute for Biology Tubingen, DE)

Christos Ouzounis (Centre for Research and Technology Hellas, GR)

Registration

The registration desk will be open from 16:00 to 18:00 on Monday, May 13th, and Tuesday morning, May 14th, from 8:00 to 9:30.

Name Badges

Participants must wear their name badges at the venue and during social events, lunches, and dinners.

Internet access

Wireless Internet is available for free.

Network: ONOMA WiFi

Password: ONOMAHOTEL

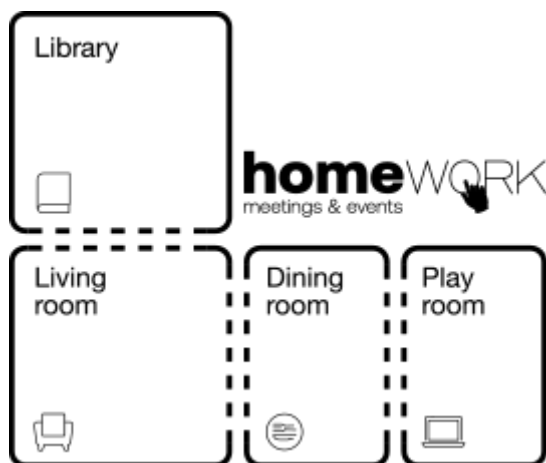
Venue rooms

The meeting will be held on the first floor at ONOMA Hotel in the living room (see picture below).

The poster presentations will be set up in the Dining room (see picture below) on Wednesday and Thursday.

Coffee breaks will be provided in the Dining room (see picture below).

Lunches and dinners will be held at the Cloud Bar at ONOMA Hotel..



Onoma Hotel First floor ground plan

Presentation guidelines

Poster guidelines

Posters should be up to **A1 size** – 59.4 cm (width) x 84.1 cm (height). Poster numbers will be communicated to authors at the Registration Desk.

Poster presenters are requested to set up their posters in the designated area before each poster session starts. Presenters are expected to be available near their posters during the assigned session. More information will be provided near the event.

Flash Poster Presentation

Selected poster abstracts should be presented as a slide presentation (5 min). Speakers are required to have their presentation ready on a USB key and contact technical staff during the coffee break before their session.

Oral presentation

Speakers are required to have their presentations ready on a USB key. Your slides should be either in .pdf, .ppt or .pptx format.

Speakers using their laptops are kindly asked to approach technical staff to clarify the technical prerequisites and coordinate the processes involved.

Presentation time

Since the meeting program is fully packed, we ask all presenters to keep to the time indicated for each presentation:

Keynotes - speaking time 35 min plus 5 min discussion

Selected talks - speaking time 25 min plus 5 min discussion

Short talks - speaking time 15 min plus 5 min discussion

Flash presentations - speaking time 5 min

Social event and dinner

Welcome Reception - Tuesday, 14 May 2024 | 19:00

The Welcome Reception will be held in the hotel restaurant. All the participants are invited to have drinks and snacks.

Social Dinner - Thursday, 16 May 2024 | 19:30

The social dinner is optional and is not included in the meal package. It will be on Thursday, May 16, 2024, at the Onoma Hotel. The cost is 42 EUR, including a buffet menu and beverages.

Program

Monday, 13 May 2024

16:00 – 18:00 Registration

18:00 – 19:30 Core Group Meeting (close session)

Tuesday, 14 May 2024

8:30 – 9:30 Registration

Conference room (Library)

9:30 – 09:40 **Opening Session**

Alexander Monzon

WGI SESSION – Primary experimental data generation

» *Keynote Lecture*

09:40 – 10:20 **What does AlphaFold reveal about the folded regions in the protein universe?**

Christine Orengo

» *Short Talk*

10:20 – 10:40 **Structural studies of mineral interactions with intrinsically disordered magnetosome-associated proteins**

Raz Zarivach

10:40 – 11:10 **Coffee Break**

» *Short Talks*

11:10 – 11:30 **Spider silk: from NMR structural studies to mechanism of formation and high-performance artificial fibres**

Kristaps Jaudzems

11:30 – 11:50 **A disordered loop in a bacterial adhesin can form a poly-proline II helix upon binding to heparin**

Dirk Linke

11:50 – 12:10 **The role of environment in amyloid aggregation**

Vytautas Smirnovas

12:10 – 12:30 **Cryo-EM structures of functional and pathological amyloid ribonucleoprotein assemblies**

Javier Garcia Pardo

12:30 – 14:00 **Lunch Break**

Hotel restaurant

14:00 – 15:40 **WG 1 meeting**

15:40 – 16:10 **Coffee Break**

» *Selected Talk*

16:10 – 16:40 **The Cys sense: Regulation of cellular proliferation and protein homeostasis through redox-switch proteins**

Dana Reichmann

» *Short Talk*

16:40 – 17:00 **Hidden flexible state(s) of sigma-1.1 domain of RNA polymerase**

Pavel Kaderavek

17:00 – 18:10 **Flash Poster Presentations**

18:10 – 19:00 **WG5 meeting and activities**

19:00 – 21:00 **Welcome reception**

Hotel restaurant

Wednesday, 15 May 2024

Conference room (Library)

WG 2 SESSION - Integrative structural biology with machine learning-based approaches

» *Keynote lecture*

09:30 – 10:10 **Sequence-Based Drug Discovery for Disordered Proteins**

Michele Vendruscolo

» *Short Talks*

10:10 – 10:30 **Weighted families of contact maps to characterize conformational ensembles of (highly-)flexible proteins**



Juan Cortes

10:30 – 10:50 **Development of machine learning algorithms for the prediction of the binding affinity between peptides and their target proteins**

Kanin Wichapong

10:50 – 11:20 **Coffee Break**

» *Selected Talk*

11:20 – 11:50 **Combined use of NMR spectroscopy and computation to understand protein phase separation**

Markus Zweckstetter

» *Short Talks*

11:50 – 12:10 **Pioneer factor Sox2 remodels nucleosomes and displaces histone H1**

Sveinn Bjarnasson

12:10 – 12:30 **Residue Interaction Network Generator v4: improved interaction recognition for all 35,000 PDB chemical components**

Alessio Del Conte

12:30 – 14:00 **Lunch Break**

Hotel restaurant

14:00 – 15:40 **WG2 meeting**

15:40 – 16:10 **Coffee Break**

» *Short Talks*

16:10 – 16:30 **FAIR Research Software as the catalyst for trustworthy AI in Life Science**

Fotis Psomopoulos

16:30 – 16:50 **Understanding specific interactions in phase-separated Germ granules by multi-scale simulations**

Vasileios Xenidis

16:50 – 17:10 **Local Energetic Frustration and Intrinsic Protein Disorder**

Gonzalo Parra

17:10 – 17:30 **Disorder of disease-related proteins – a case study of IGDD**

Jovana Kovacevic

17:30 – 19:00 **Poster Session I**

Thursday, 16 May 2024

Conference room (Library)

WG 3 SESSION - ML predictors and methods in the NGP field

» *Keynote lecture*

09:30 – 10:10 **The virtual proteome of life: What is it and why should we care?**

Andrei Lupas

» *Short Talks*

10:10 – 10:30 **Transferable deep generative modeling of intrinsically disordered protein structural ensembles**

Giacomo Janson

10:30 – 10:50 **Developing machine-learning-based amyloid predictors with Cross-Beta DB**

Valentin Gonay

10:50 – 11:20 **Coffee Break**

» *Selected Talk*

11:20 – 11:50 **Data-driven modeling of the intrinsically disordered proteome**

Giulio Tesei

» *Short Talks*

11:50 – 12:10 **Automatic Classification of Units in Tandem Repeat Proteins**

Michela Quadrini

12:10 – 12:30 **LCR-dataset: manually curated benchmark for training and fine-tuning language models to retrieve information on Low Complexity Regions and its functional annotations from scientific literature**

Sylvia Szymanska

12:30 – 14:00 **Lunch Break**

Hotel restaurant

14:00 – 15:40 **WG3 meeting**

15:40 – 16:10 **Coffee Break**

» *Selected Talk*

16:10 – 16:40 **Critical assessment of protein intrinsic disorder prediction (CAID) - Results of round 3**
Damiano Piovesan

» *Short Talks*

16:40 – 17:00 **Deciphering Allosteric Pathways**
Turkan Haliloglu

17:00 – 17:20 **Optimizing energy functions with neural networks**
Gabor Erdos

17:30 – 19:00 **Poster Session II**

19:30 – 22:00 **Social dinner**

Friday, 17 May 2024

Conference room (Library)

WG 4 SESSION - functional characterization of NGPs

» *Keynote lecture*

09:30 – 10:10 **From structure-function and globular proteins to non-globular proteins and the functional dark matter**
Christos Ouzounis

» *Short Talks*

10:10 – 10:30 **α SynPEP-DB: a database of biogenic peptides for inhibiting α -synuclein aggregation**
Carlos Pintado Grima

10:30 – 10:50 **From interaction networks to interfaces, scanning intrinsically disordered regions using AlphaFold2**
Diego Javier Zea

10:50 – 11:20 **Coffee Break**

» *Selected Talk*

11:20 – 11:50 **Blending Biology, Chemistry and AI to Enable Personalized Systems Pharmacology**
Patrick Aloy

» *Short Talks*

11:50 – 12:10 **Folding-upon-binding pathways of an intrinsically disordered protein from a deep Markov state model**
Thomas Sisk

12:10 – 12:30 **Exploring the biophysical boundaries of protein families with deep learning methods**
Miriam Poley

12:30 – 14:00 **Lunch Break**
Hotel restaurant

14:00 – 15:40 **WG4 meeting**

15:40 – 16:10 **Coffee Break**

» *Short Talks*

16:10 – 16:30 **Perturbation modelling across single-cell omic modalities**
Georgios Gavriilidis

16:30 – 16:50 **MAVISp, a modular structure-based framework for protein variant interpretation: current and future developments for supporting non-globular proteins**
Matteo Arnaudi

16:50 – 17:10 **Classification and prediction of amyloid-forming propensity of amino acid sequences**
Rafayel Petrosyan

17:10 – 17:30 **Voronoi tessellation-based analysis of 3D conformations of non-globular proteins**
Kliment Olechnovic

17:30 – 17:50 **Closing Remarks**

List of flash presentations:

- F01 Marija Vidovic: Aggregation of LEA proteins: role in desiccation tolerance
- F02 Pablo Mier: Homorepeat variability within the human population
- F03 Jan Beránek: Jumpcount: Exact confidence intervals for free energy difference estimations
- F04 Dragana Dudić: Predicting disordered proteins in *Zea mays* L.
- F05 Milana Grbic: Collective Classification Algorithms in Identifying Intrinsically Disordered Proteins within Protein-Protein Interaction Networks
- F06 Rostislav Skrabana: Side-chain propelled rings (SPURs) in globular and disordered proteins
- F07 Miguel Fernandez-Martin: Dynamic and Energetic Consequences of Disulfide Bonds in Proteins
- F08 Zoltán Gáspári: Structural ensembles of the single alpha-helix (SAH) segment of the postsynaptic protein Drebrin
- F09 Gunnar Dittmar: Systematic analysis of protein interactions using PrISMa
- F10 Mirela Lika: Exploring Disease-Associated Implications of Non-Synonymous Single Nucleotide Polymorphisms in the Human Fibronectin 1 Gene: A Comprehensive Analysis of Functional and Structural Consequences
- F11 Norbert Deutsch: Evaluating AlphaMissense Mutation Pathogenicity in Intrinsically Disordered Proteins using DisCanVis

List of Poster Presentations:

Poster session I

- F01 Marija Vidovic: Aggregation of LEA proteins: role in desiccation tolerance
- F02 Pablo Mier: Homorepeat variability within the human population
- F03 Jan Beránek: Jumpcount: Exact confidence intervals for free energy difference estimations
- F04 Dragana Dudić: Predicting disordered proteins in *Zea mays* L.
- F05 Milana Grbic: Collective Classification Algorithms in Identifying Intrinsically Disordered Proteins within Protein-Protein Interaction Networks

- F06 Rostislav Skrabana: Side-chain propelled rings (SPURs) in globular and disordered proteins
- F07 Miguel Fernandez-Martin: Dynamic and Energetic Consequences of Disulfide Bonds in Proteins
- F08 Zoltán Gáspári: Structural ensembles of the single alpha-helix (SAH) segment of the postsynaptic protein Drebrin
- F09 Gunnar Dittmar: Systematic analysis of protein interactions using PrISMa
- F10 Mirela Lika: Exploring Disease-Associated Implications of Non-Synonymous Single Nucleotide Polymorphisms in the Human Fibronectin 1 Gene: A Comprehensive Analysis of Functional and Structural Consequences
- F11 Norbert Deutsch: Evaluating AlphaMissense Mutation Pathogenicity in Intrinsically Disordered Proteins using DisCanVis
- P01 Agostina Spiga: Molecular dynamics study on evolutionary loss of calcium binding in beta-gamma crystalline proteins
- P02 Agustín Ormazábal: A comprehensive exploration of human tandem repeat RNA binding proteins: synergies between sequence and structure analysis
- P03 Aimilia Christina Vagiona: Identifying function of Intrinsically disordered proteins (IDPs) with hyperbolic mapping of human protein-protein interactome
- P04 Ana M. Melo: The role of distinct lipid compositions in the initial membrane-binding of tau protein
- P05 Andrius Sakalauskas: The major components of cerebrospinal fluid alters inhibition and aggregation pathway of amyloid- β and α -synuclein.
- P06 Carmen Stefany Neciosup Vera: Panoramic view of ML algorithms and their use in Tandem Repeat Protein tools
- P07 Dario Ferreira: Predictive Modeling of Non-Globular Proteins Using Mixed Linear Models with Fixed and Random Factors
- P08 Di Meng: Optimizing Disordered Flexible Linker (DFL) Analysis: A Comprehensive Domain-Linker-Domain Dataset Collection Approach
- P09 Emre Aktaş: Structural Dynamics and Pharmacological Interactions of SLITRK1 Protein
- P10 Haidara Nadwa: Kinetics and thermodynamics characterization of Serum Amyloid A (SAA) protein

- P11 Kristina Stevanović: Identification of the NLRP3 structural domain crucial for its activation and new bioinformatics criterion for assessment of mutations as a risk factor for inflammatory diseases
- P12 Luciana Rodríguez Sawicki: Intra-chain Interactions Shape the Evolution of HSPB1 Chaperone Proteins
- P13 Mahta Mehdiabadi: Protein Function Prediction: A Graph-Based Approach
- P14 Marc Fornt Suñé: Self-assembling protein-based nanoparticles for biomedical and biotechnological applications
- P15 Matias Chiappinelli: Exploration of the vibrational space of tandem repeat proteins
- P16 Mehdi Davari: Machine-learning for data-driven protein engineering
- P17 Mert Cihan: Exploring the Impact of Intrinsically Disordered Regions on RNA Stability and microRNA Accessibility
- P18 Oriol Bárcenas: Structural determinants of α -synuclein binding to an inhibitory peptide studied by molecular dynamics simulations
- P19 Radovan Dvorsky: Deciphering Protein-Peptide Interactions: Unveiling the Intricacies of PRM Binding to SH3 Domains
- P20 Santiago Gómes Bergna: Analysis of tandem repeat containing proteins in Alpha- and Betabaculovirus
- P21 Stefana Njemoga: Investigating conformational space of protein tau by enhanced sampling MD simulations
- P22 Zoe Manglano Artuñedo: TAC-1: Advancing towards the development of small molecules to treat tauopathies

Poster session II

- P23 Adam Polák: Combined in silico and in tube tools to study the structure of intrinsically disordered tau proteins
- P24 Andrea Antosova: Tetrahydrofurans derivatives inhibit A β amyloid fibrillation
- P25 Andrea Bartolomé-Nafría: Effects of the ionic strength on functional amyloid fibril formation

- P26 Carla Padilla Franzotti: Analysis of the impact of disordered insertions in ensembles of structural tandem repeats
- P27 D Kreil: Systematic Comparative Functional Analysis of Protein Regions Identified by Complementary Approaches to Delineating Compositional Bias, Structural Disorder, and Low Complexity
- P28 Darius Šulskis: The role of S100A1, S100A8 and S100A9 heterodimerization in modulating their amyloid aggregation
- P29 Emanuele Zippo: Chemically-driven dynamics of disordered proteins and their condensates in simulations
- P30 Federica Quaglia: Training and collaboration opportunities in DisProt, the database of Intrinsically Disordered Proteins
- P31 Guglielmo Tedeschi: Analysis and Sampling of Molecular Simulation with extended Autoencoder architectures
- P32 Hamidreza Ghafouri: DPEsembleTools, A Python package for a comprehensive analysis of conformational ensemble of intrinsically disordered protein
- P33 Ian Czarnowski: Revealing Novel Protein-Protein Interaction Modulators of Pyruvate Kinase M2 (PKM2) Enzyme using Structure-Based Drug Design (SBDD)
- P34 Jaroslaw Chilimoniuk: imputomics 2.0: comprehensive R package for missing data imputation in proteomics data
- P35 Julius Schlenzok: UdonPred: Predicting protein disorder based on novel TriZOD scores
- P36 Kemal Demirtaş: Unrevealing residue conductors driving functional motion in proteins
- P37 Kevser Kübra Kirboğa: Identification and Analysis of Cancer-Associated and Dysfunctional Non-Globular Proteins: A Bioinformatics Approach
- P38 Lautaro Rivera: Unveiling Allosteric Modulation of 14-3-3 Proteins: Implications for Neurodegenerative Disease Research
- P39 Luis Fernando Duran Armenta: Mutations in the low-complexity domain of hnRNPA2 increase its phase separation and aggregation propensity
- P40 Marcin Grynberg: Significant combinations of amino acid pairs and triplets in low complexity regions in proteins
- P41 María Leticia Ferrelli: Study of non-globular proteins present in the Occlusion Body of *Spodoptera frugiperda* granulovirus

- P42 Mariane Gonçalves-Kulik: Evaluating the stability of polyXYs within IDRs with helical structure through Molecular Dynamics Simulations and AlphaFold Predictions
- P43 Markus Haak: Farming Protein Disorder Data with TriZOD by Leveraging NMR Spectroscopy
- P44 Martyna Podlasiak: Unravelling the interactions of Nanobodies with pathologically relevant regions of Ataxin-3
- P45 Mattia Utichi: Decoding Autophagy: An Integrative Approach to study Protein-Protein Interactions
- P46 Michal Burdukiewicz: Confident dataset generation for the specific prediction and benchmarking of client, driver and negative proteins in liquid-liquid phase separation
- P47 Nazanin Farahi: Phase-separating fusion proteins drive cancer by dysregulating transcription through ectopic condensates
- P48 Ondrej Cehlar: Insights into the Alzheimer's disease-specific pre-aggregation conformation of monomeric tau proteins using antibodies DC11 and MN423
- P49 Oscar Daniel Bello: Advances in High-Purity Extracellular Vesicle Protein Isolation for Neurological Biomarker Detection
- P50 Pinar Onal: The Impact of Intrinsically Disordered Effector Regions on the developmental functions and DNA Binding Specificity of Drosophila Transcription Factor Bicoid
- P51 Tugba Onal-Suzek: A machine learning model for discovering repurposed drug candidates to inhibit the non-globular CTNNB1 protein within the Wnt/beta-catenin pathway in cancer
- P52 Valentin Iglesias: A3D Model Organism Database (A3D-MODB): a database for proteome aggregation predictions in model organisms
- P53 Viktoriia Komarysta: Putative intrinsically disordered DNA-binding proteins in algal pyrenoids
- P54 Ximena Aixa Castro Naser: Unraveling the Role of Intrinsically Disordered Proteins in Reproductive Biology

Event Venue

The Meeting will take place at the **Onoma Hotel** on the first floor..

Address: 24 Monastiriou str. P.C. 54629, Thessaloniki, Greece

GPS coordinates: 48.156322, 17.108049

Google map link: <https://www.google.sk/maps/@40.6425014,22.9291784,17z?entry=ttu>

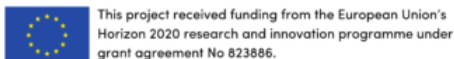
Arriving at the hotel from the Airport

Public transportation:

AIRPORT – INTERCITY BUS TERMINAL routes No 01X and 01N (Fares: 1.90 EUR single journey) to bus stop NEOS SIDIRODROMIKOS STATHMOS. From the bus stop to the hotel, it is approximately 300 m following Monastiriou Street.

Google map link: <https://maps.app.goo.gl/hYKQRkEATy378T8o9>

Co-organizing Consortia



This event is part of the activities of the COST Action ML4NGP, CA21160, supported by COST (European Cooperation in Science and Technology).