Universitat Pompeu Fabra Barcelona | Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Juan Fernández Recio
Email	juan.fernandezrecio@icvv.es
Institution	ICVV-CSIC
Website	http://www.icvv.es/english/3dbiowine
Group	Structural Bioinformatics for Wine Sciences

Project

Structural bioinformatics

Project Title:

Structural modeling of bitter taste receptors and their interactions

Keywords:

protein-protein interactions, computational docking, drug discovery, molecular modeling, taste receptors

Summary:

The interplay between genetics and environmental factors is critical in major non-communicable diseases (NCDs), which are the leading cause of death globally and a major cause of premature death. As an example, it is known that genetic variation can affect individual food preferences, which has impact on diet and health. Indeed, genetic sensitivity to bitter taste has been associated to different sensitivity to bitterness and/or linked to variable risk of alcohol dependence, obesity, nicotine dependence, longevity, miocardial infarction, or altered thyroid function (Duffy 2004; Mangold et al 2008; Campa et al 2012; Clark et al 2015). At the molecular level, bitter taste perception in humans is mediated by the 25 members of the Taste 2 receptor (TAS2R) gene family (Conte et al 2002). Each member of the TAS2R family can bind a range of compounds with different specificities, enabling the detection of tens of thousands of bitter molecules (Meyerhof et al 2010 Chem Senses). Therefore, knowing the atomic details of their binding capabilities would be important to understand the impact of these genetic variants in diet preferences and disease risk. In this project, we aim to contribute to the structural characterization of taste receptors to understand their functional mechanisms and the impact of genetic variants in health. We will model four bitter taste receptors that host variants associated to disease: TAS2R16, TAS2R38, TAS2R42, TAS2R50. Preliminary results using available models at www.gpcrdb.es show that critical residues for function gather around active site. However, these models still contain some structural errors that we will need to refine by molecular dynamics (MD). Then, we will model by docking the binding of around 100 bitter compounds to all TAS2R models and will compare the results with their known specificities (Meyerhof et al 2010). This will help to refine the modeling pipeline and will provide a theoretical framework for TAS2R binding to bitter compounds. We will also explore potential homomeric interactions of TAS2Rs by protein-protein docking in collaboration with Hugo Gutiérrez de Terán (Uppsala University). Finally, in collaboration with the groups of Masha Niv (Hebrew University of Jerusalem) and M. Purificación Fernández Zurbano (ICVV-UR), we will use our molecular models to test candidate compounds that are related to bitterness in wine, in order to build a functional model of taste perception in humans for the interpretation of genetic data affecting taste and hence diet preferences and health.

References:

Campa D, de Rango F, Carrai M, Crocco P, Montesanto A, Canzian F, Rose G, Rizzato G, Passarino G, Barale R (2012) PloS ONE 7,

e45232. Clark AA, Dotson CD, Elson AET, Voigt A, Boehm U, Meyerhof W, Nanette I. Steinle NI, Munger SD (2015) FASEB J. 29, 164-172. Conte C, Ebeling M, Marcuz A, Nef P, Andres-Barquin PJ (2002) Cytogenet Genome Res 98, 45–53. Duffy VB (2004) Appetite 43, 5-9. Mangold JE, Payne TJ, Ma JZ, Chen G, Li MD (2008) J. Med. Genet. 45, 578-582. Meyerhof F, Batram C, Kuhn C, Brockhoff A, Chudoba E, Bufe B, Appendino G, Behrens M (2010) Chem. Senses 35, 157-170.

Expected skills::

Linux, basic programming capabilities, motivation for structural interpretation of molecular mechanisms

Possibility of funding::

To be discussed

Possible continuity with PhD: :

Yes

Universitat Pompeu Fabra Barcelona

Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Jana Selent
Email	jana.selent@upf.edu
Institution	IMIM-UPF
Website	www.jana-selent.org
Group	GPCR Drug Discovery Group

Project

Structural bioinformatics

Project Title:

Deciphering the mechanism of drug action at G protein coupled receptors (GPCRs)

Keywords:

G protein-coupled receptors, molecular dynamics, data analysis, drug design

Summary:

G-protein coupled receptors (GPCRs) are the most abundant class of receptors in the human organism. They are present in almost every type of cell, and govern almost every process in the human body (i.e. cognitive and inflammatory processes or control of the cardiovascular system). Owning to their ubiquity, they are targets of more than 30% of current drugs, and every day new GPCRs are revealed to be pharmacological targets for existing diseases. GPCR drugs can either be agonist, antagonist or inverse agonists. They act by binding to the receptor and establishing transient interactions with protein residues that form the binding pocket. Those interactions alter the GPCR structure, leading to a specific downstream signalling response. However, important features responsible for a distinct drug profile (selectivity, signalling outcome, etc.) are still unclear. Uncovering those factors would provide important structural insight for further drug-design endeavours. Currently, there exist multiple GPCR structures bound to various ligands (chemicals binding to GPCRs), however a static look at ligand-receptor interactions doesn't allow to fully rationalize the signalling profile. Molecular dynamics (MD) is a novel and sophisticated technique that enables to simulate protein behaviour in a physiological environment. They offer a unique opportunity to study GPCR-ligand interactions at single atom resolution, providing insights on receptor behaviour. The development of MD techniques has been rewarded with a Nobel Prize in 2013, and the number of papers using MD is growing exponentially. In our group we have carried out a massive MD projects to unravel general principles of GPCR signalling and drug binding. With the aid of an international consortium we have simulated over 90 GPCRs crystallized with diverse ligands, amounting impressive simulation time (www.gpcrmd.org). We are looking for a motivated student that would be interested in participating in the analysis of this data. The students would be involved in analysing the generated MD data. During the project they will learn how to set up, and simulate their own biological systems. They will learn about GPCR biology, as well as about in silico drug design. To analyse the data the students will learn to write in house scripts in tcl, bash and python, as well as use several statistical methods. The student will have the opportunity to collaborate with international experts renowned in the GPCR field (members of the consortium see reference). We expect that the results of the analysis will be published in a high impact journal, and the skills acquired by the student will make him/her a valuable asset for pharma companies. The project can be extended into a PhD thesis.

References:

Rodríguez-Espigares & Torrens-Fontanals et al. GPCRmd uncovers the dynamics of the 3D-GPCRome (https://www.biorxiv.org/content/10.1101/839597v2.abstract)

Expected skills::

Experience in structural biology, python, and bash. Experience with molecular dynamics simulations is a plus. Good level of English.

Possibility of funding::

To be discussed

Possible continuity with PhD: :

Yes

Universitat Pompeu Fabra Barcelona

Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Xavier Barril
Email	xbarril@ub.edu
Institution	Universitat de Barcelona
Website	http://www.ub.edu/bl/
Group	Barril's lab

Project

Structural bioinformatics

Project Title:

Ligand optimisation for drug discovery: Use of MDmix for activity cliff prediction

Keywords:

Computer-aided drug design, structure-activity relationships, drug discovery, lead optimisation, binding free energy, molecular dynamics

Summary:

The goal of the Barril's lab is to discover bioactive molecules that bind to unexplored sites of action, exploiting novel mechanisms to achieve a therapeutic effect. To do so, we apply state of the art structure-based drug discovery methods, many of which have been developed in-house. We introduced the use of molecular dynamics with mixed solvents (MDmix) for druggability prediction,[1] as a computational counterpart of binding site detection by solvent screening.[2,3] This strategy turned out to be extremely successful and the method became widely adopted, with different adaptations (see reference [4] for a recent review). Since then, we have explored and extended the applicability of the method, describing its relationship with protein flexibility,[5] demonstrating its performance in mapping binding hot spots on protein surfaces and predicting water displaceability,[6] or as a guide in docking.[7] An open-source software was produced to help other users adopting the technique: http://mdmix.sourceforge.net Some preliminary work indicates that MDmix can also be used in predicting binding free energies of protein-ligand complexes.[7] In this project we will investigate its efficacy in predicting activity cliffs (i.e. pairs of structurally similar compounds presenting large potency difference). In medicinal chemistry, activity cliffs are crucial in systematic structure− activity relationship (SAR) analysis to identify structural modifications that determine SAR characteristics [8]. An analysis of a large set of matched molecular pairs compiled from the literature[4,5] and already available in our lab will be performed, comparing the performance of MDmix with other computational tools. This project is synergistic with other projects in our lab, and will benefit from substantial previous work and of close collaboration with other group members.

References:

J. Seco, F. J. Luque, X. Barril, Binding site detection and druggability index from first principles. J. Med. Chem. 52, 2363–71 (2009). 2. C. Mattos et al., Multiple solvent crystal structures: probing binding sites, plasticity and hydration. J. Mol. Biol. 357, 1471–82 (2006). 3. E. Liepinsh, G. Otting, Organic solvents identify specific ligand binding sites on protein surfaces. Nat. Biotechnol. 15, 264–8 (1997). 4. P. Ghanakota, H. A. Carlson, Driving Structure-Based Drug Discovery through Cosolvent Molecular Dynamics. J. Med. Chem. 59, 10383–10399 (2016). 5. D. Alvarez-Garcia, X. Barril, Relationship between Protein Flexibility and Binding: Lessons for Structure-Based Drug Design. J. Chem. Theory Comput. 10, 2608–14 (2014). 6. D. Alvarez-Garcia, X. Barril, Molecular simulations with solvent competition quantify water displaceability and provide accurate interaction maps of protein binding sites. J. Med. Chem. 57, 8530–9 (2014). 7. J. P. Arcon et al., Molecular Dynamics in Mixed Solvents Reveals Protein-Ligand Interactions, Improves Docking, and Allows Accurate Binding Free Energy Predictions. J. Chem. Inf. Model. 57, 846–863 (2017). 8. A. M. Wassermann, M. Wawer, J. Bajorath, Activity Landscape Representations for Structure−Activity Relationship Analysis. J. Med. Chem. 53, 8209–8223 (2010). 9. Y. Hu, N. Furtmann, M. Gütschow, J. Bajorath, Systematic identification and classification of three-dimensional activity cliffs. J. Chem. Inf. Model. 52, 1490–8 (2012). 10. X. Hu, Y. Hu, M. Vogt, D. Stumpfe, J. Bajorath, MMP-Cliffs: systematic identification of activity cliffs on the basis of matched molecular pairs. J. Chem. Inf. Model. 52, 1138–45 (2012).

Expected skills::

molecular dynamics, protein-ligand docking, structure-based drug discovery

Possibility of funding::

To be discussed

Possible continuity with PhD: :

To be discussed

Comments:

The group will offer a fellowship (“Beca de col·laboració”) through the Fundació Bosch i Gimpera, for a period of 6 to 12 months. The fellowship is legally capped at 617€ per month. (Subject to funds availability).

Universitat Pompeu Fabra Barcelona

Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Jordi Villà-freixa
Email	jvilla@uic.cat
Institution	Universitat Internacional de Catalunya
Website	http://mon.uvic.cat/cbbl
Group	Computational Biochemistry and Biophysics Lab

Project

Structural bioinformatics

Project Title:

Finding epitope-MHC interactions through deep learning and molecular simulations

Keywords:

Molecular simulations, MHC, Deep Learning

Summary:

Several Machine Learning methods have been established to classify antigen-MHC interactions with somehow good success. However, precise classification of the binding characteristics within those complexes is still elusive if only classification methods are used. Here we will work in a combination of state of the art deep learning tools with structure based molecular simulations.

References:

"Structure Based molecular simulations" Submitted. Martin Floor, Li Keng Jie, Luís Agulló, Jenn K. Hwang, Jordi Villà-Freixa

Expected skills::

Python, Molecular simulations, Machine Learning

Possibility of funding::

To be discussed

Possible continuity with PhD: :

Yes

Universitat Pompeu Fabra Barcelona

Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Pietro Sormanni
Email	ps589@cam.ac.uk
Institution	University of Cambridge
Website	www.ch.cam.ac.uk/chemistry-of-health/index
Group	Chemistry of Health Initiative

Project

Structural bioinformatics

Project Title:

Automated structure- and sequence-based optimisation of antibody developability potential

Keywords:

Antibody design, Biopharmacueticals developability, drug development

Summary:

Owing to their outstanding performances in molecular recognition, antibodies are extensively used in research, diagnostics, and therapeutics, with more than 90 drugs already approved in the market. However, antibody development for therapeutic applications remains a long and costly process, also because therapeutic applications often require these molecules to withstand stresses that are not present in vivo. Antibody developability is defined as the likelihood of an antibody drug candidate with suitable functionality to be developed into a manufacturable, stable, safe, and effective drug that can be formulated to high concentrations while retaining a long shelf-life. In particular, antibody developability is determined by the presence of chemical liabilities, and by key biophysical properties including thermodynamic stability and solubility. Students are invited to work at the University of Cambridge within the Chemistry of Health initiative in the Department of Chemistry to develop a computational method and associated web server for the automated prediction of mutations that improve antibody developability potential. The applicant will work in a highly multidisciplinary research team where computational method development and corresponding experimental validation are carried out side-by-side. The outcome of this research will have an impact in the emerging field of computational antibody design, and it will improve and accelerate the 'hit-optimisation' step in biopharmaceutical pipelines.

References:

Sormanni, P., Aprile, F. A. & Vendruscolo, M. Third generation antibody discovery methods: in silico rational design. Chem. Soc. Rev. 47, 9137–9157 (2018)

Expected skills::

Python programming language is required. Beneficial: web server development and some knowledge of structural biology.

Possibility of funding::

To be discussed

Possible continuity with PhD: :

To be discussed

Universitat Pompeu Fabra Barcelona

Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Oriol Gallego
Email	oriol.gallego@upf.edu
Institution	DCEXS-UPF
Website	www.gallegolab.org
Group	Live-cell structural biology

Project

Structural bioinformatics

Project Title:

Integrative Structural biology of exocytosis

Keywords:

Integrative Modeling Platform, Python, Integrative structural biology, Exocytosis

Summary:

This project aims to push the limits of integrative structural biology to resolve fundamental problems in cell biology. The student is expected to develop computational tools that strengthen a multidisciplinary team of bioinformaticians, physicists and experimentalists and that provide us with unique capabilities to resolve molecular structures. The biological question that we would like to address is exocytosis, a cellular process responsible to deliver biomolecules to the plasma membrane and extracellular space that is conserved in all eukaryotic cells. Exocytosis controls the growth of cell surface and it is directly coupled with the cell cycle and viability. However, the mechanism that regulates exocytosis is a central question in cell biology that could not be answered yet. Decades of research and the latest developments in gene editing, molecular biology and cryoEM have provided fundamental insight about exocytosis, but failed to resolve the molecular details that control this essential process. The complexity of the protein machinery involved and fast cycles of assembly-activity-disassembly have prevented full understanding of exocytosis. Recently, we developed a new method of fluorescent microscopy capable of resolving the 3D architecture of protein assemblies directly in living cells. Using this approach and computational integration of structural data we reconstructed de novo the exocytic machinery at the nanometre scale (Picco et al 2017 Cell). However, high-resolution structures and conformational dynamics necessary to understand the mechanism of exocytosis remain elusive. We offer a position for a Master student to push further integrative structural biology and that, together with our collaborators (D. Davos, CABD, Sevilla; J. Ries, EMBL, Heidelberg), works to develop the computational tools that can overcome current technical limitations. The student will use Python and the Integrative Modeling Platform (IMP, developed in A. Sali's lab at UCSF) to integrate in vitro and in cellulo datasets (i.e. live-cell imaging, cryo-EM, homology modeling, super resolution microscopy...) and to reconstruct the high-resolution structure of the supra-assembly that controls exocytosis. The student will team-up with a PhD student from our lab to explore new strategies involving Monte Carlo sampling methods and coarse-grained modeling among others. Overall, he/she is expected to contribute to a larger project aiming to resolve the mechanism of exocytosis.

References:

Picco, A., Irastorza-Azcarate, I., Specht, T., B.ke, D., Pazos, I., Rivier-Cordey, A-S., Devos*, D.P., Kaksonen*, M., Gallego*†, O., (2017) "The in vivo architecture of the exocyst provides structural basis for exocytosis." Cell 168, 400-412.e18.

Expected skills::

Expertise with Python is required. Knowledge on structural modeling or molecular dynamics will be a plus.

Possibility of funding::

Yes

Possible continuity with PhD: :

To be discussed

Comments:

High motivation for learning, team work and pushing the project forward is a must.

Master project 2020-2021

Personal Information

Supervisor	Victor Guallar
Email	victor.guallar@bsc.es
Institution	Barcelona Supercomputing Cener
Website	https://www.bsc.es/discover-bsc/organisation/scientific-structure/electronic-and-atomic-protein-modeling-eapm
Group	EAPM

Project

Structural bioinformatics

Project Title:

ML victual Screening

Keywords:

Docking, coronavirus, PELE, deep learning

Summary:

We aim at developing a Deep Docking (DD) approach to screen billions of compounds in a fast manner. It will be integrated in a hierarchical pipeline combining commercial docking techniques and our Monte Carlo scheme (PELE). The application field will include real targets a current project to screen for pancoronavirus (polypharmacology) inhibitors.

References:

https://doi.org/10.1002/minf.202000028,

Expected skills::

Previous knowledge on docking and machine learning will be a plius.

Possibility of funding::

Yes

Possible continuity with PhD: :

Yes

Universitat Pompeu Fabra Barcelona | Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Miguel Angel Santos Santos
Email	msantoss@santpau.cat
Institution	departamento de Neurologia del Hospital Sant Pau
Website	https://santpaumemoryunit.com/
Group	Unidad de Memoria HSP

Project

Structural bioinformatics

Project Title:

Multi-modal neuroimaging biomarker identification for the improvement of diagnosis and disease monitoring in primary progressive aphasia.

Keywords:

primary-progressive-aphasia, neurodegeneration, voxel-based-morphometry, diffusion-tensor-imaging, resting-state-functional-mri

Summary:

The main objective of this project is to study the clinical utility (to improve diagnosis and track disease progression) of various neuroimaging biomarkers. The secondary but equally fascinating objective is to study the neuroanatomical basis of language using primary progressive aphasia as a “lesion-model.” We hypothesize that specific neuroimaging biomarkers including structural, diffusion tensor imaging, resting-state functional MRI, and fluorodeoxyglucose positron emission tomography provide measures of brain damage that reflect differentiable pathophysiologic mechanisms. More specifically, we hypothesize diffusion tensor imaging and functional MRI will be able to capture brain damage at early disease stages when volumetric atrophy is not apparent. To this end, we will preprocess and analyze the before-mentioned neuroimaging scans and then apply various statistical methodologies to compare across diagnostic groups and study their association with other biologic and language measures. The data for this study originates from a multitude of past and ongoing projects based at the Hospital Sant Pau Memory Unit.

References:

Classification of primary progressive aphasia and its variants. Neurology. 2011 Mar 15;76(11):1006-14. doi: 10.1212/WNL.0b013e31821103e6. Epub 2011 Feb 16. Gorno-Tempini ML1, Hillis AE, Weintraub S, Kertesz A, Mendez M, Cappa SF, Ogar JM, Rohrer JD, Black S, Boeve BF, Manes F, Dronkers NF, Vandenberghe R, Rascovsky K, Patterson K, Miller BL, Knopman DS, Hodges JR, Mesulam MM, Grossman M. Features of Patients With Nonfluent/Agrammatic Primary Progressive Aphasia With Underlying Progressive Supranuclear Palsy Pathology or Corticobasal Degeneration. JAMA Neurol. 2016 Jun 1;73(6):733-42. doi: 10.1001/jamaneurol.2016.0412. Santos-Santos MA1, Mandelli ML1, Binney RJ2, Ogar J1, Wilson SM3, Henry ML4, Hubbard HI1, Meese M1, Attygalle S1, Rosenberg L1, Pakvasa M1, Trojanowski JQ5, Grinberg LT6, Rosen H1, Boxer AL1, Miller BL1, Seeley WW6, Gorno-Tempini ML1. Rates of Amyloid Imaging Positivity in Patients With Primary Progressive Aphasia. JAMA Neurol. 2018 Mar 1;75(3):342-352. doi: 10.1001/jamaneurol.2017.4309. Santos-Santos MA1,2,3,4, Rabinovici GD1,5, Iaccarino L1,6, Ayakta N1,5, Tammewar G1,5, Lobach I7, Henry ML8, Hubbard I1, Mandelli ML1, Spinelli E1,6, Miller ZA1, Pressman PS1,9, O'Neil JP10, Ghosh P1, Lazaris A1, Meyer M1, Watson C1, Yoon SJ1,11, Rosen HJ1, Grinberg L1,12, Seeley WW1,12, Miller BL1, Jagust WJ5,10, Gorno-Tempini ML1. Functional Connectivity is Reduced in Early-stage Primary Progressive Aphasia When Atrophy is not Prominent. Alzheimer Dis Assoc Disord. 2017 Apr-Jun;31(2):101-106. doi: 10.1097/WAD.0000000000000193. Bonakdarpour B1, Rogalski EJ, Wang A, Sridhar J, Mesulam MM, Hurley RS. The Sant Pau Initiative on Neurodegeneration (SPIN) cohort: A data set for biomarker discovery and validation in neurodegenerative disorders. Alzheimers Dement (N Y). 2019 Oct 14;5:597-609. doi: 10.1016/j.trci.2019.09.005. eCollection 2019. Alcolea D1,2, Clarimón J1,2, Carmona-Iragui M1,2,3, Illán-Gala I1,2, Morenas-Rodríguez E1,2, Barroeta I1,2, Ribosa-Nogué R1,2, Sala I1,2, Sánchez-Saudinós MB1,2, Videla L1,2,3, Subirana A1,2, Benejam B1,2,3, Valldeneu S1,2, Fernández S1,2,3, Estellés T1,2, Altuna M1,2, Santos-Santos M1,2, García-Losada L1,2, Bejanin A1,2, Pegueroles J1,2, Montal V1,2, Vilaplana E1,2, Belbin O1,2, Dols-Icardo O1,2, Sirisi S1,2, Querol-Vilaseca M1,2, Cervera-Carles L1,2, Muñoz L1,2, Núñez R1,2, Torres S1,2, Camacho MV4, Carrió I4, Giménez S5, Delaby C1,6, Rojas-Garcia R7,8, Turon-Sans J7,8, Pagonabarraga J2,9, Jiménez A10, Blesa R1,2, Fortea J1,2,3, Lleó A1,2.

Expected skills::

Basic knowledge in statistics, basic hands on experience with programming and imaging analysis software

Possibility of funding::

To be discussed

Possible continuity with PhD: :

To be discussed

Master project 2020-2021

Personal Information

Supervisor	Magdalena Scharf
Email	magdalena.scharf@pharmazie.uni-marburg.de
Institution	Philipps-University Marburg
Website	www.kolblab.org
Group	Peter Kolb

Project

Structural bioinformatics

Project Title:

In silico prediction of novel ligands for a chemokine receptor

Keywords:

Chemoinformatics, homology modelling, docking calculations, GPCRs, computer-aided drug design

Summary:

Your project will evolve around one member of the chemokine receptors. The ultimate goal is to find novel ligands that modulate the activity of this target. Since there are no crystal structures available for the target receptor, the first step is to prepare a three dimensional structure of it by homology modelling. This model will then be used in docking calculations to screen a large library of molecules against it.

Expected skills::

basic chemical knowledge to evaluate protein-ligand interactions

Possibility of funding::

To be discussed

Possible continuity with PhD: :

To be discussed

Universitat Pompeu Fabra Barcelona | Master in Bioinformatics for Health Sciences

Master project 2020-2021

Personal Information

Supervisor	Juan Cortés
Email	juan.cortes@laas.fr
Institution	LAAS-CNRS
Website	https://www.laas.fr/public/en
Group	Robotics and Interactions

Project

Structural bioinformatics

Project Title:

Studying how protein evolution (re-)shapes local structural preferences

Keywords:

Protein evolution, local structural preferences, structural database, protein flexibility, intrinsically disordered proteins

Summary:

The structural and dynamical properties of a protein are largely determined by its sequence, and strongly influence its function. To maintain protein functions during evolution, these properties must be robust to sequence variations (for example mutations). Nevertheless, structural/dynamical changes may be beneficial, for instance if they improve the way the protein functions or generate functional innovations. Understanding this trade-off between structural stability and malleability in evolution is of great interest for fundamental biology and for protein design. In recent years, we have contributed to better characterize the sequence-structure/dynamics relationship, particularly in the context of highly-flexible protein regions. We have constructed an extensive database of small fragments involving three consecutive residues (called tripeptides) extracted from coil regions in experimentally-determined protein structures. We have shown that this database is useful to accurately sample the conformational variability of protein loops [1] and intrinsically disordered proteins (IDPs) [2]. We have also developed an approach to characterize the structural preferences of each tripeptide sequence, and we have defined metrics to quantify the structural differences between different sequences. The goal of this project is to investigate how mutations occurring during evolution affect protein local structural propensities. To do so, the candidate will exploit our structural database and use the developed approaches and metrics. Starting from a set of currently observed proteins, s/he will infer the evolutionary history relating them and will quantify the correlation between changes in sequence and changes in local structures. We will consider several protein families. The analysis will be particularly focused on proteins in which (local or global) flexibility plays essential roles, such as antibodies, enzymes and IDPs.

References:

[1] Barozet, A., Molloy, K., Vaisset, M., Simeon, T., Cortés, J. (2020) A reinforcement learning approach to enhance protein loop sampling. Bioinformatics, 36(4):1099–1106 [2] Estana, A., Sibille, N., Delaforge, E., Vaisset, M., Cortés, J., Bernado, P. (2019) Realistic ensemble models of intrinsically disordered proteins using a structure-encoding coil database. Structure, 27(2):381-391.E2

Expected skills::

The candidate should have: background in bioinformatics and structural biology, good programming skills, familiarity with Linux, and knowledge about databases. Teamwork skills are also essential for the achievement of the project.

Possibility of funding::

Yes

Possible continuity with PhD: :

To be discussed

Comments:

The project will be co-supervised by Elodie Laine (Sorbonne Université, Paris)