PhD student in Bioinformatics / Data Science
(with application to immunology)

UIC, Barcelona, Spain

Objective

Applications are invited for the position of a PhD student, working under the supervision of Bernhard Knapp in the Department of Basic Sciences, Faculty of Medicine and Health Sciences, UIC University, Barcelona, Spain.

The work group of Bernhard Knapp (https://www.uic.es/bioinformatics) focuses on computational simulations of the (human) immune system. We provide novel insight in fundamental processes of the immune system by means of quantitative analysis and predictive models. More specifically the work focuses on how T-cell receptors (TCRs) recognise Major Histocompatibility Complex (MHC) bound peptides in different health and disease conditions as for example allergies, cancer, autoimmune diseases, or infections. Apart from TCR/peptide/MHC interactions also antibody/antigen interactions are a key research interest of the group.
We apply a wide range of computational methods for (biomedical) challenges. Examples of work include: by far the largest dataset of T-cell receptor recognition processes [1], comparison of the molecular design [2] and instability [3] of antibodies with T-cell receptors, obtaining novel insights in peptide/MHC detachment processes [4], giving accurate and reproducible predictions of peptide/MHC binding affinities [5], optimising binding affinities using Genetic Algorithms [6], and investigating protein/protein interactions [7], and development of tools [8]. Highly parallel calculations are carried out on different supercomputers.

Requirements

<u>Must criteria:</u>

- **BSc** and **MSc** in a quantitative field (e.g. computer science, statistics, math, (bio)physics, bioinformatics).
  The average of all marks of the bachelor and master should be better than **0.75** (on a normalised scale of 0 (worst) to 1 (best)). Please separately state your average mark for your bachelor and master in your CV (and show supporting documents as an attachment). Applications **not containing this information for the BSc or the MSc cannot be considered**. Unfortunately, also applicants without a master degree (or who already hold a PhD) cannot be considered
- The **MSc** must have been obtained after the 1 November 2016
- **Broad experience** and fluency in several **programming** languages e.g. Python
- Strong **statistics** knowledge (univariate, multivariate)
- Strong **analytical** skills, **creativity**, **curiosity**, **enthusiasm**, and ability to work in a **team**
- **Excellent** command of the **English** language

Ideal criteria:

- Basic knowledge of molecular biology, genetics, immunology is an advantage but not required
- Previous experience with Monte Carlo simulations is an advantage
- Knowlegde of data science techniques: Regression (e.g. ANN, Lasso, Ridge) , classification (e.g. SVM, KNN), clustering (e.g. k-means, hierarchical), heuristic optimisation (e.g. genetic algorithms), parallel computing (e.g. MPI, pySpark), SQL, noSQL is an advantage
- Knowledge of common bioinformatics techniques as Immunoinformatics, B- and T-cell epitope predictions, molecular dynamics simulations, free energy predictions, protein/ligand and protein/protein docking, virtual screening, protein structure prediction, sequence alignments, tree building is an advantage
- Knowledge of LINUX operating systems and command line operations is an advantage
- Experience with the use of high performance computing clusters is an advantage
- Previous teaching experience in English and/or Spanish is an advantage

Offer

- The PhD position is funded by the EXCELENCIA program of the Spanish ministry of economics, industry and competitiveness
- The successful candidate will start his/her work in an exciting project in the field of computational simulations of the immune system
- Access to high performance computers is available
- Funding to go to conferences is available
- The successful candidate can start immediately

Application documents (as one pdf file)

1) CV
2) Publication list (if existing) including journal impact factors
3) Letter of motivation (short, max 1 page) answering to each of the "must-criteria" as stated above. Especially state the marks average of your BSc and MSc.
4) Contact details of 2-3 references

Before you apply please keep in mind that common reasons for not being invited for an interview are: (1) not fulfilling one or multiple of the "must criteria", (2) not following the application guidelines, (3) generic copy & paste applications, (4) low level of English, and (5) or missing of the deadline.

In a first step the applicants are encouraged to pass a quick test in order to decode the application email. Please apply (in English) to:

[1][2][3][4][5][6]@uic.es

where answering the questions underneath decodes the characters 1 to 6 above (without the square brackets and lower case). (Hint: You know that you have decoded correctly if the email address has to do with the name of the head of the work group)

[1] A variable that can only be true or false (only the first character)

[2] Simplest and most common clustering algorithm (only the first character before the "-")

[3] A pointer to nothing in C/C++ (only the first character; not "void")

[4] Building blocks of proteins (only the first character)

[5] Probability for a given statistical model that (if the null hypothesis is true) that the statistical summary (e.g. mean difference between two compared groups) would be the same as or more extreme than the actual observed results (only the first character before the "-")

[6] import numpy as np

idx = np.argmax([99, 9, 1, 98, 7])

print("python!"[idx])

Applications will be reviewed continuously on a 'first come first served' basis and as soon as an appropriate candidate is found the position will be closed. However, the call will close latest by the 9th of September 2018.

Reference List

1. Knapp, B.; Dunbar, J.; Deane, C. M. Large Scale Characterization of the LC13 TCR and HLA-B8 Structural Landscape in Reaction to 172 Altered Peptide Ligands: A Molecular Dynamics Simulation Study. *PLoS Comput Biol* **2014,** *10* (8), e1003748.
2. Dunbar, J.; Knapp, B.; Fuchs, A.; Shi, J.; Deane, C. M. Examining Variable Domain Orientations in Antigen Receptors Gives Insight into TCR-Like Antibody Design. *PLoS Comput Biol* **2014,** *10* (9), e1003852.
3. Knapp, B.; Dunbar, J.; Alcala, M.; Deane, C. M. Variable Regions of Antibodies and T-cell Receptors may not be Sufficient in Molecular Simulations Investigating Binding. *J. Chem. Theory Comput.* **2017,** *accepted*.
4. Knapp, B.; Demharter, S.; Deane, C. M.; Minary, P. Exploring peptide/MHC detachment processes using Hierarchical Natural Move Monte Carlo. *Bioinformatics* **2016,** *32* (2), 181-186.
5. Wan, S.; Knapp, B.; Wright, D.; Deane, C.; Coveney, P. V. Rapid, Precise and Reproducible Prediction of Peptide-MHC Binding Affinities from Molecular Dynamics that Correlate Well with Experiment. *J. Chem. Theory Comput.* **2015,** *11* (7), 3346-3356.
6. Knapp, B.; Giczi, V.; Ribarics, R.; Schreiner, W. PeptX: Using Genetic Algorithms to optimize peptides for MHC binding. *BMC. Bioinformatics.* **2011,** *12*, 241.
7. Esmaielbeiki, R.; Krawczyk, K.; Knapp, B.; Nebel, J.; Deane, C. M. Progress and Challenges in Predicting Protein-Protein Interfaces. *Brief Bioinform* **2016,** *17* (1), 117-31.
8. Knapp, B.; Alcala, M.; Zhang, H.; West, C.; van der Merwe, P. A.; Deane, C. M. pyHVis3D: Visualising Molecular Simulation deduced H-bond networks in 3D: Application to T-cell receptor interactions. *Bioinformatics* **2018**.