PhD position: Development of Computational Protein Design Tools

We are looking for a highly motivated candidate for a three year PhD to work at the intersection of Discrete Optimization and Computational Structural Biology, on models and algorithms for Computation Protein Design (CPD). CPD's aim is to rationally design and optimize proteins in order to define, enhance, or modify the properties of enzymes to create new bio-catalysts. These may have important uses in the synthesis of other valuable molecules (food, drugs, cosmetics...) or as crucial actors in renewable bio-energy processes. As proteins, these enzymes are also eco-friendly (degradable, usable at ambient pression/temperature).

Various algorithms and models have already been defined to optimize protein designs, with different targets but all are faced with the huge discrete space defined by protein sequences and their 3D structures, defining NP-hard problems. The work will take place in an ongoing project involving a team of structural biologists specialized in molecular modeling from LISBP (INRA, CNRS, Toulouse) and computer scientists specialized in discrete optimization from INRA, Toulouse. Together, we have recently shown that recent discrete exact optimization techniques introduced to solve Weighted Constraint Satisfaction problems (or Cost Function Networks) could speed-up existing exact approaches to CPD by several orders of magnitude [1,2]. The aim of the PhD will be to bring these improvements to a new level by tackling more complex protein design models, incorporating local or global backbone flexibility as well as side-chain flexibility, possibly with new criteria mixing stability and affinity, positive and negative design targets, with an ultimate practical goal of producing an ensemble of sequences to test for activity. A computing cluster with thousands of cores can be accessed at INRA Toulouse.

Duties and eligibility: Motivated candidates with an academic degree at the Masters level in either

- Structural Bioinformatics/Molecular Modeling of proteins with a strong taste for algorithms, programming and mathematics or
- Computer Science/Discrete optimization with a strong taste for applications in Biology

are invited to apply for this position. A strong interest in both Biology and Computer Science is absolutely required. Because of the interdisciplinary nature of the project, candidates will be evaluated on a case-by-case basis and interested candidates with a background in bioinformatics with an emphasis in algorithms will also be considered. The work will include algorithm design, implementation and testing on actual protein designs, protein flexibility modeling as well as preparation of protein designs by different methods and writing of research papers. Excellent programming skills, ideally in efficient compiled langages such as C/C++, will be extremely beneficial. A good knowledge of spoken and written English is an highly valuable asset. A good knowledge of spoken French is also appreciated. If needed, training can be provided.

A PhD grant is provided for 3 years, at a current level of 1,757€/month (gross salary). Applications should include a CV, a brief description of research interests and past research experience, copies of exam-certificates and grades, a copy of your Master's thesis (or a summary) and a cover letter motivating why you are interested in the position. The candidates are encouraged to provide contact information of at most two reference persons. The evaluation of candidates will begin on June 15, 2015 and will continue until the position is filled.

Further information can be obtained from Sophie.Barbe (Sophie.Barbe@insa-toulouse.fr) and Thomas Schiex (Thomas.Schiex@toulouse.inra.fr). An overview of the recent obtained results can be found in the following references:

[1] A New Framework for Computational Protein Design through Cost Function Network Optimization
S. Traoré, D. Allouche, I. André, S. de Givry, G. Katsirelos, T. Schiex, S. Barbe, Bioinformatics, 2013.
http://bioinformatics.oxfordjournals.org/cgi/content/abstract/btt374

[2] Computational Protein Design as an Optimization Problem. D. Allouche, J. Davies, S. de Givry, G. Katsirelos, T. Schiex, S. Traoré, I. André, S. Barbe, S. Prestwich, B. O'Sullivan. Artificial Intelligence Journal, March 2014.
http://www.sciencedirect.com/science/article/pii/S0004370214000332

and related publications on the biological and optimization sides can be accessed below:

[A] Insights into lid movements of *Burkholderia cepacia* lipase inferred from molecular dynamics simulations. S. Barbe, V. Lafaquiere, D. Guieysse, P. Monsan, M. Remaud-Simeon, I. André. Proteins. 77(3):509-523, 2009.
http://www.ncbi.nlm.nih.gov/pubmed/23842814

[B] Soft Arc Consistency Revisited. MC. Cooper, S. de Givry, M. Sanchez, T. Schiex, M. Zytnicki, T. Werner. Artificial Intelligence, 2010.
http://www.sciencedirect.com/science/article/pii/S0004370210000147

Sophie Barbe
Laboratoire d'Ingénierie des Systèmes Biologiques et des Procédés
INSA de Toulouse (INRA 792 - CNRS 5504)
135 avenue de Rangueil
31077 Toulouse CEDEX 04
Sophie.Barbe@insa-toulouse.fr
Phone: (+33) 5.61.55.99.63

Thomas Schiex
Laboratoire de Mathématiques et Informatiques Appliquées de Toulouse
INRA de Toulouse (INRA 875)
Chemin de Borde Rouge - Auzeville, CS 52627
31326 Castanet Tolosan Cedex, France
Thomas.Schiex@toulouse.inra.fr
Phone: (+33) 5.61.28.54.28