2018-00472 - Post-Doc Position: Graph representation of protein 3D conformations for docking by combinatorial assembly of fragments [S]

Context et atouts du poste

Our research combines aspects of various disciplines: biophysics, structural biology, and algorithmic related to 3D coordinates, graphs, clustering and machine learning.

We have recently developed an original methodology to build atomic representations of single-stranded nucleic acids bound to a protein, using a fragment-based approach [cf. ref]. The method is uniquely capable of providing accurate structural model from only protein structure and RNA sequence. The goal of this project is to generalize this method towards proteins of approximately known structure from homology. To account for its flexibility, we will represent the protein as a graph of possible local conformations at different scales. This development will be based on molecular docking tools existing and/or developed in the CAPSID team.


Mission confiée

CONTEXT

Many protein structures have been determined experimentally and similar sequences tend to adopt similar 3D conformations, especially in conserved domains. While enabling the use of homology modelling, this still leaves room for a lot of uncertainty and variability in the protein conformations, at different scales: Several domains in a protein are linked by flexible loops and can be differently oriented; They contain themselves flexible loops with multiple possible backbone conformations; All residues can adopt different side-chain orientations.

In classical docking (i.e. modeling of a molecular assembly from the structure of its constituents), an ensemble of protein conformations can be sampled in order to increase the chances get close enough to the protein bound form. But this requires a combinatorial sampling of each flexibility scale and each protein region, and can therefore explore only a very limited part of the conformational space. In our fragment-based approach, the conformations of each binding site of the receptor can be sampled independently, and their compatibility assessed during the linear fragment assembly process.

Principales activités

Project description

We will exploit this linearity property on an archetypal domain: the RNA-recognition motif (RRM). While the RRM is estimated to be present in 1-2% of the human genome, its RNA-recognition code has not been fully cracked yet. Using the wealth of existing knowledge on RRM atomic structures, local 3D connectivity patterns between RRM regions and between RRM and ssRNA fragments will be identified. These patterns will then be integrated into the construction of a global connectivity graph of ssRNA fragments (using the approach previously developed in our group). This graph will provide an atomic-level description of the RRM-ssRNA interaction.

We will first use the approximation that loops are independent of each other, and that only spatially close rotamers are interdependent. Second, we will add as a constraints the knowledge of one or two key amino-acid - nucleotide contacts [S]. Rotamers dependencies and key contacts will be

Informations générales

- Thème/Domaine : Modélisation et commande pour le vivant
- Ville : Villers-lès-Nancy
- Centre Inria : CRI Nancy - Grand Est
- Date prise de fonction souhaitée : 01-09-2018
- Durée de contrat : 1 an, 4 mois
- Date limite pour postuler : 06-06-2018

Contacts

- Equipe Inria : CAPSID
- Recruteur : Chauvot De Beauchene Isaure / isaure.chauvot-de-beauchene@loria.fr

L’essentiel pour réussir

How to apply

Upload your CV on jobs.inria.fr; this should be a pdf file of at most 2Mo.

In addition, send the following documents to both contacts in a single pdf or ZIP file:

CV including a description of your research activities (2 pages max) and a short description of what you consider to be your best contributions and why (1 page max and 3 contributions max); the contributions could be theoretical or practical. Web links to the contributions should be provided. Include also a brief description of your scientific and career projects, and your scientific positioning regarding the proposed subject.

The report(s) from your PhD external reviewer(s), if applicable.

If you haven’t defended yet, the list of expected members of your PhD committee (if known) and the expected date of defense (the defense, not the manuscript submission).

In addition, at least one recommendation letter from your PhD advisor should be sent directly by their author(s) to both contacts.

Applications are to be sent as soon as possible.

Conditions pour postuler

The campaign is not open to local students who have not done any significant mobility.

Sécurité défense : Ce poste est susceptible d’être affecté dans une zone à régime restrictif (ZRR), telle que définie...
identified by (co)-evolution in protein sequence alignments [2]. In addition, new sequence-based predictors will be developed by machine learning applied on improved multiple sequence alignments and HMM approaches, in collaboration with the lab of Dr. Wim Vranken at the Vrije Universiteit Brussels, Belgium [7]. If the precision obtained previously is not satisfying enough, we will apply data-driven docking: (simulated) low resolution data on the geometry of the complex, that could be obtained experimentally (SAXS, NMR), will be added as additional constraints.

**Compétences**

The project is interdisciplinary: the day-to-day work involves a lot of programming on atomic representations of proteins and nucleic acids.

Candidates must have (or be about to get) a PhD degree in any of the relevant disciplines: computer science, (bio-)physics, bio-informatics or structural biology.

Strong programming skills (preferentially Python and/or C++) are required. Skills in discrete mathematics, statistics and/or knowledge of molecular structures are very desirable.

Candidates must be fluent either in French or in English.

**Avantages sociaux**

- Subsidised catering service
- Partially-reimbursed public transport
- French courses

**Rémunération**

Salary: 2653€ gross/month

---

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.