

<b>Títol</b>	Optimització de PELE per explorar conformacions de regions flexibles sense estructura secundaria
<b>Title</b>	Optimization of PELE to explore free loop conformations

<b>Directors</b>	Victor Guallar, Pablo Navarro
<b>Departament</b>	BSC, Nostrum Biodiscovery
<b>Universitat</b>	-
<b>Tutor<sup>1</sup></b>	Iberio de Pinho Ribeiro

<b>Curs</b>	2023-24
-------------	---------

<b>Summary<sup>2</sup></b>	
The Protein Energy Landscape Exploration algorithm (PELE) is a stochastic Monte Carlo method developed to explore the complex free energy landscapes associated with molecular recognition processes (e.g. protein-ligand or RNA-ligand interactions). It leverages the efficiency of C++ and advanced algorithms to explore the degrees of freedom of these systems. A remaining challenge is the study of disordered regions in proteins or RNA. This work will involve the optimization of a protocol using an inverse kinematic solver and advanced exploration/relaxation algorithms to explore the conformational space of regions that lack a regular secondary structure.	
<b>Keywords<sup>3</sup></b>	Molecular modelling, Monte Carlo, C++, algorithms

<b>Breu descripció del projecte<sup>4</sup></b>
<p>PELE és un algoritme estocàstic de Monte Carlo que serveix per estudiar interaccions entre macro-molècules com proteïnes o RNA amb lligands o altres proteïnes. L'algoritme és particularment adient per trobar poses de <i>binding o pathways</i> d'entrada i sortida d'una cavitat entre d'altres. L'èxit de la simulació depèn en gran mesura de l'habilitat de PELE aproximant els graus de llibertat del sistema real. En aquest sentit, els receptors amb regions sense una estructura secundaria clara presenten un repte per l'algoritme. Aquestes regions presenten una gran flexibilitat i una <i>varietat conformacional</i> que no és capturada pels algoritmes originals que pertorben el receptor (basats en modes normals de la proteïna) [1]. Per això es va implementar un nou algoritme anomenat <i>loop closure</i>, capaç de proposar noves conformacions per aquestes regions.</p> <p>L'algoritme, originalment concebut per estudiar <i>folding</i> de RNA [2-3], hauria de ser adaptat per l'estudi de proteïnes. A més el protocol (conjunt d'estratègies per fer servir <i>loop closure</i> i altres algoritmes per estudiar l'espai conformacional) hauria de ser optimitzat. Un cop tinguem una primera versió d'un protocol millorat es podria comparar els resultats de PELE amb els d'altres mètodes com Dinàmica Molecular. Al final del treball l'alumne haurà guanyat experiència en diferents àrees de coneixement: programació orientada a objectes (C++), desenvolupament e implementació d'algoritmes, anàlisi i visualització de simulacions avançat (Python, Pymol, ús de collective variables), modelització molecular, comunicació de resultats a la resta de l'equip...</p> <p>[1]: Borrelli, K. W., Vitalis, A., Alcantara, R. &amp; Guallar, V. PELE: Protein energy landscape exploration. A novel Monte Carlo based technique. <i>Journal of Chemical Theory and Computation</i> 1, 1304 1311 (2005)</p> <p>[2]: Mak, C. H., Chung, W. Y. &amp; Markovskiy, N. D. RNA conformational sampling: II. arbitrary length multinucleotide loop closure. <i>Journal of Chemical Theory and Computation</i>, 1198 1207 (2011).</p> <p>[3]: Mak, C. H. Loops MC: An all-atom Monte Carlo simulation program for RNAs based on inverse kinematic loop closure. <i>Molecular Simulation</i> 37, 537 556 (2011).</p>

<sup>1</sup>Si el director no és un professor de la UB o de la UPC, caldrà assignar un tutor del TFM que designarà la Comissió Coordinadora del Màster.

<sup>2</sup>Aquest "summary" és el que apareixerà a la futura pàgina web dedicada al TFM. Procureu que sigui concís i entenedor (màx. 10 línies).

<sup>3</sup>Aquestes "keywords" no només són les que apareixeran al web sinó que ajudaran la Comissió Coordinadora del Màster a assignar el projecte a un àrea concreta.

<sup>4</sup>Procureu ser concisos però proporcioneu prou informació per tal que l'estudiant i la Comissió Coordinadora del Màster es facin una idea prou acurada de en què consistirà el treball. Indiqueu 3-6 publicacions de referència en la descripció del projecte per donar una idea dels fonament, metodologia, objectius, etc.

**Competències addicionals<sup>5</sup> (opcional)**

Tasques a desenvolupar <sup>6</sup>		Cronograma (setmanes)																	
Tasca	Descripció	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18
T01	Introducció i capacitació (aprox. 60 h)	X	X	X															
T02	Adaptació de <i>loop closure</i> a proteïnes (aprox. 80h)			X	X	X	X												
T03	Optimització del protocol (aprox. 150h)						X	X	X	X	X	X							
T04	Validació (amb MD) (aprox. 80h)										X	X	X	X					
T05	Redacció de la memòria i preparar l'exposició (aprox. 80h)												X	X	X	X	X	X	X

**Observacions i comentaris**

Per a la realització del treball es preveu una dedicació d'unes quatre hores diàries durant cinc dies a la setmana, amb un horari flexible i possibilitat de teletreballar per fer el desenvolupament del projecte compatible amb les classes del màster.

Signatura (el director del TFM)

Signatura (el tutor del TFM, si s'escau)

<sup>4</sup> Enumereu breument qualsevol competència addicional a les competències genèriques enumerades en el Pla Docent del TFM (opcional).

<sup>5</sup> Feu servir només les línies que calgui. Escolliu-les de manera que donin una idea aproximada de en què consistirà el treball i la seva distribució temporal.