INTERNATIONAL JOURNAL FOR INNOVATION EDUCATION AND RESEARCH
ISSN: 2411-2933

Journal URL: https://scholarsjournal.net/index.php/ijier

DOI: https://doi.org/10.31686/ijier.vol0.iss11.3967

ISP - IFRO SLICE PROTEIN: automatizing the process of protein fragments extraction

Dyuly Kelry Costa Guimarães, Micaely Raira Linhares de Araújo, Kethelin Vitória Zaire dos Santos, Henrique Lopes de Barros, Marcio Rodrigues Miranda, Kaio Alexandre da Silva

Abstract

Macromolecules are proteins formed by amino acids joined by peptide bonds, and can be described in different structural levels: primary, secondary, tertiary and quaternary. It's possible to extract a small part of the three dimensional structure of the protein to be used as a ligand. However, the extraction of fragments by experimental methods is expensive and time-consuming. In this context, the development of a web service to extract fragments of three-dimensional proteins makes the process more assertive and less costly. The methods used for the development of the protein slicer web service were the Python programming language, and the Javascript, PHP and HTML languages are being used. And for the testing of the system, three-dimensional structures of proteins present in the RCSB Protein Data Bank (RCSB PDB) were used.

Keywords: three-dimensional structures; ligands; polypeptides;

Introduction

Macromolecules are proteins formed by amino acids joined by peptide bonds, your function is defined by the arrangement of atoms present in your three-dimensional structure (WOLYNES, 2015). Proteins can be described in different complexity levels, in a kind of a conceptual hierarchy arrangement (NELSON *et al*., 2011), usually defined in primary, secondary, tertiary and quaternary.

The primary structure is represented by the sequence of amino acids joined by peptide bonds. The secondary structure of the protein is composed of stabilized spatial arrangements of amino acids, the two most common arrangements being the alpha-helix (α) and beta-sheet (β). The tertiary structure is the protein structure entanglement, and the quaternary structure is the spatial arrangement due to the binding of more than one protein chain (NELSON et al., 2011). Currently, to perform the extraction of protein fragments to be used as receptors or ligands, only a small part of the protein is used to be made the removal of the ligand or receptor, for example, botulinum toxin (TxB) used in aesthetic treatment. TxB is a neurotoxin formed by a protein complex of biological origin, produced from the bacterium Clostridium Botulinum, which has its serotype A (BoNT/A), the most potent type of toxin, removed for use as a tool to prevent aging (DE MELLO SPOSITO, 2004). However, the experimental methods used for the extraction of fragments are costly and time consuming.

In this context, the extraction of fragments of three-dimensional proteins, through a web system, will have a lower cost and time than the use of experimental methods. Because of this, a slicing service that automates the extraction process of polypeptides using the entire three-dimensional structure is being developed.

Methodology

The development of the system will follow the methodology of development by prototyping, which is a valid technique that can be perfectly employed provided that the system to be developed has the characteristics of having dynamic queries, interacting strongly with people, or having an algorithm or combinatorial processing that needs to be developed in an evolutionary way. The creation of a prototype is a work that must be done interactively and that follows the cycle exposed in Figure 2 (Pressman, Roger S; Maxim, 2016).

As in all software development approaches, prototyping begins with requirements gathering. During this phase, customers and developers are in constant interaction, thus facilitating the gathering of system requirements and functionality. A "quick design" is then developed, focusing on the aspects defined by the user. Rapid design leads to building a prototype that is evaluated by the customer/user and is used to refine the requirements. Prototyping is an efficient Software Engineering paradigm. The prototype will be built in order to serve as a basis for requirements definition. It will then be discarded - at least in part - and the real software will be designed, taking quality and maintainability into consideration (Pressman, Roger S; Maxim, 2016).

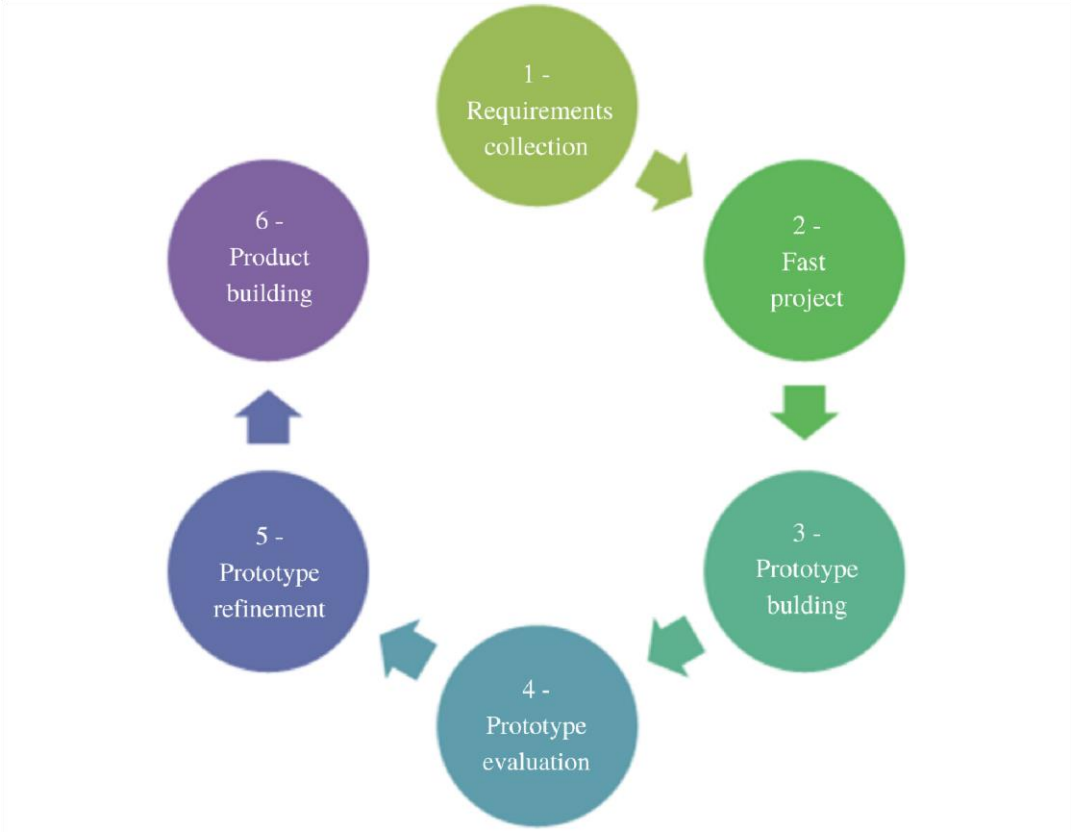

Figure 1. Prototyping Model

For the development of the slicer algorithm and the obtaining of the protein the Python programming language was used, and for the development of the web interface the Javascript, PHP and HTML languages are being used.

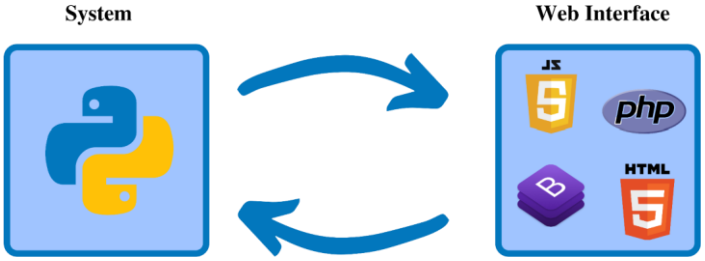

Figure 2. System Development

Figure 2 illustrates the development of the slicer system, which, as previously described, to perform the slicing service and obtain the protein, the Python programming language was used to develop the back-end, responsible for executing the service. To build the front-end, Javascript, PHP, HTML and the Bootstrap framework are being used to create the user interaction interface.

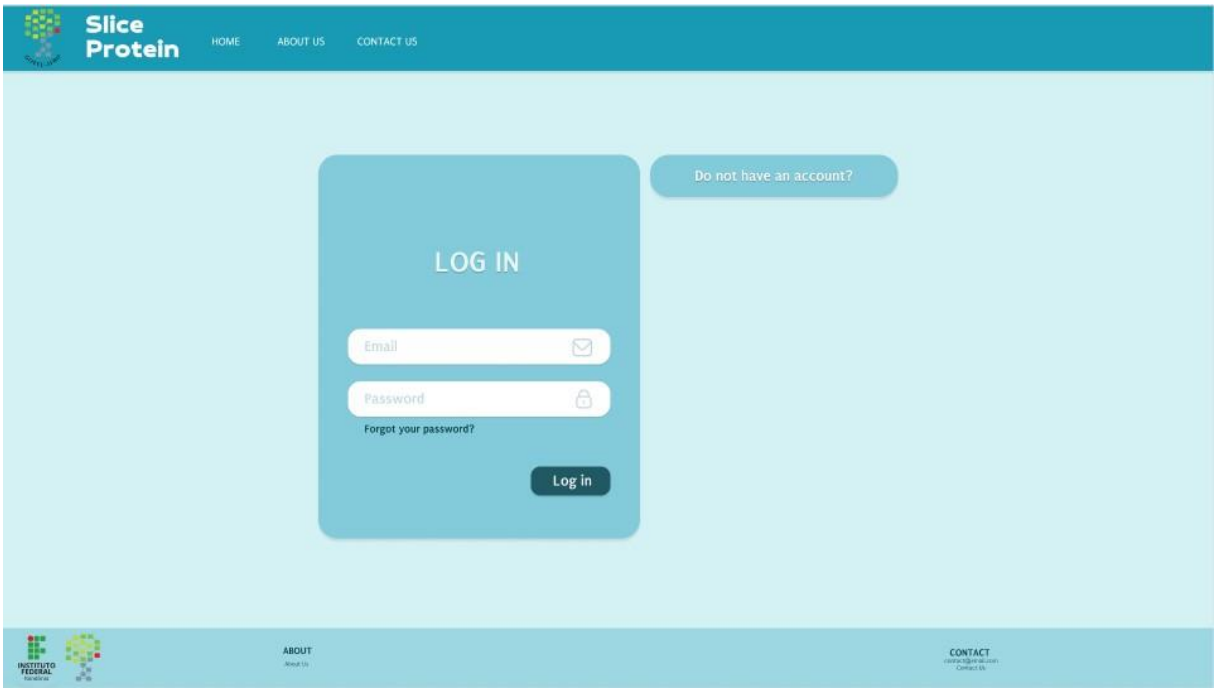

Figure 3. Login Screen Prototype

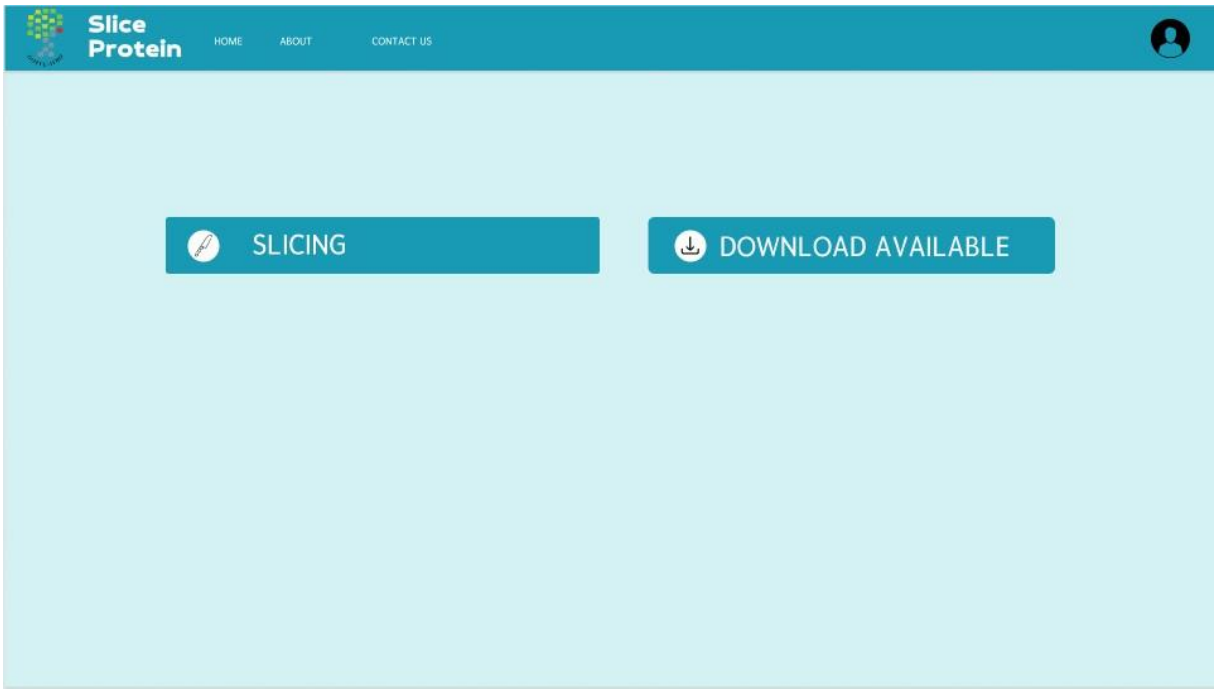

Figure 4. Home Screen Prototype

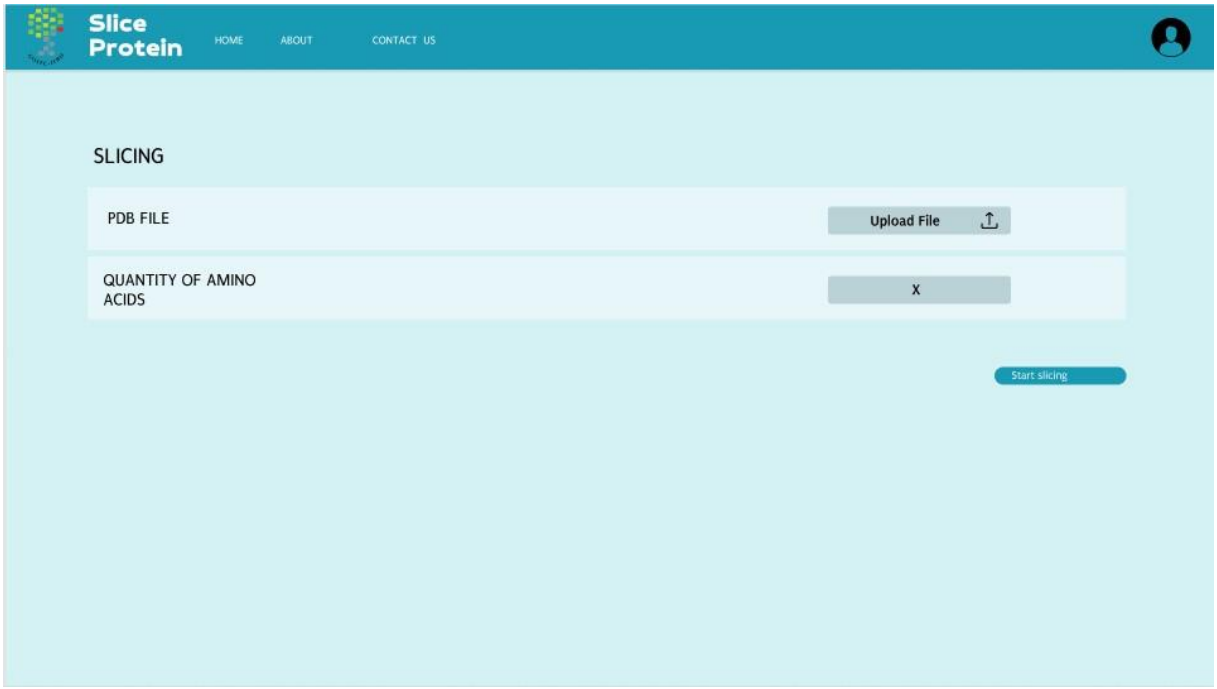

Figure 5: Slicing Screen Prototype

As shown in Figure 4, after the user logs in to the platform, the options "slicing" and "download available" will appear. To start the process, the researcher must upload the PDB file and specify the parameters for the system to generate the fragments.

Results and Discussion

The slicing algorithm was tested using three 3D structures of proteins present in the RCSB Protein Data Bank (RCSB PDB), namely: 1 GCQ, 5IEB and 5MXS.

The 1GCQ structure has 3 A, B and C chains, with 61 amino acids in the A, B chain and 70 amino acids in the C chain. It took about 00:03.24 seconds to generate 176 fragments, of which there are 54 fragments of the A chain, 55 fragments of the B chain, and 76 fragments of the C chain.

Structure 5IEB has only one chain, having 130 amino acids in chain A, and took about 00:32:89 seconds to generate 2,320 fragments, of which: 116 fragments of chain A in all twenty models.

Structure 5MXS has only one chain, containing 15 amino acids, and took about 00:07.80 seconds to generate 260 fragments, of which 13 were fragments of chain A in all twenty models.

Conclusion

The IFRO Slice Protein Web service, which is still under development, showed great results during the algorithm test, in which it was possible to identify that the slicer has fast results and many fragments generated in a short period of time.

Therefore, it's of prime importance to develop new tools that make the process of in silico prospecting agile, since they can be great allies by having lower costs and deadlines than experimental methods, and are able to accelerate and expand the scope of the process of prospecting new medicines.

Acknowledgment

The authors thank IFRO (Instituto Federal de Educação, Ciência e Tecnologia de Rondônia) and CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for the financial support.

References

Alberts, Bruce, et al. Fundamentos da Biologia Celular-4. Artmed Editora, 2006.

Borges, Michelle Jayme. "Efeito da toxina botulínica tipo a no tratamento de espasticidade de pacientes com sequelas de AVC e sua influência na funcionalidade: meta-análise." (2018).

COLHADO, Orlando Carlos Gomes; BOEING, Marcelo; ORTEGA, Luciano Bornia. Toxina botulínica no tratamento da dor. Revista Brasileira de Anestesiologia, v. 59, p. 366-381, 2009.

de Mello Sposito, Maria Matilde. "Toxina Botulínica do Tipo A: mecanismo de ação." Acta fisiátrica 16.1 (2009): 25-37.

LODISH, H. et al. Biologia celular e molecular. 7. ed. Porto Alegre: Artmed, 2014.

Nelson, David L.; Cox, Michael M. Princípios de bioquímica de Lehninger. 5 ed. Porto

Alegre: Artmed, 2011.

Oliveira, Lariza Laura de. *Uso de estratégias baseadas em conhecimento para algoritmos genéticos aplicados à predição de estruturas tridimensionais de proteínas*. Diss. Universidade de São Paulo, 2011.

RCSB PDB, RCSB Protein Data Bank: www.rcsb.org/.

Rino, José Pedro, e Nelson Studart. "Um potencial de interação para o estudo de materiais e simulações por dinâmica molecular." *Química Nova* 24.6 (2001): 838-845.

Silva, LX; Bastos, LL; Santos, LH. Modelagem computacional de proteínas. In: BIOINFO Revista Brasileira de Bioinformática e Biologia Computacional. 1. Ed. Vol. 1. Lagoa Santa: Editora Alfahelix, 2021. DOI: 10.51780/978-6-599-275326.

Sposito MM de M. Toxina botulínica tipo A: propriedades farmacológicas e uso clínico. Acta Fisiátr. [Internet]. 14 de dezembro de 2004 [citado 28 de maio de 2022];11(Supl.1):S7-S44. Disponível em: https://www.revistas.usp.br/actafisiatrica/article/view/102495

Wolynes, P.G. Evolution, Energy Landscapes and the Paradoxes of Protein Folding. Biochimie 2015, 119, 218–230.