Université Paris Cité

Noncommercial user License agreement for PEP-FOLD Software

USE OF THE PEP-FOLD SOFTWARE IS SUBJECT TO THE FOLLOWING TERMS AND CONDITIONS. IF YOU DO NOT ACCEPT THESE TERMS AND CONDITIONS, DO NOT LOAD THE SOFTWARE.

NON-PROFESSIONAL USERS ARE NOTIFIED THAT THE SOFTWARE AND THE RISKS INHERENT TO ITS USE, MODIFICATION AND/OR DEVELOPMENT AND REPRODUCTION ARE NOT COVERED BY ANY WARRANTY ON THE PART OF UNIVERSITE PARIS CITE AND THAT ITS SPECIFICITY MAKES IT COMPLEX TO OPERATE. THIS IS WHY IT IS HIGHLY RECOMMENDED THAT YOU ONLY LOAD AND/OR USE THE SOFTWARE IF YOU HAVE SPECIALIZED IN-DEPTH INFORMATION TECHNOLOGY KNOW-HOW.

UNIVERSITÉ PARIS CITÉ, a public institution of a scientific, cultural and business character, whose registered office is located at 85, boulevard Saint-Germain, 75006 Paris, France, No. SIRET 130.025.737.00011, APE code 8542Z, represented by its President, Mr. Edouard KAMINSKI, hereinafter referred to as "**UPCité**" hereby grants the right to use the PEP-FOLD software under the conditions set forth below.

SECTION 1 – DEFINITIONS

In this Agreement the following capitalized words shall have the following meanings:

"**Agreement**": refers to this agreement, its schedules and any amendments thereto.

"**Software**": refers to both the binary code of the software called "PEP-FOLD", the technical and operational features of which are described in schedule 1, and to the associated documentation which includes a user guide.

"**User**": the physical person who uses the Software.

SECTION 2 – PURPOSE

2.1 The purpose of this Agreement is the grant by UPCité to the Licensee of a non-exclusive, non-transferable, worldwide noncommercial license over the Software for the entire legal duration of the protection of the Software in accordance with the provisions of the french Intellectual Property Code (article L123-3).

2.2 This license is granted to the natural or legal person – hereinafter the "**Licensee**".

SECTION 3 – ACCEPTANCE

3.1 By loading the Software, the Licensee accepts the terms of this Agreement.

3.2 The Licensee acknowledges that the Agreement shall prevail on Licensee's general terms of purchase. The Licensee's general terms of purchase are unenforceable in whatever form.

SECTION 4 – SCOPE OF LICENSE

4.1 The user right granted in Section 2 to the Licensee, which it accepts, is limited to the right to use the Software for internal research purposes, excluding any collaborative research, in accordance with its intended use such as described in its associated documentation and with the terms and conditions set out below.

4.2 Any commercial use of the software is prohibited.

4.3 Any use of the Software that is not in compliance with its intended use is prohibited.

4.4 The User may not reproduce the Software in whole or in part, regardless of the form or format:

- except as strictly required to load, display, run, transfer or store the Software for the sole purpose of its use,
- or to make one (1) backup copy, by taking all necessary precautions to avoid its disclosure.

4.5 A single copy of the documentation may be made for internal use without UPCité's prior written authorization.

4.6 The User may not distribute, translate, adapt, arrange or modify the Software, or directly correct any anomaly. The Licensee further agrees not to reverse engineer, decompile or disassemble the Software, except if to make it interoperable, under the conditions set forth by law.

4.7 Any use not expressly authorized by the Agreement or by law is prohibited.

4.8 The Licensee expressly agrees not to assign, transfer or convey to a third party, with or without consideration, the user right granted hereunder.

SECTION 5 – INSTALLATION OF THE SOFTWARE

The Licensee shall load and install the Software at its own expense and risks.

SECTION 6 – FEE AND COMMERCIAL PURPOSES

6.1 ROYALTIES

This user license is granted free of charge for internal research purposes, excluding any collaborative research.

7.2 USE FOR COMMERCIAL PURPOSES

For any request for a license to use the Software for commercial purposes, the Licensee shall contact UPCité at the following address: partenariat.recherche.drive@u-paris.fr.

SECTION 7 - LICENSEE'S COMPLIANCE WITH ALL IMPORT REGULATIONS IN LICENSEE'S COUNTRY OF ORIGIN

7.1 Inasmuch the Software is imported by the Licensee, the Licensee is required to contact the competent national authorities so as to ensure compliance with all tax (specifically in the field of VAT) and customs (i.e. import license) regulations that may apply to the import of the Software.

7.2 Any costs in connection therewith shall be borne by the Licensee.

SECTION 8 – WARRANTY IN RESPECT OF THE SOFTWARE

8.1 The Licensee hereby recognizes that the Software is experimental software and that the current state of scientific and technical knowledge at the time of its release does not allow testing or verifying all possible uses or detecting the existence of any defects.

8.2 The Licensee hereby recognizes that the Software is supplied "as is" by UPCité, without any warranty whatsoever, express or implied, in particular without any warranty as to its merchantability, secure or innovative nature and fitness for a particular use.

8.3 UPCité does not warrant that the Software is error-free, or will operate without interruption, or that it will be compatible with the Licensee's equipment and software configuration or that it will meet the needs of the Licensee.

SECTION 9 – LIABILITY

9.1 UPCité's liability cannot be incurred by reason of (i) damages arising out of the non-performance, in full or in part, of its obligations by the Licensee, or (ii) indirect damages, even if UPCité knew of the

possibility of the occurrence of such damages. The Parties expressly agree that any financial or business loss (including without limiting to lost data, lost profits, loss of customers or orders, loss of earnings, commercial disturbances) or any lawsuit directed against the Licensee by a third party, constitutes indirect damages for which no remedies are available.

9.2 The Licensee releases UPCité from any and all liability in respect of any damages arising out of the use of the Software that are caused by the Licensee to a third party and assumes all risks inherent to the use of the Software with respect to its customers.

SECTION 10 – INTELLECTUAL PROPERTY RIGHTS

10.1 The Parties recognize that the Software is and shall remain the exclusive property of UPCité, which reserves all ownership rights on such basis. The license subject matter of this Agreement does not entail the transfer of any intellectual property right over the Software to the Licensee.

10.2 The Licensee expressly agrees:

- not to delete or otherwise modify in any manner the intel-lectual property notices or other ownership notices displayed on the Software; and
- to reproduce as is the said intellectual property notices or other ownership notices on the backup copy of the Software.

10.3 Similarly, no other right over a trademark, trade name or other distinctive sign is conferred to the Licensee by the Agreement.

10.4 The Licensee agrees not to directly or indirectly in-fringe UPCité's intellectual property rights and to take the necessary measures with respect to its staff so as to ensure their compliance with UPCité's intellectual property rights.

SECTION 11 - INFRINGEMENT

11.1 Any use of the Software outside the scope of the license grant by UPCité shall constitute infringement and war-rant proceedings being brought by UPCité against the Licen-see.

11.2 The Licensee shall notify UPCité of any act of in-fringement or unfair competition by a third party of which the Licensee has knowledge, and UPCité shall bring proceedings if it deems fit.

SECTION 12 – PUBLICATIONS

In all of the Licensee's publications concerning studies resulting from the Software, the Licensee is required to indicate that these studies were carried out through use of the Software and expressly mention Université Paris Cité and Inserm's names as the author of the Software and holder of the associated intellectual property rights.

SECTION 12 – MISCELLAENOUS

12.1 The Parties shall not be liable for any breach of one of their obligations arising out of circumstances outside of their control, such as strikes, exceptional weather conditions, acts of war, terrorism, riots, fires, natural disasters, malfunction or interruption of means of communication or telecommunication, including networks.

12.2 No failure, by either of the Parties, even if repeated, to assert one or more provisions of the Agreement, may be construed in any circumstance as implying a waiver by the Party concerned of its right to assert said provision(s) subsequently.

12.3 In the event one or more provisions of this Agreement were to conflict with a statute or legislative provision, existing or future, such statute or legislative provision shall prevail, and the Parties shall make the necessary amendments so as to comply with such statute or provision. All other pro-visions shall remain in effect.

SECTION 13 – LAW AND DISPUTE RESOLUTION

13.1 This Agreement is governed by French law.

13.2 Disagreements or disputes shall be referred to the competent courts of Paris by the more diligent Party.

Université Paris Cité

Annex 1 – Software description

PEP-FOLD is a software that will generate 3D conformations from a peptide amino acid sequence.

Context :

PEP-FOLD relies on the concept of structural alphabet [1].

A structural alphabet corresponds to a collection of conformations of short fragments that it is possible to assemble to describe accurately the structure of any complete protein. The structural alphabet in use here corresponds to a description of a protein structure as a series of fragments of 4 amino acids overlapping by 3 residues, undergoing a Markovian process of order 1. Hence, a protein of size L is desribed by a series of L-3 fragments. The alphabet has been learnt from a collection of non-redundant protein structures for which an optimum of 27 states has been identified. Using this model, it is possible to decode a protein 3 structure of L amino acids as a series of L-3 « structural alphabet letters », i.e. compress the 3D description as s 1D string of states. Once a collection of proteins decoded, it becomes possible to study the association between the structural alphabet letters and the amino acid sequence, which has been done using Support Vector Machines for PEP-FOLD.

PEP-FOLD consists of two steps :

1. The prediction of a probabilities of each of the 27 structural alphabet letters at each of the L-3 positions, from the amino acid sequence, i.e. the generation of a structural alphabet profile associated with the sequence.

2. The generation of the 3D models from the structural alphabet profile. Here, each structural alphabet letter is associated with a limited numer of representative 3D fragments that are used to rebuild the full structure. In practice, the fragments of 4 amino acids are superimposed on those assigned to the neighbour positions using an overlap of 3 amino acids (See [3-4]). The process is driven by the sOPEP force field [2], in which an amino acid is represented by 6 beads : the full backbone except the alpha-carbon hydrogen (5 beads), and one bead for the side chain, Different evolutions of the assembly process and the force field have been reported [5-9].

User interface:

Input : a fasta sequence

Output : a series of 3D models, clusterized and ranked by increasing sOPEP values.

Implementation :

The software comes as archives of 2 docker images : pyppp3-light, pepfold-core that contain the necessary executables. A python script (pepfold4.py) provides the protocol to go from sequence to 3D models.

Citing : The latest evolution of PEP-FOLD should be referred to using :

Rey, Julien, et al. "PEP-FOLD4: A pH-dependent force field for peptide structure prediction in aqueous solution." *Nucleic Acids Research* 51.W1 (2023): W432-W437.

References :

1. Camproux, Anne-Cloude, Romain Gautier, and Pierre Tuffery. "A hidden markov model derived structural alphabet for proteins." *Journal of molecular biology* 339.3 (2004): 591-605.

2. Maupetit, Julien, P. Tuffery, and Philippe Derreumaux. "A coarse-grained protein force field for folding and structure prediction." *Proteins: Structure, Function, and Bioinformatics* 69.2 (2007): 394-408.

3. Maupetit, Julien, Philippe Derreumaux, and Pierre Tuffery. "PEP-FOLD: an online resource for de novo peptide structure prediction." *Nucleic acids research* 37.suppl_2 (2009): W498-W503.

4. Maupetit, Julien, Philippe Derreumaux, and Pierre Tufféry. "A fast method for large-scale De Novo peptide and miniprotein structure prediction." *Journal of computational chemistry* 31.4 (2010): 726-738.

5. Thévenet, Pierre, et al. "PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides." *Nucleic acids research* 40.W1 (2012): W288-W293.

6. Lamiable, Alexis, et al. "PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex." *Nucleic acids research* 44.W1 (2016): W449-W454.

7. Lamiable, Alexis, Pierre Thévenet, and Pierre Tufféry. "A critical assessment of hidden markov model sub-optimal sampling strategies applied to the generation of peptide 3D models." *Journal of computational chemistry* 37.21 (2016): 2006-2016.

8. Tufféry, Pierre, and Philippe Derreumaux. "A refined pH-dependent coarse-grained model for peptide structure prediction in aqueous solution." *Frontiers in Bioinformatics* 3 (2023): 1113928.

9. Rey, Julien, et al. "PEP-FOLD4: A pH-dependent force field for peptide structure prediction in aqueous solution." *Nucleic Acids Research* 51.W1 (2023): W432-W437.