# pyPept: a python library to generate atomistic representations of peptides

Rodrigo Ochoa, J.B. Brown, Thomas Fox

Medicinal Chemistry, Boehringer Ingelheim Pharma GmbH & Co. KG 88397 Biberach, Germany
rodrigo.ochoa@boehringer-ingelheim.com

We present pyPept, a set of python scripts to manipulate and analyze peptide molecules using the BILN format to generate atomistic 2D and 3D representations[1]. The scripts allow the analysis of natural and modified peptides that are assembled based on personalized monomer dictionaries. From the line notation, the peptide construct can then be represented as an RDKit object for further prediction of properties and chemical structures.

One important task is the generation of relevant peptide conformers with correct PDB atom naming from the pyPept/RDKit object. We found that including secondary structure (SecStr) restraints during the conformer prediction is necessary. For this end, a similarity-based tool was developed to assign SecStr motifs to the peptides based on a dataset of bioactive conformers available in the PDB. The restraints are included in the distance bound matrix that is subsequently implemented by the ETKDGv3 method from the RDKit[2]. The obtained peptide conformers can embed cyclic restrictions, as well as conserve SecStr elements key for their active conformations, as shown for a few illustrative examples.

[1] T. Fox, M. Bieler, P. Haebel, R. Ochoa, S. Peters, A. Weber, **2022**, *Journal of Chemical Information and Modeling, 62(17),* 3942–3947.
[2] S. Wang, J. Witek, G. A. Landrum, and S. Riniker, **2020**, *Journal of Chemical Information and Modeling*, *60(4)*, 2044–2058.