

A tool to easily calculate pKa values and isoelectric points of proteins

Monday, 25 September 2023 15:10 (20 minutes)

PypKa is a tool developed by the Machuqueiro Lab at the University of Lisbon (UL), Portugal. It's a Poisson-Boltzmann-based pKa predictor for proteins using 3D structures as input. The tool also predicts isoelectric points and can process pdb structures to assign the correct protonation states to all residues. The impact of the PypKa cloud service is to predict pKa values of titratable sites in proteins (Reis et al. 2020). The team from the UL applied to the EGI-ACE Call in April 2021 with the intention to elevate the PypKa tool to the next level by porting it onto a scalable and easy-to-use cloud service that allows for fast pKa and isoelectric point calculations using user-provided protein structures or those obtained from the Protein Data Bank.

To better support this mission, and make the PypKa the go-to solution for the calculation of fast pKa and isoelectric point calculations, the EGI-ACE user support team allocated resources, services and provided consultancy in the framework of a dedicated Competence Centre (CC) led by a technical expert from CNRS. The Competence Centre was composed of members from appropriate service, resource and technology providers (UPV, CESGA, CNRS, and LIP). From a technical point of view, considering that the performance of the PypKa cloud service scaled up almost linearly concerning the number of vCPU cores, to support this challenge, the PypKa cloud server was deployed on the cloud resources of the EGI cloud infrastructure. Specifically:

- Universitat Politècnica de València (UPV) was involved as service provider of the Infrastructure Manager (IM) to facilitate the configuration and the set-up of the virtual and elastic cluster, whose resources scale dynamically, taking into account the number of users' requests to be served.
- LIP was involved as a cloud resource provider because of the close proximity to the user (the national cloud in Portugal).
- CESGA was involved as a cloud resource provider to complement the pool of resource capacity allocated by LIP.

The result is a cloud-based Thematic Service that is available as a web portal at <https://pypka.org/> and offers the functionality to predict Poisson-Boltzmann-based pKa values of biomolecules. Since late 2021, the portal with the scalable setup has been used by ~1,000 researchers from 15 European countries.

The technical support offered to the University of Lisbon to operate the PypKa Thematic Service will continue after the end of the EGI-ACE project with the following reallocation:

- LIP resource provider will continue to support the Thematic Service as a national cloud. The principal investigator has already received a national grant to continue using the LIP resources.
- CESGA cannot guarantee support and resources and it will be replaced with local resources provided by UL.

Primary authors: Dr FERNANDEZ, Carlos (CESGA); LA ROCCA, Giuseppe (EGI Foundation); Dr PANSANEL, Jerome (CNRS); Dr DAVID, Mario (LIP); Dr CABALLER, Miguel (UPV); Dr MACHUQUEIRO, Miguel (University of Lisbon); Dr REIS, Pedro de Brito Pires Santos (University of Lisbon)

Co-author: LAMATA, Ignacio (EGI)

Presenter: LAMATA, Ignacio (EGI)

Session Classification: IBERGRID

Track Classification: Cooperation between Iberian research communities