

Scientists at IRB Barcelona and BSC publish the world's largest data bank of proteins in motion for the pharmaceutical industry

- The new data base, which includes 1,700 proteins in motion, allows a more efficient design of drugs.
- MoDEL holds 30% of human therapeutic targets and the objective is to cover 80% in three years.

After four years of conducting intensive calculations in the supercomputer MareNostrum at the Barcelona Supercomputing Center, scientists headed by Modesto Orozco at the Institute for Research in Biomedicine (IRB Barcelona) have presented the world's largest data base on protein motions. Called MoDEL, this new database holds more than 1,700 proteins and is partially accessible through Internet to researchers worldwide. MoDEL has been developed to study the basic biology of proteins and to accelerate and facilitate the design of new pharmaceutical agents.

"Nowadays we design drugs as if the proteins against which they are to act were static and this goes a long way to explain failures in the development of new drug therapies because this is not a true scenario. With MoDEL this problem is solved because it offers the user from 10,000 to 100,000 photos per protein, and these confer movement to these structures and allow a more accurate design", says Orozco, head of the "Molecular modelling and bioinformatics" group at IRB Barcelona, director of the Life Sciences Programme of the Barcelona Supercomputing Center and full professor at the University of Barcelona. According to this researcher, several pharmaceutical companies are already using the MoDEL strategy to develop the first drugs against cancer and inflammatory diseases, which could become available this year.

A project in expansion

The scientists that develop MoDEL work from an international catalogue of static protein structures (approximately 40,000) called the Protein Data Bank (PDB). "1,700 videos of proteins from the 40,000 that make up the PDB may appear to be a small proportion, but many of the structures in the PDB are very similar. Therefore, following internationally established similarity criteria, we are representing about 40% of the proteins with a known structure".

But for Modesto Orozco the most relevant point is that MoDEL is now covering more than 30% of human proteins structures that are of pharmacological interest, that is to say, those that are potential targets of a new drug. "We obtained this datum through a very strict test so we consider that we are in fact covering more. However, MoDEL will continue to grow and this can be achieved more quickly because the system is well established". According to the researchers, the main objective is to focus on relevant proteins in human diseases and in a period of between 2 and 3 years cover 80% of pharmaceutical targets.

In order to undertake the MoDEL project, Orozco and his group are supported by resources provided by IRB Barcelona, the Barcelona SuperComputing Center, the Marcelino Botín Foundation, the *Fundación Genoma España*, the National Bioinformatics Institute and several European projects.

Reference article

Structure (Nov. 10 print issue, 2010). MoDEL (Molecular Dynamics Extended Library): A database of atomistic molecular dynamics trajectories". Tim Meyer, Marco D'Abramo, Adam Hospital, Manuel Rueda, Carles Ferrer-Costa, Alberto Pérez, Oliver CarrilloJordi Camps, Carles Fenollosa, Dmitry Repchevsky, Josep Ll. Gelpi, and Modesto Orozco.