

# Scientist helps move structural biology into 'big data' era

15 March 2016



Enrico Di Cera, M.D., professor and chairman of the department of biochemistry and molecular biology at Saint Louis University. Credit: Saint Louis University

In a recent paper published in *Nature Communications*, structural biologists detailed how a new data sharing consortium is helping scientists more quickly share and benefit from findings in their field.

Enrico Di Cera, M.D., chair of biochemistry and molecular biology at Saint Louis University, is an author on the paper and says that the Structural Biology Grid Consortium has developed a repository, the Structural Biology Data Grid, to deposit, search and download [structural biology](#)

data sets. In the current study, researchers found that the repository was effective in allowing researchers to reproduce earlier findings, letting work in the field progress.

"This is a transformative development in the field," said Di Cera. "Finally, we may take full advantage of the enormous amounts of data being generated by structural biologists."

X-ray crystallography, one of the most powerful tools in structural biology, allows researchers to determine the structure of proteins, nucleic acids and other small molecules at atomic level resolution. Understanding a protein's structure opens the door to understanding the molecular basis of diseases and developing new therapeutic strategies of intervention.

Crystallographers share their findings in academic journals and currently use standard repositories of processed datasets like the Protein Data Bank. The Structural Biology Data Grid supports archiving of raw experimental datasets using a distribution model of computing clusters. Benefits include rapid access of the original experimental data for general use and validation. With the data collection process becoming increasingly streamlined, archiving through the Structural Biology Data Grid will become mainstream.

In order to better leverage the breakthrough findings coming out of laboratories around the world, structural biologists created the Structural Biology Grid Consortium. The consortium's strategies include: curating and supporting a collection of data processing software; managing raw, experimental data sets; establishing a publication system for data sets; and integrating the storage resources of multiple research groups and institutions.

In the current study, researchers conducted a pilot study, analyzing data from the repository collection.

They found that the repository was effective in allowing researchers to reprocess data from earlier experiments, offering the opportunity to reproduce earlier findings, improve existing models, and catch possible mistakes earlier.

"The Grid started as a joint effort of top structural biology labs around the world. We are proud to be part of a great initiative that uses big data for the benefits of the entire scientific community," said Di Cera.

**More information:** Peter A. Meyer et al. Data publication with the structural biology data grid supports live analysis, *Nature Communications* (2016). [DOI: 10.1038/ncomms10882](https://doi.org/10.1038/ncomms10882)

Provided by Saint Louis University

APA citation: Scientist helps move structural biology into 'big data' era (2016, March 15) retrieved 22 February 2020 from <https://phys.org/news/2016-03-scientist-biology-big-era.html>

*This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.*