

Proteopedia: life in 3D, and teaching too

In this new instalment, the author presents a resource of great value in structural biology and, particularly, wants to share with you some ideas about how it reaches further than being just a source for information and how we can make profit of it to enrich our teaching and our students learning

Proteopedia¹⁻³ is a website with similar approach to well known sites like Wikipedia, but with a



PROTEOPEDIA
— LIFE IN 3D —

strong scientific focus towards structural biology. Its contents are a combination of automatic retrieval from databases and content contributed by users. In addition, the characteristics of a wiki are supplemented with the edition of interactive three-

dimensional models of molecular structure, which are embedded into the page of each article that makes the wiki. The subheading, *Life in 3D*, highlights this directive. The other, descriptive, subtitle outlines the philosophy of the site: *The free, collaborative 3D-encyclopedia of proteins & other molecules*.

Proteopedia's aim is to collect, organise and disseminate structural and functional knowledge about macromolecules, their complexes and their interaction with small molecules, in a format that is relevant and accessible both to scientists and students.⁴ This structural information requires models that allow to analyse the structure in diverse ways, as well as to perceive their volume and 3-dimensional nature. To achieve this, Proteopedia has developed the embedding of 3D interactive models into the pages, by including a tool (the *Scene Authoring*

Tool, or SAT) which allows any user to generate such structural models intuitively.

Technical issues

The principle of a wiki like this one is, as you well know, collaboration on a content that is accessible using a web browser. Everything is done within it: pages may be consulted but also any registered user may edit them directly, without the need for any additional software other than the browser.

Since this is a scientific environment, membership as Proteopedia user must be approved by the administrators and it must be done under the real name. Users that contribute to a page are automatically identified as such in the page footer. License for use and reproduction is Creative Commons, allowing dissemination, redistribution and change, as long as authorship is acknowledged.

The major uniqueness in Proteopedia is the use of molecular models within the pages. This was originally achieved using Jmol,⁵ which requires installation of Java in the user's computer. However, recently this has been replaced by the new variant, JSmol,⁶ that avoids Java and hence the need to configure or install any other software than the web browser; it also allows the use in mobile and touch devices (tablets) which do not support Java. Every user with a Proteopedia account may opt, as part of his or her preferences, for using Jmol-Java if desired, or else keep the default JSmol option. The Java version has a more agile behaviour, particularly with large molecules and with surfaces, than the non-Java JSmol version, which relies just on the HTML5 capabilities of the browser.

In order to partially ameliorate the limitations of the JSmol version, the *seeded pages* (those generated automatically in the system) have

«The major unique feature in Proteopedia is the use of molecular models within the pages.»

This article was first published, in English and Spanish, in SEBBM Journal, issue 180, June 2014 (the Journal of the Spanish Society of Biochemistry and Molecular Biology). It is republished here after the journal's original website was taken down.

recently been implemented with an initial simplified view of the structures, with a shorter loading time. This accelerates the initial build of the page, including the molecular model, reduces the risk of the user quitting the page when it is slow in loading, and avoids a possible initial absence of the structural view, a core element in the page conception. When the users wish so, they can request the full model, with all its details, simply by clicking on a button. The simplified model includes just the atoms that define the backbone of the molecules (alpha carbons and nucleic phosphori), enough to sketch the chain in cartoon mode. Ligands, ions, prosthetic groups... are displayed in full from the beginning. In addition, in the case of NMR-solved structures, only the first one is displayed; the other alternative conformers (often some tens of them) are only loaded if the full model is requested.

Types of Proteopedia pages

Proteopedia extends the approach of other wikis by offering a core content in addition to what the users might contribute.

Each new structure that gets published in the Protein Data Bank⁷ automatically generates a page in Proteopedia (*seeded page*). Information is similar to that in the PDB portal, but it is arranged in another way, possibly simpler and more accessible to a novel user, since it is more

Example of an authored Proteopedia page.

focused on structure and function than on features related to the experimental techniques and to connection with other databases.

On the other hand, there are authored pages, built from scratch by some users. They are often devoted to a protein of interest, or to a particular topic (like e.g. types of helices in proteins), or for instance they may be a summary of all existing records of solved structures for a certain protein. They might be described in short as a didactic unit, a tutorial for study.

Authored pages are public and may be modified by other users, following the principles of the wiki and the licence. There is, however, the alternative of using a private user space, protected from someone else's editions; this may be useful, for example, when preparing teaching material.

SAT (Scene Authoring Tools) [hide]

Interface of the scene authoring tool.

Features related to molecular structure

A prominent functionality in Proteopedia is that the usual methods to edit text are supplemented with the SAT (*Scene Authoring Tool*),⁸ devoted to edition of "scenes", i.e. customised models of the 3-dimensional structure that become part of the page. This offers a friendly interface that avoids the need to learn the use of Jmol commands in order to set the rendering style of the model, which allows to highlight the

features of interest. Additionally, and in line with the usual approach in wikis, the version history of each scene is stored, allowing one to recover past work if needed.

Among other utilities included in the seeded pages and which may also be easily built into authored pages, we can mention:

- Special green-coloured hyperlinks which, rather than taking to another page, load in the current page a customised model previously prepared using the SAT.
- A listing of the ligands that are part of the structure, allowing to highlight and focus each one of them.
- An evolutionary conservation analysis may be performed (via a connection to the ConSurf server⁹) and displayed using a colouring scheme that matches the conservation rating of each amino acid residue.¹⁰ Pay attention to this resource, that offers a significant potential for teaching.
- It is possible to include easily some multiple choice questions, that are automatically assessed.¹¹
- Citations may be inserted quickly by their id's either in PubMed (PMID) or DOI.¹²

An information media for teaching

For its conception as scientific medium, together with its multimedia distinctive feature, the possibility of authoring and collaboration and, finally, the accessible and open character, material in Proteopedia is a useful reference as a complement to lectures as well as to self-study.¹³ The possibly most useful feature is to take us –and our students– out from a static and flat view of biomolecules, imposed by printed textbooks. There is no need to explain here the relevance of understanding biomolecules as entities with a volume, both for their structure, their function and their interactions. All this is attainable if during our lectures we *pull* the internet connection to display that molecule we were talking about and we *push* our students to play with the structure, read the text, use the links... a world they are predisposed to. And all this for free, without any kind of piracy, and scientifically reliable. Ah, and if the internet connection would fail in the classroom, it is also possible to have previously saved the Proteopedia article –models included– and show it offline from our computer.

Proteopedia pages may be used as support during explanation of any protein, nucleic acid or

multimolecular complex we are dealing with in the lecture (even the whole ribosome!). Other candidate topics are the folding of the polypeptide chain, the different conformations of nucleic acids, the types of non-covalent interactions that determine tertiary structure, the concept of quaternary structure, and ligand-receptor, enzyme-substrate, or antigen-antibody interactions. In specialised courses, we can use it as a reinforcement for structural determination techniques – mainly X-ray diffraction of crystals and NMR in solution.

A way to set out activities for the students

We are reaching that most interesting and provoking part, and also the most difficult challenge: to plan tasks for the students that will lead them to using this tool as part of their training experience. Here, imagination and good teaching criteria are the rules. You can plan activities directed towards resolution of a question, or others that deal with preparing a topic for presentation to their fellow students in the classroom.¹⁴ I will just throw a few ideas that may serve as guidelines; with practice, surely you will be able to think of others.

As general advice, it is convenient to guide the students towards using the same media that the instructor has used in the lecture hall, and to provide some kind of introductory help, maybe an example of what we expect them to do. It is important, in these activities we are not used to, to explicitly and clearly state in advance the formative objectives (why we are proposing this activity), the tasks to do (what is expected from the student) and the assessment and marking criteria we will apply. If we define the task as something global to be done, we risk being ambiguous and, certainly, it will later be hard to assess and mark. To avoid this, it is convenient to propose specific questions, with a simple answer (and hence assessable), the resolution of which will require the student to use and explore the tool and the structural information.

There is room for promoting the students to do a small research about an example or a problem. It's important to highlight that models of solved structures are not just fancy illustrations but, first, they are the result of experimental work, corroborated and validated. Second, that the model contains the whole chemical information; with it one can study polarity, the hydrophobic

effect, salt bridges, disulphide bonds, ligand coordination, active centres, contacts with ligand or between subunits, even a catalytic mechanism. Another line to follow may be the study of functionally relevant conformational

changes: comparison of open and closed forms, changes upon binding of substrate, ligand or inhibitor...

Students involved in Proteopedia

The Proteopedia community provides us with some suggestions and examples of teaching implementations that involve the students in active and formative ways, such as projects for students¹⁵ or adoption in university courses.¹⁶ To stop being a passive user and become an author and publisher of new pages in Proteopedia will elicit on students a deeper knowledge of structural biology, and will enhance their scientific communication abilities. Regardless of whether they manage to make a significant contribution to the Proteopedia corpus, the effort of generating any content, as ephemeral as it might be, will often engage the students in a way that makes them think differently regarding structural data, to extract relevant perceptions and to share information productively with their mates.⁴ We will, hence, achieve some contribution to their scientific training.

Finally, there are also examples of advanced training projects that use Proteopedia as the working and publishing platform, both in secondary education¹⁷ and in university.¹⁸

If we propose to induce the students into becoming authors, even temporarily, using Proteopedia as the working environment, we may apply for creation of shared student accounts, so avoiding the need for everyone to sign up. In a typical setup, we will have a single account associated to our course, with a username and password, and inside that space one page will be created for each student. With a little discipline and good manners, each one will change just his page but everyone will be able to see each other's work. Later on, those students who wish so may apply for a regular user account, i.e. private, and use it for publishing their work in regular pages with general access, if that is the case, with no more effort than copying the content of the test page into the final one.

Table 1. Articles in BAMBED linked to Proteopedia entries

Articles published in the '*Multimedia in Biochemistry and Molecular Biology Education*' section of the journal *Biochemistry and Molecular Biology Education*, as outlines of pages built in Proteopedia. See also <http://proteopedia.org/w/Proteopedia:BAMBED>

2014 Feb	Protein kinase catalytic domain	doi:10.1002/bmb.20782
2013 Nov	Coiled-coil structure of keratins	doi:10.1002/bmb.20746
2013 Sep	Ricin	doi:10.1002/bmb.20716
2012 Sep	How we get the oxygen we breathe	doi:10.1002/bmb.20646
2012 Jul	β' subunit of bacterial RNA polymerase	doi:10.1002/bmb.20630
2011 Dec	DNA	doi:10.1002/bmb.20566
2011 Nov	Triose phosphate isomerase	doi:10.1002/bmb.20550
2011 Nov	Bovine pancreatic ribonuclease A	doi:10.1002/bmb.20568
2011 Sep	Phosphoinositide 3-kinase	doi:10.1002/bmb.20540
2011 May	Citrate synthase	doi:10.1002/bmb.20519
2011 Jan	HMG-CoA reductase	doi:10.1002/bmb.20481
2010 Nov	Ramachandran plots.	doi:10.1002/bmb.20457
2010 Sep	Large ribosomal subunit	doi:10.1002/bmb.20444

Instructors that prepare their own content

Another path, for those who feel brave, is to do ourselves what we were proposing the students to do: becoming authors. Proteopedia may be the means to build your own teaching materials related to biomolecular structure, customised to your teaching needs, with wide visibility and recognition of authorship. Some will maybe feel reluctant to this open publishing, but others may get indeed attracted by this route, which goes along with the increasing trend –I may even say fashion– of open materials and open courses, with free access.

Connections between Proteopedia and journals

Among its objectives, Proteopedia aims for integration with articles published on traditional media. To this purpose, several initiatives have been developed.

Within Proteopedia it is possible to activate a protected space, aimed at preparing material linked to journal articles, a sort of multimedia figures akin to the concept of supplementary material.¹⁹ Several journals have already joined this system.²⁰

Another, somewhat different, formula is that of *Biochemistry and Molecular Biology Education*, that has opened a section entitled 'Multimedia in Biochemistry and Molecular Biology Education' which accepts for publication brief outlines or reports of a main material published in a Proteopedia page (Table 1).

Articles published in the '*Multimedia in Biochemistry and Molecular Biology Education*' section of the journal *Biochemistry and Molecular Biology Education*, as outlines of pages built in Proteopedia. See also <http://proteopedia.org/w/Proteopedia:BAMBED>

Proteopedia is registered with ISSN 2310-6301 and an editorial committee has been appointed for selection of prominent pages, with a sound elaboration, that are then assigned a DOI.

With all this, Proteopedia, as a scientific platform, pretends to maintain an interrelationship with research and specifically with the publishing system.

Angel Herráez
Biochemistry and Molecular Biology,
Dept. of Systems Biology,
University of Alcalá
Alcalá de Henares, Spain

References and notes

- 1 (a) Proteopedia: The free, collaborative 3D-encyclopedia of proteins & other molecules. <http://proteopedia.org/>
(b) Getting started in Proteopedia. http://proteopedia.org/w/Proteopedia:Getting_Started
(c) Video guide to Proteopedia. http://proteopedia.org/w/Proteopedia:Video_Guide
- 2 E. Hodis, J. Prilusky, E. Martz, I. Silman, J. Moulton, J. L. Sussman (2008) Proteopedia - a scientific 'wiki' bridging the rift between three-dimensional structure and function of biomacromolecules. *Genome Biol.* **9**: R121. doi:10.1186/gb-2008-9-8-r121
- 3 J. Prilusky, E. Hodis, D. Canner, W. Decatur, K. Oberholser, E. Martz, A. Berchanski, M. Harel, J. L. Sussman (2011) Proteopedia: a status report on the collaborative, 3D web-encyclopedia of proteins and other biomolecules. *J. Struct. Biol.* **175**: 244-252. doi:10.1016/j.jsb.2011.04.011
- 4 J. Prilusky, W. Decatur, E. Martz (2013) Publishing in Proteopedia: The Guide. *In: Advancing Methods for Biomolecular Crystallography*; NATO Science for Peace and Security Series A: Chemistry and Biology. (R. Read, A. G. Urzhumtsev, V. Y. Lunin, eds.) pp. 277-295. Springer. doi:10.1007/978-94-007-6232-9_24
- 5 Jmol: an open-source Java viewer for chemical structures in 3D. <http://www.jmol.org/>
- 6 JSmol: an open-source HTML5 viewer for chemical structures in 3D. <http://wiki.jmol.org/index.php/JSmol#JSmol>
- 7 PDB: RCSB Protein Data Bank. An information portal to biological macromolecular structures. <http://www.pdb.org/>

- 8 Proteopedia scene authoring tools. <http://proteopedia.org/wiki/index.php/SAT>
- 9 The ConSurf server for the identification of functional regions in proteins.
<http://consurf.tau.ac.il/credits.php>
- 10 How to insert a ConSurf result into a Proteopedia green link.
http://proteopedia.org/wiki/index.php/Help:How_to_Insert_a_ConSurf_Result_Into_a_Proteopedia_Green_Link
- 11 Creating quizzes in Proteopedia. <http://proteopedia.org/w/Help:Quiz>
- 12 Proteopedia Help: Citing literature references.
http://proteopedia.org/wiki/index.php/Help:Editing#Citing_Literature_References
- 13 (a) Proteopedia primer for teaching purposes. <http://proteopedia.org/support/ProteopediaPrimer.doc>
(b) Teaching strategies in Proteopedia.
http://proteopedia.org/w/Teaching_Strategies_Using_Proteopedia
(c) Advanced Proteopedia authoring and its use in teaching.
http://proteopedia.org/w/Proteopedia:Practical_Guide_to_Advanced_Proteopedia_Authoring_and_Its_Use_in_Teaching
- 14 S. S. Jaswal, P. B. O'Hara, P. L. Williamson, A. L. Springer (2013) Teaching structure: student use of software tools for understanding macromolecular structure in an undergraduate biochemistry course. *Biochem. Mol. Biol. Educ.* **41**: 351-359. doi:10.1002/bmb.20718
- 15 Student projects in Proteopedia. http://proteopedia.org/w/Student_Projects
- 16 Adoptions of Proteopedia in college and university classes.
http://proteopedia.org/w/Proteopedia:News#Adoptions_in_College_and_University_Classes
- 17 SMART Teams Program (students modelling a research topic).
<http://proteopedia.org/w/Group:SMART:Teams>
- 18 University of Massachusetts Amherst Chemistry-Biology Interface Program (CBI Molecules).
http://proteopedia.org/w/CBI_Molecules
- 19 Interactive 3D complements in Proteopedia
http://proteopedia.org/wiki/index.php/Interactive_3D_Complements_in_Proteopedia
- 20 Proteopedia list of I3DC pages. <http://proteopedia.org/wiki/index.php/Proteopedia:I3DC>