Volume 175, Issue 2, August 2011 ISSN 1047-8477

ELSEVIER

Journal of
Structural
Biology

Spine2
complexes

interaction

Available online at www.sciencedirect.com

ScienceDirect

SPINE 2 OMPLEXES

SPECIAL ISSUE:
Spine2-Complexes EU-FP7 Project

Proteopedia: A status report on the collaborative, 3D web-encyclopedia of proteins and other biomolecules

Jaime Prilusky [a],[*],[1], Eran Hodis [b],[*],[1],[2], David Canner [c],[d], Wayne A. Decatur [e], Karl Oberholser [f], Eric Martz [g], Alexander Berchanski [c],[h], Michal Harel [i], Joel L. Sussman [c],[i],[*]

[a] Bioinformatics Unit, Biological Services Unit, Weizmann Institute of Science, Rehovot 76100, Israel
[b] Computer Science and Applied Mathematics Department, Weizmann Institute of Science, Rehovot 76100, Israel
[c] The Israel Structural Proteomics Center, Biological Chemistry Faculty, Weizmann Institute of Science, Rehovot 76100, Israel
[d] Chemistry Department, Princeton 08544-1009, NJ, USA
[e] Department of Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham, NH, USA
[f] Chemistry and Biochemistry Department, Messiah College, Grantham, PA 17055, USA
[g] Dept. of Microbiology, University of Massachusetts, Amherst, MA 01003-5720, USA
[h] Dept. of Organic Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel
[i] Structural Biology Department, Weizmann Institute of Science, Rehovot 76100, Israel

ARTICLE INFO

ABSTRACT

Proteopedia is a collaborative, 3D web-encyclopedia of protein, nucleic acid and other biomolecule structures. Created as a means for communicating biomolecule structures to a diverse scientific audience, *Proteopedia* (http://www.proteopedia.org) presents structural annotation in an intuitive, interactive format and allows members of the scientific community to easily contribute their own annotations. Here, we provide a status report on *Proteopedia* by describing advances in the web resource since its inception three and a half years ago, focusing on features of potential direct use to the scientific community. We discuss its progress as a collaborative 3D-encyclopedia of structures as well as its use as a complement to scientific publications and PowerPoint presentations. We also describe *Proteopedia*'s use for 3D visualization in structure-related pedagogy.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

With the solution of the first protein structure in 1958 by John Kendrew (Kendrew et al., 1958) came the inevitable problem of visualizing macromolecule structures. Kendrew's team approached this problem by constructing a physical model, and pictures of this model from various angles together served as a figure in the landmark publication. Computer-aided molecular visualization eventually superseded the building of physical models as the standard means to interpret and communicate a newly solved structure. However, a remarkable similarity with Kendrew's time remains: although computer-aided molecular visualizations can be easily shared electronically, when the time comes to publish, today's structural biologist still mimics the actions of Kendrew's team more than 40 years ago by capturing and submitting still images of a 3D model (now computer-generated) from various angles. Granted, many papers include stereoscopic images, but these are far from interactive, 3D representations. Model coordinates can also be downloaded and opened in the reader's molecular visualization application of choice, but the views and representations that have been carefully chosen for a figure are thus lost. Furthermore, most biologists without a background in structural biology are not familiar with molecular visualization applications. The notable exceptions to this apparent anachronism in communicating the authors' viewpoints of 3D macromolecule structures, namely "3D interactive journal figures", began in 1992 with David and Jane Richardson's Kinemage molecular illustrations (Richardson and Richardson, 1992). Currently there are several alternative approaches, including the proprietary iSee platform (Abagyan et al., 2006; Lee et al., 2009); a method for embedding interactive 3D figures within PDF files (Kumar et al., 2008, 2010); and Jmol display (Hanson, 2010; Jmol, 2010), which in various forms includes the IUCr Jmol enhanced toolkit (McMahon and Hanson, 2008), *Protein Science*'s implementation (Palmer and Matthews, 2009), Molecules in Motion (Reichsman, 2010) and *Proteopedia*'s Interactive 3D Complements [U1] (n.b. the symbols [U1], [U2] etc. refer to the *Proteopedia* URLs listed in Table 1).

The difficulties encountered in visualizing and communicating macromolecular structures motivated the creation of *Proteopedia* [U2] as a collaborative, 3D web-resource for presentation of biomolecule structures (Hodis et al., 2008). *Proteopedia* aims to allow the community of structural biologists, as well as the scientific community as a whole, to intuitively and easily communicate 3D structural annotations to a broad scientific audience. Use of the *Proteopedia* web resource is free, requires no downloads or installations (other than Java, standard on most computers) and makes contribution of structural annotations relatively easy. Interactive, 3D models of biomolecule structures are displayed, using Jmol, adjacent to descriptive text containing links. When clicked, these links evoke changes in the orientation, representation, and labeling of the 3D model, thus illustrating the points made in the text. In this status report we describe advances in *Proteopedia* since its creation almost four years ago. For the original communication describing the *Proteopedia* web resource the reader is referred to (Hodis et al., 2008). Here we detail recent progress in *Proteopedia*'s use:

- as a 3D web-encyclopedia of protein and other biomolecule structures
- for interactive visualizations to complement scientific publications and Microsoft PowerPoint presentations
- as a pedagogic tool for teaching biomolecule structure and function in the classroom

Structure communication on the web through a collaborative, 3D encyclopedia of biomolecule structures.

Table 1
Referenced *Proteopedia* URLs.

Number in text	URL	Proteopedia page/article
[U1]	proteopedia.org/w/Interactive_3D_Complements_in_Proteopedia	*Proteopedia*'s Interactive 3D Complements
[U2]	proteopedia.org	*Proteopedia*'s Main Page
[U3]	proteopedia.org/w/4hhb	4hhb
[U4]	proteopedia.org/w/Hemoglobin	Hemoglobin
[U5]	proteopedia.org/w/Proteopedia:Structure_Index	*Proteopedia*'s Structure Index
[U6]	proteopedia.org/w/Peroxisome_Proliferator-Activated_Receptors	Peroxisome proliferator-activated receptors (PPAR)
[U7]	proteopedia.org/w/Angiotensin-converting_enzyme	Angiotensin-converting enzyme (ACE)
[U8]	proteopedia.org/w/Plant_Viral_Protein_p19_Suppression_of_RNA_Silencing	Plant Viral Protein p19 Suppression of RNA Silencing
[U9]	proteopedia.org/w/Phosphoinositide_3-Kinases	Phosphoinositide 3-Kinases (PI3K)
[U10]	proteopedia.org/w/Cation-pi_interaction	Cation-π interactions
[U11]	proteopedia.org/w/Quality_assessment_for_molecular_models	Quality assessment for molecular models
[U12]	proteopedia.org/w/Hydrogen_in_macromolecular_models	Hydrogen in macromolecular models
[U13]	proteopedia.org/w/Lac_repressor	Lac repressor
[U14]	proteopedia.org/w/Citrate_Synthase	Citrate synthase
[U15]	proteopedia.org/w/Recoverin	Recoverin, a calcium-activated myristoyl switch
[U16]	proteopedia.org/w/Mechanosensitive_channels:_opening_and_closing	Mechanosensitive channels: opening and closing
[U17]	proteopedia.org/w/Forms_of_DNA	Forms of DNA
[U18]	proteopedia.org/w/Helices_in_Proteins	Helices in proteins
[U19]	proteopedia.org/w/Large_Ribosomal_Subunit_of_Haloarcula	Large ribosomal subunit of *Haloarcula*
[U20]	proteopedia.org/w/Ramachandran_Plot	Ramachandran plot
[U21]	proteopedia.org/w/HMG-CoA_Reductase	HMG-CoA reductase
[U22]	proteopedia.org/w/Proteopedia:Video_Guide	*Proteopedia*'s Video Guide
[U23]	proteopedia.org/w/Help:Contents	*proteopedia*'s help pages
[U24]	proteopedia.org/w/Proteopedia:How_to_Make_a_Page	*Proteopedia*'s How to make a page
[U25]	proteopedia.org/w/Proteopedia:Table_of_Contents	*Proteopedia*'s Table of Contents
[U26]	proteopedia.org/w/HIV_Protease	HIV-1 protease
[U27]	proteopedia.org/w/Acetylcholinesterase	Acetylcholinesterase
[U28]	proteopedia.org/w/AChE_inhibitors_and_substrates	AChE inhibitors and substrates
[U29]	proteopedia.org/w/Atorvastatin	Atorvastatin (Lipitor)
[U30]	proteopedia.org/w/Metabolic_Disorders	Metabolic disorders
[U31]	proteopedia.org/w/Pharmaceutical_Drugs	Pharmaceutical Drugs
[U32]	proteopedia.org/w/Pharmaceutical_Drug_Targets_and_Diseases	Pharmaceutical Drug Targets and Diseases
[U33]	proteopedia.org/w/Avidin	Avidin
[U34]	proteopedia.org/w/Workbenches	*Proteopedia*'s Workbench feature
[U35]	proteopedia.org/w/User:Janice_C._Telfer/Group_B_SRCR_domains	Group B SRCR domains (I3DC)
[U36]	proteopedia.org/w/User:Oleg_Kovalevskiy/Engineered_mutants_of_HlyIIR	Engineered mutants of HlyIIR (I3DC)
[U37]	proteopedia.org/w/3btp	*Agrobacterium* virulence complex VirE1–VirE2 (I3DC)
[U38]	proteopedia.org/w/Journal:JBIC:6	Sco proteins are involved in electron transfer processes (I3DC)
[U39]	proteopedia.org/w/Journal:JBIC:8	A hydrogen-bonding network formed by the B10-E7-E11 residues of a truncated hemoglobin from *Tetrahymena pyriformis* is critical for stability of bound oxygen and nitric oxide detoxification (I3DC)
[U40]	proteopedia.org/w/JBIC	List of *JBIC* I3DCs in *Proteopedia*
[U41]	proteopedia.org/w/Serine_Proteases:_A_Tutorial_of_Chymotrypsin,_Trypsin_and_Elastase	Serine proteases: a tutorial of chymotrypsin, trypsin and elastase
[U42]	proteopedia.org/w/Structural_templates	Structural templates
[U43]	proteopedia.org/w/Teaching_Scenes,_Tutorials,_and_Educators'_Pages	Central repository for 3D structure tutorials
[U44]	proteopedia.org/w/Teaching_Strategies_Using_Proteopedia#Student_Authoring_of_Temporary_Proteopedia_Pages:_Sandboxes	Student Authoring of Temporary Proteopedia Pages: Sandboxes
[U45]	proteopedia.org/w/GFP	Green fluorescent protein
[U46]	proteopedia.org/w/Triose_Phosphate_Isomerase	Triose phosphate isomerase
[U47]	proteopedia.org/w/Group:SMART:Teams	Students Modeling A Research Topic (SMART) Teams
[U48]	proteopedia.org/w/Group:SMART:A_Physical_Model_of_the_β2-Adrenergic_Receptor	A SMART Team's page: A Physical Model of the β2-Adrenergic Receptor

2. Growth and new articles

Proteopedia is a wiki based on the same software as Wikipedia (MediaWiki, 2007), with custom-added 3D structure annotation features. *Proteopedia* is composed of many individual pages (or articles), each of which can be edited by users whose names appear at the bottom of the articles they have edited. When *Proteopedia* went online in 2007, the number of user-generated articles numbered in the low tens and each was internally generated. In order to both populate the resource with useful content and to spur user contribution, we created a matching page in *Proteopedia* for each entry in the Protein Data Bank (PDB) (Berman et al., 2007; Sussman et al., 2001). These automatically-created articles display the PDB entry in rotatable, zoomable 3D alongside a two-dimensional image and the abstract of the publication describing the structure (Fig. 1A). Additional information is displayed underneath the 3D

model including coloring by evolutionary conservation as determined by ConSurf (Landau et al., 2005), links that highlight ligands and active sites within the 3D model, visualizing additional key structural features of the molecule (not visible in *Proteopedia*) in FirstGlance in Jmol (Martz, 2006), a simple macromolecular visualization tool, links to download the model coordinates, and structural and functional annotation aggregated from various resources by OCA (Prilusky, 1996). The automatically created PDB entry articles were originally intended as "seeds" that users could expand into full-fledged articles describing specific PDB entries. In parallel, other higher-tier articles that we refer to as "topic pages" describe proteins, classes of proteins or concepts in structural biology, and, in turn, link to the relevant automatically-created PDB entry pages within *Proteopedia*; for example the "Hemoglobin" [U4] article describes the hemoglobin protein using a small number of representative structures for illustration and provides
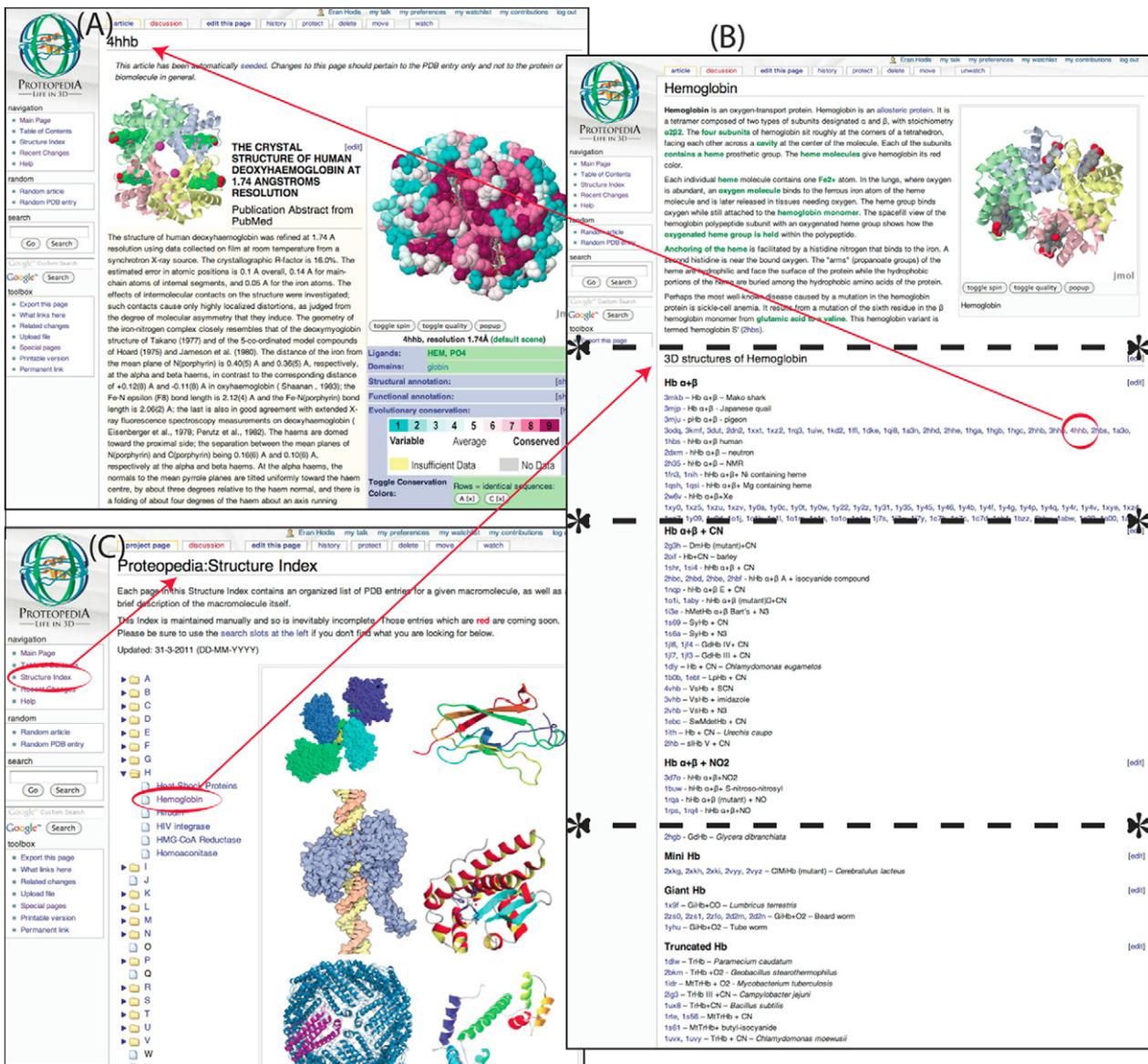

Fig.1. *Proteopedia* Topic Articles and PDB Entry Pages. (A) The *Proteopedia* PDB entry page for 4hhb [U3] is shown, with ConSurf evolutionary coloring highlighting the evolutionarily conserved residues of the 3D structure in dark magenta. The displayed page for 4hhb is representative of the pages automatically created for each of the entries in the PDB in *Proteopedia*. (B) The Hemoglobin [U4] topic article in *Proteopedia*. Dashed lines framed by asterisks represent places where screenshots from different areas of the Hemoglobin page have been pasted together to better serve the purposes of the figure, cropping out irrelevant areas of the Hemoglobin article. The top of the Hemoglobin article describes the structure of the protein using a representative 3D structure and green scene links for illustration, and the bottom of the article provides an organized list of the PDB entries for Hemoglobin. (C) The *Proteopedia* Structure Index [U5], accessible from a link on the left-hand toolbar on all pages in *Proteopedia*, provides an index linking to all of the organized lists of PDB entries in *Proteopedia*, organized by protein.

an organized list of links to the *Proteopedia* articles for the hundreds of individual hemoglobin PDB entries (Fig. 1B).

As *Proteopedia* grew, it became apparent that the expansion of automatically seeded PDB entry pages into full-fledged articles is rarely justified. *Proteopedia* users would routinely add information about a biological molecule to a PDB entry, such as descriptions of protein function, which more appropriately belong in a "topic page" titled for that biological molecule. We reasoned that a PDB entry article in *Proteopedia* should be edited only with annotations specific to that PDB entry, and not to describe the particular protein or biomolecule, which may be represented in several other PDB entry articles. Since the annotations added by users are only rarely specific to one PDB entry, we found ourselves encouraging the creation of a general "topic" article for each protein, rather than encouraging the expansion of the individual PDB entry articles that exist for each solved structure of that protein.

The result was the gradual creation of many high-quality topic articles on proteins such as peroxisome proliferator-activated receptors (PPAR) [U6], angiotensin-converting enzyme (ACE) [U7], Plant Viral Protein p19 Suppression of RNA Silencing [U8] and Phosphoinositide 3-Kinases (PI3K) [U9]. Articles not focused on a 'generic' protein also addressed concepts such as cation-π interactions [U10], Quality assessment for molecular models [U11] and Hydrogen in macromolecular models [U12]. Some users took advantage of *Proteopedia*'s (and thus Jmol's) animated display of multi-model PDB files referred to as "morphs" (generated externally, often using the Yale Morph Server (Flores et al., 2006; Krebs and Gerstein, 2000)) to show differences in conformational states such as those apparent in specific and non-specific binding of lac

repressor [U13] (Kalodimos et al., 2004; Pace et al., 1990) (Fig.2a), in free and oxaloacetate-complexed Citrate synthase [U14], in Recoverin, a calcium-activated myristoyl switch [U15] and in Mechanosensitive channels: opening and closing [U16]. Other users opted for side-by-side display of multiple 3D models whose display can be synchronized to better illustrate differences between similar structures such as B-, A- and Z-DNA [U17] (Fig.2b) as well as between alpha, 3_{10} helices and π-helices [U18]. A collaboration with the *Journal of Biochemistry and Molecular Biology Education* (*BAMBED*) publishes summaries of high-quality *Proteopedia* articles in *BAMBED* following peer review, highlighting such articles for use in pedagogy, as well as giving additional credit to *Proteopedia* article authors. To date, articles on the large ribosomal subunit [U19] (Decatur, 2010) (Fig.2c), Ramachandran plots [U20] (Oberholser, 2010), and HMG-CoA reductase [U21] (Canner, 2011) have been featured in BAMBED, with a feature on citrate synthase upcoming. It should be noted, however, that for many proteins and concepts *Proteopedia* articles remain either totally lacking or poorly developed. Resources such as the Video Guide [U22], the Help pages [U23] and the "How to make a page" [U24] article can help novice users to contribute structurally annotated articles.

## 3. Improving site organization

As the *Proteopedia* user community grew to more than 1200 registered users and contributions grew to hundreds of user-authored topic articles, a problem of organization emerged. Although a large number of articles were available, users reported
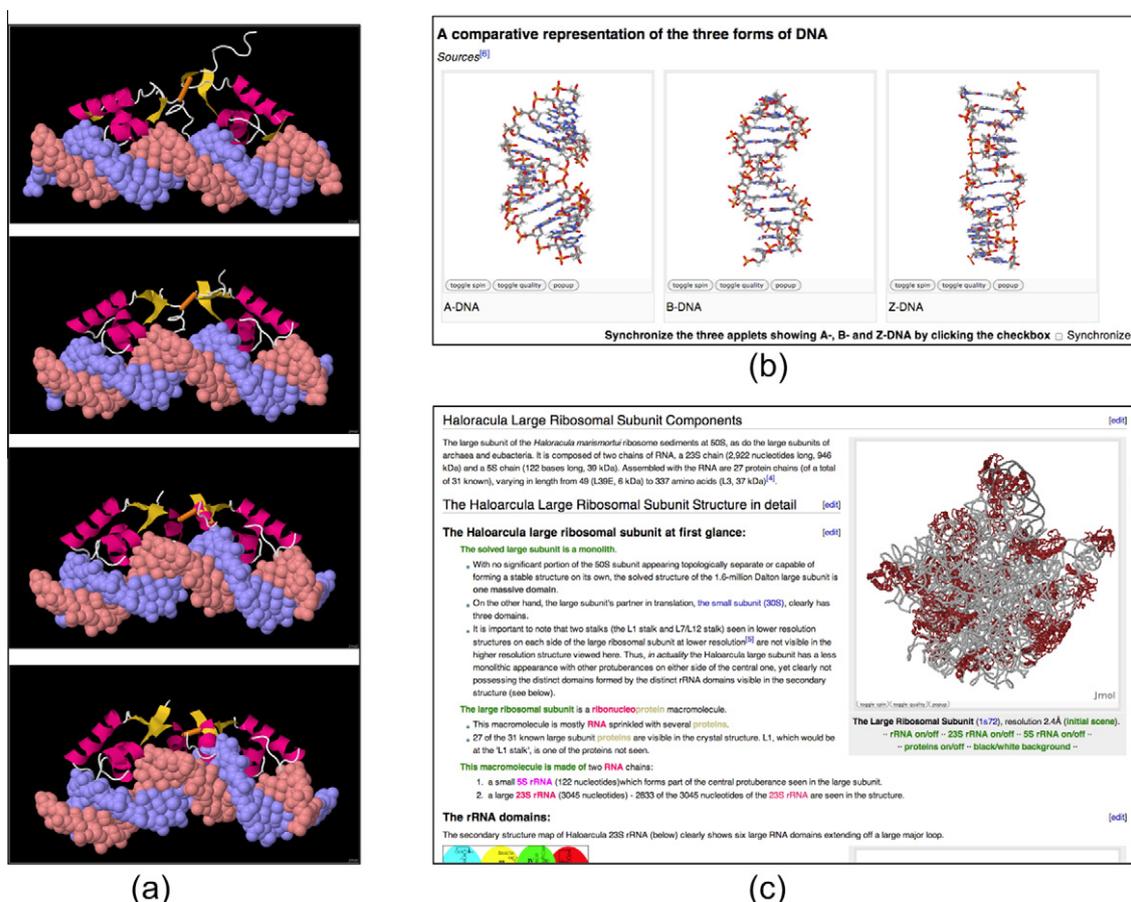
Fig.2. Examples of high quality *Proteopedia* articles made by users. (a) Portion of the lac repressor article [U13] which shows a morph of the lac repressor DNA binding domain in different conformational states as a function of specific and non-specific binding of to DNA (Kalodimos et al., 2004; Pace et al., 1990). (b) Forms of DNA [U17] – showing side-by-side display of multiple 3D models whose display can be synchronized to better illustrate differences between similar structures such as B-, A- and Z-DNA. (c) The article on the large ribosomal subunit of *Haloarcula marismortui* [U19], previously featured in BAMBED (Decatur, 2010).

difficulties in locating the articles that they were seeking and were often surprised to learn of the existence of quality articles of which they had not been previously aware. Duplicate articles started to appear as users began new articles on topics that they did not realize already had articles entitled with a slightly different phrasing of the topic name. We set out to improve site organization by three
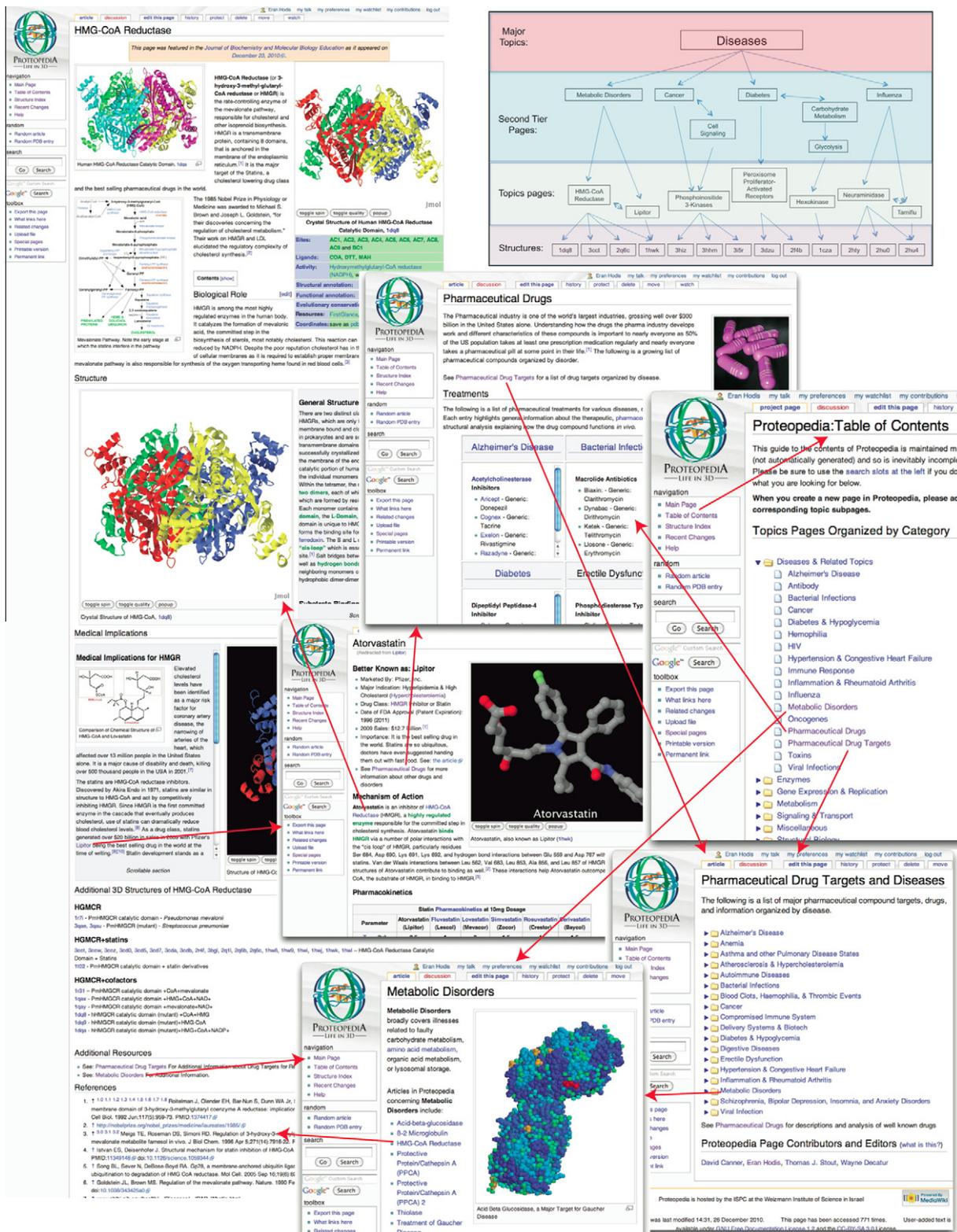
Fig.3. Hierarchical and inter-linked organization of *Proteopedia* pages. An example of the hierarchal organization of topic pages in *Proteopedia*. The image shows the *Proteopedia* HMG-CoA reductase (HMGR) [U21] article along the left-hand side of the image. The *Proteopedia* article Atorvastatin (Lipitor) [U29] and the higher-tiered organizational page Metabolic Disorders [U30] are shown as being linked to/from the HMGR article, with the inter-article links represented by red arrows in the figure. A higher-tiered organizational page Pharmaceutical Drugs [U31] is linked to from Atorvastatin, and links back to Atorvastatin as well (link not shown in screenshot used in figure). The higher-tiered organizational page Pharmaceutical Drug Targets and Diseases [U32] is shown linking to the Metabolic Diseases organizational page. The *Proteopedia* Table of Contents page [U25] is also shown, linking to several of the displayed pages, and accessible via a link on the left-hand toolbar on every *Proteopedia* page.

methods: (1) creating redirects; (2) creating links between articles; (3) creating progressively higher-tiered organizational pages, eventually leading to a Table of Contents [U25].

Creating redirects is an organizational strategy successfully employed by Wikipedia: when a user types "Fruits" into the Wikipedia search box, she is redirected to the article entitled "Fruit". Similarly, in *Proteopedia* we created redirects for many pages, and continue to create new redirects, so as to lead searches to the appropriate article. For example, "HIV-1 Protease", "HIV-1 protease", "HIV-Protease", "HIV Protease", "HIV protease" and "Hiv protease" now all redirect to the same *Proteopedia* article [U26]. Redirects funnel synonyms, acronyms, alternative spellings and pluralizations to the same article, thereby improving both search results and site organization.

Links between articles weave wiki articles deeper into the web of a wiki, allowing for a reader to quickly navigate to related articles without needing to search for them. An example of this can be seen in the 'Ramachandran plots' article, where a reader can follow a link to an article about phi and psi angles when they are mentioned in the Ramachandran plots article. This inter-article linking provides for the sort of immersive reading experience to which many readers of Wikipedia may be accustomed: often a reader begins reading one article only to find a related article fifteen minutes later for the sake of curiosity. It is our goal that *Proteopedia* readers will not only easily locate the articles that they are seeking, but also find themselves immersed in reading about structural biology.

Inter-article linking led eventually to the creation of yet higher-tiered organizational pages that list all "topic" articles related to a larger concept. The motivation for the creation of these higher-tiered pages is illustrated by the following example: A user interested in Alzheimer's disease may be unaware of the specific proteins related to the disease and may, therefore, search for "Alzheimer's disease" rather than searching for a specific protein. Like most of the new higher-tiered organizational pages in *Proteopedia*, the "Alzheimer's Disease" article is brief and could benefit from expansion. But already it improves organization by serving as a landing page for users interested in Alzheimer's disease, listing several related articles in *Proteopedia* including well-developed articles on acetylcholinesterase [U27] and on several of its inhibitors [U28]. The highest-tiered organizational pages are organized into a Table of Contents [U25] accessible from any page in *Proteopedia* via the left-hand toolbar.

The end result of this continuing effort to improve organization in *Proteopedia* is embodied by the HMG-CoA reductase (HMGR) [U21] topic article (Fig. 3). The HMGR article cohesively assimilates seven unique 3D structures into a single page, describing the important features of this enzyme, which plays a crucial role in the biosynthesis of cholesterol. All currently available HMGR structural data are associated with the page through links to *Proteopedia* articles on each of the >20 HMGR PDB entries. The HMGR article itself contains a description and a 3D depiction of the various domains and catalytic machinery of the enzyme, allowing readers to observe and interact with structural information assembled from various cited publications. Scenes describing how HMGR is inhibited by members of the statin drug class are also connected to additional descriptions of interactions and drug characteristics; links to topic articles on these drugs are provided, which in turn link to a higher-tier organizational page entitled "Pharmaceutical Drugs" [U31]. The Pharmaceutical Drugs page links to all topic articles on pharmaceutical drugs in *Proteopedia*, each of which includes a description of the relevant disease, the drug's importance, its mechanism of action in the context of 3D structure, a visualization of known drug resistance mutations and pharmacokinetic data. The HMGR topic article also links out to a higher-tier organizational page titled "Metabolic Disorders" [U30], which lists other related topic articles, and is one of the major organizational articles

listed in the *Proteopedia* Table of Contents. The 'Metabolic Disorders' organizational article is also linked to from an article titled "Pharmaceutical Drug Targets and Diseases" [U32] which, in turn, is linked to from the *Proteopedia* Table of Contents, allowing for several paths for a user to reach the same content, with mutual links between articles to facilitate this (Fig. 3). The editable nature of each article allows for simple reorganization as required. This functional approach to structural data organization differs drastically from an organization based on searching for PDB entry keywords alone, and can be particularly useful for those approaching structural biology from other fields, such as biochemistry or medicine.

Importantly, as mentioned in the HMGR example above, each topic article, after employing several representative PDB entries to describe its topic (usually a protein) through descriptive text married to illustrative, interactive 3D scenes, contains a section near the end of the article with an organized list of all solved structures of that protein or topic. In most cases these organized lists of PDB entries are curated by *Proteopedia* staff, and attempt to present the list of all relevant PDB entries in an intelligent manner, dividing the entries by type, organism, mutation, ligand or number of domains present (see Avidin [U33] for an example). These curated and organized lists of PDB entries, not available elsewhere, provide a simple way to quickly survey all the solved structures for a given protein or biomolecule. Each organized list of PDB entries is linked to from a central index called the "Structure Index" (Fig. 1C) [U5], accessible from any page in *Proteopedia* from the left-hand toolbar. With each weekly update of the Protein Data Bank, *Proteopedia* automated scripts send updates to the *Proteopedia* staff alerting them to new entries that must be added to the Structure Index.

4. Structure communication in publications and lectures

The motivation for making structural information accessible to a diverse scientific audience led to *Proteopedia*'s creation as a central repository for 3D structural annotation. The ease with which 3D structural annotations can be created live on the web on *Proteopedia* has consequently led to its use for 3D visualization in lectures and to complement articles in scientific journals.

## 5. Displaying interactive 3D structures in PowerPoint presentations

Many structural biology researchers are well aware that their lectures are made much more understandable when they present their 3D structures not through static 2D figures but through movies or interactive animations. There are several programs that provide this functionality including iSee (Abagyan et al., 2006; Lee et al., 2009), Discovery Studio Visualizer 3.0 (Accelrys, 2008) and the *Proteopedia* website. Users can take advantage of the structural annotation that they or others have added to *Proteopedia* in order to enhance lectures. Any *Proteopedia* article can be embedded into a PowerPoint presentation using the free LiveWeb plugin (http://skp.mvps.org/liveweb.htm) (Fig. 4), which permits the display of any webpage within a PowerPoint presentation. The particular *Proteopedia* article then appears, with the 3D structure and scene links fully functional and interactive, within a slide of the presentation. Since relying on an active internet connection during a live presentation is risky, we have developed an "Export Page" feature that saves an entire *Proteopedia* article, including all displayed interactive 3D models and scene links, to a user's computer (Fig. 4). Any page in *Proteopedia* can be exported by using the "Export Page" button on the site's left-hand toolbar. Although LiveWeb works only in Windows at the moment, all users may display the

exported page. Thus, Macintosh and Linux users can choose to transition between their presentation software and the exported page in order to present their 3D structures. As an additional benefit, audience members with laptops may be invited to interact with the 3D structure as the lecturer describes it. They may also return to view the related page on *Proteopedia* at a later date.
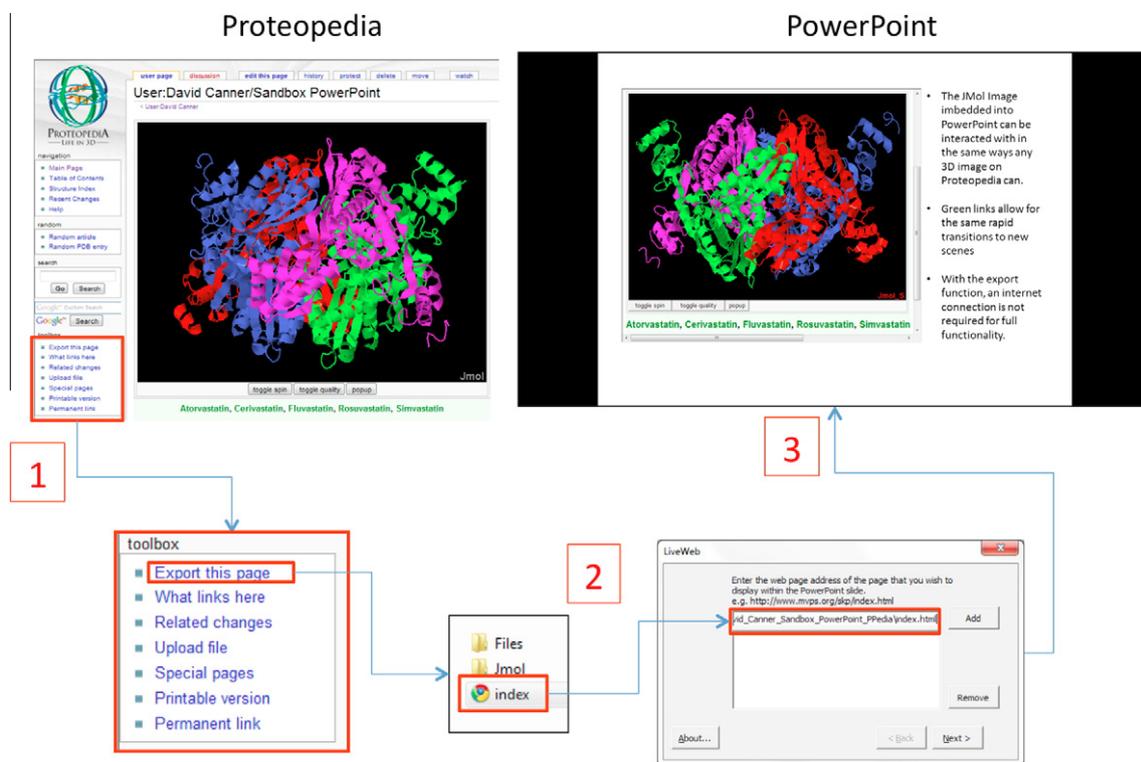
Fig.4. Displaying *Proteopedia* Pages within PowerPoint Presentations. A step-by-step schematic displaying how to embed a functional *Proteopedia* page with 3D scenes within a PowerPoint presentation. With the LiveWeb plugin installed on PowerPoint, the LiveWeb webpage button is used to add a *Proteopedia* URL or an exported *Proteopedia* page to the slide. The 3D structure and the green scene links are functional within the PowerPoint presentation.
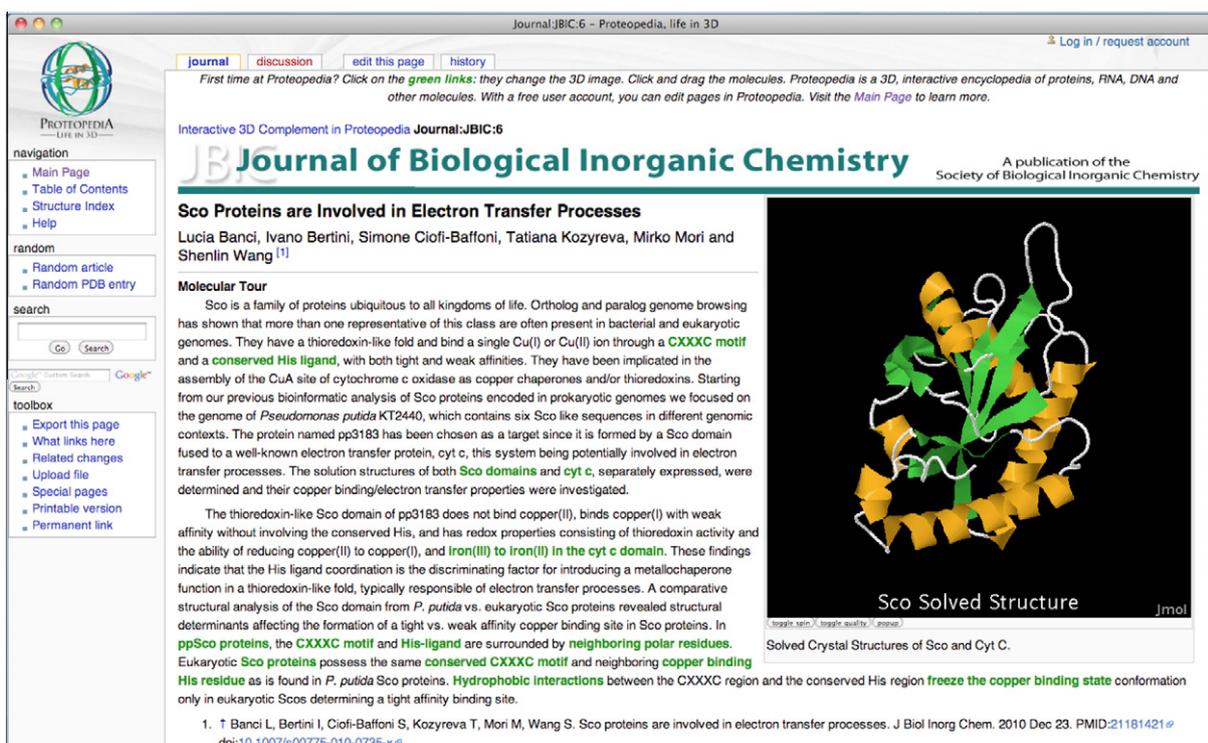
Fig.5. Interactive 3D complements to structure-related journal articles. Interactive 3D complement (I3DC) [U1] for an article in the *Journal of Biological Inorganic Chemistry* by Banci et al. on Sco proteins involved in electron transfer processes [U38] (Banci et al., 2011).

6. Interactive 3D complements to structure-related journal articles

To enable *Proteopedia* users to complement their scientific publications with interactive 3D visualizations, we have developed a system whereby users can create private articles that only they may view and edit. We expanded an existing system that allows users to create "protected" pages in their designated user-area that are publicly viewable but that only they can edit. In contrast, the new "Workbench" feature, intended expressly for the creation of articles complementing upcoming scientific publications, allows authors to hide their protected article during its development (see Workbench page [U34] for details). These hidden pages may be made accessible for viewing and editing by co-authors, referees and editors, and can quickly be made public upon publication of the accompanying paper. Such "Interactive 3D Complements" (I3DCs) [U1] to publications have been used effectively by individual research groups:

- Group B SRCR domains [U35] (Wang et al., 2011).
- Engineered mutants of HlyIIR [U36] (Kovalevskiy et al., 2010).
- Structure of the *Agrobacterium* virulence complex VirE1–VirE2 [U37] (Dym et al., 2008).

In addition, I3DCs are already appearing for structure-oriented articles in the *Journal of Biological Inorganic Chemistry* (*JBIC*), e.g.

- Sco proteins are involved in electron transfer processes [U38] (Banci et al., 2011) (Fig. 5).
- A hydrogen-bonding network formed by the B10-E7-E11 residues of a truncated hemoglobin from *Tetrahymena pyriformis* is critical for stability of bound oxygen and nitric oxide detoxification [U39] (Igarashi et al., 2011).
- Additional I3DCs for *JBIC* articles can be found at: *Proteopedia*'s JBIC [U40] page.

Such 3D complements allow all readers to view newly solved structures in a friendly interactive format that closely parallels the manner in which the authors themselves have likely observed and analyzed the novel structure.

7. Structure communication in pedagogy

Simplified 3D structure communication can play an important role in teaching students about macromolecular structure and the relationship between structure and function. Interactive, online articles describing 3D structures, such as those found on *Proteopedia* and elsewhere, such as the Macromolecular Museum (Marcey, 2011), Biomolecules at Kenyon (Slonczewski, 2009) and Molecular Visualization Resources (Martz, 2010) serve as excellent tutorials that educators can project in front of their class and that students can privately navigate at their own pace. In *Proteopedia*, educators are encouraged to share and adapt one another's tutorial
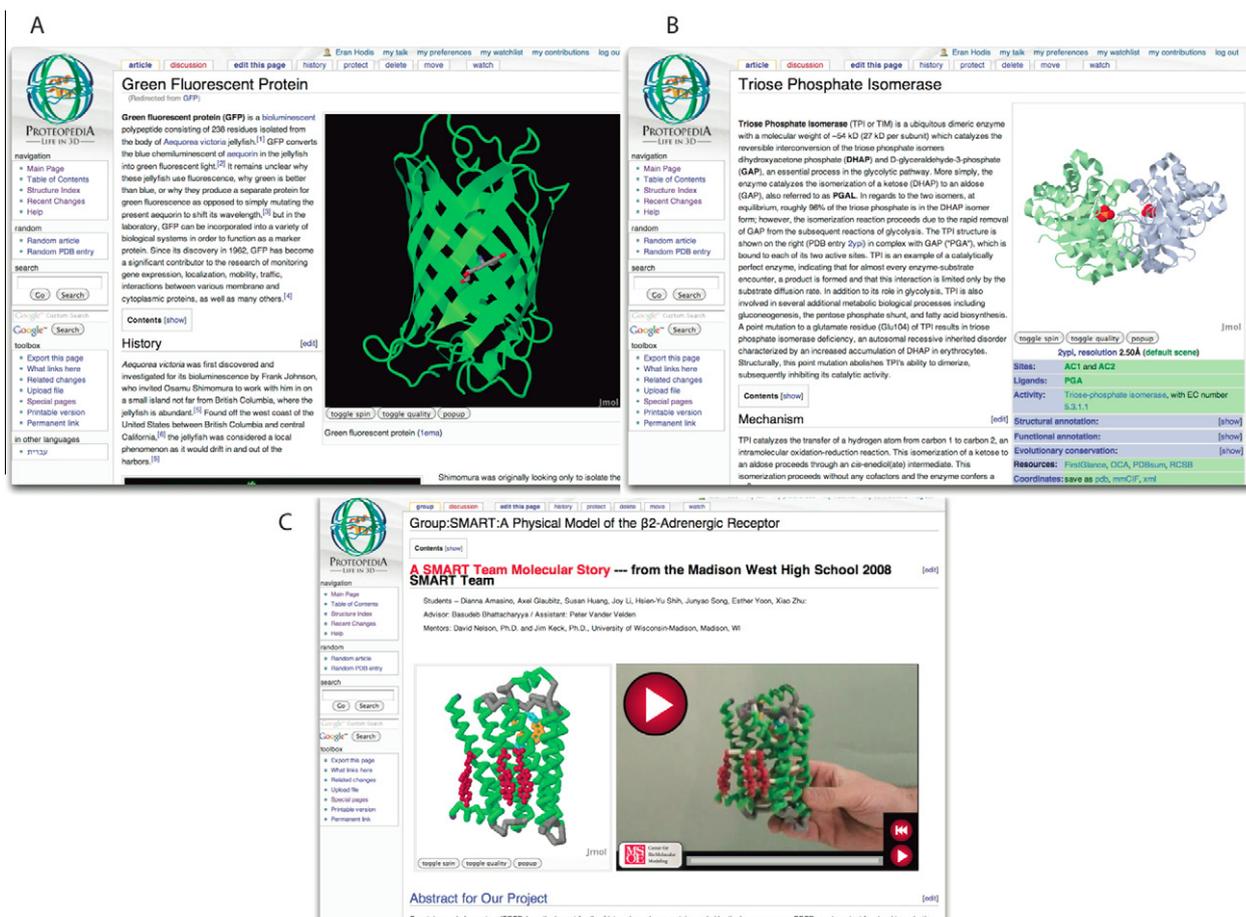
Fig. 6. *Proteopedia* use in structure-related pedagogy. (A) Green fluorescent protein (GFP) [U45] *Proteopedia* article started by Laura Carbone, who was at the time an undergraduate at Messiah College. (B) Triose phosphate isomerase [U46] *Proteopedia* article by Gregg Snider, graduate student in the Department of Biochemistry at the University of Vermont. (C) *Proteopedia* page in the reserved "Group:SMART" area for projects by the Students Modeling A Research Topic (SMART) [U47] program participants. The displayed page is entitled "A Physical Model of the β2-Adrenergic Receptor" [U48] and is authored by the Madison West High School 2008 SMART team. To the right of the interactive 3D structure is a movie the SMART program created and uploaded, which uses a physical model of the β2-adrenergic receptor to explain its structure and function.

pages, for example a tutorial on Serine Proteases [U41] and a tutorial on Structural Templates [U42]. In addition, there is a central repository for 3D structure tutorials [U43].

The ease with which structural annotations can be created on *Proteopedia* allows for students to learn about macromolecule structure by creating their own *Proteopedia* pages as well. These student-authored articles may be created either with the aim of authoring a reference article in *Proteopedia* or as a temporary "Sandbox" page [U44] marked for practice or class projects. The content of Sandbox pages is not permanent, and will be erased or replaced at a later time. Multiple Sandbox pages for use in class projects can be reserved using a semi-automatic reservation form that also allows the educator to easily place the same explanatory text at the top of each of his or her reserved sandbox pages, including the name of the class and a date when the content can be erased. *Proteopedia* has been used in many graduate and undergraduate courses, and several student-authored articles exist on the site, such as an undergraduate-student-authored article on green fluorescent protein [U45] that has since been expanded by multiple contributors (Fig. 6A) and a graduate–student–authored article on triose phosphate isomerase [U46] (Fig. 6B). In addition, groups of United States high school students exploring macromolecule structures through a program called SMART (Students Modeling A Research Topic) Teams (Herman, 2006) upload their final projects to a designated area for such projects in *Proteopedia* (Fig. 6C).

8. Conclusions

Proteopedia has evolved over almost four years from a new, untested and unpopulated resource into a collaborative environment for 3D structure communication in use by a growing number of members of the structural biology research and education communities. Its simple and intuitive display of 3D structures allows a diverse scientific audience to access and create structural annotations that are immediately accessible across the web. In addition, with *Proteopedia*'s growth, the opportunity for structural biologists to create visible, authoritative articles on their proteins, biomolecules or areas of research in the 3D encyclopedia grows as well – existing *Proteopedia* articles rank highly in Google search results, and thus are a likely destination for internet searches. Through efforts such as *Proteopedia* and the other related efforts mentioned in this article, structure communication improves past its level during Kendrew's time and allows for the fruits of structural biology research to reach a broad scientific audience not well versed in molecular visualization, from fellow biologists to students to laymen.

Acknowledgments

References

Abagyan, R., Lee, W.H., Raush, E., Budagyan, L., Totrov, M., et al., 2006. Disseminating structural genomics data to the public: from a data dump to an animated story. TIBS 31, 76–78.

Accelrys, 2008, Discovery Studio Visualizer 2.0, <http://accelrys.com/products/discovery-studio/visualization/discovery-studio-visualizer.html>.

Banci, L., Bertini, I., Ciofi-Baffoni, S., Kozyreva, T., Mori, M., et al., 2011. Sco proteins are involved in electron transfer processes. J. Biol. Inorg. Chem. (ePub).

Berman, H., Henrick, K., Nakamura, H., Markley, J.L., 2007. The worldwide Protein Data Bank (wwPDB): ensuring a single, uniform archive of PDB data. Nucleic Acids Res. 35, D301–D303.

Canner, D., 2011. Proteopedia entry: HMG-CoA reductase. Biochem. Mol. Biol. Educ. 39, 64.

Decatur, W.A., 2010. Proteopedia entry: the large ribosomal subunit of *Haloarcula marismortui*. Biochem. Mol. Biol. Educ. 38, 343.

Dym, O., Albeck, S., Unger, T., Jacobovitch, J., Branzburg, A., et al., 2008. Crystal structure of the *Agrobacterium* virulence complex VirE1–VirE2 reveals a flexible protein that can accommodate different partners. PNAS 105, 11170–11175.

Flores, S., Echols, N., Milburn, D., Hespenheide, B., Keating, K., et al., 2006. The database of macromolecular motions: new features added at the decade mark. Nucleic Acids Res. 34, D296–301.

Hanson, R., 2010. Jmol – a paradigm shift in crystallographic visualization. J. Appl. Cryst. 43, 1250–1260.

Herman, T., 2006. SMART (Students Modeling A Research Topic) Teams <http://cbm.msoe.edu/stupro/smart>.

Hodis, E., Prilusky, J., Martz, E., Silman, I., Moult, J., et al., 2008. Proteopedia – a scientific 'wiki' bridging the rift between 3D structure and function of biomacromolecules. Genome Biol. 9, R121.

Igarashi, J., Kobayashi, K., Matsuoka, A., 2011. A hydrogen-bonding network formed by the B10-E7-E11 residues of a truncated hemoglobin from *Tetrahymena pyriformis* is critical for stability of bound oxygen and nitric oxide detoxification. J. Biol. Inorg. Chem. 16, 599–609.

Jmol, 2010. Jmol: An Open-Source Java Viewer for Chemical Structures in 3D. <http://www.jmol.org>.

Kalodimos, C.G., Biris, N., Bonvin, A.M., Levandoski, M.M., Guennuegues, M., et al., 2004. Structure and flexibility adaptation in nonspecific and specific protein–DNA complexes. Science 305, 386–389.

Kendrew, J.C., Bodo, G., Dintzis, H.M., Parrish, R.G., Wyckoff, H., et al., 1958. A three-dimensional model of the myoglobin molecule obtained by X-ray analysis. Nature 181, 662–666.

Kovalevskiy, O.V., Solonin, A.S., Antson, A.A., 2010. Structural investigation of transcriptional regulator HlyIIR: influence of a disordered region on protein fold and dimerization. Proteins 78, 1870–1877.

Krebs, W.G., Gerstein, M., 2000. The morph server: a standardized system for analyzing and visualizing macromolecular motions in a database framework. Nucleic Acids Res. 28, 1665–1675.

Kumar, P., Ziegler, A., Grahn, A., Hee, C.S., 2010. Leaving the structural ivory tower, assisted by interactive 3D PDF. TIBS 35, 419–422.

Kumar, P., Ziegler, A., Ziegler, J., Uchanska-Ziegler, B., Ziegler, A., 2008. Grasping molecular structures through publication-integrated 3D models. TIBS 33, 408–412.

Landau, M., Mayrose, I., Rosenberg, Y., Glaser, F., Martz, E., et al., 2005. ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. Nucleic Acids Res. 33, W299–W302.

Lee, W.H., Atienza-Herrero, J.N., Abagyan, R., Marsden, B.D., 2009. SGC – structural biology and human health: a new approach to publishing structural biology results. PLoS One 4, e7675.

Marcey, D., 2011. The Online Macromolecular Museum, <http://www.callutheran.edu/Academic_Programs/Departments/BioDev/omm/gallery.htm>.

Martz, E., 2006. FirstGlance in Jmol. <http://firstglance.jmol.org>.

Martz, E., 2010. Molecular Visualization Resources. <http://MolviZ.Org>.

McMahon, B., Hanson, R.M., 2008. A toolkit for publishing enhanced figures. J. Appl. Cryst. 41, 811–814.

MediaWiki, 2007. <http://www.mediawiki.org>.

Oberholser, K., 2010. Proteopedia entry: Ramachandran plots. Biochem. Mol. Biol. Educ. 38, 430.

Pace, H.C., Lu, P., Lewis, M., 1990. Lac repressor: crystallization of intact tetramer and its complexes with inducer and operator DNA. PNAS 87, 1870–1873.

Palmer, A.G., Matthews, B.W., 2009. Interactive graphics return to protein science. Protein Sci. 18, 677.

Prilusky, J., 1996. OCA, a browser-database for protein structure/function. <http://oca.weizmann.ac.il/oca-docs/oca-home.html>.

Reichsman, F., 2010. Molecules in Motion. <http://www.moleculesinmotion.com>.

Richardson, D.C., Richardson, J.S., 1992. The kinemage: a tool for scientific communication. Protein Sci. 1, 3–9.

Slonczewski, J., 2009. Biomolecules at Kenyon. <http://biology.kenyon.edu/BMB/biomolecules.htm>.

Sussman, J.L., Lin, D., Jiang, J., Manning, N.O., Prilusky, J., et al., 2001. The protein data bank at Brookhaven. In: Rossmann, M.G., Arnold, E. (Eds.), International Tables for Crystallography. Volume F. Crystallography of Biological Macromolecules. Kluwer Academic Publishers, Dordrecht, pp. 649–656.

Wang, F., Herzig, C.T., Chen, C., Hsu, H., Baldwin, C.L., et al., 2011. Scavenger receptor WC1 contributes to the gammadelta T cell response to Leptospira. Mol. Immunol. 48, 801–809.