J Struct Funct Genomics (2012) 13:57–62
DOI 10.1007/s10969-012-9133-7

# The Protein Structure Initiative Structural Biology Knowledgebase Technology Portal: a structural biology web resource

**Lida K. Gifford · Lester G. Carter · Margaret J. Gabanyi · Helen M. Berman · Paul D. Adams**

Received: 18 October 2011 / Accepted: 5 March 2012 / Published online: 6 April 2012


**Abstract** The Technology Portal of the Protein Structure Initiative Structural Biology Knowledgebase (PSI SBKB; http://technology.sbkb.org/portal/) is a web resource providing information about methods and tools that can be used to relieve bottlenecks in many areas of protein production and structural biology research. Several useful features are available on the web site, including multiple ways to search the database of over 250 technological advances, a link to videos of methods on YouTube, and access to a technology forum where scientists can connect, ask questions, get news, and develop collaborations. The Technology Portal is a component of the PSI SBKB (http://sbkb.org), which presents integrated genomic, structural, and functional information for all protein sequence targets selected by the Protein Structure Initiative. Created in collaboration with the Nature Publishing Group, the SBKB offers an array of resources for structural biologists, such as a research library, editorials about new research advances, a featured biological system each month, and a functional sleuth for searching protein structures of unknown function. An overview of the various features and examples of user searches highlight the information, tools, and avenues for scientific interaction available through the Technology Portal.

**Keywords** Database · Protein · Protein production · Structural biology · Structural genomics · Technology

**Abbreviations**

| | |
|---|---|
| PSI | Protein Structure Initiative |
| SBKB | Structural Biology Knowledgebase |
| PSI–MR | PSI:Biology–Materials Repository |

## Introduction

The Protein Structure Initiative began in 2000 and during the first two phases of the initiative, PSI-1 and PSI-2 Centers focused on rapidly determining protein structures on a genomic scale and developed many tools and technologies for this purpose [1]. The Structural Genomics Knowledgebase was established in 2008 to provide a centralized access point for data and technological information garnered from the PSI effort. The Knowledgebase combined these advances with publicly available resources to become a comprehensive web resource intended to enable biological research [2]. Shortly after its inception, the Knowledgebase entered into collaboration with the Nature Publishing Group to become a gateway site to deliver editorial content as well as provide access to the PSI data and public resources. At the beginning of third phase of the initiative, PSI:Biology, the website changed its name to the Structural Biology Knowledgebase (SBKB; http://sbkb.org). The SBKB continues to offer an array of resources for structural biologists, such as a research library, structural biology updates, a featured biological system each month, and a Functional Sleuth for searching protein structures of unknown function.

L. K. Gifford · P. D. Adams (✉)
Physical Biosciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, MS 64R0121, Berkeley, CA 94720, USA
e-mail: pdadams@lbl.gov

L. G. Carter
Beamline 4.2, Stanford Linear Accelerator Center, Menlo Park, CA 94025, USA

M. J. Gabanyi · H. M. Berman
Department of Chemistry & Chemical Biology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

The SBKB is also a portal of portals, providing access to discreet websites dedicated to experimental data tracking, protocols, materials, annotation, modeling, PSI publications, and technology [1]. The PSI SBKB Technology Portal (http://technology.sbkb.org/portal/) was established during the second phase of the initiative and functions as a repository for the technological developments catalyzed by all stages of the PSI program. This web resource is comprised of over 250 tools and technology summaries representing each step of the protein structure determination pipeline, the majority of which have been developed by the PSI and are in use at PSI Centers. In addition, the Technology Portal functions as a conduit for communicating technological advances to researchers at PSI Centers and the wider scientific community, creating opportunities for collaboration among scientists. Recently, the Technology Portal home page was redesigned to present all features in one easy-to-navigate page. In this article, we will describe the features of the Technology Portal, such as searching the technology pages and accessing the online resources in the Technology Toolbox, which can be used to enable structural biological research.

## Navigating the Technology Portal

The Technology Portal can be reached directly (http://technology.sbkb.org/portal/) or by visiting the SBKB (http://sbkb.org), selecting the Methods Hub, and clicking on the Technology Portal link. In addition, information from the portal can be accessed via the SBKB using a keyword search. Once at the Technology Portal, users can access all of the technology pages and online resources through the homepage. Following is a full description of the home page content and functionalities (Fig. 1).

### Architecture of the Technology Portal

The Technology Portal website is hosted using Apache HTTP Server software [3]. This web resource is constructed using Django 1.2, a Python-based web framework [4] that queries the data stored in a SQLite database [5] and displays these data as technology pages. The open source software package, Django, has pre-developed tools for site administration and provides a user interface for managing database information content, such as entering new technologies, uploading figures, and editing existing records.

### Searching the Technology Portal

The Technology Portal home page offers two ways to search the technology pages: plain text and by experimental stage (Fig. 2a). A text search of two or more words will automatically perform an AND search; placing terms in quotation marks will produce results where the words are adjacent in the text. Entering text in the search box will perform a search of all technology pages and return a list of results with the most recently edited pages listed first, prioritizing technology pages containing the newest information. Each technology is indexed by experimental stage, so querying by individual stage in the protein structure determination process can be used when a more general search is desired. The experimental steps currently indexed on the Technology Portal are: Target Selection, Reagents, Cloning, Protein Expression, Purification, Crystallography, NMR, Function/Annotation, Modeling, and Dissemination Tools. Selecting one of these categories from a drop-down menu will return an alphabetical listing of all records tagged for that particular process (Fig. 2b). The user can browse the list of returned items containing a title and summary of each technology page and a link to the full article. Once the desired article is located, the user can click the "More" link to get a technology page that contains a description of the technology, figures, and information regarding publication, whom to contact, web links, and availability, if applicable (Fig. 2c).

### Exploring the Technology Toolbox

#### *Featured Technology*

The Featured Technology highlight is periodically updated and focuses on technologies of interest to structural biologists. This section has been used to increase awareness of a variety of specific PSI center-developed technologies; Biosync, a widely used X-ray crystallography web resource (http://biosync.sbkb.org) [6]; and the many robotic and automation techniques in use at two of the PSI Centers for Membrane Protein Structure Determination, NYCOMPS and TEMIMPS.

#### *Web-based Resources*

The Web-based Resources section allows the user to browse a list of technology pages describing web servers or web-based tools that can be employed by structural biologists to design, predict, and model results. Over 50 technology pages detail and provide links to online tools for designing [7, 8] and predicting experimental results [9–12], comparing [13, 14] and annotating protein structures [15–17], ligand binding and searching [18, 19], modeling [20–22], data analysis and management [23–27], and more. This list includes technology pages detailing databases for designing experiments [9], modeling [20, 21], and determining function [28, 29], as well as the PSI Centers for High-

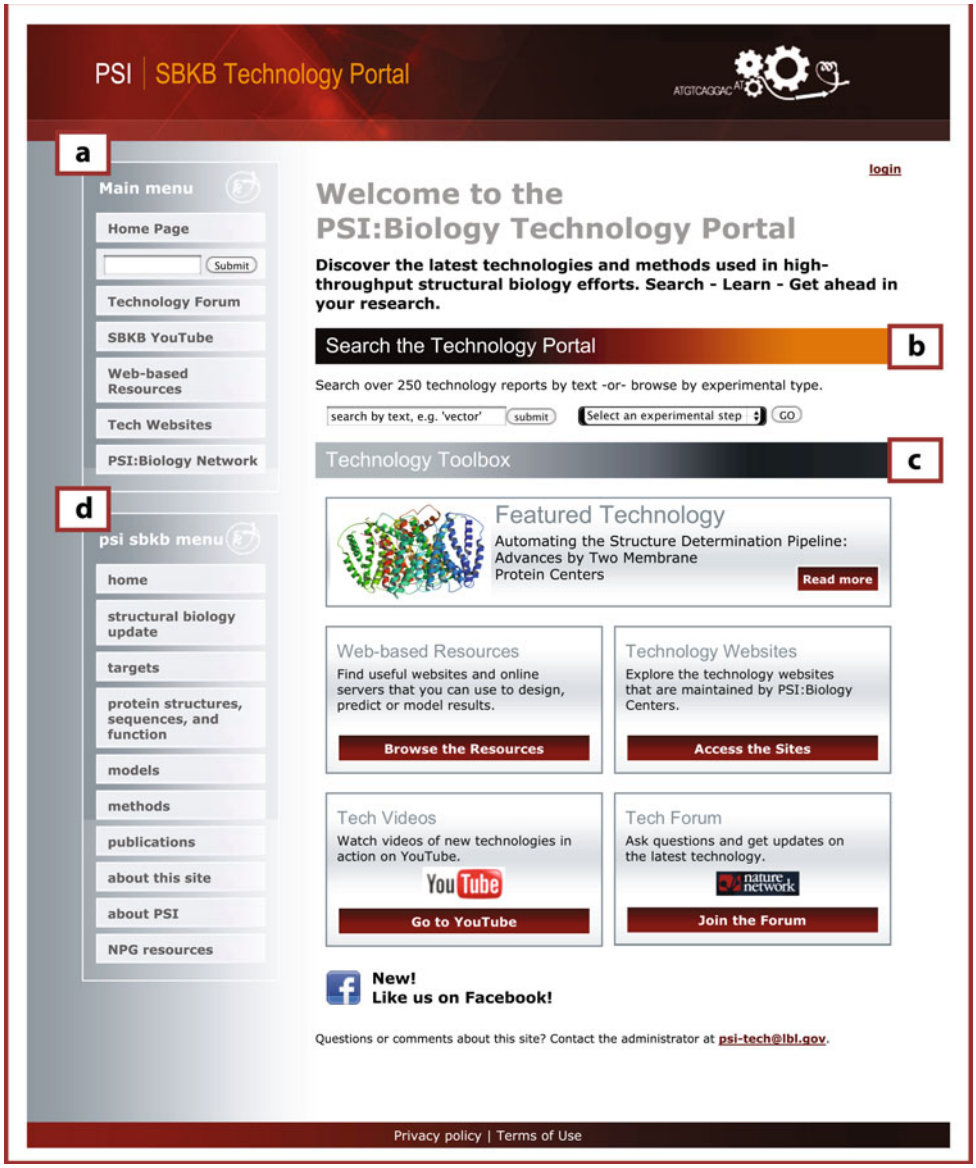


**Fig. 1** The PSI SBKB Technology Portal home page. The access point to the website has several useful features (clockwise from the *top left*): **a** Main Menu containing links to the main features of the Technology Portal and a quick keyword search box; **b** Search area allowing users to access information by keyword or experimental step; **c** The Technology Toolbox, containing a dynamic article on a featured technology, and links to web-based resources, technology pages, the Technology Portal YouTube channel, and the Technology Forum group hosted by the Nature Network; **d** The PSI SBKB Main Menu linking users directly to myriad resources that are encompassed by the SBKB. The Main Menu and PSI SBKB Menu are static and show up on all pages to allow users a shortcut to other information in the Technology Portal or SBKB

Throughput Structure Determination structure galleries [30–34].

*Technology Websites*

All four of the PSI:Biology Centers for High-Throughput Structure Determination maintain web pages describing the software, tools, and other technologies they have developed to relieve bottlenecks in protein structure determination. In addition, over half of the Centers for Membrane Protein Structure Determination include technology and methods pages on their websites that describe advances in expressing, purifying, screening, and determining structures of membrane proteins. Clicking on the link in this portion of the Toolbox returns a complete list of the technology links on PSI Center websites, allowing for easy access to more information and the opportunity to browse the technologies in the context of the full scope of distinct high-capacity structure determination projects.

*Technology Videos*

The PSI SBKB Technology Portal has established a channel on YouTube (http://www.youtube.com/user/sbkbtech) to house videos of technologies related to various stages of high-throughput protein structure determination. Presently, there is a movie of a large-scale fermenter in action [35], as well as narrated demonstrations of technical advances in crystallography [36–38] and NMR [7].

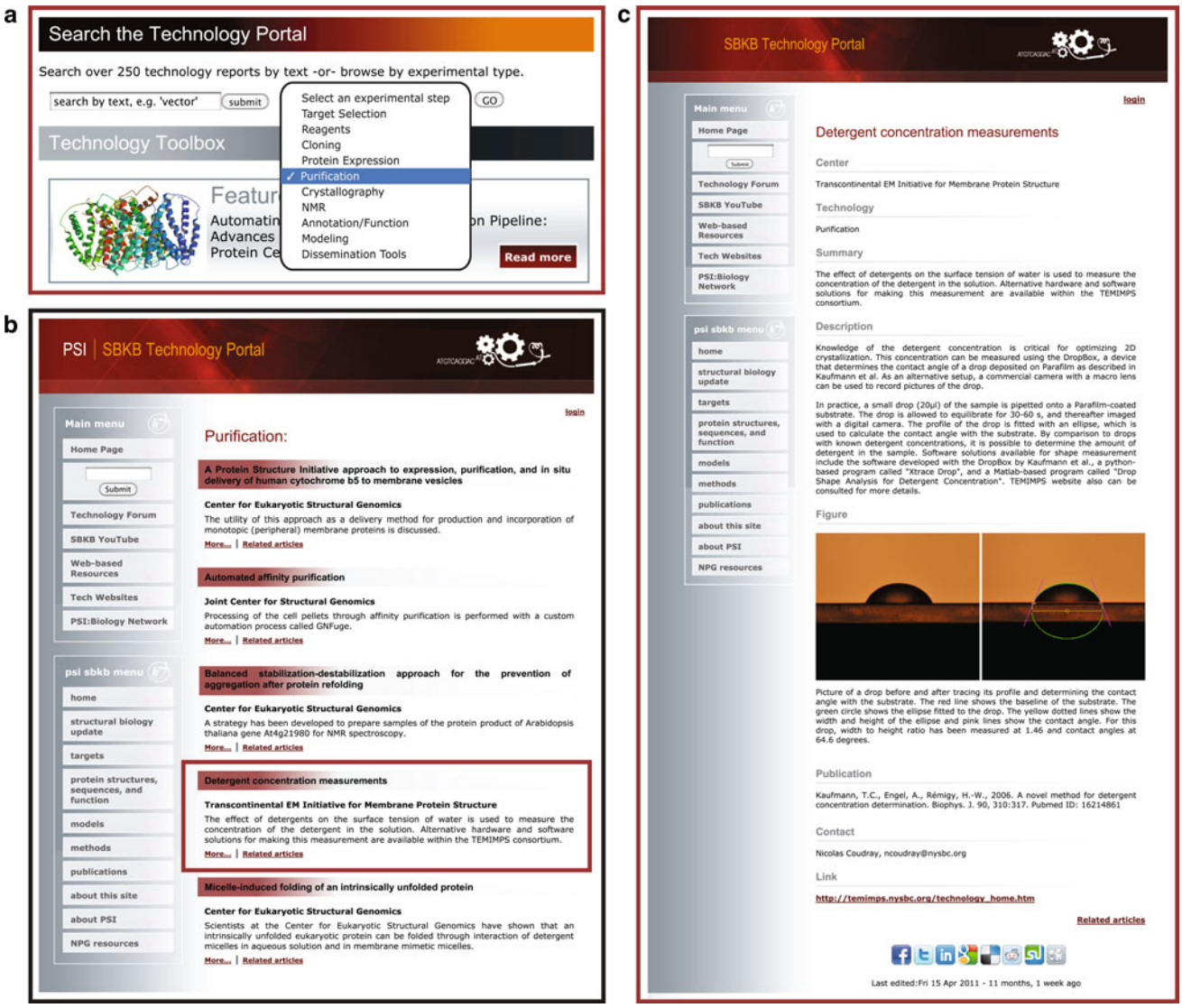


**Fig. 2** Highlight of the main search area and example of experimental step search results. **a** Close-up of the Search area from the home page showing the experimental step drop-down menu, with Purification highlighted. **b** The first page of results received when choosing "Purification" and clicking "Go" on the main page. Each experimental step results page returns an alphabetical listing of technology pages displaying the title, summary, and links to the full technology report and related articles. **c** The highlighted page is a typical technology page and contains the following sections: Title, PSI center, Summary, Description, Figure and legend, Publications, Contact information, and a Link

*Social networking*

The Technology Portal has several opportunities for social networking. The PSI SBKB Technology Portal Forum group (http://network.nature.com/groups/psikb_tech/) is hosted on the Nature Network. This forum provides a place for users of the portal to connect on a professional level, get more information, receive email updates of technology posts, and to ask the community questions about structural biology technology. The PSI SBKB Technology Portal recently established a Facebook page (http://www.facebook.com/pages/PSI-SBKB-Technology-Portal/209476552428114?sk=wall). This web page allows for the Technology Portal to communicate information about new content or other changes to the website directly to users who have "liked" the page. Participation in these social outlets makes it possible to establish a line of communication with users who have indicated interest in learning more about technologies that can be useful in their research.

In addition to these opportunities for information transfer, all technology pages can be posted to social networking accounts on Facebook, LinkedIn, and Twitter, allowing scientists to share what they think is interesting with their friends and colleagues directly from the

Technology Portal. The pages can bookmarked on Google and del.icio.us for easy access from the browser at a later date and each page can be shared or recommended to a wider audience on stumbleupon, reddit, and digg it.

Menus

The home page has two menus on the left-hand side that are accessible to the users regardless of which page they are viewing: the Main Menu and the PSI SBKB Menu (Fig. 1a, d). The links in the Main Menu provide an easy way to quickly navigate back to the home page or access the Technology Toolbox pages from anywhere in the website. Another handy feature of the Main Menu is the text search box that can be used to perform a quick keyword search from any page of the website. The PSI:Biology Network link allows users to peruse the list of PSI:Biology Centers and Consortia for High-Throughput-Enabled Structural Biology Partnerships and navigate to their landing pages housed on the SBKB.

The PSI SBKB menu is similarly available and gives users the ability to hop to the SBKB for further information or tools. Once at the PSI SBKB, the user can still access information in the Technology Portal by performing a keyword search on the SBKB home page.

## Synergy with other PSI:Biology resources

The Technology Portal works closely with two other PSI:Biology web resources by using link-outs to connect useful information between sites. The first collaboration was established with the PSI Publications Portal (http://olenka.med.virginia.edu/psi/), which contains all publication information and statistics for the more than 1,600 peer-reviewed articles that have been published by the PSI over the past 10 years [1]. If a publication listed on a technology page can be found in the PSI Publications Portal, the user has the opportunity to click on a link to search that website for further information about the reference. Reciprocally, all descriptions of articles in the Publications Portal referenced by the Technology Portal contain a link to the appropriate technology page. This allows users to access more information about a given technology quickly and easily directly from the Publications Portal.

Crosslinks have also been established with the other PSI resource center, the PSI:Biology-Materials Repository (PSI–MR, http://psimr.asu.edu/) [39, 40]. The PSI–MR stores, maintains, and distributes PSI-created protein expression plasmids and vectors. As of October 2011, the materials repository has over 50,000 PSI plasmids and 85 empty vectors available for distribution. Several of the technology pages are dedicated to describing the different types of vectors designed and used by PSI Centers [41–47]. Each one of those pages informs users that the materials are available from the PSI–MR and provides users with a link to the page that describes these empty vectors in more detail or takes them directly to an order page. In addition, at several points on the Empty Vectors page (http://psimr.asu.edu/EmptyVectors.html), the PSI–MR directs the user to the Technology Portal for more information and provides a link to the appropriate technology page.

## Conclusion

The Technology Portal is a resource dedicated to capturing and highlighting technological advances that are instrumental to enabling structural biological research. This is accomplished by maintaining a dynamic web site, interacting with PSI Centers and members of the wider scientific community, and using the web portal to disseminate knowledge and provide tools scientists can take and use in their research. We welcome feedback from the community at psi-tech@lbl.gov.

**Acknowledgments** The authors would like to thank the Protein Structure Initiative PIs, researchers, and Structural Biology Knowledgebase members for their collaboration and support of the Technology Portal. We thank Ralf Grosse-Kunstleve for server maintenance and technical support and Jeffrey Headd for helpful discussions. The Technology Portal is a resource center within the Protein Structure Initiative and is supported by grant U01GM093324 from the National Institute of General Medical Sciences. This work was supported in part by the US Department of Energy under Contract No. DE-AC02-05CH11231.

## References

1. Gabanyi MJ et al (2011) The structural biology knowledgebase: a portal to protein structures, sequences, functions, and methods. J Struct Funct Genomics 12(2):45–54
2. Berman HM et al (2009) The protein structure initiative structural genomics knowledgebase. Nucleic Acids Res 37(Database issue):D365–D368
3. Apache HTTP Server Project (2011) Available from: http://httpd.apache.org/
4. Django: a Python web framework (2005) Available from: http://www.djangoproject.com
5. SQLite 3: SQL database engine. (2004) Available from: http://www.sqlite.org/
6. Kuller A et al (2002) A biologist's guide to synchrotron facilities: the BioSync web resource. Trends Biochem Sci 27(4):213–215
7. Rossi P et al (2010) A microscale protein NMR sample screening pipeline. J Biomol NMR 46(1):11–22
8. Everett JK, Acton TB, Montelione GT (2004) Primer Prim'er: a web based server for automated primer design. J Struct Funct Genomics 5(1–2):13–21
9. Bowers PM et al (2004) Prolinks: a database of protein functional linkages derived from coevolution. Genome Biol 5(5):R35

10. Price WN 2nd et al (2009) Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. Nat Biotechnol 27(1):51–57
11. Goldschmidt L, Cooper DR, Derewenda ZS, Eisenberg D (2007) Toward rational protein crystallization: a web server for the design of crystallizable protein variants. Protein Sci 16(8): 1569–1576
12. Overton IM, van Niekerk CA, Barton GJ (2011) XANNpred: neural nets that predict the propensity of a protein to yield diffraction-quality crystals. Proteins 79(4):1027–1033
13. Gipson B, Zeng X, Stahlberg H (2007) 2dx_merge: data management and merging for 2D crystal images. J Struct Biol 160(3): 375–384
14. Gipson B, Zeng X, Zhang ZY, Stahlberg H (2007) 2dx–user-friendly image processing for 2D crystals. J Struct Biol 157(1): 64–72
15. Ellrott K et al (2011) TOPSAN: a dynamic web database for structural genomics. Nucleic Acids Res 39(database issue):D494–D496
16. Krishna SS et al (2010) TOPSAN: use of a collaborative environment for annotating, analyzing and disseminating data on JCSG and PSI structures. Acta Crystallogr Sect F Struct Biol Cryst Commun 66(Pt 10):1143–1147
17. Weekes D et al (2010) TOPSAN: a collaborative annotation environment for structural genomics. BMC Bioinform 11426
18. Capra JA, Laskowski RA, Thornton JM, Singh M, Funkhouser TA (2009) Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. PLoS Comput Biol 5(12):e1000585
19. Kumar A et al (2010) Ligands in PSI structures. Acta Crystallogr Sect F Struct Biol Cryst Commun 66(Pt 10):1309–1316
20. Pieper U et al (2009) MODBASE, a database of annotated comparative protein structure models and associated resources. Nucleic Acids Res 37(database issue):D347–D354
21. Pieper U et al (2011) ModBase, a database of annotated comparative protein structure models, and associated resources. Nucleic Acids Res 39(database issue):D465–D474
22. van den Bedem H, Lotan I, Latombe JC, Deacon AM (2005) Real-space protein-model completion: an inverse-kinematics approach. Acta Crystallogr D Biol Crystallogr 61(Pt 1):2–13
23. Zolnai Z et al (2003) Project management system for structural and functional proteomics: sesame. J Struct Funct Genomics 4(1): 11–23
24. Baran MC et al (2006) SPINS: a laboratory information management system for organizing and archiving intermediate and final results from NMR protein structure determinations. Proteins 62(4):843–851
25. Baran MC, Moseley HN, Sahota G, Montelione GT (2002) SPINS: standardized protein NMR storage. A data dictionary and object-oriented relational database for archiving protein NMR spectra. J Biomol NMR 24(2):113–121
26. Bertone P et al (2001) SPINE: an integrated tracking database and data mining approach for identifying feasible targets in high-throughput structural proteomics. Nucleic Acids Res 29(13): 2884–2898
27. Goh CS et al (2003) SPINE 2: a system for collaborative structural proteomics within a federated database framework. Nucleic Acids Res 31(11):2833–2838
28. Nederveen AJ et al (2005) RECOORD: a recalculated coordinate database of 500+ proteins from the PDB using restraints from the BioMagResBank. Proteins 59(4):662–672
29. Huang YJ et al (2008) Targeting the human cancer pathway protein interaction network by structural genomics. Mol Cell Proteomics 7(10):2048–2060
30. Northeast Structural Genomics Consortium. NESG structure gallery available from: http://nmr.cabm.rutgers.edu:9090/gallery/jsp/Gallery.jsp
31. New York Structural Genomics Research Consortium. NYSGRC structure gallery available from: http://www.nysgrc.org/nysgrc/result.html
32. Joint Center for Structural Genomics. JCSG structure gallery. Available from: http://www.jcsg.org/prod/newscripts/structure_gallery/gallery.cgi
33. Midwest Center for Structural Genomics. MCSG structure gallery available from: http://olenka.med.virginia.edu/mcsg/phtml/structure_viewer.phtml
34. Elsliger MA et al (2010) The JCSG high-throughput structural biology pipeline. Acta Crystallogr Sect F Struct Biol Cryst Commun 66(Pt 10):1137–1142
35. Kreusch A, Lesley SA (2004) High-throughput cloning, expression, and purification technologies. In: Grandi G (ed) Genomics, proteomics, and vaccines. Wiley Press, UK, pp 171–184
36. Gerdts CJ et al (2006) Time-controlled microfluidic seeding in nL-volume droplets to separate nucleation and growth stages of protein crystallization. Angew Chem Int Ed Engl 45(48): 8156–8160
37. Zheng B, Gerdts CJ, Ismagilov RF (2005) Using nanoliter plugs in microfluidics to facilitate and understand protein crystallization. Curr Opin Struct Biol 15(5):548–555
38. Phizackerley RP, Cohen AE, Ellis PJ, Miller MD, Deacon AM (2002) An automated system to mount cryo-cooled protein crystals on a synchrotron beamline, using compact sample cassettes and a small-scale robot. J Appl Crystallogr 35:720–726
39. Cormier CY (2010) Protein structure initiative material repository: an open shared public resource of structural genomics plasmids for the biological community. Nucleic Acids Res 38(Database issue):D743–D749
40. Cormier CY et al (2011) PSI:Biology-materials repository: a biologist's resource for protein expression plasmids. J Struct Funct Genomics 12(2):55–62
41. PSI SBKB Technology Portal (2009) High-throughput protein expression. Available from: http://technology.sbkb.org/portal/page/1/
42. PSI SBKB Technology Portal (2009) LIC (Ligation-independent cloning) vectors. Available from: http://technology.sbkb.org/portal/page/36/
43. PSI SBKB Technology Portal (2009) New vectors for co-expression of proteins. Available from: http://technology.sbkb.org/portal/page/52/
44. PSI SBKB Technology Portal (2009) Customized expression vector platform. Available from: http://technology.sbkb.org/portal/page/98/
45. PSI SBKB Technology Portal (2009) Structural genomics methods applied to production of TEV protease. Available from: http://technology.sbkb.org/portal/page/107/
46. PSI SBKB Technology Portal (2009) A combined approach to improving large-scale production of tobacco etch virus protease. Available from: http://technology.sbkb.org/portal/page/164/
47. PSI SBKB Technology Portal (2009) One-plasmid tunable co-expression for mycobacterial protein-protein interaction studies. Available from: http://technology.sbkb.org/portal/page/167/