

nature Structural biology

may 1999
vol. 6 no. 5

molecular form & function

Would you use an e-print server?

The community is buzzing with discussions about the possibility of developing an electronic preprint, or e-print, server for life science research — a single site to house all types of research papers, one that would be open to the public over the internet¹⁻³. Those pushing for its establishment, most notably Patrick Brown of Stanford University and David Lipman of the National Center for Biotechnology Information, envision that such a site could completely replace traditional avenues of publication in the next few years. Not everyone agrees that replacing the traditional journal system is feasible or desirable, and thus an e-print server that complements instead of replaces academic journals may be a more likely possibility. Currently, a proposal for the establishment and curation of a life science e-print site is under development, and the planning phase is being encouraged by Harold Varmus, director of the National Institutes of Health.

The physical sciences community already has an e-print server, the Los Alamos National Laboratory e-print archive (<http://xxx.lanl.gov/>). Anyone can upload an article onto this site and although papers are grouped into categories, there is no peer review or editing. However, revised versions of papers (with additional data) can be posted at any time, and it is possible to note that a previously posted version of the paper has been published in a peer-

reviewed journal. Paul Ginsparg created this archive in 1991, and its use has increased dramatically in the last few years — monthly submissions are currently over 2,000. With this site as a role model, Lipman and Brown hope to make their vision a reality in the very near future.

This talk of a 'journal revolution' comes on the heels of remarkable activity on the web. In the past five years, nearly every traditional print journal (including *Nature Structural Biology*) has developed a web site. Although the sites vary widely in appearance and features, some minimal standards for quality and utility have been set: on-time posting of the latest issue; public access to the table of con-

tents and, for subscribers, the full text of all printed materials (current and archived); search capabilities; and the option of printing PDF files (despite high levels of web usage, a large proportion of readers still seem to prefer reading printed articles). In addition to the proliferation of individual web sites, the amount of electronically available scientific data, such as genome sequences and structural coordinates, has increased dramatically, and the need for consolidation of these data has never been greater. Furthermore, many types of images (for example, extremely large sequence comparisons, or 'movies' of migrating cells or rotating protein models) would be difficult or impossible to accommodate in print, yet many journal web sites are not equipped to accept such data. (Although at *Nature Structural Biology*, we do encourage submission of 'web specials' that would add value to published papers.) Thus, the idea of a web site that would tie all of these aspects of biology together and allow a wider

What do you think about the proposed e-print server?

Please send comments by email to nsb@natureny.com

(with the subject heading 'e-print feedback').

variety of images seems, on the surface, quite appealing. But will such a site really serve its purpose? It could be argued that a 'mega-site' could actually make it more difficult for scientists to obtain pertinent, high-quality information. Two questions are relevant to this issue: will there be any form of quality control of material posted to the e-print server, and will everyone readily contribute papers?

A major concern about such a site, especially in the burgeoning molecular and biomedical field, is information overload. The established system of peer-reviewed journals provides a filtering device that endeavors to minimize this problem. Editors of a high-quality traditional journal go to considerable lengths to be a proxy for their readers by selecting papers for appropriateness with regard to the journal's repertoire, and by attempting to ensure quality and originality through rigorous, anonymous peer review. Editors can insist on — and have the resources and authority to police — things such as accessibility of the primary data, for example, by requiring structural coordinate deposition into the Protein Data Bank. Furthermore, journal editors can demand, prior to acceptance, that papers be made more accessible to a wider audience. No one would argue that the current system of editing and peer review is perfect, but if carried out by all parties with the best of intentions, then it is an extremely useful — and value-adding — process.

Although specific details of the proposal are not yet available, the life science e-print server may have some form of simple editing, but it will most likely not be of the type currently performed by editors of traditional academic journals. Most journals require substantial revisions of manuscripts that have been anonymously peer-reviewed. It is anticipated that sheer volume would preclude such a practice on the e-print server and that instead the server staff may simply group papers into categories. It has been suggested that only certain papers of particular interest would be peer reviewed, openly, with comments posted on the web and signed by the reviewers. Such open peer review — in which some reviewers may not feel comfortable revealing their true opinions — is presumably less rigorous, although some might argue more fair, than anonymous peer review. As a result, the quality of items posted is likely to vary widely, with a naive reader having little information, especially for papers with no reviews, on which to judge the validity of the results or the discussion. Thus, it is possible that the egalitarian nature of the site may actually be detrimental to its goal of distributing useful, high quality information.

For a life science e-print server to be successful, researchers will have to see the service as a legitimate form of scientific communication and will have to contribute and access papers. The usefulness of the site as with any type of journal, print or electronic, is likely to lie in large part with the quality of papers it attracts, and with its organization. It has been suggested that the e-print postings could be linked to Medline, making it easier for researchers to target relevant papers that have been printed either in journals or only on the e-print server. However, the range and quality of papers appearing only on the e-print server may be limited. It seems worth considering that biomedical research communities may be fundamentally different from the physical sciences communities in certain ways that would make a life science e-print server less attractive for submissions. For example, for better or for worse, competitiveness is probably more prevalent in the biomedical community. Thus, despite the good intentions of the organizers of the site to reduce the competitive nature of science, some of the most current and preliminary results may not be posted. Relevant to this is whether or not papers posted only on an e-print server that has limited peer review would be considered valid publications, in the traditional sense, by funding agencies and potential employers.

Structural biologists are traditionally viewed as more up-to-date on computer-related developments, and thus may be more interested in using an e-print server. In fact, there have been limited proposals recently, such as in the field of structural genomics, to use the web to share pre-publication information more readily^{4,5}. On the other hand, structural biology has been and continues to be a competitive enterprise, a reality that could perhaps limit posting of new structural determinations onto the site.

We do not yet know details of the proposal for the e-print server. However, they could be released very soon, and we are interested in *Nature Structural Biology* readers' views on these issues. Do you see distinct advantages to such a site? Do you have any particular concerns? Would you submit your own papers to the e-print server? Would you routinely visit the site? We encourage you to submit comments by email to nsb@natureny.com (with the subject heading 'e-print feedback') and look forward to continued discussion of these issues.

1. Marshall, E. *Science* **283**, 1610–1611 (1999).
2. Butler, D. *Nature* **397**, 91 (1999).
3. Editorial, *Nature Genet.* **21**, 341–342 (1999).
4. Editorial, *Nature Struct. Biol.* **5**, 1019 (1998).
5. Sali, A. *Nature Struct. Biol.* **5**, 1029–1032 (1998).