

Reminiscences of Phil Bourne (1953–2026) and the IUCr

Brian McMahon^{a*} and John R. Helliwell^b

^aInternational Union of Crystallography, 5 Abbey Square, Chester, United Kingdom, and ^bThe University of Manchester, Brunswick Street, Manchester M13 9PL, United Kingdom. *Correspondence e-mail: bm@iucr.org

Keywords: Phil Bourne; obituary.

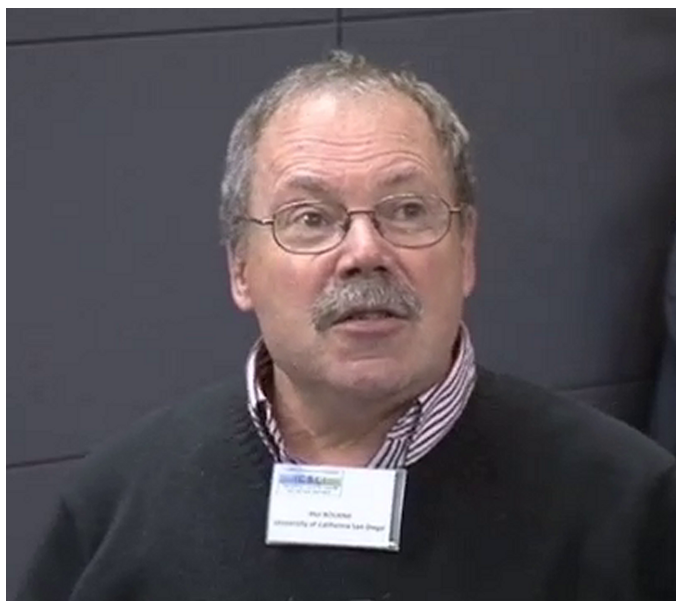


Figure 1
Phil Bourne in full flow at the 2010 Paris workshop on interactive journals.

Philip Bourne's death on 8 March 2026 at the age of 72 has prompted a large number of obituaries and appreciations (Turner, 2026; Mather, 2026; RCSB, 2026). In a wide-ranging career as data scientist and science administrator, he was constantly pushing the envelope of how the electronic age could benefit scientific information, using the latest tools in database management and curation, publishing, large-scale computing, visualization and social media.

Phil's early training was as a crystallographer. His PhD thesis from Flinders University in South Australia dealt with the binding of nucleic acid components to *cis*-platinum drugs, and it stimulated both his interest in biological structures and in the power of computational crystallography to solve, understand and eventually visualize three-dimensional macromolecules. This was to lead to his abiding interest in the management and interrogation of large volumes of structural data, and his early involvement in the novel field of bioinformatics (Bourne, 2017).

The combination of all these skills brought him into contact with a group actively developing more powerful ways of using the information stored in the Protein Data Bank (PDB) archive of biological macromolecular structures. When management of the PDB moved to Rutgers University, partly on the strength of software tools that Phil had helped to develop, he became an Associate Director of the PDB. Working at the San Diego Supercomputing Center (SDSC), he continued to develop powerful tools and workflows to improve database management, literature integration and visualization (RCSB, 2026).

Phil's subsequent career at the US National Institutes of Health, where he was leader of the Big Data to Knowledge Program, and the University of Virginia, where he was Founding Dean of the School of Data Science, are more fully described elsewhere (Mather, 2026; Mura *et al.*, 2023; National Institutes of Health, 2026; Mura, 2023; Bourne, 2021, 2026). However, throughout his career, he was always fascinated by the importance of structural biology, and cherished his early career in crystallography.

Our interactions with him as representatives of the International Union of Crystallography (IUCr) were most active in the early to mid 1990s.

Guy and Eleanor Dodson convened a 1993 workshop at the University of York to explore how best to develop the PDB, as the collection of archived protein structures was growing apace. Challenges of scale and complexity beset the existing storage and discovery platform, and also the established PDB file interchange format. Hopes were high that the novel CIF format adopted by IUCr journals for chemical structures could be extended and overcome the constraints of the PDB file. However, database specialists at the workshop considered that the way in which the items stored in a CIF file were defined was not rigorous enough to accommodate the requirements of a relational database platform, which was considered essential to manage the future scale and complexity of a next-generation archive. Yet the CIF idea still held great promise, and its adoption for protein structures alongside chemical structure descriptions would be important for the interoperability between these domains that subsequent history has shown to be so significant. How could the interchange format and database requirements be reconciled?

Phil hosted a follow-up workshop at Tarrytown in upper New York State six months later, to revisit the problem. In the intervening period, Phil had worked with Paula Fitzgerald, Helen Berman and Keith Watenpaugh to compile a list of data items (additional to those used in the core chemical CIF dictionary) necessary for describing a macromolecular structure determination experiment and its result. John Westbrook had taken the dictionary definition language that was used for core CIF, and extended it to make explicit the data typing and relations between different item definitions. He tightened the category groupings of related data items into a form that could be translated directly to relational database tables. The result was a specification that would fully meet the needs of the proposed new PDB that was, in due course, established at Rutgers University by Helen and John, with support from Phil at San Diego, and which today still flourishes as the primary archive site of the Worldwide Protein Data Bank.

This brilliant technical result did, however, result in a slightly different dialect of CIF (mmCIF, subsequently extended further as PDBx/mmCIF), and it required a third workshop in Brussels a year later to resolve the tensions between the chemical and macromolecular communities that had arisen as a result of the potential schism in moving away from the original CIF formulation. That workshop initiated a lengthy process of revision of the original CIF standard to reduce friction in the essential interoperability between chemical and macromolecular structures. That revisionist project will come to completion this year with the formal publication of the second edition of *International Tables* Volume G, *Definition and exchange of crystallographic data*.

All three of these workshops were crucial for the future of mmCIF and the PDB, but the one in Tarrytown stands out for the intensity of its intellectual interactions together with the cordiality of the personal relationships in the relaxed setting of the Hudson Valley in the fall. There is no doubt that Phil



Figure 2
Phil (right) in characteristic good humour performing the *Two Crystallographers from Yorkshire* sketch with Anthony Nicholls.

Bourne's enthusiasm and commitment, and his warm personality, contributed immeasurably to its success.

Once located in San Diego, Phil and his team worked on innovative software to facilitate the new data management workflows that the re-engineered PDB required. But he also became increasingly interested in the dissemination not just of data but of knowledge. SDSC became an early mirror site (at a time when international bandwidth was still a pinch point in information transfer) of the IUCr's comprehensive web pages for crystallography, established in collaboration with Howard Flack.

Phil was a member of the IUCr Commission on Crystallographic Computing between 1993 and 1999, serving as its Chair from 1996. In that role he ran, with Keith Watenpaugh, a very successful macromolecular crystallography computing school at Bellingham, Washington State. For the first time, the full proceedings of such a computing school were published on the still-fledgling World Wide Web (Bourne & Watenpaugh, 2017), combining not only conventional written accounts, but also links to other online resources. Remember that, commonplace as this is nowadays, it was a revolutionary initiative at the time. Not that Phil was ignorant of the importance of history in shaping the development of the future. His after-dinner entertainment with Anthony Nicholls at that workshop captures his sense of fun, but also his perspicacity and grounding in the significance of earlier pathfinders (Nicholls & Bourne, 2017; Fig. 2).

Around the same period, Phil started to become interested in visual representations of structures, and then in the question of how to link dynamic visualizations with traditional journal publications. He visited the IUCr offices in 1999 to demonstrate a tool to integrate structural data with visualizations and annotation, thus allowing easy authoring of an annotated structure exploration tool. While this project remained at the prototype stage, it led to a standalone Windows desktop implementation (Yang *et al.*, 2003) and informed the approach to integrated structure visualization followed by IUCr journals in their 3D-view visualizations that accompany every structure report, and also the Proteopedia platform (Prilusky *et al.*, 2007) that allows publication of annotated 'molecular tours', independently or as complements to journal articles.

In the following decade Phil became immersed in a number of initiatives in the scholarly publishing arena. Spurred by the desire to make scientific knowledge as widely accessible as possible, he became in 2005 the founding Editor-in-Chief of the open-access journal *PLoS Computational Biology*, and initiator of the *Ten Simple Rules* series in that journal. This was to become an extensive series of brief career development guides, touching on best practice in publishing, reviewing and disseminating scientific research. In 2007, he launched the now-dormant Scivee (*Scivee: Making science visible*; <http://www.scivee.tv>), a ‘scientific YouTube’ for authors to provide video presentations of their published work.

By 2010, when invited to an ICSTI workshop that we ran in Paris on *Interactive journals and the record of science* (McMahon, 2010), he offered a range of topics he wanted to talk about, including ‘semantic enrichment during authoring and post-publication’, ‘the interplay between databases and literature’, ‘the role of rich media in scientific discourse’ and ‘the need for author disambiguation and credit’, illustrating his great breadth of interest and involvement. In the end, his presentation, entitled *Experiences with rich media in the dissemination and comprehension of science*, was a thoughtful analysis of the direction of travel of scholarly communication, how it could benefit from still-evolving new media and tools, and how best to retrieve and retain the necessary knowledge from the ever-growing mountain of information and data that threaten to overwhelm the modern scientist. The coda to his presentation was a wish list of what a publisher should do for the scientist – to become ‘a contractor for all aspects of scholarly output’, providing a record and an appropriate level of access control for an idea, a hypothesis, an experimental description and data collection, as well as the final report. As modern tools of electronic laboratory notebooks, code workbooks and raw data repositories become increasingly integrated with the best scholarly publishing platforms, we see how insightful he was, and how much further we still have to travel to satisfy his ideal.

We also recall his attendance at the International Data Week 2016 meetings in Denver, where he was an enthusiastic advocate of the NIH initiative for preprints in general and bioRxiv in particular. He swept aside any worries that might arise about a lack of underpinning data on the basis that having a citable record was better than having no preprint at all.

After Phil left the University of California at San Diego we had little direct interaction with him, but our paths would often cross at conferences with publishing or data as their theme, and he was always supportive of IUCr initiatives in both those fields. He had great respect for the IUCr’s efforts to innovate within publishing and data management while always maintaining the highest standards, and we in turn are indebted to him for guiding us in the best direction to do so.

We remember Phil for his brilliance, for his enthusiasm, his ability to inspire, but most of all for his warm human qualities. We shall miss him.

References

- Bourne, P. E. (2017). *PLoS Biol.* **15**, e2002041.
- Bourne, P. E. (2021). *PLoS Biol.* **19**, e3001165.
- Bourne, P. E. (2026). *Enhancing Biomedical Research Through Sustainable Data Science Strategies at NIH*. <https://www.slideserve.com/paloma-bradley/data-philip-e-bourne-ph-d-associate-director-for-data-science-national-institutes-of-health>.
- Bourne, P. E. & Watenpaugh, K. (2017). Editors. *Crystallographic Computing 7, Proceedings from the Macromolecular Crystallography Computing School*, August 17–22, 1996, Western Washington University. <https://www.iucr.org/resources/commissions/computing/schools/school96>.
- Mather, M. (2026). *In memoriam: Philip Bourne, founding dean of the School of Data Science*. <https://news.virginia.edu/content/memorial-philip-bourne-founding-dean-school-data-science>.
- McMahon, B. (2010). *Inf. Serv. Use*, **30**, 1–16.
- Mura, C. (2023). *Interview with Dr. Philip E. Bourne, Data Science Honoree*. <https://datascience.virginia.edu/news/interview-dr-philip-e-bourne-data-science-honoree>.
- Mura, C., Candelier, E. & Xie, L. (2023). *Biomolecules*, **13**, 181.
- National Institutes of Health (2026). *In memoriam: Philip Bourne, founding dean of the School of Data Science*. https://www.linkedin.com/posts/national-institutes-of-health_in-memorial-philip-bourne-founding-dean-activity-7437500948924530688-VP2X.
- Nicholls, A. & Bourne, P. (2017). *The two crystallographers from Yorkshire sketch*. <https://www.iucr.org/resources/commissions/computing/schools/school96/banquet-humour>.
- Prilusky, J., Sussman, J. L. & Herraez, A. (2007). *Proteopedia*. <https://proteopedia.org>.
- RCSB (2026). *Remembering Philip E. Bourne*. <https://www.rcsb.org/news/feature/69b1615f8557ae0f261bc714>.
- Turner, S. D. (2026). *Remembering Phil Bourne. Personal reflections on a legendary scientist, leader, builder, and mentor*. <https://blog.stephenturner.us/p/remembering-phil-bourne>.
- Yang, P., Craig, P. A., Philip, E. & Bourne, P. E. (2003). *Features of Bioeditor*. <https://users.sdsc.edu/~pcraig/bioeditor/features.htm>.