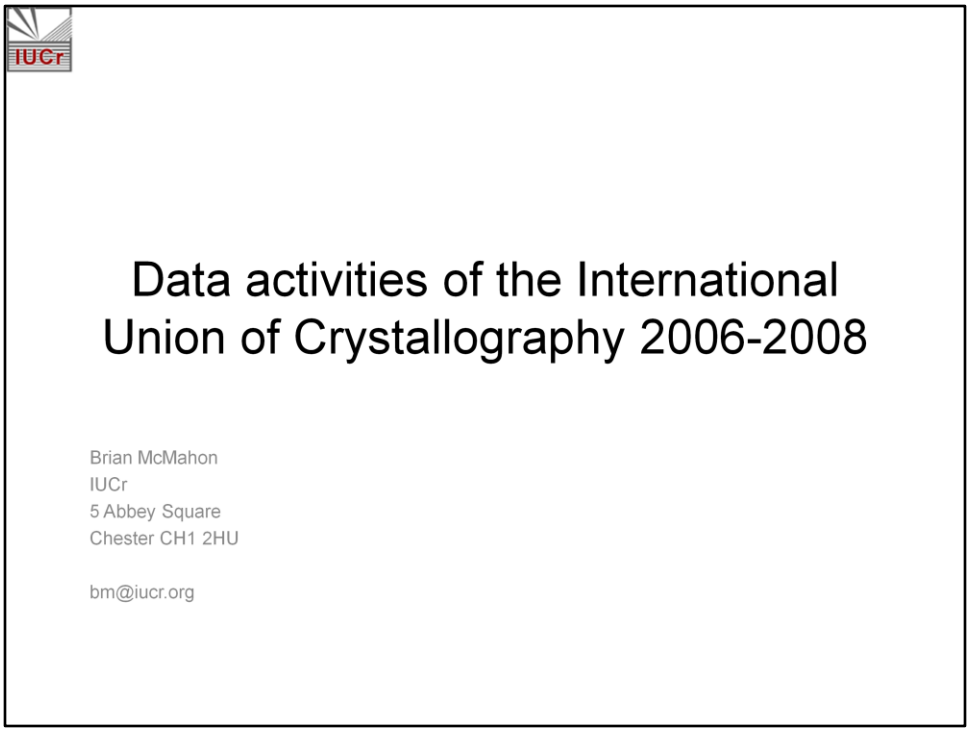
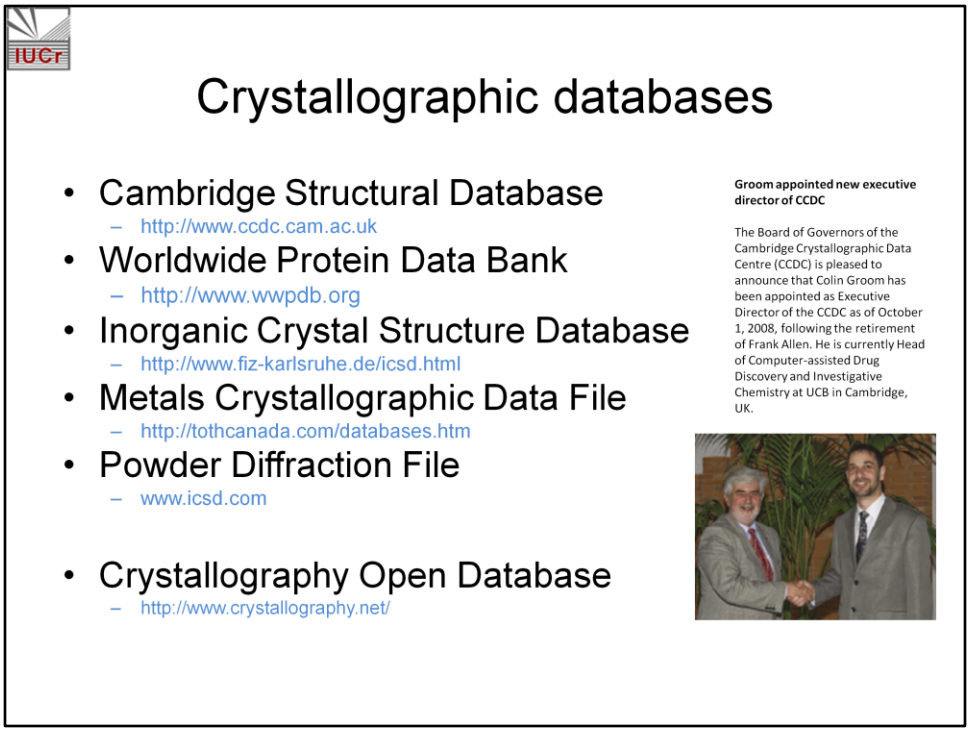


The International Union of Crystallography (IUCr) is a scientific union adhering to the International Council for Science (ICSU). Its objectives are to promote international cooperation in crystallography and to contribute to all aspects of crystallography, to promote international publication of crystallographic research, to facilitate standardization of methods, units, nomenclatures and symbols, and to form a focus for the relations of crystallography to other sciences.



Several independent databases exist that store and manage the results of crystal structure determinations and other non-structural data. Among the most important are

" Cambridge Structural Database for organic and metal-organic small-molecule structures and oligonucleotides (CSD) (www.ccdc.cam.ac.uk/);

" Worldwide Protein Data Bank for protein and nucleic acid structures (PDB) (www.wwpdb.org/) - a federation that includes the RCSB Protein Data Bank in the United States, the European Bioinformatics Institute MSD, the Japanese PDBj and BioMagResBank in the US;

" Inorganic Crystal Structure Database for inorganic materials (ICSD) (www.fiz-karlsruhe.de/icsd.html)

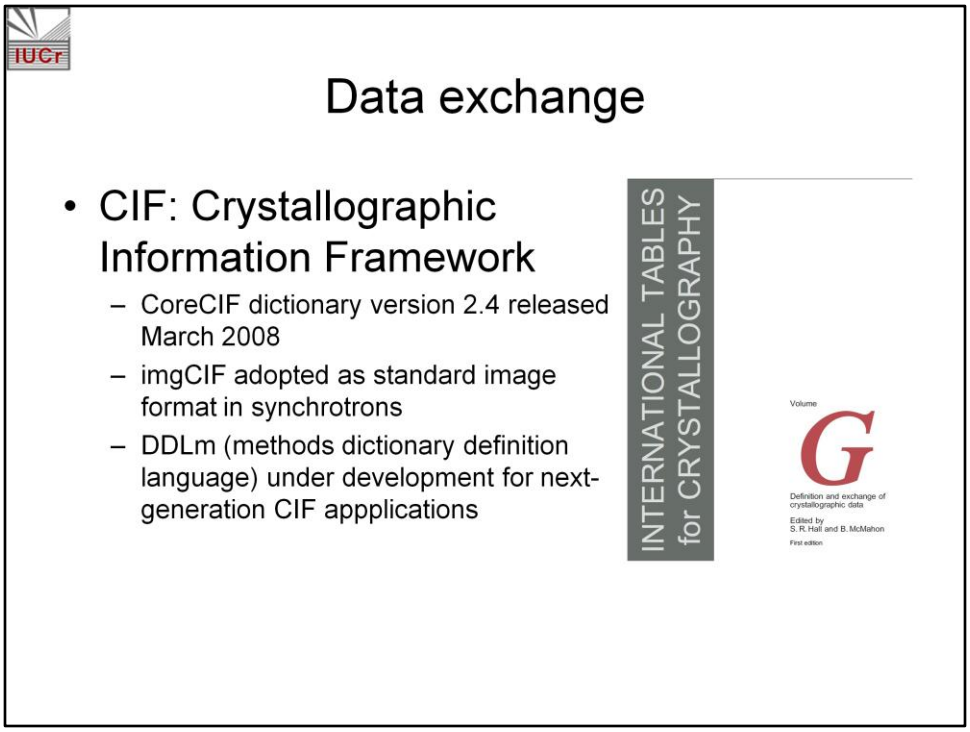
" Metals Crystallographic Data File for metals (CRYSTMET) (tothcanada.com/databases.htm);

" Powder Diffraction File (www.icsd.com).

These databases are curated by independent organisations, but the IUCr monitors their development through a standing Database Committee (CCD) that reports directly to the Union's Executive Committee.

Dr Colin Groom became Executive Director of CCDC on 1 Oct. 2008 upon Dr Frank Allen's retirement.

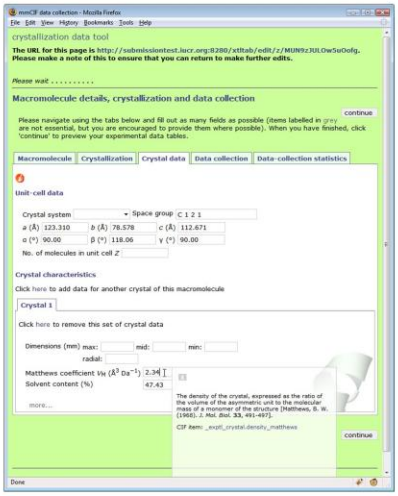
The Crystallography Open Database, set up in 2003, also exists as an independent open-access database of over 70,000 small-unit-cell structures, many of which are not published in the research literature.



Development continues on the Crystallographic Information Framework (CIF), the standard for archiving and exchanging crystallographic data. Version 2.4 of the core CIF data dictionary was released in March 2008; this dictionary defines items in the area of small-molecule structural science. Reviews of the CIF project were presented at an open meeting during the IUCr Congress in Osaka, Japan, during August 2008.

Macromolecular CIF (mmCIF)

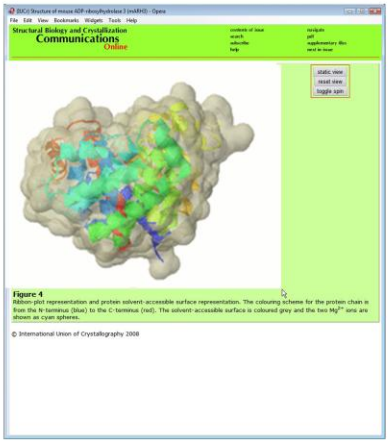
- Basis of data model for protein structures deposited in wwPDB
- Extensions for protein production, NMR, cryo-electron microscopy
- Streamlining of data deposition/publication
- Capture of crystallization conditions/protocols



Work continues between the IUCr and the RCSB Protein Data Bank to streamline deposition and publication of macromolecular structures, and to capture experimental data regarding protein crystallization, through development of required macromolecular CIF (mmCIF) data items for journal publication. The IUCr has developed a prototype web application for authors of research articles in biological crystallography to compose data-rich articles by extending or creating mmCIF data sets.

Enhanced figures

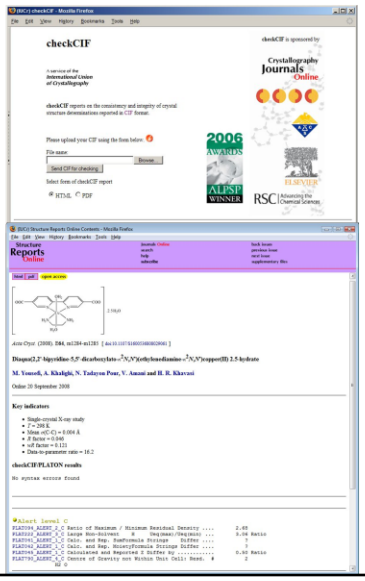
- Structure visualization using Jmol
  - Toolkit for authoring enhanced figures
    - Multiple views
    - Dynamic scripts
    - Allows reader to interact with model
  - Completely integrated with submission/review system
  - Completely integrated with journal production workflows
  - Static image for archiving illustration
  - Web and standalone applications
  - Shortlisted for ALPSP Publishing Innovation Award 2008



The IUCr's CIF editor (publCIF) for authors of small-molecule structure reports has been extended to allow the creation of interactive visualizations of crystal and molecular structures using the open-source application Jmol. A parallel web service allows authors to create interactive visualizations as integral parts of published structure report articles. This service was shortlisted for the 2008 Award for Publishing Innovation of the Association of Learned and Professional Society Publishers (ALPSP), and was demonstrated during the CODATA 2008 Conference.

Data validation

- checkCIF
  - Reports on data integrity and self-consistency
  - Deviations from crystallographic/chemical norms
  - Free web service
  - Used prior to article submission
  - Incorporated in submission/review system
  - Reports published in online *journal Acta Cryst. E*
- Sponsored by other publishers/databases
- Possible provision of checking services to other publishers

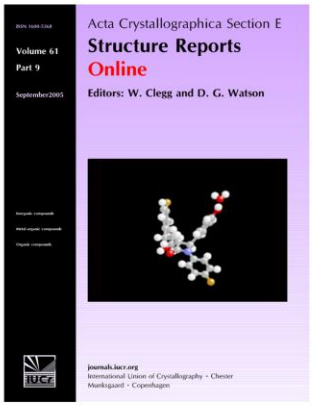


All structural data sets published in IUCr journals have since 1990 been checked for internal consistency by software capable of reading CIF submission or deposit files directly. The checking procedures are published on the web, and a public service to return a standard report on structures subjected to these checks has been established at http://checkcif.iucr.org. This service continues to function as a community standard for reviewing and assessing the consistency and quality of small-molecule and inorganic structure determinations. The criteria for issuing validation alerts in cases of suboptimal data quality or deviation from recognised chemical norms are under active review by an IUCr Working Group.

There have been initial discussions on providing services for checking crystallographic data to a number of other publishers.

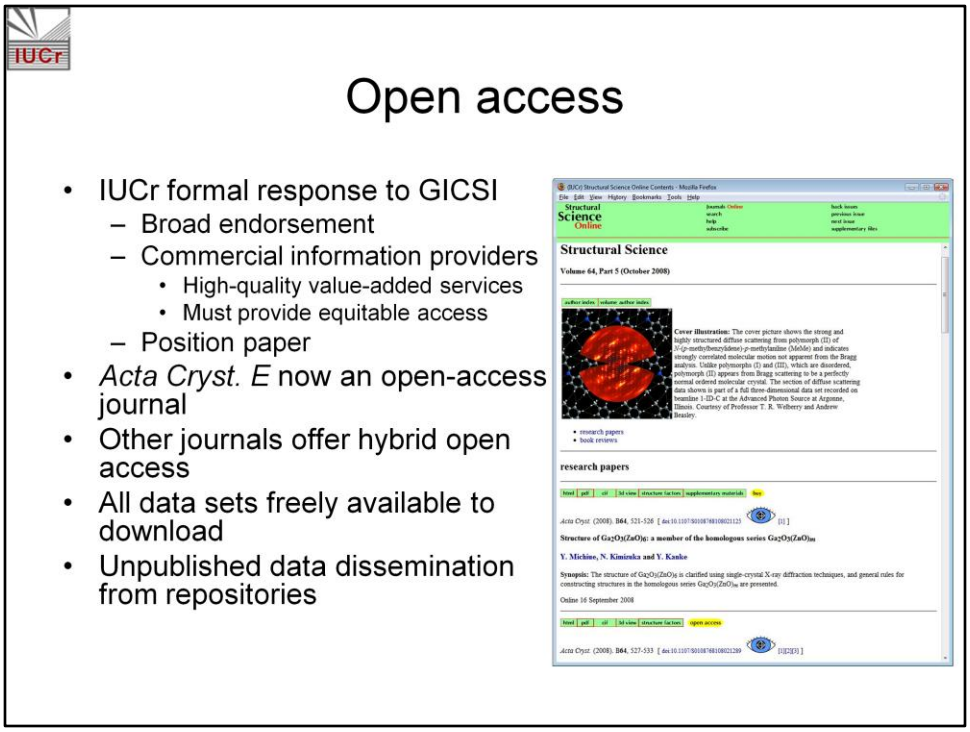
Electronic publishing

- Acta Crystallographica Section E: Structure Reports Online
  - Open-access since January 2008
  - Author fee $150
  - 300 articles/month
  - 'Publication package'
    - Brief peer-reviewed summary report
    - More extensive commentary, figures, tables
    - Enhanced interactive figures
    - checkCIF validation report
    - Full structural data, structure factors, Rietveld profiles
- Other IUCr journal titles
  - Free access to CIF, structure factors, other supplementary data
  - All structures processed by checkCIF during peer review



The IUCr publishes seven primary research journals in crystallography, and an eighth covering the technology, instrumentation and uses of synchrotron radiation. The online-only journal Structure Reports Online was launched in 2001 to disseminate short reports of individual crystal and molecular structures, and has grown rapidly to become the IUCr's largest journal in terms of numbers of articles and pages published. In the current climate of diminishing library budgets, it was decided that the only economically viable method of sustaining the journal's growth was to convert it to an author-pays open-access title from January 2008. After an initial drop in the submission rate, submissions have recovered well and the journal remains popular. The routine nature of the articles and their heavy dependence on reporting structural data allow the journal production workflow to be highly automated, and the open-access fee to authors is considerably lower than in most other publications.

The IUCr continues to collaborate with the Cambridge Crystallographic Data Centre and the Inorganic Crystal Structure Database to check new submissions for prior publication.



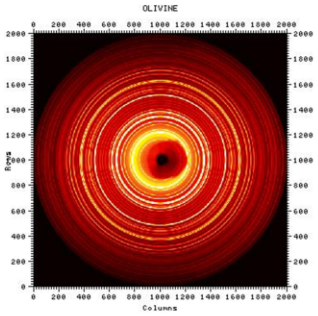
The IUCr responded in 2006 to the initial prospectus for the Global Information Commons for Science Initiative (GICSI) with a formal statement and a position paper. Both remain on the IUCr web site and inform the Union's policy directions in data management. During the course of the last two years, the decision was taken to operate one of the IUCr's journals as an open-access title (see section on Electronic Publishing above).

Experimental data sets associated with articles in all the IUCr journals are freely available for download as supplementary files; access to them is not restricted to journal subscribers.

Collaboration continues between the IUCr and a number of service crystallography facilities who are working with national scientific computing grid funding agencies to collect and store data sets. A one-day workshop on New routes to crystallographic data publication took place during the IUCr Congress in August 2008 to further this project. Work is beginning to analyse the requirements for archiving and providing access to crystallographic data managed within a federation of subject-specific repositories and databases.

Data preservation/archiving

- Journal publications
  - Supplementary data depositions with DOIs
  - Archiving arrangements with KB (articles), PubMedCentral
- Curated databases
- Laboratory/institutional repositories
  - E-Crystals
  - ReciprocalNet
  - Cambridge Chemistry Department
- Experimental data deposition
  - Structure factors mandatory in IUCr journals
  - Structure factors mandatory in PDB
  - Recent literature and workshop demands for archiving image data
    - *J. Appl. Cryst.* (2008). **41**, 659 *Of crystals, structure factors and diffraction images* (Jovine, Morgunova & Ladenstein)
- Proposed IUCr White Paper: *Towards the Crystallographic Archive*



The IUCr continues to encourage the deposition of derived and experimental data sets with journals and databases. For its own journals, deposition of a structural CIF and structure factors or Rietveld profile data is mandatory for all published structures. The Protein Data Bank has also mandated the deposition of structure factors. Demands are now originating within the community for deposition of primary image data [e.g. Jovine et al. (2008), J. Appl. Cryst. 41, 659], with the possibility that some portion at least of the primary data should be archived for long-term access and reuse. IUCr journals self-archive publications and data, but are working with the Dutch National Library and PubMedCentral to provide satellite archives. Arrangements are in place for archiving articles, and studies are under way to establish how best to archive the associated data sets. All data sets associated with journal publications are now registered with digital object identifiers (DOIs) under the central CrossRef registry (http://www.crossref.org).

Studies are beginning to establish a high-level description of the requirements for providing a distributed archival architecture across a federation of crystallographic data holders.