CIF Operations and Applications at the CCDC

Frank H. Allen, Clare F. Macrae, Sam Motherwell, Lucy H. Purkis, Gregory P. Shields, Robin Taylor, *CCDC, 12 Union Road, Cambridge CB2 1EZ, UK*. E-mail: allen@ccdc.cam.ac.uk

Most major journals require CIF deposition to the CCDC, usually during the submission process, and more than 98% of raw input to the CSD now arrives in CIF form. The CCDC maintains a Supplementary Data Archive of deposited CIFs and, after publication, individual CIFs are made freely available via a simple Web-based request form. The CCDC program enCIFer is available for Web download to check, edit and visualize CIF format, syntax and content, while the CCDC visualiser, Mercury, is fully CIF-enabled, and the most recent version will represent atomic displacement parameters, a feature that will also be incorporated into enCIFer.

We are now developing software tools that assist the processing of CIFs to the CSD: (a) We use knowledge-based and bond valence sum data to establish crystal connectivity, and then assign chemical bond types algorithmically. The algorithm has a success rate of 86.4% when validated against a test set of 1104 structures, including a significant proportion of challenging metal-organics. (b) Heuristic analysis of CIFs permits resolution of disordered molecules/ions into their discrete components, using atomic occupancy factors, together with any CIF 'group' and 'assembly' fields that may be available. (c) We are generating chemical diagrams directly from the CIF using a variety of software tools and measures of chemical similarity with existing CSD structures.

Keywords: CIF applications, cambridge structural database, data processing and visualisation