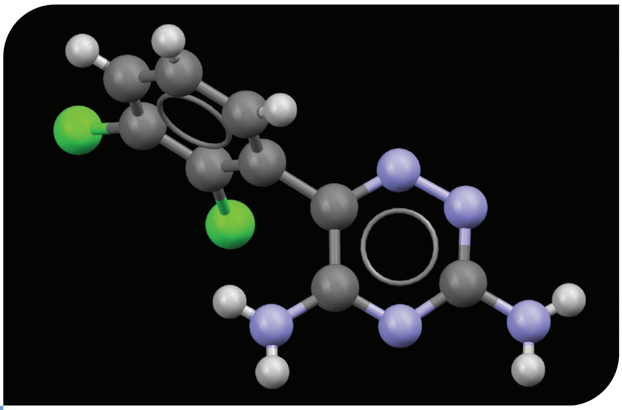
www.ccdc.cam.ac.uk



The Cambridge Structural Database System

Pharmaceutical Discovery

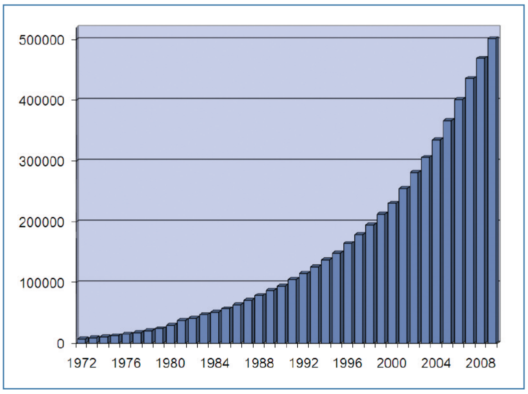
A powerful and highly flexible suite of software, the Cambridge Structural Database System enables you to unlock the knowledge contained within more than 500,000 crystal structures.

Cambridge Crystallographic Data Centre

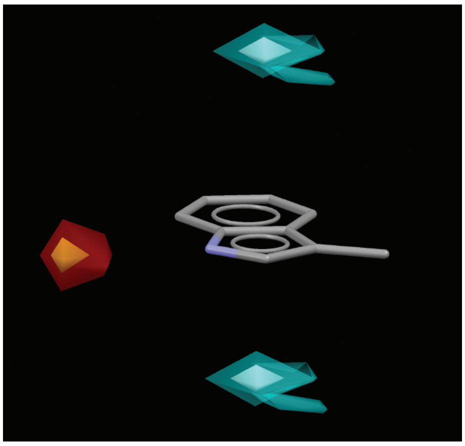
The Cambridge Structural Database (CSD)

Comprehensive and fully curated the CSD is the world's repository of experimentally determined small molecule structures.

- **World's repository:** Established in 1965, the CSD [1] is the world's repository for small-molecule organic and metal-organic crystal structures. Containing the results of over half-a-million X-ray and neutron diffraction analyses this unique database of accurate 3D structures has become an essential resource to scientists around the world.
- **Comprehensive coverage:** With comprehensive and fully retrospective coverage of the published literature you can have full confidence that your CSD searches are returning all crystal structure matches. The CSD also contains directly deposited data that are not available anywhere else.
- **Fully curated:** Each crystal structure undergoes extensive validation and cross-checking by expert chemists and crystallographers to ensure that the CSD is maintained to the highest possible standards. Also, each database entry is enriched with bibliographic, chemical and physical property information, adding further value to the raw structural data. These editorial processes are vital for enabling scientists to interpret structures in a chemically meaningful way.
- **Continually updated:** The CSD is continually updated with new structures (>40,000 new structures each year) and with improvements to existing entries. With regular web-updates and early online access to newly published structures you can keep fully up-to-date of the latest research.
- **Build proprietary databases:** In addition to searching published crystal structures, companies can build their own secure databases of proprietary structures that are then searchable either independently of, or in conjunction with, the CSD.



The CCDC archived the 500,000th crystal structure to the CSD on December 10th 2009. The half-millionth entry was the anti-convulsant drug Lamotrigine (refcode EFEMUX01). The histogram shows the growth of the CSD from 1970 to present.

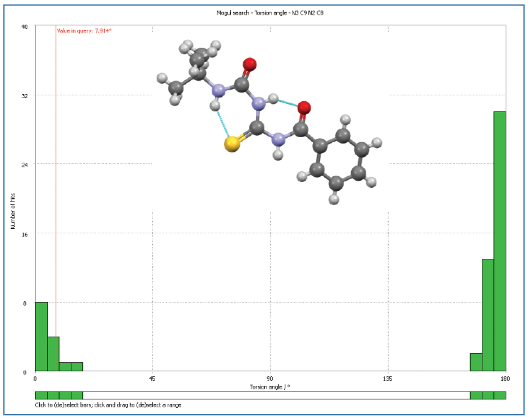


IsoStar-generated contour maps of the interaction of indol-3-yl with OH contact groups (coloured red) and NH contact groups (coloured cyan), both PDB-based data. Note the differences in the preferred orientations of interaction: OH prefers to interact with the indolyl N atom whereas the NH groups prefer electrostatic interactions above and below the ring.

Assess Intermolecular Interactions

The CSD contains millions of non-bonded contact interactions and information on both common and less common interacting groups can be extracted.

- **Definitive source of information:** Crystal structure data are essential for identifying and optimising interactions that are key to molecular recognition in biological systems [2]. Indeed, CSD analyses have been fundamental to our understanding and control of intermolecular interactions of all types, particularly weaker interactions such as C-H•••X [3], X-H•••π [4] and dipole-dipole interactions [5].
- **Quickly visualise preferred interaction geometries:** Access hundreds of pre-processed interaction geometries for commonly observed donor, acceptor, aromatic and aliphatic groups, taken from the CSD and PDB [6]; see how drug-like substructures interact with proteins [2] or within crystal structures in the CSD and use these preferences to guide your drug design project or validate potential new drugs.
- **Carry out more tuneable searches for less common interacting groups:** Tailor searches for non-bonded interactions to your specific requirements. Visualise standard contact distances and angles and their dependency on each other. View statistics relating to the contact(s). Validate experimentally determined or modelled contacts [7, 8].
- **Apply information about discrete contacts across an entire binding site or around a small molecule:** Use information on experimentally observed contacts to identify favourable binding sites, hotspots and pharmacophores [9].



A Mogul search of a thiourea torsion in CSD entry ADEZEP shows it to be unusual. Inspection of the structure in Mercury reveals two intramolecular interactions and it is these that have induced the unusual torsion observed in Fig 1. A crystal packing feature search using the Materials module of Mercury shows this motif to be found in ten other crystal structures.

Validate Molecular Geometries

Use experimentally observed conformational preferences in the crystalline state to validate your molecular design

- **A reliable source of low-energy conformations:** The validity of using conformer distributions in the crystalline state is well established. Studies comparing the conformational energies calculated by *ab initio* methods to observed CSD torsion distributions routinely show good correlation and a systematic study of organic substructures shows that torsion angles with high strain energy (>1 kcal/mol) are rarely observed in crystal structures [10]. In addition, it has been extensively demonstrated that small molecule crystal structure conformations correlate well with protein-bound conformations [11, 12].
- **Validate models:** The diverse range of functional and linker groups in the CSD means it is the ultimate resource to check ligand geometries in experimentally determined protein-ligand complexes and model-derived ligand geometries. The CSD System allows you to quickly validate bond, angle, torsion and ring geometry.
- **Analyse conformation:** Validate modelled ring geometries against similar structures in the CSD and establish the influence of substitution on the ring(s). Use the CSDS to identify preferred acyclic torsional geometries and to estimate barriers to rotation [13]. Check computationally derived torsion potentials against the data in the CSD. Verify the geometric validity of new chemistry by comparing your modelled compounds to similar structures in the CSD or conversely, use the CSD to check conformer quality.
- **Bioisosteric replacement:** Use the CSD to source substituents or groups that are analogous to those in your active compound with a view to using them as scaffold or link replacements [14]; make use of the CSD's low strain geometries [10].

Case Studies and Educational Resources

Case studies illustrating how the CSD System can be applied to a diverse range of research problems are available from the CCDC website.

References

[1] F. H. Allen, *Acta Cryst.*, **B58**, 380-388, 2002.

[2] C. Bissantz, B. Kuhn, M. Stahl, *J. Med. Chem.*, **53**, 5061-5084, 2010.

[3] R.Taylor, O. Kennard., *J. Am. Chem. Soc.*, **104**, 5063-5070, 1982.

[4] M. Nishio, *CrystEngComm*, **6**, 130-158, 2004.

[5] R. Paulini, K. Müller, F. Diederich, *Angew. Chemie, Int. Ed.*, **44**, 1788-1805, 2005.

[6] I. J. Bruno, J. C. Cole, J. P. M. Lommerse, R. S. Rowland, R. Taylor and M. L. Verdonk, *J. Comput. -Aided Mol. Des.*, **11**, 525-537, 1997.

[7] I. J. Bruno, J. C. Cole, P. R. Edgington, M. Kessler, C. F. Macrae, P. McCabe, J. Pearson and R. Taylor, *Acta Cryst.*, **B58**, 389-397, 2002.

[8] C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek and P. A. Wood, *J. Appl. Cryst.*, **41**, 466-470, 2008.

[9] J. W. M. Nissink, R. Taylor, *Org. Biomol. Chem.*, **2**, 3238-3249, 2004.

[10] F. H. Allen, S. E. Harris, R. Taylor, *J. Comput.-Aided Mol. Des.* **10**, 247–254, 1996.

[11] H.-J. Boehm, G. Klebe, *Angew. Chem., Int. Ed.*, **35**, 2588–2614, 1996.

[12] G. Klebe, T. Mietzner, *J. Comput.-Aided Mol. Des.* **8**, 583–606, 1994.

[13] K. A. Brameld, B. Kuhn, D. C. Reuter, M. Stahl, *J. Chem. Inf. Model.*, **48**, 1-24, 2008.

[14] CCDC case studies: www.ccdc.cam.ac.uk.

The Cambridge Structural Database System

The CSDS is an integrated suite of tools comprising the following software components:

The Cambridge Structural Database
comprehensive of the published literature and fully curated

PreQuest
build searchable databases of proprietary structures

ConQuest
search and retrieve information from the CSD

WebCSD*
access newly published structures online

Mercury
visualise and explore extended molecular networks

Materials Module of Mercury**
investigate and analyse solid forms

Vista
interactively analyse numerical data retrieved from the CSD

Mogul
instantly validate molecular geometries

IsoStar
rapidly assess intermolecular interactions

* Available to unlimited licence holders ** Add-on component of the CSD System

Evaluations

To request an evaluation of the Cambridge Structural Database System, please contact **admin@ccdc.cam.ac.uk**

Supported platforms

The Cambridge Structural Database System is supported on Windows, Mac and Linux platforms. For a full list of supported operating systems visit the CCDC website.

Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, UK
www.ccdc.cam.ac.uk • Email: admin@ccdc.cam.ac.uk • Tel: +44 1223 336408

Registered in England No. 2155347 • Registered Charity No. 800579