

Chemical Database Service (CDS) NEWSLETTER

Issue 18 Oct 2003



See <http://cds.dl.ac.uk/links/news1.html> for this and previous editions of our newsletter

A summary of the main articles

CDS Road Show starts in October 2003. See the article on page 2.

See page 3 for news on DETHERM and the 2003 CDS user forum.

Screening compounds catalogues and the new ACD Finder 2.0 service pack 6.0 are also explained on page 3.

See page 4 for the latest update to CrystalWeb.

ISIS/Base users will need to update their Hview in order to access the new ChemInform data (page 5).

For more information on our planned data trials of Accelry's Failed Reaction database and the HNMR data, see page 6.

CDS hardware upgrade and what it means to our users. (Page 7).

Once again Telnet users are advised to upgrade to Secure Shell (Page 8).

The EPSRC logo consists of the letters 'EPSRC' in a bold, black, sans-serif font, enclosed within a rectangular border.

The CDS is funded by the Engineering and Physical Sciences Research Council

Inside this issue:

CDS Road Show	2
Detherm	3
CDS User forum 2003	3
Screening Compounds Catalogues	3
ACD Finder	3
CrystalWeb Update	4
Update to ChemInform	5
New Data Trials	6
Hardware Replacements	7
Upgrade from Telnet to secure shell	8
Databases available from the CDS	9
General Enquiries	10

CDS Databases

ISIS: A chemical reaction information management system.

SPEC: SpecInfo is a multi-technique spectroscopic database

ICSD: Inorganic Crystal Structure Data File.

CRYSTMET: It consists of Crystal structure data for metals, alloys and intermetallics

CDIF: Crystal Data Identification File. Consist of Crystal class and unit cell data

ELYS: Electrolyte Solutions Database.

CSD: The Cambridge Structural Database. Consist of crystallographic data for organic and organometallic compounds

IsoStar: A database of non-bonded interactions.

Special points of interest:

- CDS Road Show
- Detherm
- ACD Finder
- CrystalWeb Update
- ChemInform Update
- New Data Trials
- Telnet

CDS Road Show

Starting in Oct. 2003 CDS will be embarking on a tour of Universities. The road show will include poster displays, face to face CDS discussion and handing-out of CDS flyers. The posters will cover the Service, with special emphasis on newly available Physical Chemistry datasets and CrystalWeb (a web interface to our Crystallography data). The plan is for the tour to kick-off at local universities such as Liverpool and Manchester.



The Computational Science and Engineering (CSE) department at Daresbury laboratory will also be on the road, starting later this year. CSE will be visiting about 12 universities over a short period and CDS will also be present on this road show as part of CSE. The idea is that this will expose CDS to a broader range of potential users.

If you would like CDS to visit your institution as part of your induction of new postgraduate intake or if you simply think a visit from CDS will be beneficial to your institution, contact me (David Osa-Edoh) via email d.osa-edoh@dl.ac.uk or call 01925603163

CDS Road Show, starts in Oct.2003

Detherm

Detherm's funding application was incorporated into our service funding renewal application. The funding application is currently at the final stage of the assessment process and if our grant is renewed by the EPSRC, a re-launch date for Detherm will then be announced.

CDS User forum 2003

CDS's annual user meeting took place on the 16th of Sept. 2003. Users were introduced to the service's new development and the future direction of the service. To view the powerpoint presentation from the meeting see :-

http://cds.dl.ac.uk/CDS_forum2003/

Screening Compounds Catalogues

Three main catalogues have been removed from ACD. They are **SALOR** (Sigma-Aldrich Library of Rare Chemicals), **MAYBRIDGE** and **BIONET**. This is because most of the chemicals in these catalogues are not 'off the shelf' compounds and are more intended for high throughput screening. This does, however, mean that the remaining compounds in ACD should be readily available compounds.

All three databases of screening compounds are available as a separate database. A total of 171,710 compounds are searchable by Formula, Molecular Weight, LogP, Number of Donors, Acceptors and Rotational Bonds. In addition, the compound Name is available for BIONET and MAYBRIDGE, Lipinski rule of 5 number is available for SALOR and Melting Point is available for BIONET.

ACD Finder

The latest version of the Finder for use with PC's is version 2.0 SP6 and contains extra boxes for searching and viewing data such as LogP, Number of Donors, Acceptors and Rotational Bonds.

H acceptors	H donors	MW(frag)	Torsional d.f.	Calc. logP	Violations

ACD Finder has also been enhanced to allow searching by catalogue number. To get the full benefit of this release and enhanced ACD content ISIS/Base users will need to download ACD FINDER 2.0, service pack 6 from <http://cds.dl.ac.uk/cgi-bin/down/isis.html>

CrystalWeb update

CrystalWeb Hit Display - Microsoft Internet Explorer provided by DL

Address: <http://cds.dl.ac.uk/cgi-bin/cweb/cwd4>

CrystalWeb Hit Display

Next Goto hit Hits per page 1 Display All data New query Bibliographic search

Hit number 1 of 940 Cambridge Structural Database Refcode JOFWEG	Reference	Cell data																
2alpha 3alpha 4beta-3 4- Dihydro-3-hydroxy-2-methyl- 4-(2-oxopyrrolidin-1-yl)-2H- 1-benzopyran-6-carbonitrile C15 H16 N2 O3	D. R. Buckle D. S. Eggleston C. S. V. Houge-Frydrych I. L. Pinto S. A. Readshaw D. G. Smith R. A. B. Webster, Perkin 1, (1991) p2763	<table border="1"><thead><tr><th colspan="2">Lengths</th><th colspan="2">Angles</th></tr></thead><tbody><tr><td>a</td><td>9.938 (3)</td><td>alpha</td><td>90</td></tr><tr><td>b</td><td>8.411 (3)</td><td>beta</td><td>90.85 (2)</td></tr><tr><td>c</td><td>15.685 (5)</td><td>gamma</td><td>90</td></tr></tbody></table>	Lengths		Angles		a	9.938 (3)	alpha	90	b	8.411 (3)	beta	90.85 (2)	c	15.685 (5)	gamma	90
		Lengths		Angles														
a	9.938 (3)	alpha	90															
b	8.411 (3)	beta	90.85 (2)															
c	15.685 (5)	gamma	90															
System: Monoclinic Space group: P21/c Space group number: 14 R-factor= 0.044 Z=4																		

Display structure using Chime
Download structure using Jmol

Chemical Markup Language Download

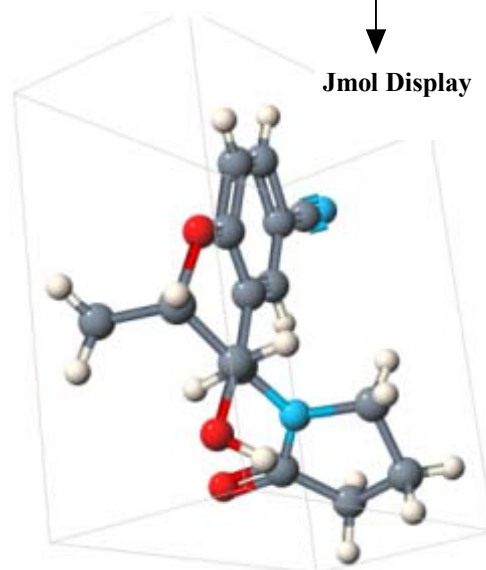
Download structure

Select file format: MDL molfile Platform: PC

Download

- MDL molfile
- CDS xr file
- PDB file
- SheX file
- XYZ file
- Alchemy file
- FDAT file for HyperChem/ViewerLite
- Sybyl mol2
- IUCr CIF(MIF) file
- CML version 1 file**
- CML version 2 file

Jmol Display



- Data from the September 2003 update of CrystMet has now been added.
- Structures can now be displayed using the Jmol applet in place of the Chime plugin.
- Structures can now be downloaded in Chemical Markup Language.
- Data from the July 2003 update of CSD has now been added

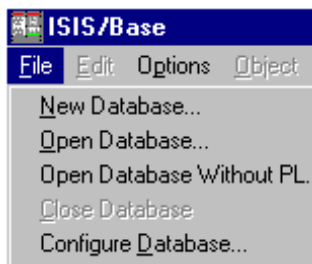
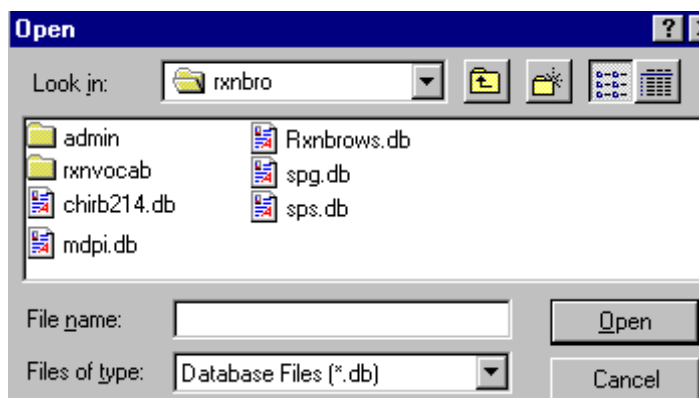
Update to ChemInform

The ChemInform Reaction Library has been updated. The update includes new data from the weekly compendia of ChemInform, 2003, issues 1-26. This adds a total of 36,478 reactions and 49,995 molecules to the ChemInform database. Users using ReactionWeb will automatically pick up this new data. ISIS/Base users who downloaded their Reaction Browser files before the above date **MUST** update their reaction browser file to take account of this new data.

Checking whether you are accessing the complete ISIS data

(1) Start ISIS/Base. (2) At the top of the ISIS/Base Window click on the File menu and select Open Database. Mac Select the RXN Browser file from the RXN Browser folder. PC Select the file rxnbrows.db from the rxnbro directory. SGI Select the file all.db. (3) Enter your username and password to access the ISIS/Host server - the Browser Window will appear. (4) Click on the expert query box in the Browser Window - you may need to scroll through the Window to see this box. (5) Move back to the top of the ISIS/Base Window, click on the Search menu and select Retrieve All. ISIS should then retrieve a hitlist of **1,234,852** reactions (10/09/03) and this will be displayed in the Browser Window. If you get fewer reactions **you are not accessing all of the available data.**

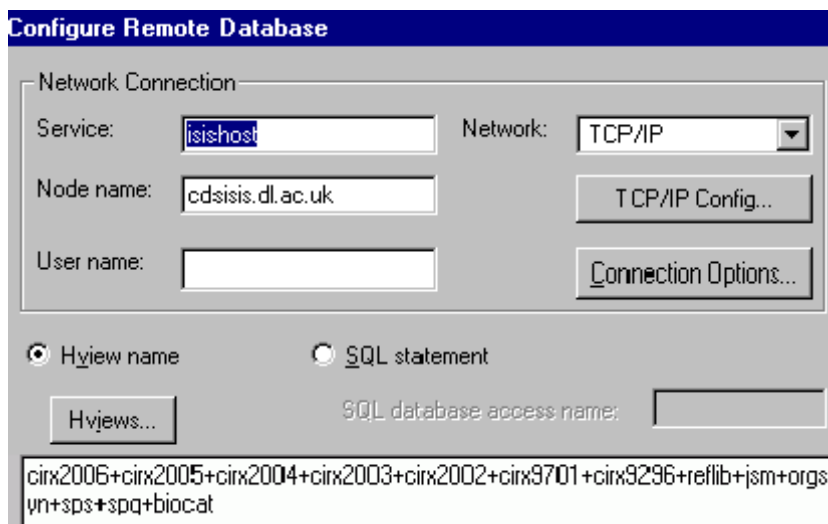
How to configure Reaction Browser



1. Open up ISIS/Base and choose **Configure Database...** from the **File** menu.

2. Locate the required reaction browser file in the ISIS directory and open it

Configure your hview now to future proof it to 2006



3. Updates to the ChemInform database (CIRX) occur twice a year, but updates to the Hview list need only occur once a year. Simply type in the additional hview in the appropriate box. The latest CIRX database should have the hview cirx2004, cirx2005 & cirx2006. If your Hview list does not correspond to the example on the left, then update your list. This update will see you through to 2006.

New Data

Failed Reactions (Database trial starting soon)

Accelrys' Failed Reactions offers a unique resource for the synthetic chemist - the chance to avoid repeating other people's mistakes!

The database contains reactions, abstracted from the literature, which did not proceed as expected. These reactions fall into three categories:

- Unexpected Products
- Further Reactions
- No reactions where one was expected

Although Failed Reactions can be used independently, it attains its full potential when used in conjunction with other synthetic chemistry databases. This allows you to search for a desired product, then see examples of both successful and unsuccessful experiments. To facilitate working, great care has been taken in developing the user interface forms to clearly distinguish failed reactions from those in other databases.

All the references are abstracted by expert PhD chemists with experience in seeking out this type of data, so you can be sure of the relevance and quality of the information. Failed Reactions contains over thirteen thousand reactions and this will double over the next two years or so. Regular updates ensure that the database continues to offer timely access to new pitfalls and synthetic failures as well as expanding the pool of backlog data.

HNMR Data

HNMR Collection No. 2, Ed. 1 and HNMR Collection No. 3, Ed. 1 (from Chemical Concepts)

Both databases contain fairly up-to-date data, since the spectra have not been extracted from the literature but are collected directly from research groups.

HNMR Collection No. 2, Ed. 1

15,837 full spectra, peak lists & structures (all assigned)

HNMR Collection No. 3, Ed. 1

22,503 full spectra, peak lists & structures (all assigned)

This dataset is currently on order and we expect to have it available to users very shortly.

Hardware Replacement

The current CDS server has been in place since 1998, with upgrades. However, now there is no longer a sensible upgrade path for this system and since MDL and Oracle no longer support the IRIX platform it is time to change. New hardware has been bought and configured and we intend to switch over to the new system on Monday 10th November.

The switch is intended to cause as little disruption as possible to users, though a few changes are unavoidable and these are listed below. Most services will be unchanged and there will be no need to re-register on the new server or to change your password. The old server will be left running for a period in case of difficulties.

The main changes which will impact users are:

- ISIS users will need to download new database access files (<http://cds.dl.ac.uk/cgi-bin/down/isis.html>)
- The SpecInfo X-Windows interface will not be available and all access will be via the SpecSurf web interface.
- User will need to log on to the machine cds.dl.ac.uk.
- The main server runs Red Hat Linux rather than IRIX.
- Users will need to copy over any old files they require from the old machine. (http://cds.dl.ac.uk/cds/news_and_highlights/hr.html)

The new CDS system is made up of the following machines:

Main server – 2 processor Linux box

Dell PowerEdge 2600, 2 3.06GHz Xeon processors, 2GB RAM, 108GB disk

Secondary server – 2 processor Sun

Sun Fire 280R, 2 1.2GHz UltraSPARC III processors, 2GB RAM, 144GB disk

Oracle server – single processor Linux box (already in use)

Dell PowerEdge 2600, 1 2.4GHz Xeon processor, 1GB RAM, 85GB disk

Upgrade from Telnet to Secure Shell (SSH)

What is Telnet?.....

The term "TELNET" refers to the remote login that is possible on the Internet because of the TELNET Protocol. It gives the user the ability to be on one computer system and do work on another over the Internet. For example, when you "telnet in from your department to CDS", you are using telnet to connect from one computer (your computer) to another computer (CDS's computer) which is in a different location. Once you have established your telnet connection, you then log in to the CDS computer and execute commands remotely on that computer through your telnet interface.

The use of this term as a verb, as in "Telnet to a host" means to establish a connection across the Internet from one host to another. Usually, you must have an account on the remote host to be able to login to it once you've made a connection

Alternative, why upgrade and what you should do.....

Secure Shell (SSH) enables you to connect securely to computers on the Internet that are running SSH. Older protocols such as Telnet, FTP, rsh, rlogin, and rexec transmit everything in plain text over the network for all to see. This means that someone illegally monitoring network traffic can pick up your password. If this happens it is potentially a disaster, since someone with your username and password can gain access to your accounts and all files on your disk space. You need to upgrade to SSH to protect your password. SSH encrypts the information being transmitted, protecting your privacy, passwords, and other sensitive information.

For information on obtaining and installing SSH clients for Windows (TTSSH and PuTTY), Macintosh (Nifty telnet and MacSSH) and Unix (SSH and OpenSSH) see <http://cds.dl.ac.uk/links/news11701.html>

SSH provides a more secure computing environment than Telnet

Databases available from the Chemical Database Service

A brief description of all the databases currently available from the CDS at Daresbury.

ISIS: A chemical reaction information management system allowing search, retrieval and display of molecules, reactions and their associated data. Currently contains over 1,200,000 searchable reactions. Specialist Databases available include Protecting Groups, Solid Phase Synthesis, BioCatalysis and Chiral Separations.

SPEC: SpecInfo is a multi-technique spectroscopic database package intended to cover NMR, IR and mass spectra. The package is designed to aid the chemist in spectral interpretation and structure elucidation problems. The database currently contains 99,059 ¹³C NMR; 999 ¹⁵N NMR; 856 ¹⁷O NMR; 2,183 ³¹P NMR; 1,825 ¹⁹F NMR; 26,000 ¹H NMR; 20,898 infra-red spectra and 128,000 mass spectra.

CRYSTMET: Metals Data File. Crystal structure data for nearly 75,000 metals, alloys and intermetallics.

CDIF: Crystal Data Identification File. Crystal class and unit cell data for over 237,000 crystal structures.

ELYS: Electrolyte Solutions Database. Thermodynamic and transport property data such as density, viscosity and diffusion coefficients. Currently contains about 11,000 entries.

In addition there is the Available Chemicals Directory (ACD), which contains supplier information on over 235,978 different compounds from over 639 different suppliers and National Cancer Institute Database. This database is produced using data and structures from the National Cancer Institute (NCI) Developmental Therapeutic Program (DTP).

CSD: The Cambridge Structural Database. Crystal structure data for over 296,000 organic and organometallic compounds. Currently available under Cambridge's Quest and ConQuest retrieval software and our in-house software (CSSR). VISTA is a Visual STATistics package which reads the tables files which are when derived geometric parameters are requested with a CSD search. Mercury offers comprehensive facilities for visualising crystal structures in three dimensions.

Screening Compounds Catalogues

Made up of three catalogues (removed from ACD). They are **SALOR** (Sigma-Aldrich Library of Rare Chemicals), **MAYBRIDGE** and **BIONET**.

Both VISTA and Mercury can also be called from within ConQuest.

IsoStar: A database of non-bonded interactions taken from the CSD, PDB and theoretical calculations.

ICSD: Inorganic Crystal Structure Data File. Contains about 67,200 inorganic structures - the companion file to the Cambridge organic file. ICSD is now accessible via CrystalWeb on the CDS website.

Acknowledging the Service

The results of research using the Chemical Database Service may be published through the normal academic channels provided the following acknowledgement is quoted:

"We wish to acknowledge the use of the EPSRC's Chemical Database Service at Daresbury."

The following publication should also be cited:

The United Kingdom Chemical Database Service, Fletcher, D.A., McMeeking, R.F., Parkin, D., *J. Chem. Inf. Comput. Sci.* (1996), **36**, 746-749.

General Enquiries

Chemical Database Service (CDS)
Daresbury Lab
Daresbury
Warrington
WA4 4AD
01925603162
Fax:01925603031
Cdsbb@dl.ac.uk

**CDS website: [http://
cds.dl.ac.uk/cds](http://cds.dl.ac.uk/cds)**

CDS
ON-LINE

(CDS) provides on-line access to a variety of quality databases in the field of Chemistry, plus support, training and advice.

Access:

The service is available **free of charge** to UK academic research groups for non-commercial work. Each individual user will be issued with a unique ID. It is not our policy to allow shared IDs.

Remit: To provide Chemical Database Service and support to UK academic community as well as helping to maintain/improve the service according to the service level agreement with the EPSRC

Information

For ISIS specific problems:

Dr. D. Parkin
Email: D.Parkin@dl.ac.uk
(01925)603162

For SpecInfo specific problems:

Dr. D.A. Fletcher
Email: D.A.Fletcher@dl.ac.uk
(01925)603492

For crystallography specific problems:

Dr. R. F. McMeeking
Email: R.F.McMeeking@dl.ac.uk
(01925)603669

World Wide Web Site

Information about CDS, including online help and documentation is available over the World Wide Web at the CDS website, URL: <http://cds.dl.ac.uk/cds> Web based interfaces to selected databases are also available from this site.

Documentation:

Most documentation is available online and some can be downloaded from the CDS web site. Printed manuals can still be ordered from us at the above address.

cds (main service machine):

Internet name: cds.dl.ac.uk

Comments:

All comments, questions and suggestions about this newsletter should be sent to:
David Osa-Edoh
Email: D.Osa-Edoh@dl.ac.uk
(01925)603163

The Chemical Database Service