Chemical Database Service (CDS) NEWSLETTER

Issue 15 Feb 2002

A summary of the main articles

The full CDS trial of the Detherm database started on the 21st of September 2001 and will run to the 16th of March 2002. The details of its subject and data/property coverage are given in the article on page 2.

The National Cancer Institute database (NCI) is the latest addition to our database portfolio. See the article on page 3 for more details.

Crystal Web is our new web-based interface for bibliographic searching of the Crystallography databases. See the article on page 4 for more details.

ChemInform RX has been updated, making the loading and searching of data faster. More information on this update can be seen on page 5. The new

CDS mailing list and the forthcoming CDS user meeting are also discussed on page 5.

CDS provides training courses free of charge to its users/potential users. See the article on page 6 for more information on how to arrange a CDS training course.

EPSRC

CDS Databases

ISIS: A chemical reaction information management system.

SPEC: SpecInfo is a multi-technique spectroscopic database

ICSD: Inorganic Crystal Structure Data File.

MDF: Metals Data File. It consists of Crystal structure data.

CDIF: Crystal Data Identification File.

ELYS: Electrolyte Solutions Database.

Detherm: Is one of the world's largest

thermophysical property databases.

CSD: The Cambridge Structural Database. Consisting of Crystal class and unit cell data.

IsoStar: A database of non-bonded interactions.

Special points of interest:

- *Detherm trial extended*
- *New CDS mailing list*
- *National Cancer Institute (NCI) database*
- *Cheminform*
- *CDS training*
- *CDS User meeting*

The Detherm Database

The Detherm database provides thermophysical property data for **around 120,000 systems** (about 19,200 pure compounds and 92,500 mixtures). Currently there are some **3.75 million** data sets containing the following properties: -

Phase Equilibrium Data (e.g. Vapour-Liquid-Equilibria, Liquid-Liquid-Equilibria, Solid-Liquid-Equilibria, Activity Coefficients at Infinite Dilution, Gas Solubilities, and Vapour Pressures)	Transport Properties (e.g. Viscosities, Thermal Conductivities, and Diffusion Coefficients)
Excess Properties (e.g. Excess Enthalpies, Excess Heat Capacities)	**PVT-Data** (e.g. Densities, Compressibility, Critical Data)
Caloric Properties (e.g. Specific Heat Capacity, Enthalpy, Entropy)	**Electrolyte Data** (e.g. Conductivities, Transference Numbers)
Interfacial Properties (e.g. Surface Tension)	**Azeotropic Data**

Total: *211* different Properties

Contents of a Dataset

Identification of substance(s)
- Systematical name, synonyms, sum formula, CAS-No., descriptors.

Numerical data
- Data, description of data, errors, remarks.

Literature sources
- Author, Journal, Volume, Year, Page, Coden, abstract.

Technical aspects

The Detherm database runs under a client/server system, similar to the ISIS software. The data is stored inside an Oracle database running on one of the CDS servers, whilst the client software used to access the data runs on your local PC.

You will be able to download the software, free of charge, from the CDS website in a similar fashion to the ISIS client. The Detherm client software only runs under Windows, and this situation is not likely to change.

More information about the database is available on the Dechema website at the following URL: http://www.dechema.de/f-infsys-e.htm?englisch/dbMain.htm and from our own website at http://cds.dl.ac.uk/detherm.html
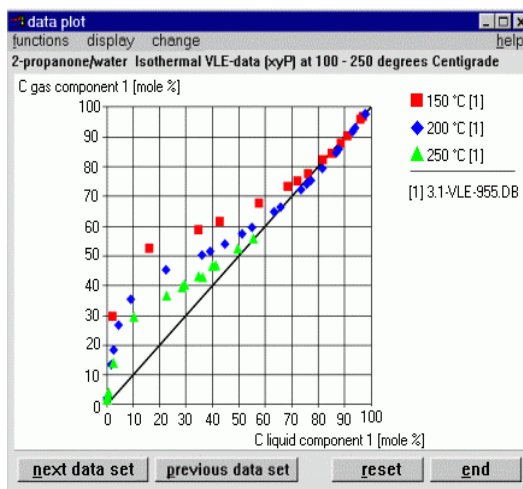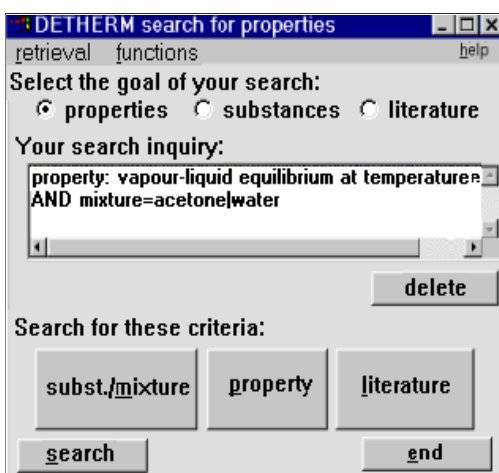
"The full CDS trial of the Detherm database started on September 21st and will run to 16th March 2002"

National Cancer Institute Database

This database is produced using data and structures from the National Cancer Institute (NCI) Developmental Therapeutic Program (DTP).
The database consolidates structures and data from the publicly released:

> NCI Database (126,705 compounds)
> Plated Compounds Database (139,735 compounds)
> Cancer Screened Database (37,330 compounds)
> AIDS Screened Database (42,687 compounds)

The database is available via ISIS/Base or over the Web (http://cds.dl.ac.uk/nci)

2(1H)-Quinoxalinone, 3-(bromomethyl)-

NCI:	Y	Plated:	Y	Cancer:	Y	AIDS:	Y
NSC Number:	144168	CAS Number:	62235-61-4	Mol. Wt.:	239.071	Formula:	$C_9 H_7 Br N_2 O$

CANCER

	Panel Name	Panel Number	Cell Line Name	Cell Line Number	Max Tests	Cell Tests	Conc. Units	Hi Conc. (log)	Result (-log)	Sensitivity (log)
GI50	LNS	1	NCI-H522	3	3	3	M	-4	6.195	1.0001
TGI	LNS	1	NCI-H226	13	3	2	M	-4	4.428	-0.335397
LC50	CNS	12	U251	9	3	3	M	-4	4.286	-0.0636324

AIDS

	No. of Tests	Efficacy	Conc. Unit	Hi Conc. (ug/ml)	Midpoint (ug/ml)	Hi Conc. (molar)	Midpoint (molar)
EC50	2	>	M	47.8142	47.8142	0.0002	0.0002
IC50	2	=	M	47.8142	0.676571	0.0002	2.83e-06

Activity CI (CI=Confirmed Inactive; CA=Confirmed Active; CM=Confirmed Moderately Active)

For the AIDS Database, results of the screening tests are evaluated and placed in one of three Activity Conclusion categories:

CA = Confirmed active;
CM = Confirmed moderately active;
CI = Confirmed inactive

The Chime Pro plug-in for web browsers allows you to change various display options for the molecule such as wireframe or spacefill, or show Van der Waals radii as a dot surface.

Crystal Web

CrystalWeb is our new web-based interface to structural information. It accesses the main crystallographic databases available from the CSD (Cambridge Structural Database), Inorganic Crystal Structure Database, CrystMet Database and Crystal Data Identification File).

It is not intended to be a replacement for programs like Quest or ConQuest and only has a limited functionality by comparison.

The current, unpolished, version allows bibliographic searching, and has a limited structural display using the Chime plugin. An important facility in CrystalWeb is automatic one click access from a retrieved entry to the primary electronic literature using the MDL LitLink system.

CrystalWeb is under active development at present and future plans include cell data and formula searching as well as

enhanced retrieved structure manipulation.

Try out CrystalWeb at the following URL (you will need your CDS username and password)

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

We would greatly appreciate comments on the current version as well as suggestions for improvements.

CrystalWeb

Bibliographic details search form

Click on individual field names to get help on searching that field.

Search [All Databases ▼]

[Author(s)] Surname [_____] First initial [__]
[and ▼] Surname [_____] First initial [__]
[Exact surname search ▼]

[Journal] [_____] [Journal coden] [_____]

[Year] [= ▼] [____] [Volume] [____] [Page] [____]

[Collection code] [____] from [Cambridge Structural Database ▼]

[Compound name] [_____]

[Search] [Clear form]

ChemInform RX update

ChemInform RX is a comprehensive reaction database that covers the current advancements in synthetic research taken from the weekly compendia of ChemInform since 1991.
The database is updated every six months and consists of separate annual releases. The databases currently contain almost 700,000 reactions and over 800,000 molecule entries.

However, every 5 years the separate annual databases are now added together to form one database, hence CIRX92 - CIRX96 and, more recently (January 2002), CIRX97 - CIRX2001 are now CIRX9296 and CIRX9701 respectively. This corrects some data and also makes loading and searching faster.

In order to access all the CIRX databases the following list of Hviews should be in the Browser

file :-
    cirx2002+cirx9701+cirx9296
Users may wish to tidy up their browser configuration file, but do NOT NEED to do so. As long as cirx96 and cirx97 are present in the list, you will only pick up the combined CIRX9701and CIRX9296 databases.

New CDS JISCMailing List

A JISC (Joint Information System Committee) mailing list for CDS (CDS-UK) was established in December 2001. In order to join the mailing list, it can be accessed via www. jiscmail.ac.uk/archives/cds-uk. html or via the JISCmail homepage (www.jiscmail.ac.uk) and then entering the mailing list name (CDS-UK) in the box provided.
JISCmail is a service running on an NT system using Listserv software. This service permits UK academics and support staff to join, create and control

electronic discussion lists.
The purpose of the CDS-UK list is to provide a forum for discussion of topics of interest to users of the EPSRC Chemical Database Service (CDS) and other interested individuals.

CDS intend to use CDS-UK to announce significant new Service enhancements, and float ideas for future developments. We also welcome comments and suggestions from the community at large. It is hoped that users of the Service will feel free to use the CDS-UK list to discuss points of general interest about the CDS and other chemical database systems, but specific technical database queries, any access problems, and the should be send to CDS Helpdesk (E-mail: cdsbb@dl.ac.uk, phone 01925 603162).

CDS User Meeting

This year's CDS user meeting is planned to take place on Wednesday July 10th at Daresbury laboratory. Topics for discussion this year include:

- The long term future of the Detherm physical property database on CDS.
- Website redevelopment
.

- The new crystallographic web interface, CrystalWeb including planned developments.
- Future training strategy.

The user meeting is an ideal opportunity to meet the CDS team, learn about future plans and to discuss any aspects of the Service. User representatives are

"CDS user meeting is planned to take place on Wednesday July 10th"

particularly encouraged to attend, though all users are welcome.
As in previous years, we will be able to cover the travelling expenses of attendees. If you are interested in attending please contact us at cdsbb@dl.ac.uk.

How do I arrange a CDS training course?

In order to enhance its service and accessibility, CDS conducts training/introductory courses free of charge to its users and potential users in the UK academic community, throughout the year.

These courses are largely one-day events in collaboration with University staff in various Higher Education Institutions throughout the UK, or courses can be held at Daresbury. Courses are run on a variety of topics, namely:

- One hour demonstrations of any or all of the databases.

- Introductory courses giving an overview of the service.

- Courses covering the Crystallographic, Spectroscopy, Organic Synthesis and Physical Chemistry areas.

- Specialist courses on particular systems, such as QUEST, ISIS and SpecInfo

Course structure is flexible to meet the needs and resources of the University and can take the form of individual 1/2 - 1 hr sessions or a set of lectures, demonstrations and exercises throughout the day to a group.

If you are interested in arranging a course for your department/ group, the procedure is to contact your group representative or head of department (if you have no representative).

Group representative or departmental heads can contact CDS via email: cdsbb.dl.ac.uk or phone: 01925603162.

When you contact CDS to arrange a course, we need to ascertain the following:

- The number of trainees (a minimum of six)

- The type of course your group or department is interested in (i.e. available database demonstration or service overview)

- Experience level of the trainees with regards to using CDS

"CDS conducts training/ introductory courses free of charge to its users and potential users in the UK academic community"

- Your domain of chemistry

- Your available training facilities (such as I.T suite and multimedia projectors)

- How to contact your I.T. support department. Since their co-operation will be needed in setting-up the training suite as well as dealing with any technical problems.
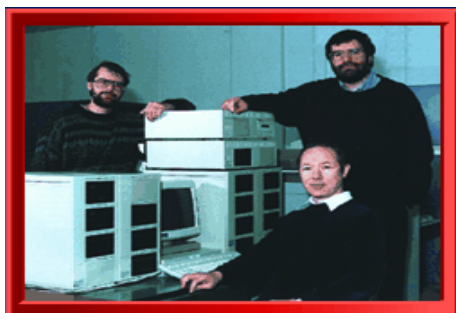
The experts

Providing chemical database service and support to UK academic community

Some Database Search Tricks

Searching multiple databases in SpecInfo

The SpecInfo system now contains 5 different databases and both proton and carbon-13 NMR spectra are contained in more than one database. Ideally, spectral searches and CNMR predictions should use all available data, rather than just the data from one database. This can be achieved by creating a set of databases to search in.

To create a set, select the Edit Database Sets option from the Edit menu on the Main Window. A list of available databases appears; holding down the shift key, select the databases you require and then click on the Add to Sets button. You will be prompted for a name for the set. Enter one and click Save. The set will not be accessible until you exit and log back into SpecInfo. After logging back in, the database set will be available and can be selected to search or run predictions against just like a single database.
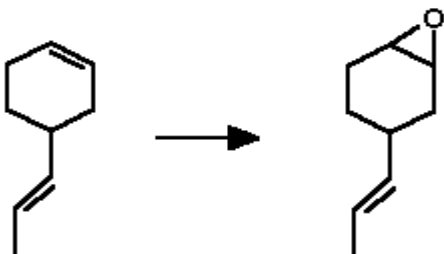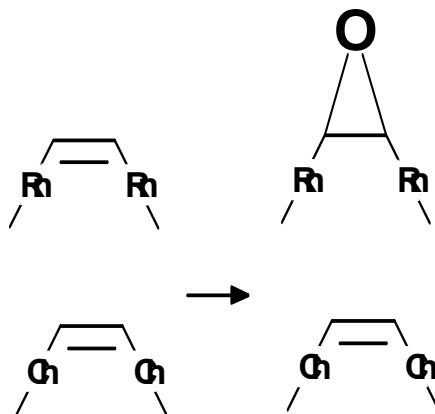

A multi-technique spectroscopic database system, covering NMR,

Tips for searching reaction data

Rather than search for whole molecule transformations, use partial structures and fragments. For instance, if you were interested in the reaction: -


Write the query as: -

Where Rn stands for **Ring** and Ch stands for **Chain**

(Bond Topology menu)
(The -Rn-=-Rn- can be replaced by a cycloalkene ring for more specific searches)

Make sure each of the two fragments of the reactant and the products are grouped into one molecule (highlight them then choosing **Group** from the **Object** menu).

Drawing shortcuts

TYPE	RESULT		TYPE	RESULT	
Ph	**Ph**	- valid for searching	**[5r]**		
+ph	**Ph**	- valid for searching	**66zn**	^{66}Zn	- valid for searching
***ph**			**cu2+**	Cu^{2+}	- valid for searching
+*ph					

Two common CDS user queries and their solutions.

1. Dear CDS, I am having problems transferring my drawn molecule from ISIS/Draw to the molecule/diagram box in the organic reaction database web interface.

There are two ways of transferring your drawing. The first is by clicking on the transfer mode button in ISIS/Draw and the second is by copying your drawing and then pasting it into the web interface. To paste your drawing into the molecule box of the web interface, right click the molecule box on the query form on the web interface, scroll to Edit in the menu, then click on paste in the resulting submenu. If the paste option is greyed-out, open ISIS/Draw on your desktop or double click on the molecule box in query form on the web interface. Click "Options" in the ISIS/Draw tool bar and scroll to settings. Tick "Copy Mol/Rnxfile to Clipboard" option box, if it is not already ticked (under the General tab).

If you are still unable to transfer your drawing, then check that you have ChimPro plugin installed and not the freely available Chime plugin. Go to http://cds.dl.ac.uk/cds/webaccess.html
Under organic synthesis in the webaccess page, right click on the rotating molecule and scroll down and click the "about" option on the menu. This will reveal which of the two plugins you have installed. You can download the ChimePro plugin from the above address.

2. Dear CDS, how can I change my not-so-memorable CDS password ?….

Note: Each password must have at least six characters. Only the first eight characters are significant. Each password must contain at least two alphabetic characters and at least one numeric or special character. In this case, ``alphabetic'' means upper and lower case letters. New passwords must differ from the old by at least three characters.

User A

As a new CDS user you will only be prompted to reset your password if you log-on to the CDS server directly. You can do this by clicking on the telnet button on our website (http://cds.dl.ac.uk/cds).

User B

If you wish to change your CDS password, click the button at our website (above address) and log-on as usual. Once you have logged-on, at the cds3% prompt use the "passwd" command. You will then be prompted to enter your old (current) password, then your new password, followed by a confirmation of your new password.

"check that you have ChimPro plugin installed and not the freely available Chime plugin."

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 1,041,000 searchable reactions. Specialist Databases available include Protecting Groups, Solid Phase Synthesis, BioCatalysis and Chiral Separations. In addition there is the Available Chemicals Directory (ACD) which contains supplier information on over 380,000 different compounds from over 651 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).

Integrated scientific information systems

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 102,459 ^{13}C NMR; 999 ^{15}N NMR; 856 ^{17}O NMR; 2,183 ^{31}P NMR; 1,825 ^{19}F NMR; 10,600 ^{1}H NMR; 20,898 infra-red spectra and 128,000 mass spectra.

CSD: The Cambridge Structural Database. Crystal structure data for over 245,000 organic and organometallic compounds. Currently available under Cambridge's Quest and ConQuest retrieval software and our in-house software (CSSR). VISTA is a **VI**sual **STA**tistics package which reads the tables files which are automatically generated by QUEST when 3D

parameters are defined for a CSD search.
Mercury offers comprehensive facilities for visualising crystal structures in three dimensions

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

ICSD: Inorganic Crystal Structure Data File. Contains about 61,775 inorganic structures - the companion file to the Cambridge organic file. ICSD is now accessible via the CDS website.

MDF: Metals Data File. Crystal structure data for nearly 61,000 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 10,000 entries.

Detherm: Is one of the world's largest thermophysical property databases of pure compounds and compound mixtures. For

instance, in the field of vapor-liquid-equilibrium data, it contained more than 95 % of data published worldwide.
It contains:-

3.75 Million data sets for around 120,000 systems (over 19,000 pure substances and 92,500 mixtures) covering more than 200 properties

"Detherm: Is one of the world's largest thermophysical property databases of pure compounds and compound mixtures"

General Enquiries

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

CDS website: http:// cds.dl.ac.uk/cds

CDS ON-LINE

The Chemical Database Service (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.
Mission: 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 number: 193.62.124.35
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