

CLRC

Chemical Database Service

NEWSLETTER

Issue 10

March 1999

This is the last month for old users, who only have ids on cds1, to re-register on the new server. After March 31 the re-registration procedure will be stopped and any remaining old users will need to register as a new user in order to use the service. See page 10 for more details.

SpecInfo has been upgraded to version 3.2 and a proton NMR dataset is now available. This is a fairly major upgrade and a variety of features have been added or improved. See the article starting on page 2 for more information.

**This year's user meeting will be on July 7 at Daresbury.
All users are invited.**

A new browser for accessing the Available Chemicals Directory using ISIS is now available. It is called ACD finder and it makes extracting pricing information or supplier's addresses much simpler. Details of how to install and use ACD finder can be found in the article on page 8.

A web interface to the Inorganic Crystal Structure Database (ICSD) is now accessible from the CDS website. This interface has essentially the same search functionality as the command driven version on cds, but has a number of additional display features, such as the calculation of powder diffraction patterns. It is described on page 6.

Contents

SpecInfo.....	2	Re-registration.....	10
ICSD web interface.....	6	CDS website	10
ACD finder	8	Databases available	11
User meeting	9	Useful information	12

The Chemical Database Service is funded by the
Engineering and Physical Sciences Research Council

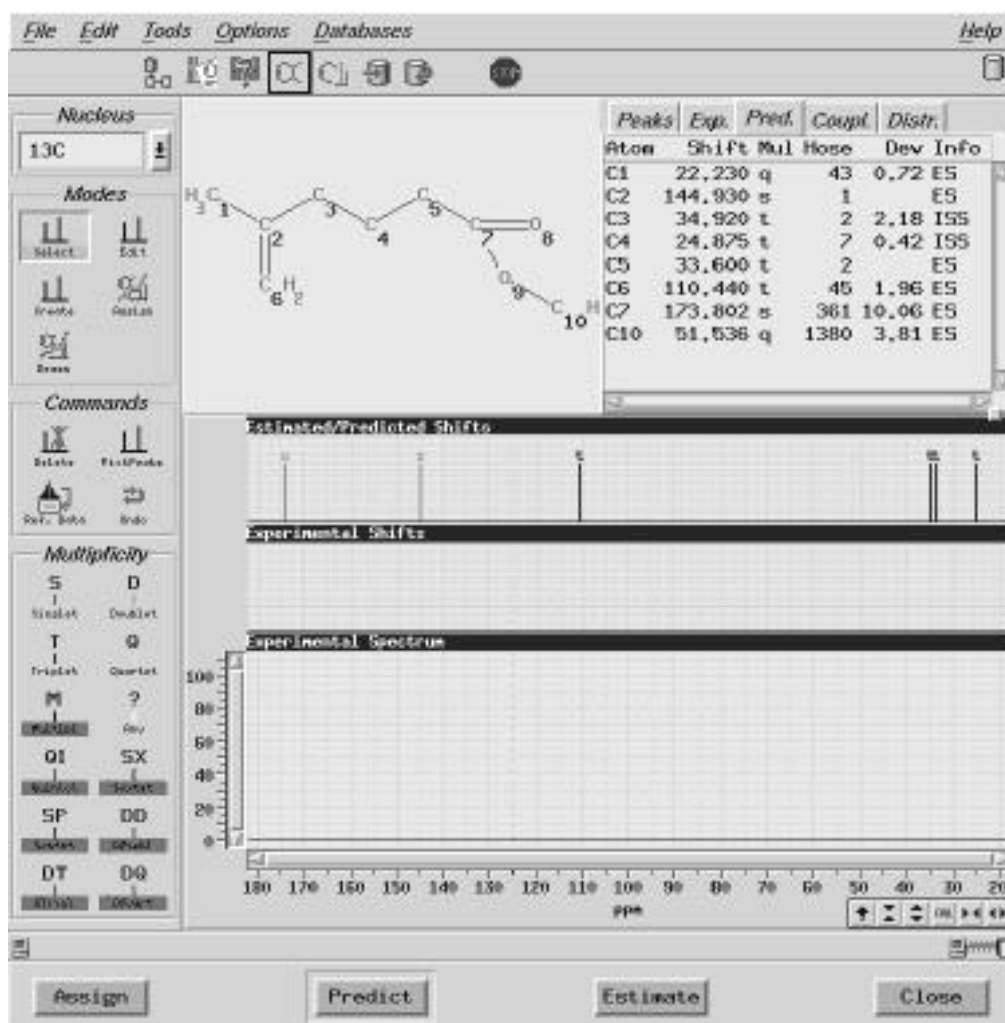
EPSRC

SpecInfo

SpecInfo has been upgraded to version 3.2. This is a major upgrade and there are a significant number of changes and improvements to the program. In addition a new dataset of 2500 proton NMR spectra is now available.

The most noticeable change is the need for users to logon to the SpecInfo server before running any searches or predictions (see below). This is now necessary because the new version stores a user's hitlists within the SpecInfo database management system. Individual user accounts are therefore required. It has the advantage that your hitlists are automatically kept between sessions, unless you delete them yourself.

Once running the program works much like the previous version, though some of the screens look a little different. Documentation (a user guide) is available online from the CDS website at URL <http://www.dl.ac.uk/CDS/cds.html>.



C-13 NMR prediction in SpecInfo v3.2

New and improved features

New data:

- 2500 proton NMR spectra dataset.

New spectrum prediction modes:

- Two modes, estimation and prediction, are now available for the calculation of spectra from structures.
 - Prediction is a statistical method based on the contents of the attached database (as in previous versions of the program).
 - Estimation is an increment based spectrum calculation which does not depend on database contents.
- Proton NMR spectra estimation (but not prediction) from structures.

New search routines:

- New NMR spectra search modes for partial structure searches and mixture analysis.
- HOSE code (atom environment) sub-structure searches.
- Mass spectra search routines (no data yet).
- A variety of search parameters are now adjustable from the options menu.

Other features:

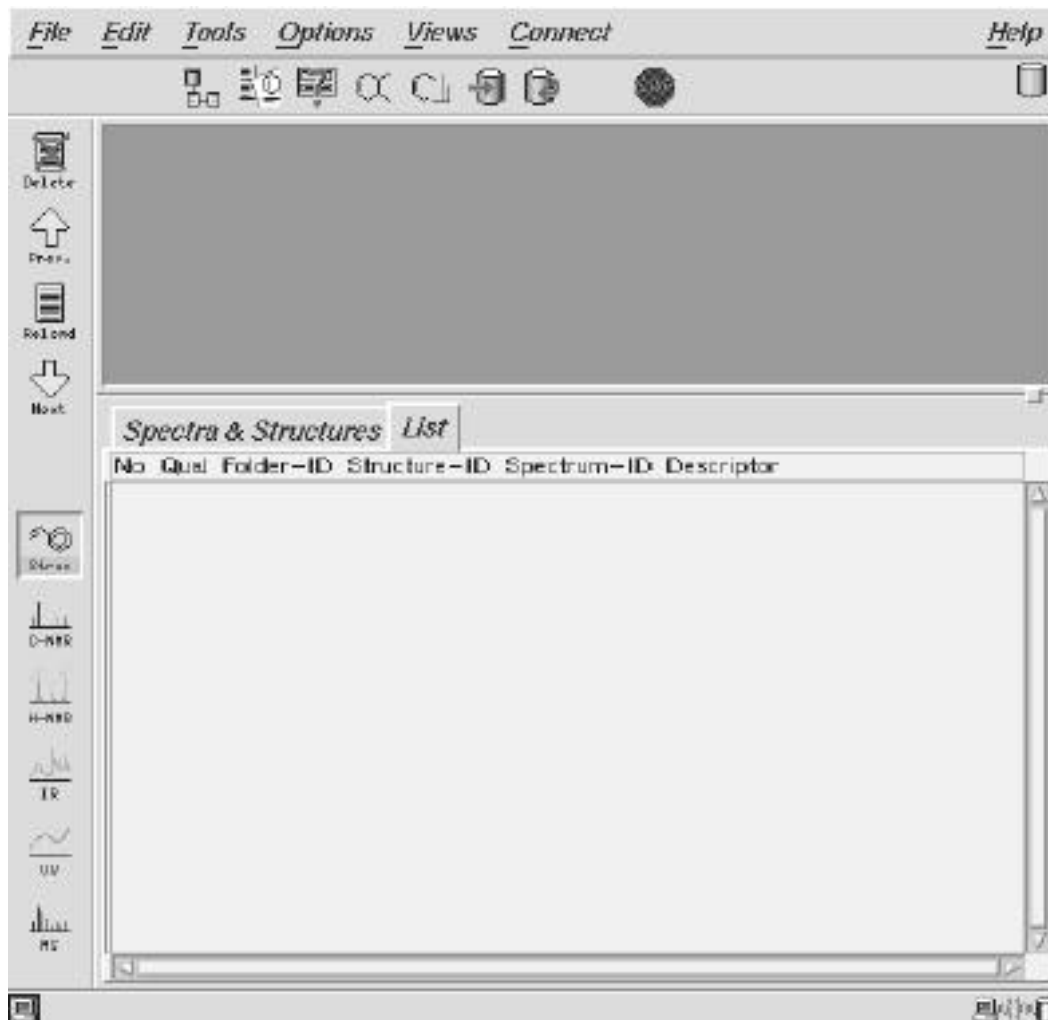
- You can now export spectra as JCAMP format files.
- Structures can be loaded and saved in MDL molfile format.
- Hitlists are now kept between sessions.
- A peak picking routine, for preparing a peak list from a full spectrum, has been added.
- The hitlist display has been improved.
- A new online user guide is available on the CDS website.

NMRP utility

- This is a utility to build NMR peak position files for import into SpecInfo for spectrum searching. The program takes a list of peak positions (in ppm) as its input, either from a file or from the keyboard, and produce a suitable formatted file for SpecInfo.

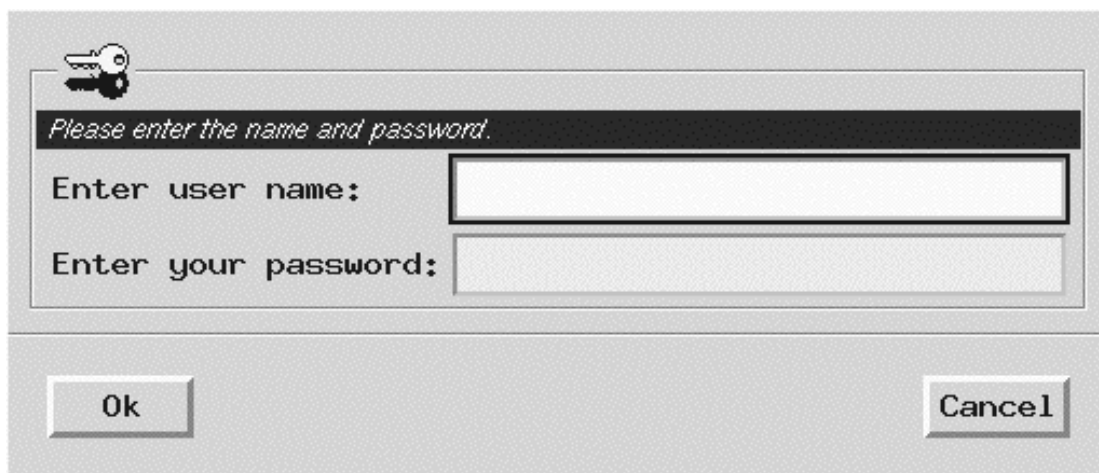
Starting SpecInfo

SpecInfo v3.2 is still an X-Windows application, so you need an appropriate display and the DISPLAY environment variable needs to be set. Typing spec then starts the SpecInfo client and brings up the following window:



The client can be used for drawing structures and preparing queries, but before performing any searches (including spectrum predictions) you need to logon to the SpecInfo server.

To logon to the SpecInfo server you need to click on the Connect menu, located on the menu bar at the top of the screen, and select the logon option. The following dialog box will then appear.



You use your usual CDS id with the password *specinfo* (in lower case). Do not use your usual password for logging on to the CDS machine. Once you have connected to the server, the icon at the bottom right of the screen changes from a

disconnected screen



to a connected one.



Future features

SpecInfo is still being improved and developed by Chemical Concepts. The following features are in development, though timescales are approximate.

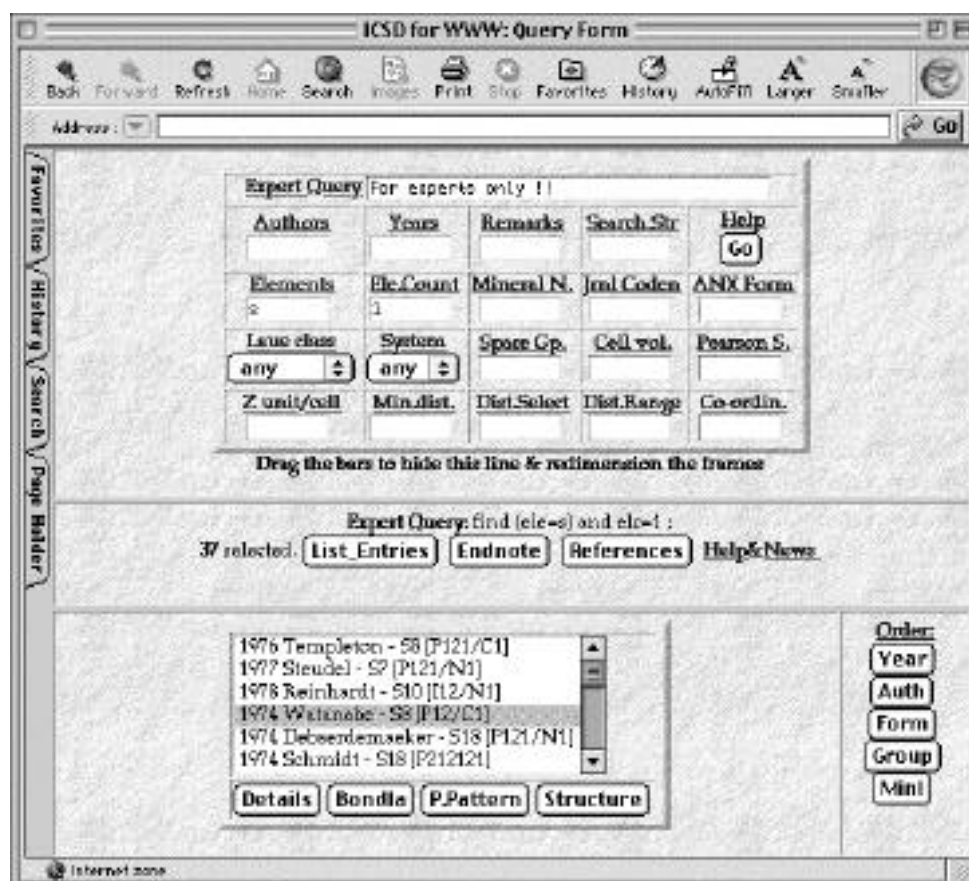
- A full online illustrated guide should be made available by Chemical Concepts in the next month or two.
- We intend to obtain mass spectra dataset in the near future, probably the NIST-EPA-NIH database of some 64,000 spectra.
- A web based client should be available in the next couple of months. It will have limited functionality compared with the normal interface, but may be sufficient for general use.
- Version 3.3, which will probably be available next year, will have prediction as well as estimation of proton NMR spectra and prediction of coupling constants.

ICSD web interface

A web based interface to the Inorganic Crystal Structure Database (ICSD) is now available. The interface is produced by Dr. Hewat of the ILL-Grenoble and works with the CRYSTIN search software. It is accessible from the CDS website via the ICSD link.

Only registered CDS users are eligible to use the interface, moreover users need to access the site from a machine registered in the ac.uk domain, otherwise only a small subset database will be searched.

The search features available are similar to those available for ICSD on cds.



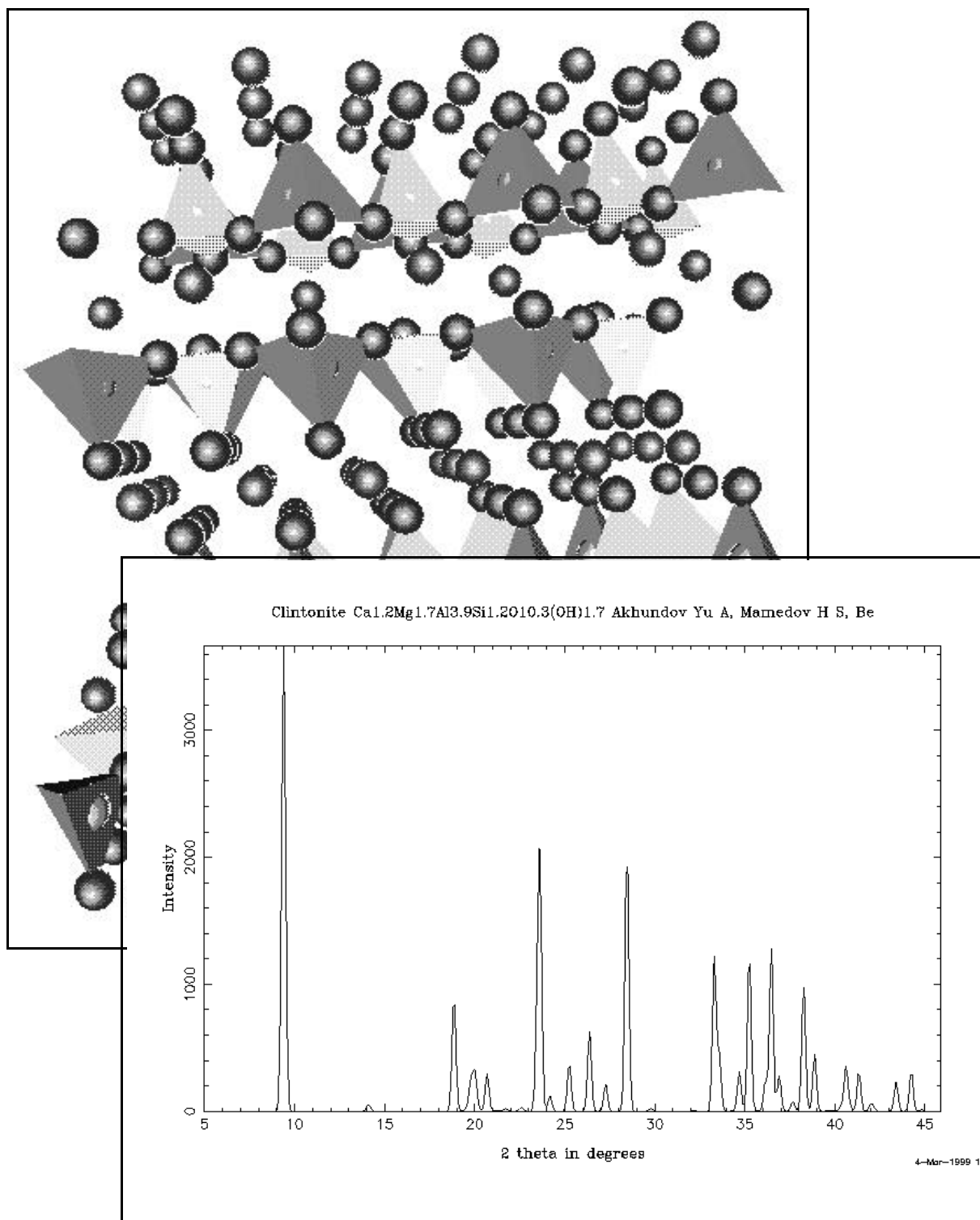
A picture of the ICSD web interface showing the results of a search for allotropes of elemental sulphur is shown above. The search criteria are for the element S with a restriction of only a single element being in the formula.

Once a record has been retrieved there are a number of different display options, including:

- Bibliographical details and database entry.
- Calculated bond lengths and angles.

A calculated powder diffraction pattern, produced in PostScript format by the University of Geneva's Lazy PulverIx program.

The crystal structure, produced by the xtal-3D crystal structure visualisation software by Marcus Hewat, and displayed as a VRML file.



**VRML model of the clintonite mineral structure and
calculated X-Ray powder diffraction pattern**

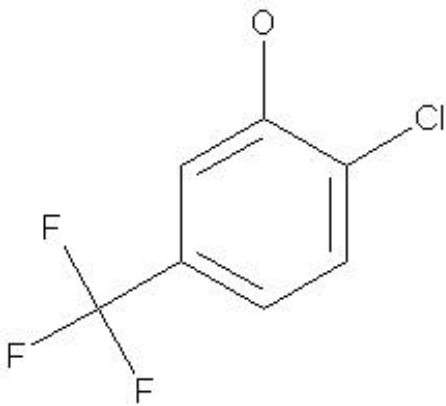
ACD and the ACD finder

The Available Chemicals Directory (ACD) is essentially a database for finding chemicals and their suppliers. Currently there are over 1/4 million unique substances from 543 suppliers and the database is updated every 6 months.

A search can be conducted by structure or substructure, identifiers such as name, CAS registry number, formula, molecular weight or even by supplier.

The new ACD finder browser makes information like prices, catalogue entries and the name, address and contact numbers of the suppliers in ACD much easier to find.

Once an entry has been retrieved, simply click on one of four different tabs to see the appropriate form and information.

Structure	Prices	Catalog	Address
ACD 98.2	CAS# 40889-91-6	MF: C ₇ H ₄ Cl F ₃ O	MW: 196.5550
MFCD00019995	Name: 4-CHLORO-3-HYDROXYBENZOTRIFLUORIDE		
Suppliers: ABCR ALDRICH APOLLO FLROCHEM INDOFINE-F INTRCHM-FLUR LANCASTER MARSHALL OAKWOOD PFALTZ-BAUER SALOR	Structure: 		

Structure data display in ACD finder

Structure	Prices	Catalog	Address			
ACD 98.2	CAS# 40889-91-6	MF: $C_7H_4ClF_3O$	MW: 196.5550			
MFCD00019995	Name: 4-CHLORO-3-HYDROXYBENZOTRIFLUORIDE					
Information						
Supplier	Catalog #	Amount	Units	Currency	Price	Purity
ABCR	F F01930W	1	G	USD	39.60	98%
ABCR	F F01930W	5	G	USD	110.00	98%
ALDRICH	F 36,580-7	1	G	USD	18.60	96%
ALDRICH	F 36,580-7	5	G	USD	66.50	96%
APOLLO	F * PC1930W	1	G	UKL	16.00	98%
APOLLO	F * PC1930W	5	G	UKL	48.00	98%
FLROCHEM	F 4474	1	G	UKL	18.00	98%
FLROCHEM	F 4474	5	G	UKL	50.00	98%
INDOFINE-F	F 09-586	5	G	USD	160	98%
INTRCHM-FLUR	F 56971G	1	G	FF	288.00	98%
INTRCHM-FLUR	F 56971F	5	G	FF	800.00	98%
LANCASTER	F 9942	1	G	USD	17.50	98%

Price data display in ACD finder

More information about ACD finder and examples of these forms can be seen on the CDS website at <http://www.dl.ac.uk/CDS/acd.html>.

The ACD finder can be downloaded from the ISIS/Client download area on the CDS website. It is available for the PC and Mac only.

User meeting

This year's user meeting will be held on Wednesday 7th July at Daresbury. All users are invited to attend and Department reps are particularly encouraged to attend.

Current topics for discussion this year include:

- The introduction of web based interfaces to the databases (ICSD, SpecInfo, PDB, ISIS).
- Future funding of the service.

As in previous years we will be able to cover the traveling expenses of attendees.

Re-registration

CDS users who have not yet re-registered on the new machine need to do so before the end of March. After this date the re-registration mechanism will stop and any remaining people will need to register as a new user if they wish to continue using the service. Users who have cds1 ids but have forgotten the id or password should contact UIG at the address given on the back of this newsletter.

Both re-registration and new user registration can be performed via the web by following the Becoming a User link from the CDS website home page at URL:

<http://www.dl.ac.uk/CDS/cds.html>

If you do not re-register you will no longer be on the Newsletter and other mailing lists.

CDS website

The CDS website, at URL <http://www.dl.ac.uk/CDS/cds.html>, has been reformatted and updated in order to make it easier for users to find relevant information.

There is a new hierarchical menu system with a pop-up list locator and important sections for users, such as online documentation, are now clearly linked from the home page.

There have been a number of technical changes, such as the change from frame based to table based pages, and we would appreciate any comments on whether this has improved the site (or not).

Summary of 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 819,000 searchable reactions and supplier information on 255,000 different compounds.
- 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 ^{13}C NMR; 999 ^{15}N NMR; 856 ^{17}O NMR; 2,183 ^{31}P NMR; 1,825 ^{19}F NMR; 2,500 ^1H NMR and 20,898 infra-red spectra. It does not, at present, contain any mass spectra.
- CSD** The Cambridge Structural Database. Crystal structure data for 190,000 organic and organometallic compounds. Currently available under both Cambridge's QUEST retrieval software and our in-house software (CSSR). The QUEST software has a number of features not available in CSSR, such as 3D searching.
- IsoStar** A database of non-bonded interactions taken from the CSD, PDB and theoretical calculations.
- ICSD** Inorganic Crystal Structure Data File. About 47,500 inorganic structures - the companion file to the Cambridge organic file.
- MDF** Metals Data File. Crystal structure data for nearly 56,000 metals, alloys and intermetallics.
- PDB** The Brookhaven Protein Data Bank containing bibliographic and coordinate details for proteins and other biological macromolecules. There are currently over 8,000 coordinate sets.
- CDIF** Crystal Data Identification File. Crystal class and unit cell data for over 237,000 crystal structures.
- FNMR** A databank of 6,000 ^{19}F NMR spectra and coupling constants.
- ELYS** Electrolyte Solutions Database. Thermodynamic and transport property data such as density, viscosity and diffusion coefficients. Currently contains about 10,000 entries.

Information

General queries:

Computer Help Desk/User Interface
Group (UIG)
Daresbury Laboratory
Daresbury, Warrington
WA4 4AD
(01925) 603351
Email uig@dl.ac.uk

Hours are:
Mon-Thur 9:00-12:00, 13:00-17:00
Fri 9:00-12:00, 13:00-16:00

Otherwise 24 hour answer phone.

General database queries:

Chemical Database Service (CDS)
Room C18, Daresbury Laboratory
Daresbury, Warrington
WA4 4AD
(01925) 603162
Email [cgsbb@dl.ac.uk](mailto:cdsbb@dl.ac.uk)

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) 603162

For QUEST specific problems:

Dr. R. F. McMeeking
Email R.F.McMeeking@dl.ac.uk
(01925) 603162

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://www.dl.ac.uk/CDS/cds.html>

Registration of new accounts:

You can now register online for a CDS
account from the CDS website.

Documentation:

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

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.

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:

Dr. D.A. Fletcher
Email D.A.Fletcher@dl.ac.uk