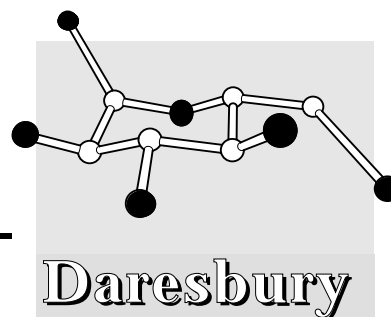# Chemical Database Service

NEWSLETTER

Issue 4   November 1995

There have been a number of significant changes to the service recently, as new hardware and software becomes available. All logins to the service now have to go via cds1, even though the package you use may be running on a different machine. This means that you no longer logon to DLVB directly. However this system will allow us to transparently add new machines and remove old ones without affecting the way a package is accessed.

The big software change is the introduction of the ISIS client/server system to access the organic reaction data. It is the replacement for REACCS , which will eventually be withdrawn. Currently you can access over 530,000 reactions via ISIS including the recently acquired ChemInform database, as opposed to only 240,000 via REACCS. See the articles later in this newsletter for more details about acquiring and using the ISIS system.

Users are reminded that the ISIS and REACCS data should not be used for commercial work.

In the very near future we will make the new version of SpecInfo (version 3.1) available. It is an X-Windows only package and a vast improvement on the old version. The old version will remain available for a 3 month overlap period, before being withdrawn.

Online help is available via the command "ois" on cds1 and works with both X and text displays.

Contents

Some CDS communications, including this newsletter, can be sent by electronic mail instead of paper. If you would prefer this method of communication, please send us your CDS username and a current email address.

The full collection of CDS Newsletters is available online via the OIS.

Organic reaction data via ISIS

ISIS (Integrated Scientific Information System) is a client/server system to access the organic reaction data. This system supersedes the REACCS software and currently accesses all of the REACCS data, (except for the Synopsys Protecting Groups database) plus the ChemInform database (a total of 530,000 reactions, compared with 240,000 available via REACCS). Like REACCS, ISIS is available for academic use only (the memorandum signed for REACCS allows use of ISIS under the same conditions). The data accessible via REACCS is now "frozen", i.e. the databases will not be updated, and users are advised to switch to ISIS as soon as possible. REACCS will continue to run in parallel with ISIS until the end of October 1996.

The ISIS system consists of two main components, ISIS/Host (the "server") which runs on cds1 at Daresbury, and ISIS/Desktop (the "client") which runs on the user's local machine. ISIS/Host accept queries from clients, searches the databases and returns hit information to the clients. ISIS/Desktop is used to prepare search queries and to display and manipulate search results; it consists of two parts, ISIS/Base, which deals with requests to the server and display of hits, and ISIS/Draw a chemically intelligent, presentation-quality chemical drawing package. Currently ISIS/Desktop runs on PCs, Macs and Silicon Graphics computers and is available free of charge to CDS users.

Databases accessed by ISIS

The following databases are available via ISIS:

THEILHEIMER (Synthetic Methods of Organic Chemistry)
CLF (the Current Literature File)
CHIRAS (Asymmetric Synthesis)
METALYSIS (Transition metal-mediated transformation)
ORAC CORE (Established Literature)
REACCS-JSM (Journal of Synthetic Methods)
ORGSYN (Organic Syntheses)
CHC (Comprehensive Heterocyclic Chemistry)
CHEMINFORM (Current Awareness)

The first five databases are combined into a single database called REFLIB. Currently users have access to a total of 530,000 searchable reactions, growing at a rate of 60,000 per year.

In addition ISIS allows access to ACDRX (Available Chemicals Directory), a database of suppliers of chemicals worldwide. This currently contains about 175,700 different compounds from 220 suppliers.

Advantages of ISIS

The ISIS system provides a number of advantages to users compared with REACCS, including:

- A more user-friendly interface
- Simplified query generation
- Customisable display of information
- Scrollable display boxes
- Simplified printing of results
- Clustering of hits into like reactions for ease of browsing (not SGI version)
- Cut and paste both text and graphics into other packages
- Presentation-quality structure drawing package

Getting your copy of ISIS/Desktop

The following section describes how to acquire your own copy of ISIS/Desktop. If you have any problems or difficulties, please do not hesitate to contact us at the address given on the back of this newsletter.

The ISIS/Desktop software is stored on cds1 in the directory /ftp/isis. This contains three sub-directories ibm, mac and sgi which contain the software for the IBM-PC, Macintosh and Silicon Graphics versions, respectively. Each sub-directory contains a file called Getting_ISIS_Desktop, which gives instructions for downloading the software. Before users are permitted to download ISIS/Desktop, they must sign a licence agreement (this is different to the form to sign to use ISIS/REACCS). This form can be downloaded as a PostScript file from /ftp/ISISD_form.ps on cds1 or it can be obtained from the Computer Help Desk at the address shown on the back of this newsletter.

System requirements for running ISIS/Desktop are given below. All computers will need an Internet connection in order to interface with ISIS/Host running at Daresbury.

IBM-PC
A PC with a 80386 or higher microprocessor
Microsoft Windows 3.1
4 MB RAM
7 MB hard disk space
VGA monitor, mouse
WinSockets-compliant TCP/IP networking software

Macintosh
A Macintosh with a 68020, 68030 or 68040 microprocessor
System 6.0.7 or later
5 MB RAM
7 MB hard disk space
Colour monitor, mouse
VersaTerm Telnet Tool 1.1.9 (or higher) and MacTCP or
TCPack Connection Tool 2.0 and MacTCP

Silicon Graphics
       IRIX 5.2 or later
       16 MB RAM
       21 MB hard disk space

ISIS/Desktop may be used on a cluster of computers from a file server, though it has not been specifically designed to work under such circumstances. The only restriction is that only users registered with the CDS to use ISIS may use the software.

An experimental ISIS/Desktop server for X-Windows users who do not have access to a machine which can run ISIS/Desktop is also available. Interested users should contact Dave Fletcher at the address given on the back of this newsletter.

Documentation and help

There is no printed manual for the ISIS system, however a variety of online help and information sources are available and this will increase with time. The following are currently available:

Both ISIS/Base and ISIS/Draw contain online help facilities.

The CDS WWW server has a guide to connecting to Daresbury and conducting searches with ISIS at URL "http://www.dl.ac.uk/CDS/isis.html".

A one page guide to using ISIS is available for downloading in a variety of formats from the directory /ftp/one_page_guides/ISIS on cds1.

An ISIS/Draw tutorial is available for downloading in PostScript format from the directory /ftp/isis/docs on cds1.
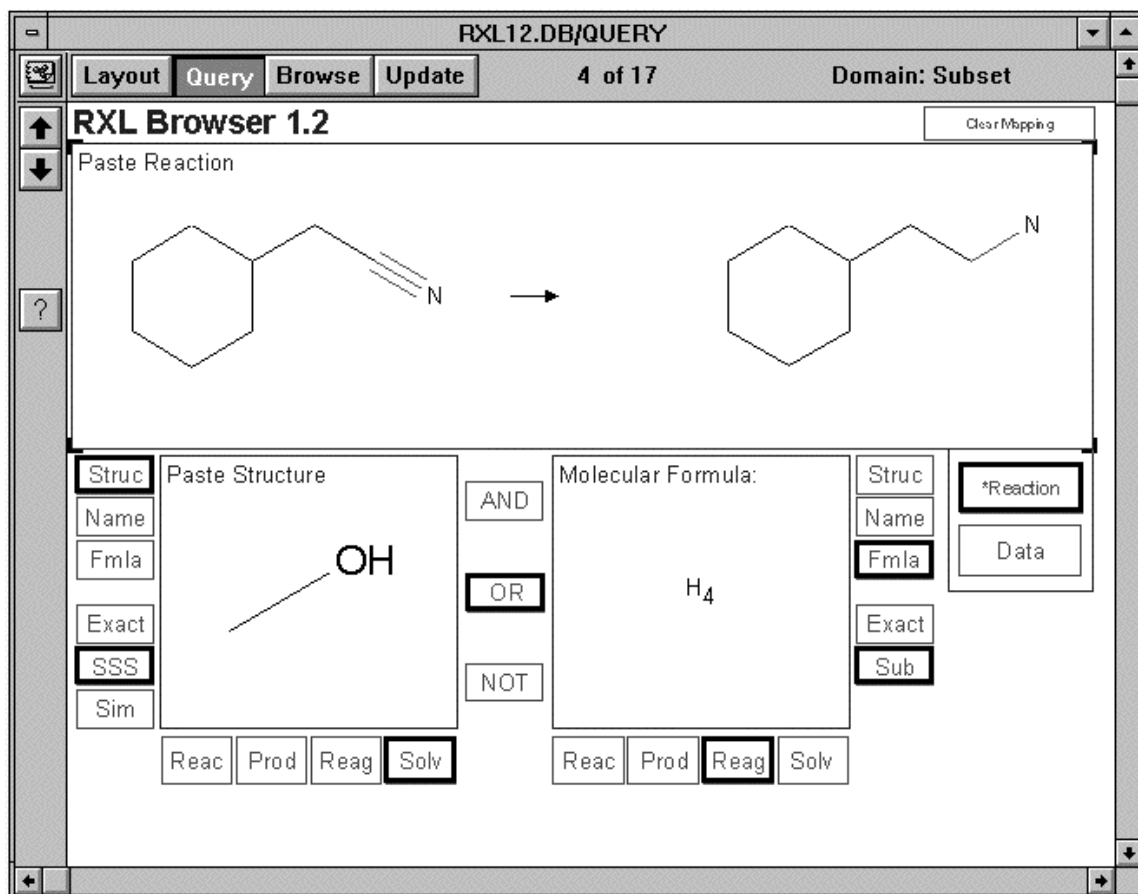
CDS Discussion List

We have set up an email discussion list for CDS users. The purpose of the list is to provide users with a forum to discuss topics pertaining to the service. Messages sent to the list are forwarded by email to everyone who subscribes to the list. If you would like to subscribe to the list send the following text in the body of an email message to majordomo@dl.ac.uk.

       subscribe cds-list

Once you have subscribed to the list you can post messages to the list by sending them to cds-list@dl.ac.uk.

Using ISIS

An ISIS session (reaction searching) is started by double clicking on the browser.db file (RXL Browser icon on a Mac) and entering your cds1 username and password when prompted. You are then presented with the Browser query form.

The above figure shows a typical search using the RXL Browser (PC version). The query structures have all been drawn using ISIS/Draw (double clicking on a structure or reaction box in the form takes you directly to ISIS/Draw), and the formula was typed in after clicking on the Fmla box. The form is setup to search for reactions where the nitrile is reduced to an amine and that either the solvent contains oxygen (e.g. alcohol or ether - bottom left hand box) OR the catalyst is a hydride (such as NaBH4 or LiAlH4 - bottom right hand box). The search is started by choosing the By Form option from the Search menu. When the search is complete, the first hit is displayed with the Cluster form shown below.

Previous Cluster | Go To Cluster # | Next Cluster | Current Cluster: 4 of 8 | Solvent
Save Cluster | Print Cluster | Recluster | Position within Cluster: 2 of 2
| | Find Other Examples with Same Cluster Value | Restore Cluster List | View Scheme

RefLib 94.1 — RREF00152009

Chiral — Chiral

NaBH4 (Excess) i-PrOH

60.0%

Cluster Value: C3H8O   Variation: 1 OF 1

The reactions in a hitlist are clustered depending on their similarity at the reaction centre. The hitlist can also be reclustered based on Author, Solvent, Catalyst, Journal or Journal Year. The arrows on the left of the form are used to scroll through the hits. Alternative views of the data for this reaction can be seen by clicking the buttons on the left hand side.

su - shows summary data for each reaction record.
d1 - shows the reaction structure and chemical data for each reaction record.
d2 - shows the reaction structure and the literature reference data for each reaction record.
va - displays all the literature references and the reagent and solvent information for variations of the reaction.

Printing from ISIS

Single hits or the entire hitlist can be printed directly from ISIS using the Print option from the File menu. When printing the whole list you have the option to specify how many hits are to be displayed on each page.

eXodus for Mac setup for access to cds1

X-Windows has become one of the best ways to access the Chemical Database Service. A number of utilities and packages (such as SpecInfo version 3.1) are available via X-Windows only. One option for the PC or Mac user is an X-Windows emulator. We would recommend Vista eXceed for the PC and MacX or eXodus for the Mac. MacX and Vista eXceed were covered in Newsletter 2, and information about using them to access the service is available online on cds1 through the ois command. eXodus is an X-Windows emulator for Macs, which is slightly more difficult to set up. Below are instructions for configuring eXodus for access to the Chemical Database Service via cds1.

1. Set the screen to a rootless display using the Screen Editor option on the Settings menu.
2. Choose Client Editor from Clients menu
3. Build a new client, use TCP/IP communications
4. Enter Host (cds1.dl.ac.uk or 148.79.169.113) and username
5. In Command box put:

    setenv DISPLAY <IP>:0;xterm -ls

where <IP> is your IP or Internet address.

To logon, choose client from Clients menu. You will be prompted for your password. If this is going to be your normal method of access to cds1 you should add the line
        more /etc/motd
to your .login file, otherwise you will never see the login messages.

Vista eXceed icons

Vista eXceed is an X-Windows emulator for the PC. Use of the program to access the Chemical Database Service was covered in Newsletter 2, which can be read online via "ois". CDS users may find it useful to know that there is an option to setup icons which automatically startup the program and log the user onto cds1. Instructions for setting up an icon are given below.

Click on the Xconfig(/w) icon followed by the communication icon and set the MODE to PASSIVE.
Click on the Window Mode icon and set Window Mode to multiple.
Click on the Xstart icon and fill in the dialogue box with your Daresbury user ID (lower case) and host (cds1.dl.ac.uk). Enter the following command line.

setenv DISPLAY <IP>:0;xterm -ls

where <IP> is your IP or Internet address.

    The Start Method should be set to Rexec (TCP/IP) and Program Type to X-Window.

    Use the File menu to save your settings and then Install to build the icon.

Crystallographic changes

    There have been a number of recent changes to the crystallography section of the service, detailed below. Full details of the crystallography packages and utilities available are available online via "ois".

QUEST and rasmol

    When using QUEST to search the Brookhaven Protein Databank (questp or questa options) an active link to rasmol is available. If you start up a rasmol session beforehand, then PDB structures will be automatically loaded into rasmol when found during a search. Note that this feature is only available to X-Windows users.

New version of Vista

    The Vista program has been completely rewritten for the October 1995 release of the Cambridge Structural Database. At the heart of Vista 2.0 is a scrollable table spreadsheet. The parameters form the columns of the spreadsheet with the fragments (represented by their refcodes) forming the rows. One of the major new features developed for Vista 2.0 is the ability to create new parameters by application of mathematical functions to the parameters defined in the original Quest3D search. Another new feature is the ability to export tables of data. This operation will allow users to write selected columns of data to an ASCII text file in a form which is then suitable for import into other statistical or spreadsheet packages. The PostScript editor now includes a "Preview" command, and Encapsulated PostScript files (EPSF) can be created, for import into graphics and word-processing programs. The old version of Vista is still available via the command "ovista".

File format converters

    Two new file format converters are available, cam2shx and xr2shx. These utilities convert from FDAT and XR format to Shelx format.

QUEST batch jobs

A batch system is an efficient way to make use of the available resources on cds1 and does not involve your watching over the execution of a job. You can logoff once you have set up the batch job. The jobs you submit to the batch system are placed on a queue and are run when resources become available.

To set up a batch job

- Start a fresh quest session using questx in place of quest
  ```
  questx -j query
  ```
- Prepare the query in quest
- Click on save refc from the Search menu (hit refcodes will then be saved in a separate file)
- Start the search

Quest will then exit with the information to run the search saved in the file query.jnl

To start the batch job

```
batch <<!
quest -j answer -if query.jnl -of output
!
```

The search will then be run when system load level permits. The results are saved into the usual journal file answer.jnl and the list of refcodes of the hits are saved in answer.gcd. The normal screen output goes into the file output. In addition, some initialisation output will appear as mail to your account.

Viewing results

To view the results of the search, start up quest
```
quest -j view -v answer.gcd
```
Select the search menu
Select hitall
Select start

You can now scroll through the list of hits. If you wish to highlight your search substructure, you will need to repeat your original search on the database subset. This can be done by starting up quest with the following command and restarting the search.

```
quest -j view -v answer.gcd -ini query.jnl
```

Online documentation

Documentation is available online for most of the CDS packages via the online information server. Simply type "ois" to start and select the online documentation option. Currently the following manuals are available online:

Brookhaven Protein Databank (PDB)
Cambridge Structural Database System (QUEST, PLUTO and VISTA) - including the new Vista V2.0
Crystal Structure Search and Retrieval (CSSR)
Crystal Data Identification File (CDIF)
Inorganic Crystal Structure Data File (ICSD)
Metals Data File (MDF)
SpecInfo
Rasmol
Xmol
Babel
Pluton
CDS Users Guide
Guide to UNIX on cds1

Documentation can also be downloaded from cds1, for local printing, from the directory /ftp. The following items are currently available:

cds_id.ps	CDS account application form (PostScript)
CSD_form.ps	CSD authorisation form (PostScript)
FNMR_form.ps	FNMR authorisation form (PostScript)
ISISD_form.ps	ISIS/Desktop authorisation form (PostScript)
PDB_form.ps	PDB authorisation form (PostScript)
REACCS_form.ps	ISIS/REACCS authorisation form (PostScript)
CDS_packages.ps	List of CDS packages available (PostScript)
CDS_packages.rtf	List of CDS packages available (Rich Text Format)
CDS_Users_Guide.ps	CDS Users Guide (PostScript)
CDS_Users_Guide.rtf	CDS Users Guide (Rich Text Format)
unix_guide.ps	Guide to UNIX on cds1 (PostScript)

Single sheet guides or quick reference cards are available for the following packages in a variety of formats.

CSSR, ICSD, ISIS, MDF, PDB, QUEST, RASMOL and SpecInfo

The sheets can be found in the directories /ftp/one_page_guides/name, where name is the package in question.

A tutorial for ISIS/Draw (in PostScript format) can be found in three parts in /ftp/isis/docs.

Report on 2nd Annual User Forum

The second CDS User Forum was held at Daresbury on May 24th. It was a useful meeting and many of the changes and improvements to the service described in this newsletter were discussed.

As a direct result of requests made at the meeting a CDS discussion list has been started (see page 4 for more details) and basic one or two page tutorial/quick reference guides have been made available for many of the CDS packages. These guides can be downloaded from the directory /ftp/one_page_guides on cds1 and are available in a variety of formats (PostScript, Word, ASCII and RTF).

The training program was discussed and it was stressed that the training courses provided by the CDS could be anything from a simple demonstration to an advanced single package course. Better publicity was requested for courses and we hope to be able to provide this by having regular courses at some sites. ISIS training courses will be available soon and SpecInfo courses will resume once version 3 is available next month.

It was brought to our attention that there was some confusion over use of the word *academic* in terms of eligibility to use the service. In this context it includes graduate students, not just people on academic grades.

Suggestions were made to improve attendance at future meetings. These included earlier publicity for the event and moving the date to later in the year. Early July was the most popular option, therefore the next meeting has been arranged for July 10th 1996.

Training courses

The following training courses are planned:

Date	Venue	Course
9 th January 1996	Daresbury Lab	ISIS
10 th January 1996	Daresbury Lab	SpecInfo
11 th January 1996	Daresbury Lab	QUEST

If you would like to attend one of these courses or would like a course to be held at your site, please contact the Chemical Database Service directly.

Information

General queries and registration of new accounts:

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 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

Documentation:

Documentation is available through the online orderdoc utility or from UIG at the above address.

Access:

The service is available free of charge to UK academic research groups for non-commercial work, through Janet or Internet.
Individual user names are issued on request to UK academics.

cds1:

Internet number: 148.79.160.113
Internet name: cds1.dl.ac.uk
DTE address: 00000100026012
NRS name:
        UK.AC.DARESBURY.CDS1

DLVB:

Direct access to DLVB is no longer possible; it is only accessible via cds1.

Comments:

All comments, questions and suggestions about this newsletter should be sent to:

Dr. D.A. Fletcher
Chemical Database Service (CDS)
Room C18, Daresbury Laboratory
Daresbury
Warrington
WA4 4AD
(01925) 603162

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