Chemical Structure Association Trust

ISSUE 3

NEWSLETTER

SPRING 2003

Personalia

Bill Town has retired from his position as Managing Director, ChemWeb, and has established a new business called Kilmorie Consulting (www.kilmorie.com), which helps organisations make the best use of their information resources, develop new information services or products, and bring them to market. Bill is also following his passion for travelling the world and visited Botswana in December to experience the eclipse of the sun and for a safari trip. Bill retains an active interest in the CSA Trust, ACS CINF, IUPAC and the RSC. He continues as a Visiting Professor at the Department of Information Studies at the University of Sheffield, and is to join the Board of Governors of the Cambridge Crystallographic Data Centre.

Sponsorship Secretary John Holliday will maintain the register of donors. Please send donations to him: Dr John Holliday, Department Of Information Studies, University of Sheffield, Sheffield, S10 2TN, UK. Tel: +44-(0)114-222-2685; e-mail: J.D.Holliday@sheffield.ac.uk

---

Dr. Frank Allen is to be presented with the Herman Skolnik Award at the Fall 2003 ACS meeting in New York. He was appointed Executive Director of the Cambridge Crystallographic Data Centre in October 2002, following the retirement of Dr David Hartley.

Rob Brown and Moises Hassan have left Accelrys and joined Scitegic in San Diego. Rob is now Senior Director of Cheminformatics Marketing. He will have responsibility for the continuing enhancements and evolution of the Pipeline Pilot product for cheminformatics applications. Rob was Senior Director of Cheminformatics Marketing at Accelrys.

Rob Scoffin is now working for CambridgeSoft in the UK. He and Katriona are proud parents of Daniel, born in August 2002.

Another summer arrival was Jacqueline Kaye Dennis, born July 27, 2002, and daughter of Michael and Kim.

Charles Oppenheim has been awarded the Association of Learned and Professional Society Publishers' 2002 award for service to publisher–library relations.

Brian McClintock (formerly selling ChemSymphony) is now Sales Director for Entropia Europe Ltd.

Ronald Lees, C. Chem., FRSC, retired head of division at the Laboratory of the Government Chemist, died June 3, 2002, aged 68. CSA members will remember him for his excellent contributions to the major LGC chemical nomenclature meeting in 1981.

Mike Lynch has offered to prepare a booklet based on the CSA Archives. If you have any relevant material, please contact Mike on m.lynch@sheffield.ac.uk

Inside...

Personalia

CSA Trust AGM Report

CSA Final AGM Report

CSAT Dinner

RSC–CIG Report

ExemplarChem 2002

The CaMeL and the JUMBO

PDR AGM

CSA Trust Grant Applications

Company News

Events 2003

Product News

Sheffield Course

Contact Details

Next Copy Date

The CSA is Dead, Long Live the CSA Trust!

The CSA Trust AGM was attended by 17 of the 30 Trustees, plus two CSA members as observers. Chairman Guenter Grethe summarised the main activities in 2002:

a. Agreement of the CSA and CSA Trust to integrate activities and progress in 2002 towards this aim.
b. The first CSA Trust Mike Lynch Award to Peter Willett, presented by Mike himself.
c. The success of technical meetings that involved the CSA Trust.
d. Establishment of the process for CSA Trust Grants.
e. Successful move of the Web site to Sheffield University in the UK.
f. Creation of a new CSA Trust logo and letter head.

The Chairpersons of new Committees were agreed in order to manage the Trust following the incorporation of the former CSA activities. Budgets and Finance: Geoff Downs; CSA Trust Grants: Bonnie Lawlor; CSA Mike Lynch Award: Rene Deplanque; Public Relations: Janet Ash; Meetings and Training: Peter Nichols; Fundraising: Guenter Grethe. In addition Guenter Grethe and Clive Weeks would solicit nominations from, and consult with, Trustees for the Ernie Hyde Award. Joanne Witiak and Georg Schultheiss (retiring shortly from FIZ Karlsruhe) both tended their resignations. Andreas Barth of FIZ Karlsruhe was elected as a new Trustee and all other Trustees, whose 3-year term of office had ended, were re-elected.

The Trust fund currently stands at around £52,000. Donations to the Trust, instead of the CSA membership fee, will cover the Newsletter production. The CSA account final balance will be transferred to the CSA Trust, to be earmarked for future conferences and training.

The new Budgets and Finance Committee would consider carefully the option of taking on ownership of the Noordwijkerhout Conference. There are many aspects to consider before we can formulate a formal proposal. The 2003 CSA Trust AGM will be held on Monday 1 December 2003.

Clive Weeks

In what was probably the shortest CSA AGM in its history, the CSA was formally wound up. The meeting was held at the Linnean Society on Monday, 2nd December and unlike most such meetings, rather than concentrating on the events of the past year, this one was much more forward-looking.

At the 2001 AGM, the meeting voted in favour of the proposal: 'In order to pursue the objects of the CSA and particularly its educational functions, be it resolved that the activities of the CSA can be transferred to the CSA Trust, and that the CSA Executive be authorised to transfer the remaining funds to the CSA and wind up the CSA prior to the end of 2002'. A select committee was set up to oversee the incorporation and most of our activities in 2002 have been devoted to this task. These activities culminated in the winding up of the CSA and the transfer to the CSAT.

One very pleasurable task was the presentation of the Ernie Hyde Award, which I am delighted to say will continue under the auspices of the CSA Trust. The prize was given to John Holliday for his services as Membership Secretary and setting up and maintaining the CSA/CSAT website.

PWL Nichols

CSAT Annual Dinner

What was the Annual CSA Dinner has now become the CSAT Annual Dinner. What's more, tradition was upheld by holding it on the Monday evening before the London OnLine meeting in the sumptuous surroundings of the Cadogan Hotel, Sloane Street. Some 20 or so members and guests were present and it was a great pleasure to see a number of past CSA Executive Committee members there for the 'Farewell to the CSA' evening, following the winding up at the afternoon's AGM.

The evening commenced with pre-dinner drinks, at which point your Chairman took the opportunity to demonstrate his barman skills (I got fed up with waiting for someone to serve!). Then we withdrew to the elegant dining room. The serious part of the evening then got underway, when we were treated to a lovely Christmas dinner, the first of the season, with copious amounts of wine flowing throughout.

A very convivial and enjoyable evening was had by all, capped by a most wonderful rendition of the CSA's 'anthem', entitled 'Information Progress', written by Frances Barker (ex-RSC) for the formation of the CSA in 1982, and sung to the tune of 'Clementine' (copies of anthem available from ash@euronet.nl). I must compliment Janet Ash, Peter Willett and David Walsh in particular for their wholehearted and enthusiastic contribution to this musical conclusion to our Dinner.

PWL Nichols
CSA Chairman (retired!)

RSC–CIG Meeting: Information at the Chemistry/Biology Interface

As the boundaries between the disciplines of chemistry and biology become increasingly indistinct, new initiatives are being developed to stimulate information exchange between chemists and biologists. The Royal Society of Chemistry (RSC) Chemical Information Group held a one-day meeting on this topic at Burlington House, London, in November 2002.

Following a welcome by Don Parkin, Stephen Baker gave a paper entitled 'Are you a chemist or a biochemist?' Stephen is working at the chemistry/biology interface in the Centre for Chemical Biology at Leicester University, and his title related to a question posed by a potential sponsor. The department started in 1988 as the Centre for Research in Molecular Recognition, and its aim is to promote collaborative research which bridges the two disciplines, and to be identified as an international centre of excellence. The group does not have its own building, but has a committee with representatives in computer modelling, protein interaction, molecular enzymology, structural biology and synthetic chemistry. Further details can be found at http://www.le.ac.uk/chemistry/chembiol/index.html.

Thomas Oldfield, from the European Bioinformatics Institute in Hinxton, near Cambridge, then discussed the Macromolecular Structure Database (MSD). Protein data are collected from many different sources, and thoroughly checked before being stored as a relational database, so that complex search queries can be carried out. The MSD is one of three groups that collect macromolecular structural data, but is the only one that offers detailed cleaning of the data and storage as a relational database. Information on the MSD can be found at http://www.ebi.ac.uk/msd/index.html.

Simon Edwards is the Life Sciences Manager at the RSC. His role involves looking at the interface between chemical sciences and biosciences, and working closely with the chemical biology community. Simon discussed the problem of falling numbers of undergraduates studying chemistry in the UK, at a time when medicine and related subjects are showing an increase. The RSC is trying to broaden its remit to include key interfaces between chemistry and other sciences, such as biomedicine, biomaterials and pharmacology. Workshops have been set up in priority areas. Key challenges include: identifying the people who constitute the chemical biology community; forming partnerships with Research Councils and learned societies; and developing a programme of activities in the chemical biology area. Further information can be found at http://www.rsc.org/science/lifescience.htm.

The morning sesson closed with a talk from Lesley Thompson, Programme Manager of the Life Sciences Interface Programme within the EPSRC (Engineering and Physical Sciences Research Council). UK research funding is organised through Research Councils, and the Life Sciences Interface was formed in 1999 to fund high quality research and to provide training at the life sciences/physical sciences interface. The work involves collaboration with other EPSRC programmes and other Research Councils. Further details can be found at: http://194.66.183.26/website/default.aspx?CID=560&ZoneID=3&MenuID=410.

After an excellent lunch, the afternoon session began with a very different lecture from Denise Harrison (Sheffield Science Librarian), who had recently moved the Sheffield University Chemistry Branch library into the main library. Anyone planning to do the same would do well to contact Denise for some advice, as she managed to overcome a large number of hurdles and the antagonism of all the staff, students, secretaries, other librarians and porters, to achieve her aim (which was forced upon her by a change of use of the original Chemistry library building).

Simon Alcock (Product Development Manager of Derwent's Life Sciences Division) talked about the Derwent Biological databases on STN, including GENESEQ and BIOTECHABS. An interesting statistic was that the number of sequences added to the GENESEQ in 2001 more than doubled the size of the database. Details of these databases can be found at http://www.derwent.com/geneseq/ and http://www.stn-international.de/stndatabases/databases/biotecha.html.

Barry Dunne gave a lively and stimulating talk on the CAS coverage of biological sciences. He described the biochemistry coverage, GenBank data, cell and other published biosciences information within Chemical Abstracts, STN, SciFinder and ChemPort. He also gave an interesting analysis of CAS sequence data from three major pharmaceutical companies, from which their relative research activities in this area could be compared. Information on CAS and the biological sciences can be found at http://www.cas.org; http://www.cas.org/stn.html; http://www.cas.org/SCIFINDER/bio/index.html; http://www.cas.org/chemport/index.html.

The RSC–CIG is to be congratulated on bringing together some interesting speakers, and providing an excellent day which, as always, was extremely good value (£25 including lunch, with a reduction for RSC members).

ExemplarChem 2002

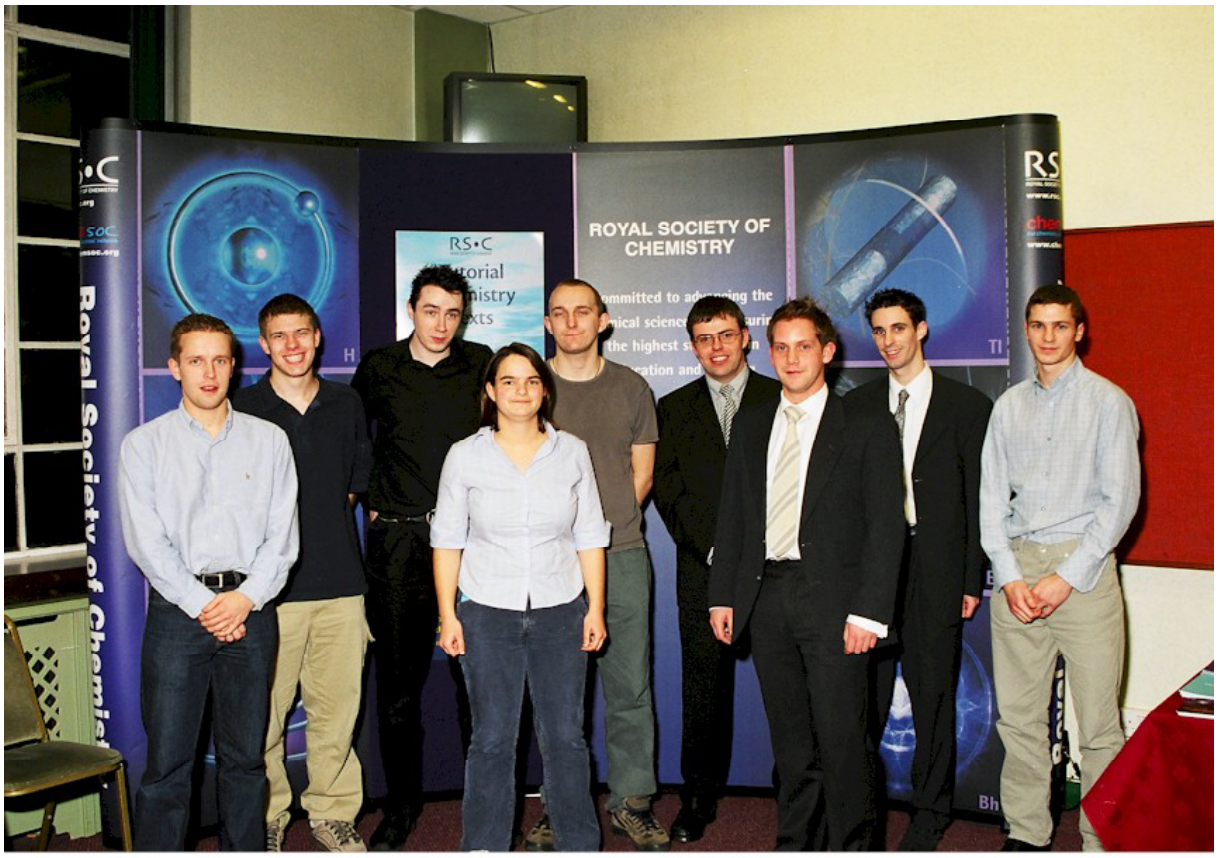
2002 is the third year of ExemplarChem, an internet exhibition of exemplary project work in chemistry, and the prizewinners' meeting on 29th November was once again organised by the Royal Society of Chemistry and held at the Scientific Societies lecture theatre in London. The competition attracts increasing numbers of entries. This year there were 43 entries from 19 different departments, and eight of the submissions were selected for prizes. Over 70 people attended the meeting, including participants, local co-ordinators, sponsors, members of the RSC, and other speakers.

There were four interesting lectures before the prize-giving. After an introduction from Sean McWhinnie of the RSC, Martyn Melvin from Evotec OAI gave an outline of the drug development process, and explained what Evotec OAI are looking for when they interview students for jobs in the company. Requirements include good communication skills, the ability to share knowledge, good IT skills and enthusiasm – all of which were also required from students entering the voluntary ExemplarChem competition.

Following a splendid lunch, Paul May from Bristol University (known to many for his 'Molecule of the Month') described Bristol University's experience of ExemplarChem. How does one inspire students to enter a voluntary competition after their exams, when the other options include having a holiday or staying in bed all day? It is clear that the techniques used by Paul May worked (which included convincing the students that it was fun and not a chore, but I guess that Paul's enthusiasm was a leading contributory factor), as there were a total of 70 entries from Bristol, three of which won prizes. Comments from entrants included 'It was the best thing we did in Chemistry all year' and 'I learnt more chemistry in those two weeks than in the rest of the year.' Certainly some of the advantages were there for all to see – those that entered had world visibility and a chance to be really creative. One previous ExemplarChem prize winner has been offered a job as a direct result of winning the prize.

Roger Gladwin then described the work of the Learning and Teaching Support Network (LTSN) Physical Sciences, which is based at Liverpool. The Centre provides information and advice to support learning and teaching in chemistry, physics and astronomy. It also provides funding for projects that show innovation in learning, teaching or assessment, such as a recent project on 'Developing business and entrepreneurial skills for physics students'. Full details about the work of the centre can be found at http://dbweb.liv.ac.uk/ltsnpsc/default.htm.

From left: Ian Grant (CSA Trust prizewinner) with Janet Ash and Paul May (Bristol University co-ordinator)



The ExemplarChem winners 2003. From left (chemistry unless stated): Mark Anderson (Oxford), James Crabb (Bristol), Ian Grant (Bristol), Kirsty Levasseur (Bristol), Anthony Dixon (Leeds), Liam Hook (Hull), Doug Bailey (Loughborough, Chemical Engineering), Rob Gill (Loughborough, Chemical Engineering) and Tim Smith (Warwick)

Before the prizegiving, Tony Murcott from Einstein TV gave a talk on 'Education – the killer of communication.' Einstein TV is a digital channel devoted entirely to leading edge science and technology, which aims to inform the public about science in an intelligent way. It is essential to entertain to get attention, and to put information over using terminology that the public can understand. Success in the media comes from inspiring people, and Tony's talk certainly showed great enthusiasm, and may have inspired some of the students present to look for a career in the media.

There were eight prizes awarded, thanks to this year's ExemplarChem sponsors: the CSA Trust, ISI, Pfizer, RSC, CAS and Evotec OAI. The overall winner received an additional prize, sponsored by the RSC. This was awarded to Mark Anderson from Oxford University for 'A Group Theory Tutorial – an Interactive Teaching Package'. As the department producing the overall winner, Oxford University received a cheque for £1000. Mark Anderson was also awarded the CAS prize. Ian Grant from the University of Bristol won the CSAT prize for his impressive entry on potential energy surfaces, which can be viewed at http://www.chemsoc.org/ExemplarChem/entries/2002/grant/about.html. The prize sponsored by Pfizer went to Kirsty Levasseur of the University of Bristol for 'Antiretroviral Drugs in the Treatment of HIV'. James Crabb of Bristol University won the prize sponsored by RSC Education in Chemistry for 'Alchemy'. Prizes also were awarded to Liam Hook, (University of Hull, 'The Third Phase', prize sponsored by ISI), Anthony Dixon (University of Leeds, 'Proteomics', prize sponsored by RSC Scientific Affairs), Rob Gill and Doug Bailey (University of Loughborough, 'The Removal of Chromium (VI) Ions from Aqueous Solutions using Solvent Impregnated Resins', prize sponsored by RSC Communications) and Tim Smith, (University of Warwick 'Ion Channels in Biological Membranes', prize sponsored by Evotec OAI). A full list of entries and winners can be found at http://www.chemsoc.org/ExemplarChem/lastyearhtm.

What was a most enjoyable day concluded with a wine reception and a photo session of prizewinners and sponsors. The students who attended and contributed to ExemplarChem should be proud of their efforts and hopefully felt that they have made a positive step forward in their careers in the chemical sciences.

The CaMeL and the JUMBO

David Bradley

Transparency in sharing information across the Internet has become essential. Today, a spectroscopist's data-stream flows as readily as the outpourings of the Human Genome Project. There is, however, a group of scientists that could have become marginalised but for the pioneering work of a handful of their number – the chemists. The problem involves the meaningful representation of what chemists talk about. It is easy to scribble down a molecular structure and add notes, but standard graphics programs lack chemical intelligence, so atom connectivity and other properties are not represented. Even chemical drawing packages are only capable of structural representation. There are countless nuances that might be associated with a structure: its molecular weight, the distribution of isotopic ratios of its carbon atoms, its spectral lines, or its X-ray data. How can this additional information be brought to hand transparently?

Not getting the hump

Thankfully, an answer has been forthcoming: CML (chemical mark-up language).

eXtensible Mark-up Language (XML) provides a universal format for structured documents and data on the Web and so offers a way for scientists and others to carry a wide range of information types across the net in a transparent way. All that is needed is an XML browser.

Yet there remains a barrier. How can the seeming simplicity of XML carry all those atomic coordinates, spectra and connectivity information in a transparent manner? The clue lies in the semantics: it is not called eXtensible for nothing. In recent years, a pioneering group of chemists has developed a chemical language system under the XML format to allow chemical information to be transported easily and retrieved and displayed usefully.

CML represents a sea change in the management of molecular information. It has been described as HTML for molecules, but it is so much more, having the scope to span disciplines from the smallest inorganic molecules, carbon monoxide, water and ozone, to vast macromolecular structures, such as polymers, proteins and DNA. It can even handle quantum chemistry. Its 'HTML for molecules' label hints at its potential to bring together disparate chemical information sources into a coherent and structured document format without loss of chemical information.

Chemical Internet pioneers Peter Murray-Rust, Henry Rzepa and Christopher Leach first introduced the concept of a Chemical Mark-up Language at the American Chemical Society meeting in August 1995. They realised that molecular modelling programs, such as MOPAC, generate interesting chemical information only in human readable form. The output is not 'self-defining' and was prone to change between versions. By making a chemical program 'CML aware' Leach showed how CML could add information of chemical value.

Mark-up languages are nothing without their translator, and Nottingham University virtual chemist Peter Murray-Rust, now at the Unilever–Cambridge University Molecular Informatics Centre, soon revealed a CML browser written in Tcl/Wish. Gradually, the language's formalisms evolved and with the emergence of the Java platform-independent programming system in 1996, Murray-Rust had evolved new species – JUMBO – a CML browser.

Importantly, the parallel evolution of XML means that XML now supports 'documents' and 'data' seamlessly. So, CML can handle many aspects of chemical informatics, data-handling and publication.

Chemistry journal articles such as those of Rzepa, Murray-Rust and Michael Wright (*New Journal of Chemistry*, 2001, 618-634) specified precisely within the body of the article text, chemical information. A CML-aware program, such as JUMBO, automatically recognises the molecules and displays them appropriately. Unlike conventional document formats which can be slow, error-prone, and non-reusable, an XML/CML article is information rich and portable.

The CML website provides examples of its handling capabilities including data files such as the International Union of Crystallography's CIF and the Protein Data Bank PDB formats. It can handle compound data cards, such as those being produced by the SELFML project to associate molecules and mixtures with their physicochemical properties and Materials Safety Data Sheets (MSDS). The system also allows for the lossless interconversion of various older formats, such as the Mol file, Sybil MOL2, JME, XYZ, SMILES, PDB and CIF.

An exciting recent development in the CML world is the creation of a sub-site on SourceForge.net, the Open Source development site. SourceForge.net provides a market for free software as well as services for developers including project hosting, version control, bug and issue tracking, project management, backups and archives, and

communication and collaboration resources. CML is application and vendor independent, so proprietary formats are no longer a problem. Rzepa explains that the CML project on SourceForge.net will carry all the machinery needed for users to make use of CML. This will include XML libraries and definitions, style sheets and Java-based tools.

Among the earliest CML-aware applications was Peter Ertl's JME chemical editor. The Sourceforge projects, JChemPaint editor and Jmol also have CML support and more recently there is Beda Kosata's (Prague Institute of Chemical Technology) BKChem, which bring us full circle. The various packages allow one to create bond-by-bond drawing of molecules, generate structures from templates for common carbon rings, add arrows, apply rich text, and align, scale and rotate molecules.

CML software is not only being developed in academia. Chemaxon's Marvin is a Java-based chemistry package that comes with an editor, conversion utilities for CML and although professionally developed is still free. MarvinSketch/Swing, for instance, is an applet for editing and visualising molecules on a web page, while its companion MarvinView/Swing can be used for viewing molecules.

Murray-Rust and his colleagues are collaborating with Dan Zaharevitz (at the US National Cancer Institute) to create 250,000 molecules in CML. This will form the core of an open molecular resource, which will collect both experimental and theoretical molecular data. Increasingly we will be able to ask 'What does this molecule do?' and 'Why is it important?', and get a comprehensive answer from a CML document. It will ultimately then be only a small step to answering questions such as 'Where can I find a molecule with certain properties?' and a chemically savvy search engine will be born.

David Bradley is a freelance science writer based in Cambridge, UK. He can be reached through his chemistry news website at www.sciencebase.com

Links

http://www.xml-cml.org/
http://www.cas.org/meetings/chemidxml/index.html
http://cml.sourceforge.net/
http://www.chemaxon.com/marvin/

Rzepa's New J. Chem article is available online at http://dx.doi.org/10.1039/b008780g

225th ACS National Meeting, New Orleans, LA, March 23–27 2003

Joint CSA Trust/CINF Symposium

Monday March 24th

Current Status of XML in Chemistry

B.A. Vickery, Organiser

1.30 pm	Compressed Chemical Markup Language (CCML) for Compact Storage and Inventory Applications **M. Karthikeyan, D. Uzagare, S. Krishnan**
2.00 pm	New Chemical Information Interchange Standards Based on CML: a Submission for the Object Management Group **M.A Miller, S.S. Markel, J.C. Esteva, W.L. Sharp**
2.30 pm	Novel applications of XML in Chemistry **P. Murray-Rust, H.S. Rzepa**
3.00 pm	The Family of XML Languages in Chemistry **H.S. Rzepa, P. Murray-Rust**

Information and Knowledge Management in the International Pharmaceutical Industry

Members of the Pharma Documentation Ring (P-D-R) represent the scientific information departments of the leading R&D-based pharmaceutical corporations. Over 45 such delegates from 11 different countries in Europe and North America attended the 44th AGM in October 2002 in Newmarket, UK, ably hosted by GlaxoSmithKline.

In his address, P-D-R President Sandy Mullen (Bayer AG) noted that Chemical Abstracts Services (CAS) register a new substance every 5 seconds. The availability of biosequence information had increased dramatically in recent years. Derwent took 20 years to reach 1 million sequences but only 9 months to add the second million to Geneseq. Later, Matt Toussant (CAS) reported on the growth in bioscience content in CAS databases. Over 40 per cent of the 23 million references in the CA database now related to bioscience topics, and the proportion of biomolecules had grown from about 10 per cent of all substances in 1995 to 45 per cent in 2002. The CAS Registry File itself now has more than 42 million substances. The P-D-R was asked to provide qualified input with regard to adding value to CAS's work in the sequence area.

Sandy Mullen also commented on the development of the P-D-R website (www.p-d-r.com). New publicly accessible features include the Pharmaceutical Company ranking tables (2001 and 2002) from Wood Mackenzie, PharmaProjects, Current Patents and overview articles from Pharma Marketletter. A new daily webspider service from Reg123 offers the very latest press releases from the P-D-R member websites. The 'members only' area of the website has additional functionality and content.

Four strategic topics were discussed:

- Measuring what we are doing
- Data and text mining
- Marketing information services
- Database futures

These more practical, pragmatic sessions proved particularly popular, there being a healthy and open discussion among delegates on the issues faced, the different approaches tried out in member companies, and an indication as their success or otherwise.

In the first session, there was a presentation from an external speaker from Elsevier Science. Peter Shepherd is the manager of the Counter Project, which involves the collection and reporting of statistics on the use of STM electronic journals. Peter explained that information managers in the private and public sectors were actively involved in the specification of statistical parameters that will help improve understanding of the ways in which electronic journals were being used in their organisations. This will assist the information managers to provide the electronic journal services of most value to their organisations. This was followed by two presentations by P-D-R company delegates on the results of internal quantitative studies on the use of information in their organisations. This useful sharing of non-confidential information, one of the most important aspects of P-D-R meetings, enabled delegates to take away some new and creative ideas on how to analyse and promote their information and KM activities.

Data and text mining of patents and the scientific literature is strong interest to P-D-R companies and recent reports suggest that this software sector will grow to a $3 billion business by 2003. The market place is still relatively immature with more than 200 vendors providing a wide range of offerings.

In the company reports sessions, it was very apparent that several companies are undergoing major internal restructuring as a result of further merger and acquisition activity. The larger companies are putting considerable efforts into establishing global licences with scientific and technical publishers and database providers. An ever increasing number of electronic journals are being licensed and made available to staff at all sites, particularly via the exploitation of linkage technology (i.e. database records to e-articles). Some companies had also evaluated the direct and indirect financial benefits resulting from the activities of the information services function.

A common theme was the significant merger and acquisition activity among information providers. It was noted that some of the larger publishers had acquired a number of smaller ones and that in terms of major database providers Thomson Scientific had acquired Current Drugs and Wila Verlag, and MicroPatent had acquired Aurigin. Some internal restructuring had also been observed and several information vendors had also formed strategic relationships with partner organisations.

This was a highly successful and interactive meeting and a survey of attendees indicated a very high level of relevant and quality presentations by P-D-R members and invited speakers. The meeting continues to be regarded by most participants as the major annual event of its type for the international pharmaceutical industry.

The 45th P-D-R AGM will be held on 24th–26th September 2003, near Thame, Oxfordshire, UK and will be hosted by Wyeth. The P-D-R web site can be found at www.p-d-r.com.

CSA Trust Grants for 2004

Purpose of the Grants:

The Grant Program aims to provide funding for the career development of young researchers who have demonstrated excellence in their education, research or development activities that are related to the systems and methods used to store, process and retrieve information about chemical structures, reactions and compounds. Grants will be awarded up to a maximum of one thousand U.S. dollars ($1,000) each. Grants are awarded for specific purposes, and within one year each grantee is required to submit a brief written report detailing how the grant funds were allocated.

Who is Eligible?

Applicant(s), age 35 or younger, who have demonstrated excellence in their chemical information related research and who are developing careers that have the potential to have a positive impact on the utility of chemical information relevant to chemical structures, reactions and compounds, are invited to submit applications. While the primary focus of the Grant Program is the career development of young researchers, additional bursaries may be made available at the discretion of the Trust. All requests must follow the application procedures noted below and will be weighed against the same criteria.

What Activities are Eligible?

Grants may be awarded to acquire the tools necessary to support research activities, or for travel to collaborate with research groups, to attend a conference relevant to one's area of research, to gain access to special computational facilities, or to acquire unique research techniques in support of one's research.

Applications must include the following documentation:

1. A letter that details the work upon which the Grant application is to be evaluated as well as details on research recently completed by the applicant;
2. The amount of Grant funds being requested and the details regarding the purpose for which the Grant will be used (e.g. cost of equipment, travel expenses if the request is for financial support of meeting attendance, etc.). The relevance of the above-stated purpose to the Trust's objectives and the clarity of this statement are essential in the evaluation of the application);
3. A brief biographical sketch, including a statement of academic qualifications;
4. Two reference letters in support of the application. Additional materials may be supplied at the discretion of the applicant only if relevant to the application and if such materials provide information not already included in items 1-4. Three copies of the complete application document must be supplied for distribution to the Grants Committee.

Deadline for Applications:

Applications must be received no later than October 17, 2003. Successful applicants will be notified by December 19, 2003.

Three copies of the application documentation should be forwarded to:

Bonnie Lawlor, CSA Trust Grant Committee Chair,
276 Upper Gulph Road, Radnor, PA 19087, USA.

E-mail submissions, if complete, may be forwarded to the Grant Committee at chescot@aol.com.

Company News

BCI and OmniViz get together

BCI Ltd. and OmniViz Inc. have announced a strategic relationship that brings together advanced chemical informatics analysis with innovative visualisation-based decision support approaches.

BCI has internationally recognised expertise in the use of clustering tools for the analysis of large chemical structure databases. BCI Ltd. directors Dr Geoff Downs and Dr John Barnard are renowned for their innovative work on clustering. Omniviz Inc. is the leading provider of integrated analysis and information visualisation to the pharmaceutical, biotechnology, and agricultural chemical industries. Its proprietary technology is the only software that integrates functional genomics, chemistry, toxicology, clinical, and marketing data with documents analysis.

The OmniViz software yields high-level data visualisations, enabling a single framework for decision support that can encompass all relevant data. The addition of the ability to cluster directly on chemical structures extends the power of OmniViz to work with multiple data types.

BCI anticipates that the collaboration with OmniViz will result in a uniquely powerful system that can overcome the significant informatics challenges currently facing industry, such as high throughput screening data analysis and compound selection on very large datasets. More information can be found on their website at www.bci.gb.com.

EVENTS 2003

March

1–31 March: The virtual Electronic Computational Chemistry Conference, ECCC9, entirely on the Internet, at http://eccc9.cooper.edu.

12–14 March: 2003 MDL European Users' Group Meeting, Royal Lancaster Hotel, London, UK. Contact: David Hughes, +44 (0)1276-701526, e-mail: d.hughes @mdl.com

17–18 March: Information Policies in a Complex World – Ownership, Access, E-legality, Security, Technology, Asset Management, EUSIDIC Spring Meeting, Renaissance Hotel, Karlsruhe, Germany. Contact: EUSIDIC Secretariat, c/o FIZ Karlsruhe, 76344 Eggenstein–Leopoldshafen, Germany. Tel: +49-7247-808-403; fax: +49-7247-808-114; e-mail: eusidic@ fiz-karlsruhe.de; http://www.eusidic.org

23–27 March: 225th American Chemical Society National Meeting, Ernest N. Morial Convention Ctr., New Orleans. Contact ACS Meetings, 1155 16th St., N.W., Washington, D.C. 20036-4899, USA. Tel: 1-800-27-5558, fax +1-(202)-872-6128, e-mail: natlmtgs@acs. org. This meeting includes the CSA Trust/CINF symposium; see page 7

24–27 March: From High-Throughput to High-Output: Can 'Omics' Assist Computational Drug Design? 3rd International Workshop on New Approaches in Drug Design & Discovery, Schloss Rauischholzhausen, near Marburg, Germany, organised by Gerhard Klebe and Hans-Joachim Boehm. Further information can be found at http://www.agklebe.de/workshop2003

31 March-3 April: IBC Life Sciences 7th Annual Drug Discovery Technology Conference and Exhibition Contact Michelle Bryant, Mortimer House, 37-41 Mortimer Street, London W1T 3JH. Tel: +44-(0)207-017-5069; fax: +44-(0)207-453-5422; email: michelle.bryant@ informa.com; http://www.drugdisc.com/europe/

April

2–4 April: Computational Approaches to the Understanding of ADMET Properties and Problems, MGMS Oxford Meeting series, Brasenose College, Oxford. http://bellatrix.pcl.ox.ac.uk/~mgms/admet_main.htm

May

4–8 May: Better Discovery Through Integration, MDL User Conference, Hilton San Diego Resort, 1775 East Mission Bay Drive, San Diego, CA 92109. Website: http://www.mdl.com/events/events.html#us-ugm. Please contact: conference2003@mdl.com

June

17-20 June: A Practical Introduction to Chemoinformatics, University of Sheffield, Sheffield, UK. See page 12 for details

29June – 3 July: 2nd European Symposium on Combinatorial Sciences in Biology, Chemistry, Catalysts and Materials, European Society of Combinatorial Sciences (ESCS), Bella Center, Copenhagen. Contact ESCS, Gamle Carlsberg Vej 10, Valby, DK-2500 Denmark, e-mail: tshj@novonordisk.com, http://www. crc.dk/escs/

August

10–15 August: Chemistry at the Interfaces, 39th IUPAC Congress and 86th Conference of The Canadian Society for Chemistry, Ottawa, Canada. Contact National Research Council Canada, Conference Services Office, Building M-19, Montreal Road, Ottawa, Ontario, K1A 0R6, Canada. Tel: +1-(613)-993-0414; fax: +1-(613)- 993- 7250; e-mail: iupac2003@nrc.ca; http://www.nrc.ca/ confserv/iupac2003/

10–15 August: Drug Discovery Technology World Congress, The Hynes Convention Center, Boston, MA, USA. http://www.drugdisc.com/section.asp

September

12–16 September: International Conference of Computational Methods in Sciences and Engineering (ICCMSE 2003) Kastoria, Greece http://www.uop.gr/ ~iccmse/ or http://kastoria.teikoz.gr/~iccmse/ Contact: Secretary ICCMSE, 26 Menelaou Street, Amfithea Paleon Faliron, GR-175 64, Athens, Greece; fax: +30210-94-20-091; e-mail: iccmse@uop.gr

7-11 September: 226th ACS National meeting, New York, NY, Javits Convention Ctr., New York City. Contact ACS Meetings, details as for March meeting

October

19–22 October: International Chemical Information Conference & Exhibition, Nîmes

December

1 December: CSAT AGM, Linnean Society, Burlington House, Piccadilly, London, CSAT Annual Dinner

2–4 December: Online Information 2003, Olympia, London, UK. Contact Learned Information Europe Ltd, Woodside, Hinksey Hill, Oxford, OX1 5BE, UK. Tel: +44-(0)1865-388000; fax: +44-(0)1865-736354; http://www.online-information.co.uk.

PRODUCT NEWS

STN Easy

STN International has significantly enhanced STN Easy for Intranets. The new capabilities in STN Easy for Intranets offer new options that support a wide range of user preferences and expertise:

1) pre-defined and editable search strategies from over 80 databases on STN Easy
2) the ability to post search results from the more than 200 sci-tech databases on STN.

The website is at http://www.cas.org.

Tripos releases SYBYL 6.9/ UNITY 4.4.

The software update features five major scientific technology upgrades that do the following:

- Design diverse but representative practical combinatorial libraries
- Streamline optimisation of lead compounds within a protein active site
- Compute topological index descriptors based on molecular structure
- Analyse molecular hydropathy and hydropathic interactions
- Perform rapid molecular electrostatic calculations

OptDesign assists in the design and redesign of combinatorial libraries, allowing the most efficient use of combinatorial robots by weighing competing objectives such as molecular diversity and reagent cost.

Real-time Automated Combinatorial Heuristic Enhancement of Lead compounds (RACHEL) generates new chemical structures within a protein-binding site to optimise binding and other properties (e.g. ADME) by building fragments onto a predefined molecular scaffold according to user-defined rules.

Molconn-Z, the third new module, calculates non-empirical topological and electronic descriptors for use in QSAR studies. The fourth, hint!, is a natural free-energy force field based on experimental data, and allows the modeling of hydrophobic interactions, solvation energy, entropy and other noncovalent effects. The fifth is ZAP, a new Poisson-Boltzmann solver that increases computational speed and accuracy over other methods and includes novel computational capabilities.

The SYBYL 6.9/UNITY 4.4 platform and associated applications have been configured to run on both SGI hardware and HP workstations and servers. A full LINUX release is planned for the first quarter of 2003.

MDL Information Systems

MDL Information Systems, Inc. (MDL) and Current Drugs Ltd. are formalising an agreement to index the Investigational Drugs Database (IDdb) into DiscoveryGateSM, MDL's web-based service that integrates, indexes, and links the essential content sources required for effective drug discovery.

MDL has also announced the release of MDL Core Interface, the middle-tier component of MDL Discovery Framework. MDL Core Interface provides session/task management, security, database access, object storage, and other fundamental services. A common application programming interface (API) supports Microsoft- and Java-based development tools and tight integration with Oracle technology. The open API also enables developers to focus on building value-added, integrated applications that enhance the research process and enable scientists to collaborate effectively in discovery projects. Operating within MDL Discovery Framework, MDL Core Interface provides the integrating foundation for all of MDL's new products, including DiscoveryGateSM, MDL's new web-based research environment.

Cerius2 4.8 from Accelrys

Three new scientific application modules are included in this release to address key issues in combinatorial library design, QSAR analysis, and structure-based drug design. The Cerius2 components being released include:

- C2.LibX - a novel approach to designing synthetically feasible, small focused libraries from large virtual combinatorial libraries.
- C2.NNet - This new regression method uses back propagation neural networks with categorical, numeric or 2D fingerprint data. The algorithm provides many unique and novel features including automated selection of misclassification penalties, an improved BFGS minimiser that works without user-adjustment of parameters on a wide variety of problems, and sensitivity analysis for determining the importance of each variable.
- C2.AutoLudi - an automated de novo design with optional minimisations and descriptor screening of intermediates and final output that can be used where molecules are invented in phases, where a single molecule is created, and where R-group substituents for specified sites are combinatorially enumerated.

Further information can be found on the Accelrys website at http://www.accelrys.com

Contact Details

Chair of Trustees

Guenter Grethe
Tel: +1-(510)-357-2222, ext 1430
e-mail: G.Grethe@mdl.com

Treasurer

Geoff Downs
Tel: +44-(0)1909-515444
e-mail: geoff@bci.gb.com

Secretary

Clive Weeks
Tel: +44-(0)1252-674890
e-mail: Drcliveweeks@aol.com

Sponsorship Secretary

John Holliday
Tel: +44-(0)114-222-2685
e-mail: J.D.Holliday@sheffield.ac.uk

Newsletter Editor

Frances Daniel
Tel: +44-(0)1949-844837
e-mail: francesdaniel@bottesford65.fsnet.co.uk

Next Copy Date

April 30 2003

A Practical Introduction to Chemoinformatics

17-20 June 2003

University of Sheffield, Sheffield, UK

This intensive short course is intended to provide an introduction to the major aspects of chemoinformatics, with particular emphasis on applications in modern drug discovery. The course will comprise a mixture of hands-on workshops, lectures, and informal discussions. A key focus will be on the underlying rationale/theory, rather than teaching which buttons to press in a particular piece of software.

Two key learning objectives are:

> What is the most appropriate method(s) to use for a particular problem?
> Why does a particular method give a particular result?

Topics covered

- 2D databases and database searching (substructure search, virtual screening, property searching, similarity searching)
- Diversity and compound selection (clustering, DBCS, cell-based approaches)
- Moving into 3D: experimental data sources (small-molecule crystallographic data: CSD, intermolecular interactions: IsoStar, SuperStar, Relibase)
- Computational methods for 3D (introduction to molecular mechanics and conformational analysis)
- 3D databases (database searching, pharmacophore generation)
- Combinatorial libraries (molecular descriptors, diversity vs focused library design, lead- and drug-like concepts, enumeration and monomer selection)
- Analysis of high-throughput screening data (data mining, data visualisation)

Course Venue

The Department of Information Studies at the University of Sheffield is recognised as one of the world's leading centres in chemoinformatics. The course will take place in the Department's new chemoinformatics research laboratory. A variety of commercial software packages will be used for the hands-on elements of the course. En-suite accommodation will be provided in a University Hall of Residence.

Course Tutors

The course tutors are Professor Peter Willett and Dr Val Gillet from the University of Sheffield, Dr Andrew Leach from GlaxoSmithKline and Dr Robin Taylor from the Cambridge Crystallographic Data Centre. They have many years of practical experience in the field, in both academic and industrial environments. There will be ample opportunity for participants to discuss their own chemoinformatics problems with the tutors.

Registration

For more details please contact:

Dr Val Gillet
Department of Information Studies
University of Sheffield
Sheffield S10 2TN, UK
E-mail: v.gillet@sheffield.ac.uk