This issue contains reports about two meetings co-sponsored by the Trust: the joint CINF/CSA Trust Symposium at the ACS meeting in Anaheim in March (see page 2), and the Sheffield Conference on Chemoinformatics held in April, co-sponsored with the MGMS (see page 4). We are very grateful to all the companies who make Corporate Donations to the Trust (listed on page 12), and would like to acknowledge particularly the five companies who donated funds for bursaries for the Sheffield conference: AstraZeneca, Cambridge Crystallographic Data Centre (CCDC), GlaxoSmithKline, Hampden Data Services, and Pfizer. CCDC, who have supported the Trust over many years, have this issue’s Sponsor’s Spotlight, on page 9.

**Report from the Chair**
Guenter Grethe on the CSA Trust/CINF symposium at the Spring ACS meeting

**3rd Joint Sheffield Conference**
All the news from the CSA Trust/MGMS Chemoinformatics Conference

**Chemical Information from $^{85}$Ac to $^{40}$Zr**
Doug Veal sees RSC/CIG spread its wings into the area of chemical data

**Events**
What’s on, when and where in the world of chemical information in 2004

**People and Places**
Andrea Twiss-Brooks chairs CCAS; Carol Duane sets up consultancy; Julian Hayward joins LHASA

**Obituary**
Wendy Warr celebrates the life of Angela Haygarth-Jackson

**IPI-ConfEx**
Lisbon was the location for a new patent conference. Bill Town reports

**IPI Award 2004**
Dr Jacques Michel is the recipient of this year’s International Patent Award

**CSIRO Molecular Science**
David Winkler outlines the work of the prestigious Australian research institute

**Sponsor’s Spotlight**
Focus on the research of the Cambridge Crystallographic Data Centre

**Chemoinformatics in China**
Xiaoxia Li on the 24th Chinese Chemical conference

**Contact Us**
Contact details of the key Trustees and Newsletter Editor

www.csa-trust.org
The Chemical Structure Association (CSA) Trust and the Division of Chemical Information (CINF) held their annual joint symposium during the 227th National Meeting of the American Chemical Society in Anaheim on 29 March 2004. The one-day symposium on ‘Research Collaboratories, Virtual Laboratories, and Grid Computing’ was organised by Wendy Warr and Guenter Grethe and featured 13 speakers from academia and industry. Good attendance at all sessions indicated the high level of interest information scientists have in this timely topic.

The University of Southampton in the UK seems to be a hotbed for grid computing. Three speakers from that university, all participants in the Comb-α-Chem Pilot Project, discussed several applications of grid computing within this project. Jeremy Frey and Michael Hursthouse described respectively the development of a grid service for interaction with the UK EPSRC National Crystallography Service, and the feasibility of a Grid-enabled chemical structure and property environment. Jonathan Essex gave two presentations on applications of distributed computing to biomolecular simulations, such as protein-ligand docking and molecular dynamics simulations. In connection with his first presentation on the BioSimGRID project he also discussed security issues associated in general with distributed computing.

With ever-increasing amounts of data to be managed in the discovery process, it is only logical that grid computing would become a fertile field. Graham Richards, Oxford University, recipient of the 2004 ACS Award for Computers in Chemical and Pharmaceutical Research, talked about virtual screening involving some 2.5 million PCs in over 220 countries, equivalent to about 250,000 years of CPU time. He was followed by Ference Darvas of ComGenex, Inc., talking about OpenMolGRID, a Grid-based large-scale drug design system.

Chris Crafford from United Devices described UD’s grid computing technology applied to the drug discovery process and showed several examples. Several research organisations and pharmaceutical companies are using this technology. The development and implementation of grid-enabled, automated predictive QSAR modelling using a combinatorial approach was discussed by Alex Tropsha, of the University of North Carolina. The program ‘Structure-activity relationships for the design of molecules’ (STARDoM) using IBM’s WebSphere system is implemented on the university’s CarolinaBioGrid through a Globus Toolkit.

Yong Zhang, from Cambridge University, discussed semantic grid computing as used in the WorldWideMolecularMatrix. This collaborative effort with Peter Murray-Rust and Henry Rzepa applies knowledge-based computing for the web to chemistry. The implementation uses XML-CML for molecules and properties and allows very rapid retrieval of molecules (50 milliseconds from a database of 250,000 molecules). Calculations of molecular properties also proceeded very fast using parallel computing. The development of a web-based grid-computing environment for research and education in computational science and engineering was described by Thanh Truong, University of Utah. The system, Computational Science and Engineering Online (CSEO), as presently applied to multi-scale modelling of complex reacting systems provides researchers with state-of-the-art tools and allows access to data not available locally. The development of a personal computing environment on the grid for molecular design was the topic of a talk by Umpei Nagashima, from the National Institute of Advanced Industrial Science and Technology, Japan. He discussed two products, Molworks and Gaussian Portal, that together represent a high-throughput desktop virtual laboratory.

The Collaboratory for Multi-scale Chemical Sciences (CMCS) as described by James Myers, of Batelle Pacific Northwest National Laboratory, represents an adaptive informatics infrastructure for multi-scale chemical science and enables the flow of information across physical scales and disciplines. Advanced collaboration and metadata-based data management technologies helped to develop a portal that provides distributed research support, data discovery and other capabilities. IBM’s efforts in the area of multi-organisational collaboration among Life Science Communities were discussed by Gary Benesko. Virtual Research Park (VRP) is a web-based research environment that facilitates joint R&D and collaboration extending beyond any one enterprise or geography.

The broad variety of presentations by experts in their respective fields provided attendees with an excellent overview of this important topic.

Abstracts can be found on the CINF website at http://www.acschem.org/cinf/meetings/227nm/227cinfabstracts.htm.

Guenter Grethe

---

**Chemical Structure Association Trust AGM 2004 and Annual Dinner**

The 2004 CSA Trust Annual General Meeting will be held on Monday 29 November at 2pm in the Roscoe Room, the SCI, 14-15 Belgrave Square, London

The meeting will be followed by the Annual Dinner (venue to be confirmed). Further details will be announced later in the year.
People and Places

Andrea Twiss-Brooks has been appointed as Chair of the ACS Joint Board-Council Committee on Chemical Abstracts Service (CCAS). Her term of service will be 2004–2006. Andrea encourages CINF members (and interested end-users they identify) to contact her with comments, concerns or suggestions about CAS products or policies. For more information about CCAS, visit http://www.chemistry.org/committees/ and choose the appropriate link.

After two decades in custom chemical manufacturing and pre-clinical drug development, Carol Duane is now CEO of her own start-up consulting business, advising the pharmaceutical outsourcing industry about marketing practices.

Julian Hayward has joined LHASA Ltd as Company Manager and will oversee company strategy, organisational change, and the day-to-day running of the business. Philip Judson has been appointed to the new position of Chief Scientist. Chris Hardy has recently joined the company to manage the ITIC project, a major initiative for the development of a new ‘industrial strength’ toxicology database.

Andy Berks has moved from Merck to IVAX Pharmaceuticals.

Following the sale of Libraria to Sertanty, Barry Bunin is running his own consultancy business, advising the pharmaceutical outsourcing industry about marketing practices.

Terry Stouch has moved from BMS and is now Head of Computational Chemistry at Lexicon Pharmaceuticals.

Hugh Look has been appointed Senior Consultant at Rightscom, the digital strategy consultancy. He previously maintained an independent practice while developing a partnership with Rightscom but is now a full-time employee focusing on the company’s projects for the European Commission, as well as developing contacts within UK government.

Obituary: Angela Haygarth-Jackson

Angela Haygarth-Jackson was a pioneer in the new discipline of “information science” and also had the distinction of achieving senior management status within ICI (Imperial Chemical Industries) at a time when women managers were few and far between in the chemical industry.

She was the first woman to hold the office of President of the Institute of Information Scientists, in 1983–84. Appointed OBE in 1984 for services to information science during her presidential year, she was an active member of many national and international committees including the Royal Society Scientific Information Committee, the British Library Advisory Council, and the International Users Council of the Chemical Abstracts Service Division of the American Chemical Society. Although not a chemist, in 1993 she was awarded fellowship of the Royal Society of Chemistry.

Angela Ray Haygarth-Jackson was born in Heaton Mersey in Cheshire in 1929, and educated at Cheltenham Ladies’ College and Manchester University, where she was awarded an MSc in Botany. Deciding not to pursue further research, she joined ICI Pharmaceuticals Division in 1956 as an information scientist in the Techno-Commercial Department then based in Blackley, Manchester. She transferred to Alderley Park, Cheshire, in 1957, and started to develop the division’s information services as a member of the library staff. In 1968 she was promoted to Head of Literature Services Section with responsibility for the three libraries and technical information services that supported ICI Pharmaceuticals’ international business. When, in 1984, Information Services Section was formed from the merger of literature services and data services, she headed an even larger group.

Over the years, as new technology became available, Haygarth-Jackson and her colleagues implemented dramatic advances in information processing techniques. This had an impact not just on the company’s pharmaceuticals division: she was also a member of the management team which co-ordinated activities within the ICI Group Information System.

There were very few women of Haygarth-Jackson’s managerial rank in ICI in the early 1980s. A handful of them used to lunch together occasionally in what their male colleagues called the “Ms Mafia”, a term that amused Haygarth-Jackson. She was by no means a feminist in the 1970s sense, but she was a mentor and example to the many women who worked for her; a caring manager who valued her staff and developed them to their maximum potential.

She lectured widely on information science, and in 1979 undertook a three-week tour of Australasia to give nine lectures on new technologies and resources available in literature services. In 1986 she spent three weeks in China advising on library and information science matters. She was an external examiner of the MSc course in Information Studies for the Department of Information Studies at Sheffield University.

Continuing to serve the profession in her retirement, she was the editor of Training and Education for Online (1989). She was chairman of the editorial board of the Journal of Documentation for many years, finally resigning at the end of 2001.

Those who had never met Angela Haygarth-Jackson, in particular Americans, were often advised that meeting her would be like meeting the Queen. Perhaps there was a slight physical resemblance, but the greatest similarity was in the effortless social graces, hospitality and genuine interest in people. Angela was certainly not regal in the “not amused” sense: she enjoyed recounting the tale of her chauffeur-driven car following the dustcart into Buckingham Palace when she was invested as OBE. She lived in the family home in Bowdon, Cheshire, for 40 years. She was a do-it-yourself enthusiast and did her own painting and wallpapering. The tapestries displayed in the house are a tribute to her embroidery skills. She was keen on the countryside and an ardent supporter of the Royal Society for the Protection of Birds.

Wendy Warr


Reprinted by permission from The Independent, Obituaries, 2 June 2004
The third Joint Conference on Chemoinformatics was held at the Octagon Centre in Sheffield in April. The conference was organised by the CSA Trust, the Molecular Graphics and Modelling Society and the Department of Information Studies, University of Sheffield. The CSA Trust awarded bursaries to Shane Weaver of the University of Leeds, Ingrid Socorro-Gutierrez, from the Chemistry Department, University of Cambridge and John McNeany, from the Department of Physical Chemistry, School of Chemistry, University of Nottingham. In the following reports, they share their conference experiences with us.

Sheffield’s Millennium Galleries were the venue for an ‘exceptional dinner’

On first arrival at the Octagon centre, it was evident that we were at a well organised, comfortably spacious venue ready for some interesting and exciting talks and poster presentations.

The conference began with a brief introduction from event organiser Val Gillet. After a warm welcome to Sheffield, the conference continued with the afternoon session dedicated to ligand-based screening. In particular my attention was drawn to the talk presented by Mark Mackey from Cresset BioMolecular Discovery on the use of molecular fields for ligand-based virtual screening. By analysing ligand electron clouds their algorithm generates ‘field points’ which then represent the ligand’s surface and shape which is presented to the protein cavity during binding. They resolved the problem of comparing complicated field point patterns through representation by a collection of 1D data composed of distance and topology for each field point pair within the pattern. Results from a commercial study using this technique were impressive in comparison to traditional virtual screening methods. I look forward to further validation studies using their method. In a domain with such a knowledgebase of experience it is always refreshing to see presentations using novel techniques.

That evening the poster session was an excellent chance to converse with students and speakers, as well as see demonstrations by commercial exhibitors.

Thursday morning’s session consisted of screening compound collections design and QSAR. The presentation by Kevin Cross (LeadScope) entitled ‘Prediction Model Building Based on Classifying Compounds by Structural Features’ was of particular interest. Instead of using slower MCS techniques they have built a structural feature library containing 27,000 entries. By using these as well as new features generated from the training set they can be ranked with respect to the biological response of their originating compounds. The method was evaluated using the PTP-1B dataset and we were taken through their parameter optimisation process and shown the powerful predictive power of the technique.

Data mining was the subject of the final session on Friday. Christos Nicolaou gave a step-by-step description of the automated decision support for screening developed at Bioreason. The method efficiently detects structural classes before performing an R-group analysis for each class individually. This data is then processed by an in-house algorithm for extracting meaningful rules for preliminary screening of large compound datasets.

Overall it was a thoroughly enjoyable and informative experience.

Shane Weaver
The conference was divided into several sessions and took place in a friendly scientific atmosphere. The first session on 'Ligand-based Virtual Screening' covered a variety of methods based on different molecular descriptors and similarity searching techniques for the optimisation of virtual screening. The day ended with a poster exhibition and a buffet.

The second day was divided into three parts. The first consisted of three lectures on 'Screening Collection Design', with a talk on a quantitative model relating chemical structural similarity to biological activity among the other absorbing presentations.

Part two was on 'Model Building/QSAR' including the description of a new approach for building predictive models based on structural features using 2-D molecular descriptors, and the presentation of a new virtual screening method providing an accurate prediction of molecular properties that can be used to optimise the screening process.

The final session, 'Structure-base Drug Design', began with the description of CavBase, a method for detecting binding site similarities. Another presentation discussed a method for complexity analysis, which offers a fast and effective ranking technique for the elimination of structures with unfeasible molecular motifs. The day ended with an exceptional dinner at the Millennium Galleries, an elegant venue right in the heart of the city.

Friday saw seven lectures on data mining, including the presentation of a new algorithm for automated creation of 2D structural formula of molecules. Also, new automated approaches for the extraction of structure-activity relationship rules were presented. The session ended with the description of a new method for the identification of relevant sub-structures in screening data.

Finally, I would especially like to thank the CSA Trust for their bursary, and all the organisers for the opportunity of presenting a poster, with some results from my PhD project at Cambridge University. Thank you also to the rest of the attendees for a very enjoyable meeting.

Ingrid Socorro

As a final year PhD student, the third Joint Conference on Chemoinformatics gave me an invaluable insight into the current state of industrial research in the field of computational approaches to drug discovery. The speakers covered a wide range of topics, from structure based drug design, to virtual screening, in detail and with enthusiasm. This highlighted the pace at which chemoinformatics is progressing.

There were several international speakers and delegates, demonstrating the global nature of this field of research, in addition to underlining the necessity of communication and collaboration. The talk presented by Martyn Ford showed what can be achieved through large-scale collaboration in the information age.

The poster sessions were an excellent showcase for the many posters presented. I had the opportunity to present my own work in poster form and to discover the many research interests of the delegates. During the poster sessions, I was able to discuss my work with industrial representatives and academic researchers. As always at meetings in Sheffield, the conference dinner was excellent, in the grand setting of the Millennium Galleries.

John McNeany


Sheffield PhD student Jayshree Patel (centre) with her poster at the conference’s popular poster session. Left: Ifat Noreen (MSc student, Sheffield); right: Sally Mardikian (PhD student, Sheffield).
This three day conference was held for the first time in March of this year, and was billed as the first annual conference of its type. One measure of the success of this first conference was the fact that over 270 patent information professionals from more than 25 countries travelled to Lisbon to participate in the meeting. General opinion among attendees was that the meeting was well organised and that it filled an unsatisfied need.

During the meeting, it was announced that the winner of the 2004 IPI Award (see below) was Jacques Michel, formerly responsible for DG 1 at the European Patent Office, who initiated a number of significant projects, including BACON and esp@cenet.

Over 30 papers were presented during the three days and there were six pre- and post-conference workshops. The keynote address was given by Dr Schoch-Gruebler (BASF) who described the users’ needs as seen from a chemical industry viewpoint, examined some of the forces shaping the landscape of the patent information market and questioned just how fruitful these recent developments may be. The landscape is continually changing and confronting patent offices, database producers, hosts and information professionals with new challenges.

From the papers delivered at the meeting both by information producers and patent professionals, it was clear that the patent information business is vibrant and undergoing rapid evolution.

New database products from new players (such as PatBase from Minesoft) are challenging the market leaders. The different market positions of traditional hosts and web-based services are starting to make a significant impact on the ways in which patent information is exploited by industry, as budget constraints force patent professionals to opt for lower-priced but lower functionality offerings.

Bill Town  
President, Kilmorie Consulting  
www.kilmorie.com

IPI Award 2004

Dr Jacques Michel, former Vice President of the European Patent Office (EPO), is the recipient of the International Patent Information (IPI) Award 2004.

Dr Michel studied for a doctorate in Physical Sciences at Paris University and from there started his career as Scientific Attaché to the French Embassy in Washington DC. Later, at the Secretariat of State for Research and the Ministry of Industry he became involved with information systems managing scientific and technical data. He directed the creation of an online information service, which was to become Questel, and also oversaw the implementation of chemical structure information management using the DARC system. In the 1980s, he became a Director at CNRS (Centre National de la Recherche Scientifique) and then moved to head up Questel, first in the US in Washington D.C., and then as Managing Director of the Questel parent company in France.

In the late 1980s, Dr Michel moved to the EPO in the Hague to become Vice President of Directorate General 1 where, under his leadership, the BACON (BAckfile Conversion) project was implemented to create a unique worldwide collection of patent documents in bitmap format. He also led the pioneering project that designed, built and implemented a new process and system for patent examination, the EPOQUE system, which is today used by 4,000 patent examiners. From EPOQUE there came esp@cenet, yet another very significant service.

The Award is presented in association with the International Patent Information (IPI) Institute and is sponsored by Technology & Patent Research (TPR) International Inc.

The IPI Award has been established to pay tribute to one individual each year who, through their career as a dedicated patent information professional, has made a significant positive impact on patent information. It is intended to highlight the patent information profession in general and how the profession is valued by others. The IPI-Institute is an organisation which builds upon this theme and has been formed to work alongside national patent information user groups and associations for the benefit of the profession.
This meeting, organised by the RSC Chemical Information Group, was held at Burlington House, Piccadilly, London, on March 30 2004. For this meeting, the CIG spread its wings into the area of chemical data, covering the whole spectrum of chemistry – organic, inorganic and physical.

The first presentation, ‘How Inspec solves chemical search problems’, was a double act between Eva Dimmock and Susan Bates of Inspec. They reviewed the organisation of Inspec and its products, emphasising the quality of their databases. They showed how chemical searches could be carried out and where chemical data can be found within Inspec files.

Next, Chris Gibson of Manchester University spoke about the Knovel service: ‘Knovel – beyond e-books’. Although he is only a user, he clearly demonstrated his enthusiasm for the service. It provides very comprehensive coverage of physicochemical data and a lot of work has gone on behind the scenes to make the data readily accessible and usable. The tools include unit conversions, and many graphical and statistical analysis tools. Results can be exported for further processing, for example, in spreadsheets.

Robert Scoffin of CambridgeSoft then spoke about the Merck Index. The current (13th) edition covers over 10,000 monographs, and is available in paper, on CD-ROM and via the internet. The content retains a pharmaceutical bias, reflecting the historical origin of the Merck Index. This edition is taking full advantage of the electronic age to offer improved searching, for example, by full text and substructure search.

‘The materials IT revolution’, the final paper of the morning, was given by Patrick Coulter of Granta Design. Granta provide software for a number of organisations supplying data services. He showed samples of different materials, and illustrated again how access was being improved by the use of electronic search tools, and the ability to consolidate data across different databases, including the ability to search for materials with a particular set of defined properties.

The afternoon session began with Rainer Poerschke of Springer talking about Landolt–Börnstein. They have data on about 100,000 organic and 100,000 inorganic compounds, totalling 300 volumes, and all evaluated by experts. Their focus is on the data, primarily in printed form, rather than the technology of access, but they do now provide CD-ROMs. Full-text search and index-supported navigation are also available.

Don Parkin of the Daresbury Chemical Data Service then spoke about 'Inorganic and thermophysical property information for the UK academic'. The CDS provides access to a number of physicochemical databases, including ICSD, DETHERM and CRYSTMET. The data cover 4.9m datasets, comprising over 400 properties of 24,000 pure compounds and 105,000 mixtures. They aim to provide ease of searching and dynamic use of data.

Mark Leach of Salford University then gave a presentation on 'The Chemical Thesaurus 3.2'. This is a different approach. Mark is seeking to build from scratch a normalised, relational database covering materials and their transformations, starting from the periodic table. This is essentially an on-going research project that is evolving into a chemical thesaurus. The information is freely available on the web, and collaboration is welcomed.

Finally, Fiona Macdonald of Chapman and Hall/CRC Press spoke on ‘Trying to make sense of inorganic chemistry – the Dictionary of Inorganic Compounds’. The 84th edition of the Handbook of Chemistry and Physics contains data on about 4k inorganic compounds. Compounds are selected for the Dictionary of Inorganic Compounds on the basis of being important and/or interesting. Fiona reviewed the problems of dealing with these, e.g. nomenclature, formulae and structure searching, and how these problems are being dealt with.

Overall, this successful meeting demonstrated clearly how access to and use of data is being greatly improved by the use of computerised techniques and provided an excellent overview of the current state of the art.

**Doug Veal**

Some of the presentations are available on the CIG website: [http://cds.dl.ac.uk/cds/CIG/powerpnt.html#mar04](http://cds.dl.ac.uk/cds/CIG/powerpnt.html#mar04)

---

**One-Day Training Course on Chemical Information for the Non-Chemist**

Royal Society of Chemistry, Burlington House, Piccadilly, London

Tuesday 12th October 2004

Organised jointly by the CSA Trust and the RSC-CIG

This one-day training course will include a series of presentations that will provide an impartial and critical overview of Chemical Information provision.

These presentations will be given by specialists from both academia and industry.

The major objectives of the course are:

- To show why chemical information is different to other types of information
- To describe how chemical structure information is handled
- To demonstrate how to search for specific and generic chemical structures and chemical reactions
- To review relevant databases

The cost of the training course will be around £20 for members and £30 for non-members. Further details will be available shortly.

Contact:
Doug Veal: doug@dovertonltd.freeserve.co.uk
http://cds.dl.ac.uk/cds/CIG/cig.html#Future
CSIRO Molecular Science is a research division of CSIRO Australia, one of the world’s largest scientific research organisations. CSIRO Molecular Science has particular expertise in chemical and biological sciences. We aim to assist the development of industries in the chemical, plastics, pharmaceutical and human health sectors. Our core science areas are:

- **Biomaterials** – we are developing synthetic materials compatible with biological systems for use in medical devices such as pacemaker leads, cardiovascular devices and contact lenses.
- **Bioactive Molecule Discovery** – we design and synthesise novel molecules with specific application in the pharmaceutical, veterinary drug and crop care industries. CSIRO was involved in developing the influenza drug Relenza, and the insecticide Cycloprothrin.
- **Polymer Materials** – we devise high-performance polymeric and composite materials with improved functionality for use in areas as diverse as drug delivery, aviation, polymer electronics, photonics, self-healing and self-sensing materials.
- **Biocatalysis** – we use biological entities such as bacteria to convert low value commodities to high value products by processes such as fermentation.
- **Sensors and Diagnostic Devices** – we are developing products and processes that detect molecular change for use in applications such as overt and covert authentication devices and disease prediction and detection.

The Division works with a broad range of local and multinational business and research collaborators to develop and innovative materials, products, processes and services. Several companies have been set up to exploit our research results.

CSIRO Molecular Science contains many discipline-diverse research groups. One such group (consisting of Professor Frank Burden, Dr Mitchell Polley, a recent PhD student Darryl Jones, and led by Professor David Winkler) studies computational methods of molecular design, with a particular interest in bioactive molecules. The group uses traditional and in-house methods to understand and predict the behaviour of bioactive molecules. They have studied *de novo* ligand design based on crystal structures of the rice blast melanin biosynthesis enzyme trihydroxynaphthalene reductase, the insulin receptor ectodomain, the IGF ectodomain, and the edysone receptor complex. A major research theme is the development of new, robust quantitative structure-property modelling methods. They have developed robust methods of structure-property mapping using Bayesian regularised neural nets, eigenvalues of molecular matrices (developed as BCUT diversity metrics by Pearlman), novel molecular descriptors based on similar theory, and new parsimonious methods of variable selection based on artificial relevance determination (ARD) and other Bayesian methods. Recent research has looked at using these new methods to build robust, predictive models of important ADME properties such as phase II metabolism, partitioning across the blood–brain barrier, acute toxicity, and intestinal absorption. Recently, the group has been looking deeper into the use of methods and tools of complex systems science for molecular design, and has received funding to establish the CSIRO Emerging Science Centre for Complexity in Drug Design, studying emergence in chemistry and biology, using tools such as agent-based modelling methods, artificial intelligence and genetic methods. The aim is to find new approaches to molecular design, using complexity, rather than reductionist methods. The group are running a symposium on ‘Complexity and Related Computational Methods in Bioactive Discovery’ at the Pacifichem 2005 congress in Honolulu in December of that year.

**David Winkler**

---

**CSIRO Molecular Science**

**CSA Trust Grants for 2005**

Grants are available for the career development of young researchers who have demonstrated excellence in their education, research or development activities related to the systems and methods used to store, process and retrieve information on chemical structures, reactions and compounds. Grants will be awarded up to a maximum of $1,000 each. Grants are awarded for specific purposes, and within a year each grantee has to write a brief report on how the funds were allocated.

**Who is eligible?**

Eligible applicants are age 35 or younger, who have demonstrated excellence in their chemical information related research and whose careers will potentially have a positive impact on the utility of chemical information relevant to chemical structures, reactions and compounds. Additional bursaries may be awarded at the Trust’s discretion. All requests must follow the application procedure and will be judged by the same criteria.

**What activities are eligible?**

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

**Applications must include the following:**

- A letter that details the work upon which the application is to be evaluated and details on research recently completed;
- The amount requested and the details about the purpose (e.g. cost of equipment, travel expenses). The relevance of the purpose to the Trust’s objectives and the clarity of this statement are essential in the evaluation of the application;
- A brief biographical sketch, with a statement of academic qualifications;
- Two reference letters in support of the application.

**Procedure for submission**

Four copies of the documentation should be forwarded to:

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

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

The deadline for applications is October 15, 2004. Successful applicants will be notified by December 17, 2004.
The Cambridge Structural Database (CSD) and its companion software and knowledge bases are widely acknowledged as indispensable in the investigation and prediction of:

- Molecular dimensions
- Conformational preferences
- Intermolecular interactions.

The Cambridge Structural Database, the world’s repository of validated crystal structure data for organic and metal organic molecules, currently contains over 320,000 entries. The database provides comprehensive coverage of all published crystal structures determined by x-ray and neutron diffraction analysis and also includes personal communications data not available elsewhere. Tools are supplied for creating in-house databases that are fully searchable, either independently or in conjunction with the CSD itself.

The search interface ConQuest provides a full range of text/numeric search options in addition to more complex functionality, such as chemical substructure and geometrical searching. The latter can be used to investigate non-bonded contacts, including pharmacophoric interactions. Hitlists can be exported into Mercury, which offers extensive facilities for crystal structure visualisation, and the results of geometric searches may be exported to Vista for manipulation and graphical display.

The CSD System also includes extensive libraries of extracted molecular and intermolecular geometry information organised according to molecular fragments and functional groups. These ‘knowledge bases’ provide rapid and direct access to essential data for use in modelling experiments, including conformational validation and protein ligand docking. Mogul contains bond lengths, valence and torsion angles, displaying results in the form of histograms with associated summary statistics. IsoStar is a compilation of intermolecular interactions derived from the CSD and the Protein Data Bank (PDB). Data are displayed in the form of scatterplots or as contoured maps. In both cases, individual data points can be traced, via hyperlinks, to the structures in which they are observed experimentally.

Protein-ligand docking with proven accuracy and reliability using GOLD

GOLD uses a genetic algorithm (GA) to dock flexible ligands into partially flexible protein binding sites: torsion angle distributions from the CSD can be used to restrict ligand conformations. The program has been validated against a large test set of diverse protein-ligand complexes in the PDB, and is widely regarded as a market leader for accuracy and reliability. GOLD offers a choice of scoring functions: GoldScore, ChemScore or, alternatively, user defined scoring functions. GOLD can be operated in serial or parallel mode, using PVM (supplied for UNIX) or proprietary distributed computing platforms. GA parameters are optimised for virtual screening applications.

Predicting interaction hotspots using Superstar

Superstar uses information on intermolecular interactions contained in IsoStar in an empirically based approach to identifying regions within protein cavities or around molecules where selected functional groups are likely to interact favourably (areas of high binding propensity). Results are displayed as contoured maps, and positions of maximum propensity can be represented as pharmacophore points. Superstar has been extensively validated using protein-ligand complexes in the PDB.

Powerful searching and data analysis for protein-ligand complexes using Relibase+

Relibase+ features detailed analysis of superimposed ligand binding sites, ligand similarity and substructure searches, such as 3D searches for protein-ligand and protein-protein interaction patterns. It is based on structures in the PDB and the individual user’s proprietary collection. It enables investigation of crystallographic packing effects around ligand binding sites, which may be relevant to extension of these data to in vitro and in vivo behaviour. Also included is detailed information about bound water molecules in these regions. Relibase+ is supplied with a tool for detecting cavities in protein surfaces and generating descriptors of their shape and chemical characteristics, and Reliscript, a command line interface which allows access to enhanced search methods from within the Python scripting language environment.
## EVENTS

### June–July

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Description</th>
<th>Location</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>27 June – 3 July and 30 June – 3 July</td>
<td>6th Australia–Japan Joint Symposium on Drug Design and Discovery and 9th Molecular Modelling Workshop (MM2004)</td>
<td>Sydney, Australia</td>
<td>E-mail: <a href="mailto:Renate.Griffith@newcastle.edu.au">Renate.Griffith@newcastle.edu.au</a></td>
</tr>
</tbody>
</table>

### August

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Description</th>
<th>Location</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>22–26</td>
<td>228th ACS National Meeting</td>
<td>Philadelphia, PA, USA</td>
<td>E-mail: <a href="mailto:natlmtgs@acs.org">natlmtgs@acs.org</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="http://www.acs.org/meetings/">http://www.acs.org/meetings/</a></td>
</tr>
</tbody>
</table>

### September

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Description</th>
<th>Location</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>5–10</td>
<td>Euro QSAR 2004: 15th European Symposium on Quantitative Structure Activity Relationships</td>
<td>Istanbul, Turkey</td>
<td>E-mail: <a href="mailto:armoria@euro-qsar2004.org">armoria@euro-qsar2004.org</a></td>
</tr>
<tr>
<td>7–10</td>
<td>Towards Accurate Calculation of Biomolecular Recognition and Reactivity: A conference in honour of Professor Ian Hillier, MGMS 2004 International Meeting, University of Manchester, UK</td>
<td></td>
<td><a href="http://pharmacy.man.ac.uk/rab/mgms/">http://pharmacy.man.ac.uk/rab/mgms/</a></td>
</tr>
</tbody>
</table>

### October

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Description</th>
<th>Location</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Chemical Information for the Non-Chemist, joint CSA Trust/RSC-CIG one-day training course, Royal Society of Chemistry, Burlington House, Piccadilly, London</td>
<td></td>
<td>E-mail: <a href="mailto:doug@dovertonltd.freeserve.co.uk">doug@dovertonltd.freeserve.co.uk</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="http://cds.dl.ac.uk/cds/CIGcig.html#Future">http://cds.dl.ac.uk/cds/CIGcig.html#Future</a></td>
</tr>
<tr>
<td>17–20</td>
<td>The 2004 International Chemical Information Conference and Exhibition, Annecy, France</td>
<td></td>
<td>E-mail: <a href="mailto:contact@infonortics.com">contact@infonortics.com</a></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td><a href="http://www.infonortics.com/chemical/index.htm">http://www.infonortics.com/chemical/index.htm</a></td>
</tr>
</tbody>
</table>

### November

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Description</th>
<th>Location</th>
<th>Contact Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>29</td>
<td>CSA Trust AGM and Annual Dinner, SCI, Belgrave Square, London</td>
<td></td>
<td><a href="mailto:Drcliveweeks@aol.com">Drcliveweeks@aol.com</a></td>
</tr>
<tr>
<td>30 Nov–2 Dec</td>
<td>Online Information 2004, Olympia Grand Hall, London</td>
<td></td>
<td><a href="http://www.online-information.co.uk/">http://www.online-information.co.uk/</a></td>
</tr>
</tbody>
</table>

### 2005

### March

<table>
<thead>
<tr>
<th>Date</th>
<th>Event Description</th>
<th>Location</th>
<th>Contact Information</th>
</tr>
</thead>
</table>
The 24th National Meeting of the Chinese Chemical Society was held in Changsha, the capital city of Hunan Province, from April 24 to 27, 2004. The conference attracted over 1700 registrations, and for the first time over 60 commercial companies participated in the exhibition. Among the 21 sessions and 3 forums, the session on chemoinformatics was organised by the Committee of Computer Chemistry, Chinese Chemical Society. The session was chaired by Professor Kaixian Chen, of the Shanghai Institute of Materia Medica, Chinese Academy of Sciences and Professor Xiaojie Xu, from Peking University. The 20 invited speakers were representative of the active groups in China, including Hong Kong. Among the topics covered were QSAR, including docking and ADMET prediction, molecular modeling in multiple scale, new generation of molecular force field, chemometrics (especially in the analysis and finger prints of traditional Chinese medicine), chemical data mining using the Support Vector Machine method, and databases of reactions.

The session attracted an audience of about 80 during the two and half days of presentations. There was a demo unit during the session. The sponsors of the session, the vendors of Accelrys and Tripos in China, introduced their products. Some participants demonstrated their programs and databases in the unit. Particularly impressive were a program, MASTER, a database, TCMD and a website, ChIN. MASTER is a program of experimental design, optimisation and machine learning, developed by Professor Nanyi Chen’s group, at Shanghai University, just going commercial. TCMD (Traditional Chinese Medicines Database) is a database on PC with over 10,000 compounds and over 4600 TCM plant species for drug design, application of natural products, and study of traditional Chinese medicines. TCMD is now commercially distributed in China. TCMD comes from Professor Jiaju Zhou’s group, at the Institute of Process Engineering, Chinese Academy of Sciences. Professor Zhou is the major author of ‘Traditional Chinese Medicines–Molecular Structures, Natural Sources and Applications’, first published by Ashgate in 1999, second edition, 2003. The website ChIN (http://chin.csdl.ac.cn) has been a directory of Internet chemistry resources since 1997. ChIN is now a database of more than 10,000 links and descriptions of chemistry web pages and sites, and serves as the chemistry portal of CSDL (the Chinese National Science Digital Library). These demos were a good example of how Chinese research in computers and chemistry is being developed into practical tools.

Xiaoxia Li

New name for Pharmacopeia
Pharmacopeia Inc. have announced that on May 11, 2004, its stockholders approved changing the company’s name to Accelrys, Inc. The new name is part of a corporate restructuring that included the spin-off of the company’s drug discovery business, Pharmacopeia Drug Discovery, Inc. Accelrys is now a company solely focused on delivering scientific software solutions to the life sciences and materials research industries.

In connection with the spin-off, the company announced today that its board of directors has appointed Mark Emkjer, previously the Accelrys president, to serve as president and chief executive officer of the company. Mr. Emkjer has been president of Accelrys since December 2002, having previously served as president and chief operating officer of Sunquest Information Systems, Inc. and as president and chief executive officer of pace Health Management Systems, Inc.

New owner for ChemWeb
Elsevier Ltd have announced that ChemWeb and the Alchemist will now be operated by ChemIndustry.com, Inc., the leading comprehensive directory and search engine for chemical professionals. The transfer of ChemWeb.com to the new ownership is nearly complete. Users are encouraged to manage their user profile as before.

Any questions, comments or requests for further information may be sent to transition@chemwebmail.com.

BioMedNet is to close on June 30. Details of what is happening to the various sections of BioMednet are outlined at www.bmn.com/general/faq#one

SPRESIweb version 2.0
InfoChem’s SPRESIweb 2.0 provides powerful new capabilities. The synthesis strategy design tool Synthesis Tree Search (STS) is fully integrated, and the system offers internet access to 4.5 million compounds and 3.5 million reactions from 540,000 references including 151,000 patents. Over 20 million facts such as chemical, physical or biological property data and keywords abstracted from the primary literature are also searchable. Structure searching has been extended by ‘Isomer search’, ‘Parent search’ and ‘Flex match’. SPRESIweb also features structure substructure searching and search in the ‘basic index’ which covers principal data fields. More information can be found at www.spresi.com/.

ACD/Labs Version 8.0
ACD/Labs Version 8.0 software has been officially released. Highlights of the Version 8.0 software release are:

- augmented usability
- expanded vendor compatibility
- processing of hyphenated techniques
- enhanced prediction capabilities.

More information can be found at www.acdlabs.com/products/new

New Organic Reactions
Organic Reactions, from Wiley InterScience, is a comprehensive new database devoted exclusively to important synthetic reactions. The focus is on reaction limitations, interfering influences, the effects of chemical structure, and the selection of experimental conditions. Each reaction is presented with information on reaction conditions, products and yields where available, and is fully referenced to the primary literature. New data will be added on a continuous basis. More information can be found at www.interscience.wiley.com/db/or.

Pharmaceutical Substances
Thieme’s online database Pharmaceutical Substances is now offered by FIZ Karlsruhe via STN. Updated twice a year, the database aims to offer a complete reference guide to every pharmaceutical compound of significance. Over 2300 active pharmaceutical ingredients are searchable online. More information can be found at www.stn-international.de/archive/pressroom/pressreleases/ps_en.html
By Popular Request

A paper published in the *Journal of Organic Chemistry* in 2002 by four researchers from Yale was the most requested scientific article of 2003, according to the Science Spotlight free web service of Chemical Abstracts Service (CAS). The article discusses a highly active, air-stable palladium catalyst for cross-coupling reactions, useful in a wide variety of synthetic organic chemistry transformations, especially in ‘green’ (environmentally friendly) chemistry applications. Researchers requested the article’s full text through several CAS services, including SciFinder, SciFinder Scholar and STN information services.

Users of CAS electronic services submit a Real-Time Document Request (RDR) for the electronic full text of a desired article or patent. Since 2001, CAS has been tallying RDRs to determine the most sought-after journals and articles. The authors, journal editors and publishers with the highest 2003 RDR scores were recognised during a CAS Science Spotlight ceremony held during the ACS national meeting in Anaheim, California. Those honoured for the highest RDR ratings were:

- Dr John F. Hartwig, Yale University, and co-authors (with their current affiliations) Dr Quinetta D. Shelby, Chicago State University; Noriyasu Kataoka, Ajinomoto Co., Inc. and Dr James P. Stambuli, Stanford University; for ‘Air Stable, Sterically Hindered Ferrocenyl Dialkylphosphines for Palladium-Catalyzed C-C, C-N, and C-O Bond-Forming Cross-Couplings’, *Journal of Organic Chemistry* (2002), 67(16)
- Dr C. Dale Poulter, editor of *Journal of Organic Chemistry*
- Robert D. Bovenschulte, President, ACS Publications Division, representing the publisher of the journal receiving the most RDRs.