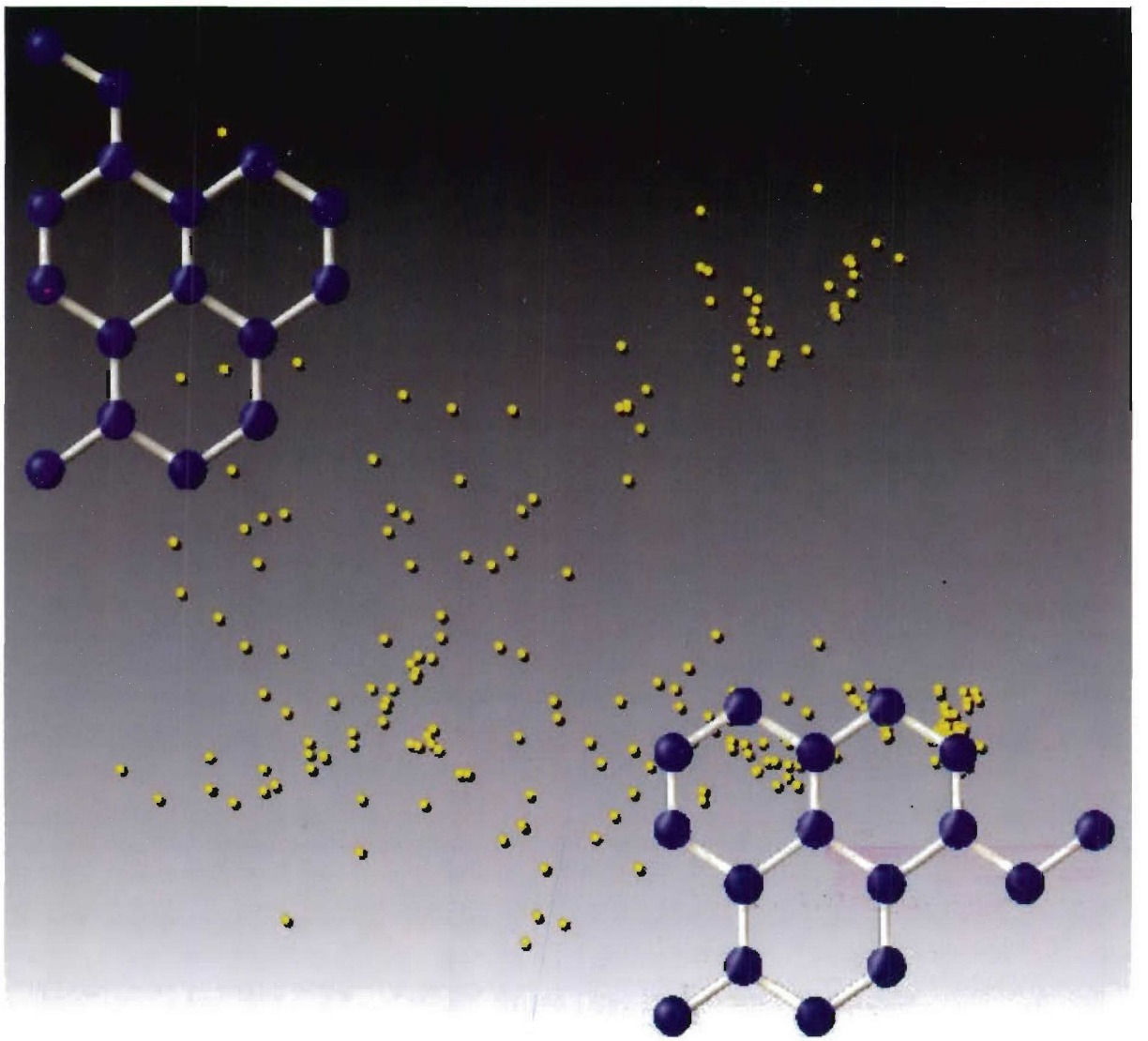
Fourth Indo-US Workshop on Mathematical Chemistry

With Applications to Drug Discovery, Environmental Toxicology, Cheminformatics and Bioinformatics

January 8–12, 2005
University of Pune
Pune, Maharastra, India



Workshop Series Co-Chairpersons:
Subhash C. Basak - University of Minnesota, USA
Dilip K. Sinha - University of Calcutta, India

Panels from Ellora & Ajanta, India

University of Pune, India

Subhash C. Basak
Workshop Series Co-Chairperson (USA)
Natural Resources Research Institute
University of Minnesota Duluth
5013 Miller Trunk Highway, Duluth, MN 55811, USA
Tel: (218) 720-4230
sbasak@nrri.umn.edu

Indira Ghosh
Convener, Bioinformatics Center
Institute of Bioinformatics & Biotechnology
University of Pune, Ganeshkhind
Pune 411 007, India
Tel: +91 20 2569 0195
indira@bioinfo.ernet.in

REPORT DOCUMENTATION PAGE

Form Approved

OMB No. 0704-0188

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
29-04-2005	Conference Proceedings	8 – 12 January 2005

4. TITLE AND SUBTITLE

Fourth Indo-US Workshop on Mathematical Chemistry

Book of Abstracts

5a. CONTRACT NUMBER
FA520905P0173

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Conference Committee

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Pune
Ganeshkhind
Pune, Maharashtra 411007
India

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AOARD
UNIT 45002
APO AP 96337-5002

10. SPONSOR/MONITOR'S ACRONYM(S)

AOARD

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

CSP-051028

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for Public Release. Copyrighted: Government Purpose Rights License.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The workshop covers topics, including but not limited to: Mathematical characterization of molecular structure and chemical similarity, Applications of Mathematical Chemistry in molecular design, drug discovery and approaches in natural product chemistry, Emerging applications of Mathematical Chemistry in Bioinformatics & Proteomics Proteomics and Genomics.

15. SUBJECT TERMS

Mathematical Chenistry

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Tae-Woo Park, Ph.D.
U	U	U	UU	62 and a CD	19b. TELEPHONE NUMBER *(Include area code)* +81-3-5410-4409

Fourth Indo-US Workshop on Mathematical Chemistry

With Applications to Drug Discovery, Environmental Toxicology, Chemoinformatics and Bioinformatics

January 8–12, 2005
Pune, Maharashtra, India

Organized by:
Natural Resources Research Institute,
University of Minnesota Duluth,
Duluth, Minnesota, USA

And

University of Pune, Pune,
Maharashtra, India

NRRI

Special thanks to the following for their support to the Fourth Indo-US Workshop:

ACCELRYS

DEPARTMENT OF SCIENCE AND TECHNOLOGY (DST), GOVERNMENT OF INDIA

ASTRAZENECA RESEARCH FOUNDATION, BANGALORE

UNIVERSITY OF PUNE, PUNE, MAHARASHTRA, INDIA

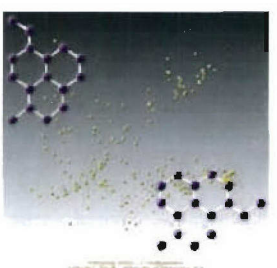
CHEMICAL ABSTRACTS SERVICE (CAS)

The University of Minnesota Duluth

COUNCIL OF SCIENTIFIC AND INDUSTRIAL RESEARCH (CSIR), GOVERNMENT OF INDIA

NATURAL RESOURCES RESEARCH INSTITUTE, University of Minnesota Duluth, MN, USA

DEPARTMENT OF BIOTECHNOLOGY



INDO-US WORKSHOP ON MATHEMATICAL CHEMISTRY SERIES

The Organizers of Fourth Indo-US Workshop on Mathematical Chemistry, which held on January 8–12, 2005 at University of Pune, Pune, Maharashtra, India,

wish to thank Air Force Office of Scientific Research, (AFOSR)*

and

Asian Office of Aerospace Research and Development (AOARD)

for their contribution to the success of this conference

Air Force Office of Scientific Research, (AFOSR)*

Asian Office of Aerospace Research and Development (AOARD)*

"AFOSR/AOARD support is not intended to express to imply endorsement by the U.S Federal Government"

THE GUEST HOUSE IN-CHARGE, INTER-UNIVERSITY CENTRE FOR ASTRONOMY AND ASTROPHYSICS (IUCAA)

THE GUEST HOUSE IN-CHARGE, NATIONAL CENTRE FOR CELL SCIENCE (NCCS)

THE GUEST HOUSE SUPERVISOR, NATIONAL CHEMICAL LABORATORY (NCL)

THE DIRECTOR, STATE ELIGIBILITY TEST (SET) OFFICE

THE PRINCIPAL, MAHARASHTRA STATE INSTITUTE OF HOTEL MANAGEMENT AND CATERING TECHNOLOGY

UNIVERSITY OF PUNE, PUNE, MAHARASHTRA, INDIA

THE VICE-CHANCELLOR, UNIVERSITY OF PUNE
THE REGISTRAR, UNIVERSITY OF PUNE
THE ESTATE OFFICER, UNIVERSITY OF PUNE
THE HOUSE MANAGER, UNIVERSITY OF PUNE
THE HEAD, LALIT KALA KENDRA, UNIVERSITY OF PUNE

Workshop Series Co-Chairpersons

Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, USA

Dilip K. Sinha, Calcutta University, India

Organizing Committee

Committee Chairman—*Ashok Kolaskar*, Vice Chancellor, Univeristy of Pune, Pune, India

Convenor—*Indira Ghosh*, Institute of Bioinformatics & Biotechnology, University of Pune, Pune, India

Bioinformatics Center, University of Pune, Pune, Maharashtra, India
Sunita Ahire, Anirban Banerji, Rajendra Bendugade, Jaydutt D. Bhalshankar, Shriram Bhosale, Vijaya Dahiphale, Rajan Doiphode, Sunita Jagtap, Mahesh Joshi, Rahul Khandare, Pramod Khandekar, Pranali Khoje, Urmila Kulkarni-Kale, Vishwanath Limbkar, Sanjay Londhe, Sunitha Manjari, Shubhada Nagarkar, Janaki Ojha, Om Prakash Pandey, G. Prabhakar, Vivek Kumar Singh, Anup Som and Raghavendra Wadke

Indian Institute of Science, Bangalore, Karnataka, India
N. Balakrishnan and Saraswati Vishweshwara

Inst. of Bioinformatics & Biotechnology, Univ. of Pune, Pune, Maharashtra, India
Subhash Chaudhary, Sangeeta V. Sawant and Vaishali Shah

National Chemical Laboratory, Pune, Maharashtra, India
S. Krishnan and Sourav Pal

Natural Resources Research Inst., Univ. of Minnesota Duluth, Duluth, MN, USA
Subhash Basak, Brian Gute and Michael J. Lalich (Institute Director)

Texas A&M University at Galveston, Galveston, Texas, USA
Alexandru Balaban and Douglas Klein

University of Pune, Pune, Maharashtra, India
Sridhar Gadhre and Pandit Vidyasagar

Manish Bagchi, Indian Institute of Chemical Biology, Kolkata, West Bengal, India
Kanika Basak, St. Xavier Computer Center, Kolkata, India
Alok Bhattacharya, Jawaharlal Nehru University, Delhi, Delhi, India
Vijay Chandru, Strand Genomics, Bangalore, Karnataka, India
S.P. Gupta, Birla Institute of Technology and Science, Pilani, Rajasthan, India
B. Jayaram, Indian Institute of Technology Delhi, Delhi, India
Eluvathingal D. Jemmis, Hyderabad Central University, Hyderabad, A.P., India
V.M.Kulkarni, Poona College of Pharmacy, Bharati Vidyapeeth Deemed University, Pune, India
V.S. Parmer, University of Delhi, Delhi, India
Milan Randic, Drake University, Des Moines, IA, USA
Chandan Raychaudury, Accelrys, Banagalore, Karnataka, India
Anil K. Saxena, Central Drug Research Institute, Lucknow, India
T.P. Singh, All India Institute of Medical Science, Delhi, Delhi, India
Dilip K. Sinha, University of Calcutta, Calcutta, India

SCIENTIFIC ADVISORY COMMITTEE

SCHEDULE OF EVENTS

TIME	ACTIVITY	LOCATION
SATURDAY, JANUARY 8		Namdeo Hall
8:00 am – 9:00 am	On-site registration	
9:00 am – 10:30 am	**WORKSHOP INAUGURATION** Prof. C.N.R.Rao (Chief Guest) Ashok Kolaskar Dilip Sinha Subhash Basak	
10:30 am – 10:50 am	High Tea	
10:50 am – 12:10 pm	**GENERAL SESSION 1** N. Balakrishnan and Haruo Hosoya Co-chairs	
10:50 – 11:10	[O17] Milan Randić, National Institute of Chemistry, P.O.B. 3430, 1001 Ljubljana, Slovenia. Four-color map representation of DNA sequences and their numerical characterization.	
11:10 – 11:30	[O63] Dr.B.Jayram, IIT, New Delhi, India Gene To Drug - A Pathway In Silico	
11:30 – 11:50	[O27] Subhash C. Basak, Natural Resources Research institute, University of Minnesota Duluth, Duluth, MN 55811, USA, Mathematical structural invariants: Developments and applications.	
11:50 – 12:10	[O10] Scott D. Kahn, Accelrys Inc, 9685 Scranton Road, San Diego, CA 92121. Computer-aided nanotechnology: Bridging lengthscales and scientific disciplines.	
12:10 pm – 2:00 pm	Lunch	
2:00 pm – 3:20 pm	**GENERAL SESSION 2** Gerald Niemi Co-chairs	
2:00 – 2:20	[O62] Indira Ghosh, Bioinformatics Center, University of Pune, India In search of new scaffold from data mining of Antituberculosis compounds.	
2:20 – 2:50	[O15] Kunal Roy, Drug Theoretics and Cheminformatics Lab, Jadavpur University, Kolkata 700 032, India Exploring 3D-QSAR of thiazole and thiadiazole derivatives as potent and selective human adenosine A_3 receptor antagonists.	
2:50 – 3:20	[O55] Guillermo Restrepo, Laboratorio de Química Teórica, Universidad de Pamplona, Ciudad Universitaria, Colombia; Carrera Bogotá, Colombia. Trees (dendrograms and consensus trees) and their topological information.	

TIME	ACTIVITY	LOCATION
3:20 pm – 3:40 pm	Break	
3:40 pm – 5:00 pm	**GENERAL SESSION 3** Saraswathi Vishveshwara, & Co-chairs	Namdeo Hall
3:40 – 4:00	[O64] Sathyapriya.R and Saraswathi Vishveshwara. Molecular Biophysics Unit, Indian Institute of Science, Bangalore. Karnataka. spriya@mbu.iisc.ernet.in Interaction of DNA with clusters of amino acids in proteins.	
4:00 – 4:20	[O19] Haruo Hosoya, Ochanomizu University (Emeritus), Bunkyo-ku, Tokyo 112-8610, Japan. Graph-theoretically derived extended Hückel rule can explain the aromaticity and unusual properties of condensed polycyclic hydrocarbon networks.	
4:20 – 4:40	[O01] Apurba K. Bhattacharjee, Department of Medicinal Chemistry, Division of Experimental Therapeutics, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA. 3D pharmacophore model for insect repellent activity and discovery of new repellent candidates.	
4:40 – 5:00	[O51] Sudhir A. Kulkarni, VLife Sciences Technologies Pvt. Ltd. 1 Akshay, 50 Anand Park, Aundh, Pune 411 007, India. Three dimensional QSAR based on k-nearest neighbor method.	

SUNDAY, JANUARY 9

Namdeo Hall

9:00 am – 10:20 am	**GENERAL SESSION 4** Krishnan Balasubramanian & Dr. Sourav Pal, Co-chairs	
9:00 – 9:20	[O20] K. Balasubramanian, Institute for Data Analysis & Visualization, University of California Davis, Livermore, CA 94550; University of California, Chemistry and Material Science Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550; and Glenn T Seaborg Center, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720. Graph theory and high-level quantum models for predictive toxicology and graph algorithms for proteomics.	
9:20 – 9:40	[O58] Ashok Deshpande, Berkeley Initiative in Soft Computing (BISC)-Special Interest Group (SIG)-Environment Management Systems (EMS), Former Deputy Director NERRI, India. River water quality fuzzily described: A case study.	
9:40 – 10:00	[O52] R. Bruce King, Department of Chemistry, University of Georgia, Athens, Georgia 30602, USA. Topological aspects of aromaticity in inorganic compounds.	

TIME	ACTIVITY	LOCATION
10:00 – 10:20	[O33] Chizuru Muguruma, Faculty of Liberal Arts, Chukyo University, 101 Tokodachi, Kaizu-cho, Toyota 471-0393, Japan. New approaches to the liquid-solid phase transition of Lennard-Jones fluids.	Namdeo Hall
10:20 am – 10:40 am	Break	
10:40 am – 12:00 pm	**GENERAL SESSION 5** Kevin Geiss &. K.N.Ganesh Co-chairs	
10:40 – 11:00	[O23] Yenamandra S. Prabhakar, Medicinal and Process Chemistry Division, CDRI, Lucknow-226 001, India A high dimensional QSAR study on the Aldose Reductase inhibitory activity of some flavonoids: Topological descriptors in modeling the activity.	
11:00 – 11:20	[O09] Vladimir A. Palyulin, Department of Chemistry, Moscow State University, Moscow, 119992, Russia. Computer-aided design of topological indices	
11:20 – 11:40	[O05] Marjana Novič, National institute of chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia. Chemometrics approach to estrogen receptor binding	
11:40 – 12:00	[O65] Dr.Uddhavesh Sonawane , CDAC, Pune University Campus Significance of High Performance Computing in Bioinformatics	
12:00 pm – 1:30 pm	Lunch	
1:30 pm – 3:10 pm	**GENERAL SESSION 6** S.P Gupta Co-chairs	
1:30 – 1:50	[O49] Marjan G. Vracko, National institute of chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia. Self organizing map and counterpropagation neural networks in structure-property modelling: Examples from environmental science and drug design.	
1:50 – 2:10	[O22] Alain J.M. Carpy, LPTC, UMR 5472 CNRS, University of Bordeaux I, 351 Cours de la Libération, Talence , France. Handling tautomerism and geometrical isomerism in drug design and synthesis: The case of phenylpyruvic acid derivatives.	
2:10 – 2:30	[O30] Jacques R. Chretien, BioChemics Consulting, Innovation Center, 45074 Orleans Cedex 2, France. The Molecular Experimental Design (MED) concept: application to the selection and use of molecular descriptors in early ADME prediction.	
2:30 – 2:50	[O37] Rainer Brüggemann, Ute Simon, Leibniz - Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany. Partial order: A general tool in ecotoxicological and ecosystems hazard assessment.	

TIME	ACTIVITY	LOCATION
2:50 – 3:10	[O40] Ray Hefferlin, Southern Adventist University, Collegedale, Tennessee 37315, USA, hefferln@southern.edu. Predictive transformations between the graphical "databases" of S. A. Shchukarev.	Namdeo Hall
3:10 pm	Meeting adjourned until Monday morning	

MONDAY, JANUARY 10

TIME	ACTIVITY	LOCATION
9:00 am – 10:20 am	**GENERAL SESSION 7** Akbar Khan Co-chairs	Namdeo Hall
9:00 – 9:20	[O24] A. Nandy, Environmental Science Programme, Jadavpur Univeristy, Jadavpur, Kolkata 700068, Index of variance of dinucleotide frequencies and evolutionary relationships in gene families.	
9:20 – 9:40	[O42] Amandeep S. Sidhu, Faculty of Information Technology, University of Technology Sydney, Australia. Protein ontology project.	
9:40 – 10:00	[O41] Akbar Khan, Defense Threat Reduction Agency-CBM, 8725 John J. Kingman Road,Fort Belvoir, USA. Proteomics and bioinformatics to design countermeasures against infectious threat agents.	
10:00 – 10:20	[O08] Tarun K. Bansal, Department of Biological Sciences, BITS-Pilani, 333031, India. Computational genomics 1: Identification and characterization of point mutation in coding region of a gene translating Arginine Vasopressin hormone.	
10:20 am – 10:40 am	Break	
10:40 am – 12:00 pm	**GENERAL SESSION 8** Dr. Ashok Deshpande, Co-chairs	
10:40 – 11:00	[O14] Satyam Priyadarshy, RKR Group, 13507 Apple Barrel Ct., Herndon, VA 20171, USA. Effect of geometrical fluctuations on DNA electron transfer rates: A computational study.	
11:00 – 11:20	[O18] Lech P. Schulz, Adam Mickiewicz University, Department of Chemistry, Grunwaldzka 6, Poznań, Poland. Fuzzy set based signal analysis of structured experimental and computational data in chemistry.	
11:20 – 11:40	[O47] Jessica J. Kraker, University of Minnesota, School of Statistics, Minneapolis, Minnesota, 55455, USA. Quantitative structure activity relationships: Model fitting and leave-out-clusters cross-validation.	

TIME	ACTIVITY	LOCATION
11:40 – 12:00	[O16] Vinay A. Wadekar, Chemical Engineering & Process development Division, National Chemical Laboratory, Pune-411008, India. Feature extraction of gene expression data using curvilinear component analysis (CCA) and self-organizing map (SOM).	Namdeo Hall
12:00 pm – 2:00 pm	Poster Lunch	
	POSTER PRESENTERS [P01] Mohammad Shahidul Islam [P02] Pradipta Bandyopadhyay [P03] Matevz Pompe [P07] Brian D. Gute [P08] Ambrish Roy [P09] Philip Prathipati [P10] Brian D. Gute [P13] Ramanathan Natarajan [P17] Christian Matson [P18] Ramanathan Natarajan [P20] Ramanathan Natarajan [P21] Ramanathan Natarajan [P22] Milan Randic [P23] M. Mohamed Naseer Ali [P28] Javendra B. Bhonsle [P30] Anand G. Fadnis [P31] Jessica J. Kraker [P34] Brian D. Gute [P35] Yenamandra S. Prabhakar [P38] Sharan Shetty	SET Bhavan
2:00 pm – 3:20 pm	**GENERAL SESSION 9** Dr. Subhash Padhye, Co-chairs	Namdeo Hall
2:00 – 2:20	[O45] Ramanathan Natarajan, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811, USA. Quantitative structure-activity relationship (QSAR) modeling of juvenile hormone activity of alkyl (2E,4E),3,7,11-trimethyl-2,4-dodecadienoates.	
2:20 – 2:50	[O02] Amit Kulkarni, Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121, USA, amit@accelrys.com. ADME/Tox: Address critical industry needs.	
2:50 – 3:20	[O57] Vellarkad N. Viswanadhan, Molecular Structure and Design, I, Amgen Center Drive, MS 29-M-B, Thousand Oaks, CA 91320, USA. Virtual high throughput screening in drug design and development.	
3:20 pm – 3:40 pm	Break	

TIME	ACTIVITY	LOCATION
3:40 pm – 5:00 pm	**GENERAL SESSION 10** Dr. Nimse, Co-chairs	Namdeo Hall
3:40 – 4:00	[O29] Chandan Raychaudhury, Accelrys K.K., 12th Floor, Discoverer, International Tech Park, Whitefield Road, Bangalore 560066, India. Molecular docking using LigandFit and LibDock.	
4:00 – 4:20	[O43] Abhilash Thakur, Softvision College of Biotechnology and Science, Indore India 452001. QSAR study on 1,4–dihydropyridine-based calcium channel antagonist: Dominating role of molar refractivity.	
4:20 – 4:40	[O12] Rini Roy, Department of Microbiology, Lady Brabourne College, P-1/2, Suhrawardy Avenue, Kolkata – 700017, West Bengal, India. Designing of less toxic to non target organophosphorus pesticide through multiple correlation analysis (MRA), a part of QSAR.	
4:40 – 5:00	[O04] Daniel S. Sem, Chemical Proteomics Facility at Marquette, Marquette University, PO Box 1881, Milwaukee, WI, 53201, USA. Chemical proteomic anti-target studies: from drug design to pollutant screening.	
7:30 pm – 10:00 pm	Banquet hosted by Indo_US workshop Series	

TUESDAY, JANUARY 11

TIME	ACTIVITY	LOCATION
9:00 am – 10:20 am	**GENERAL SESSION 11** Moiz Mumtaz Co-chairs	Namdeo Hall
9:00 – 9:20	[O11] Anil K. Saxena, Medicinal and Process Chemistry Division, Central Drug Research Institute, Lucknow-226001, The current scenario of in silico drug discovery research.	
9:20 – 9:40	[O60] A. Kalaiselvan, P. Venuvanalingam, Department of Chemistry, Bharathidasan University, Tiruchirappalli-620024 Jahn-Teller distortion in the transition state observed in the deamination of aziridines by difluoroamine: More insights from theory	
9:40 – 10:00	[O54] Steven Trohalaki, The Anteon Corporation, 5100 Springfield Pike, Dayton, OH 45431-1231, USA. Toxicity QSARs for halogenated aliphatics and melting-point QSPRs for ionic liquids.	
10:00 – 10:20	[O38 Melek Türker Saçan, Boğaziçi University, Institute of Environmental Sciences, 34342, Bebek, İstanbul, Turkey. Physical-chemical properties of PCDD/PCDFs and phthalate esters: A QSPR study.	

TIME	ACTIVITY	LOCATION
10:20 am – 10:40 am	Break	
10:40 am – 12:00 pm	**GENERAL SESSION 12** Dr. S. Krishnan Co-chairs	Namdeo Hall
10:40 – 11:00	[O07] S. Natarajan Balasubrahmanyam, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore, India. A computational evaluation of bond order and charge distributions in isomeric aminotroponimium ions and their benzo analogues.	
11:00 – 11:20	[O06] B. A. Bhongade, Department of Medicinal Chemistry, College of Pharmacy, J. N. Medical College, Belgaum-590 010, Karnataka, India. A comparative selectivity analysis on trypsin-like serine protease inhibitors employing CoMFA methodology.	
11:20 – 11:40	[O53] D.K. Sinha, Department of Applied Mathematics, 92 A.P.C Road, Kolkata-700009, India. On topological bond orders of non alternant molecular graphs: A computational approach.	
11:40 – 12:00	[O39] Moiz M. Mumtaz, Computational Toxicology Laboratory, Division of Toxicology, Agency for Toxic Substances and Disease Registry, Atlanta, GA 30333, USA. Prediction of biologic partition coefficients and binding affinities using SAR models.	
12:00 pm – 2:00 pm	Poster Lunch	SET Bhavan

POSTER PRESENTERS

[P04] Milan Randic
[P05] Milan Randic
[P06] Dhanshri C. Juvale
[P11] Milan Randic
[P12] Melek Turker Sacan
[P14] Ragothaman Yennamalli
[P15] Milan Randic
[P16] Varsha Kodali
[P19] Milan Randic
[P24] Marjana Novic
[P25] Christian T. Matson
[P27] Anshuman Dixit
[P29] Manish C. Bagchi
[P32] Tarun Jha
[P33] Marjan G. Vracko
[P36] Rajeev Kalia
[P37] Ramanathan Natarajan
[P39] Apurba K. Bhattacharjee
[P40] Manish K. Gupta
[P41] Vijay K. Agrawal
[P 42] V.K. Katiyar

TIME	ACTIVITY	LOCATION
2:00 pm – 3:20 pm	**GENERAL SESSION 13** Dr. P. Vidyasagar, Co-chairs	Namdeo Hall
2:00 – 2:20	[O34] Matevž Pompe, University of Ljubljana, Faculty of Chemistry and Chemical Technology, Aškerčeva 5, 1000 Ljubljana, Slovenia. New developments in variable connectivity indices.	
2:20 – 2:40	[O26] C. Suguna, CCMB, Hyderabad 500 007, India. Mathematical modelling of simple biochemical pathways.	
2:40 – 3:00	[O28] Reena Gollapudy, VLife Sciences Technologies Pvt. Ltd. 1 Akshay, 50 Anand Park, Aundh, Pune 411 007, India. Modeling and interactions of Aspergillus fumigatus lanosterol 14-α demethylase enzymes 'A' and 'B' with azole antifungals.	
3:00 – 3:20	[O32] Prabhat K. Sahu, Dept of Chemistry & Biochemistry, National Chung Cheng University, Chia-Yi, Taiwan. Net-sign identity information index: A novel approach towards numerical characterization of chemical signed graph theory.	
3:20 pm – 3:40 pm	**MEETING OF THE INTERNATIONAL SOCIETY OF MATHEMATICAL CHEMISTRY**	
7:30 pm –	Dinner hosted by Accelrys	

WEDNESDAY, JANUARY 12

TIME	ACTIVITY	LOCATION
9:00 am – 10:20 am	**GENERAL SESSION 14** Dr. Deelep Deobagkar, Co-chairs	Namdeo Hall
9:00 – 9:20	[O35] Dheeraj Mandloi, Institute of Engg. & Technology, Devi Ahilya University, Indore 452017, India. Novel development of quantitative structure-activity relationship for anti-bacterial activity of newly synthesized Mannich bases using NMR chemical shift as a molecular descriptor.	
9:20 – 9:40	[O36] Ramamurthy Jayalakshmi, Department of Bio-Informatics, Thanthai Hans Roever College, Perambalur, Tamil Nadu, India 621 212. A novel method of blocking malarial transmission through insect vectors.	
9:40 – 10:00	[O50] Pranav Sharad, IIIrd Year, B.Tech. Bioinformatics, Dept. of Biotechnology, Vellore Institute of Technology, Vellore. Software on pair wise sequence alignment: Drug design approach towards SARS.	
10:00 – 10:20	TBD	

TIME	ACTIVITY	LOCATION
10:20 am – 10:40 am	Break	
10:40 am – 12:00 pm	**GENERAL SESSION 15** V.M Kulkarni Co-chairs	Namdeo Hall
10:40 – 11:00	[O46] S. Kumaran, Department of Pharmacy, Birla Institute of Technology and Science, Pilani-333031, India. Quantitative structure-activity relationship studies on benzodiazepine hydroxamic acid-based inhibitors of matrix metalloproteinases and tumor necrosis factor-α converting enzyme.	
11:00 – 11:20	[O48] S.P. Gupta, Department of Chemistry, Birla Institute of Technology and Science, Pilani-333031, India. Role of hetroatoms in hetrocyclic drug actions: A quantitative aspect.	
11:20 – 11:40	[O25] Shailza Singh, Indian Institute of Technology, Hauz Khas, Delhi-110016, India. Lactone analogues as protein kinase C (PKC) ligands in prostate cancer: An in silico approach to drug-design.	
11:40 – 12:00	[O44] Mamta Thakur, Department of Chemistry, Softvision College of Biotechnology and Science, Indore India 452001, India. QSAR study on inhibition of E. coli by sulfonamides: A case study with no role of positive hydrophobic term.	
12:00 pm – 1:30 pm	Lunch	
1:30 pm –	CLOSING REMARKS & SITE SELECTION FOR THE 5TH INDO-US WORKSHOP Subhash Basak	Namdeo Hall

ORAL PRESENTATIONS

[O01] A three-dimensional in silico pharmacophore model for inhibition of *Plasmodium falciparum* cyclin dependent kinases and discovery of different classes of novel Pfmrk specific inhibitors

Apurba K. Bhattacharjee, Jeanne A. Geyer, Cassandra Woodard, April K. Kathcart, Daniel A. Nichols, Department of Medicinal Chemistry, Division of Experimental Therapeutics, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA; Sean T. Prigge, Department of Molecular Microbiology and Immunology, Malaria Research Institute, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA; Zhiyu Li, Bryan T. Mott and Norman C. Waters, Department of Medicinal Chemistry, Division of Experimental Therapeutics, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA.

The cell division cycle is regulated by a family of cyclin-dependent protein kinases (CDKs) that are functionally conserved among many eukaryotic species. The characterization of plasmodial CDKs has identified them as a leading antimalarial drug target in our laboratory. We have developed a three-dimensional QSAR pharmacophore model for inhibition of a *Plasmodium falciparum* CDK, known as Pfmrk, from a set of fifteen structurally diverse kinase inhibitors with a wide range of activity. The model was found to contain two hydrogen-bond acceptor functions and two hydrophobic sites including one aromatic-ring hydrophobic site. Although the model was not developed from X-ray structural analysis of the known CDK2 structure, it is consistent with the structure-functional requirements for binding of the CDK inhibitors in the ATP binding pocket. Using the model as a template, a search of the in-house three-dimensional multiconformer database resulted in the discovery of sixteen potent Pfmrk inhibitors. The predicted inhibitory activity of some of these Pfmrk inhibitors from the molecular model agree exceptionally well with the experimental inhibitory values from the *in vitro* CDK assay.

[O02] ADME/Tox: Address critical industry needs

Amit Kulkarni, Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121, USA, amit@accelrys.com.

Increasingly important in drug discovery is the need for assaying a potential drug compound for its pharmacological properties. Hence to be an effective drug, a compound must possess appropriate ADME/Tox properties necessary for making it suitable for use as a therapeutic agent. However despite the vast amount of in vitro and in vivo measurements obtained during the pre-clinical phase of drug development, only 10% of new chemical entities progress to commercial launch. Given the high cost and low throughput nature of ADME/Tox testing, the use of computational ADME/Tox has emerged as a viable alternative. Computational algorithms have been deployed to address therapeutic drug requirements in the areas of human intestinal absorption, metabolism, drug-drug interaction, serum transport of drugs, blood-brain barrier penetrants, aqueous solubility etc. The application of computational algorithms early in the drug discovery process shows promise of improving the quality of pre-clinical drug candidates, there by lowering the high cost typically associated with the follow-up lead optimization protocols. This talk will address research in the computational design of models for predictive ADME.

[O03] Calculations on solvents and co-solvents of fullerenes and single-wall carbon nanotubes

Francisco Torrens, Institut Universitari de Ciència Molecular, Universitat de València, Dr. Moliner 50, E-46100 Burjassot (València), Spain.

Solubility of single-wall carbon nanotubes (SWNT) is investigated in a variety of solvents, finding a class of non-hydrogen-bonding Lewis bases that provide good solubility. Investigated solvents are grouped in three classes. A molecular modelling study of linear/cyclic (1→4)-linked oligosaccharides containing 1–6 D-glucopyranose (Glc*p*) units provides a clear conception of their overall conformations, contact surfaces and cavity proportions. This gives a first estimation of their inclusion properties. Electron affinity (EA) of Glc*p* suggests the formation of

colloids of negatively charged SWNTs in water. Dipole moment μ for linear Glcp_n increases smoothly until $n = 4$, in agreement with a 18-fold helix. Provisional conclusions follow. (1) Solubility of SWNTs has been investigated in a variety of solvents, finding a class of non-hydrogen-bonding Lewis bases that provides good solubility. Solvents have been grouped in three classes. SWNTs in some organic solvents are positively charged, while in water/Triton X are negative, which is explained on the basis of dielectric permittivity/EA. (2) Lipophilicity characteristics of *cyclo*Glcp_n point towards hydrophobic effect as important for Glcp_n-inclusion complexes with hydrophobic molecules by incorporation into the hydrophobic channel. (3) EA of Glc*p* and *cyclo*Glcp_{3-4} suggest for all Glcp_n formation of colloids of negatively charged SWNTs in water. μ for linear Glcp_n increases smoothly with n until $n = 4$, in agreement with a 18-fold helix. (4) Variation of $\Delta H_f°/n$ for Glcp_n-linear, I_n^{z-} and 3n-crown-n ethers shows that shorter chains are more stable than larger ones are. *Cyclo*Glcp_n are less stable than linear Glcp_n are. Glc$p$·I_n^{z-} are competitive with amylose (Amy)·I_n^{z-}. (5) I_n^{z-}/SWNT$^-$ are suggested to form inclusion complexes with cyclodextrin/Amy. Guests/hosts are expected unperturbed in complexes. An expansion of the central channel is proposed for helical Glcp_n·SWNT.

[O04] Chemical proteomic anti-target studies: from drug design to pollutant screening

Aurora Costache, Phani Pullela, Purna Kasha and Daniel S. Sem, Chemical Proteomics Facility at Marquette, Marquette University, PO Box 1881, Milwaukee, WI, 53201, USA.

Chemical proteomics is devoted to the proteome-wide characterization of protein-ligand interactions, often for the purpose of predicting the biological effects of drugs or pollutants, based on interactions with protein targets and antitargets. One of the primary anti-target families to avoid in drug design efforts is the cytochromes P450 (CYPs). With regard to amine-containing compounds, such as the psychotherapeutic drugs, CYP2D6 is the major anti-target of concern. To aid researchers involved in drug design efforts, we have analyzed 4,000 commercially available amines in terms of their likelihood of binding to CYP2D6, based on a docking study using Autodock. This structural information is combined with a chemical property database, and disseminated freely by the Chemical Proteomics Facility at Marquette (CPFM). Issues of ionization state, energy scoring, conformational flexibility, inhibition versus metabolism, and the need for experimental validation will be presented.

Anti-targets are a concern not only in drug design, but also in the process of identifying dangerous pollutants, whose release into the environment may need to be regulated. Specifically, endocrine disruptors cause developmental defects in aquatic wildlife, with a primary route of action through binding to the estrogen receptor (ER). Environmental release of these compounds is being increasingly regulated in the United States. To allow better prediction of which pollutants are likely to bind to ER and behave as endocrine disruptors, we have built homology models for the major ER isoforms in zebrafish, a previously reported "aquatic sentinel" for endocrine disruptors. Docking studies have been performed, and tools have been developed to allow experimental validation of ER binding events in the developing zebrafish embryo. It appears that estradiol is binding to primordial germ cells, as they migrate to their final destination during sexual development.

[O05] Chemometrics approach to estrogen receptor binding

Marjana Novič, National institute of chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia.

In silico tools for prediction of receptor binding affinity of compounds and their potential biological activity diverse into two directions: one is based on the investigation of ligand – receptor complex formation with all accompanying quantum mechanical and semi empirical calculation methods, while the other is actually discovering information from a huge amount of data on chemical structure (descriptors) on one side and properties (binding affinity and/or biological activity) on the other. Two main different directions are often yielding complementary information, one being on micro scale (i.e. docking procedure calculates the difference in calculated binding affinities for different compounds, preferably congeneric for a reliable comparison), the other considered on macro scale, obtained from bulky information on structural and property parameters. The latter one is addressed here as chemometrics approach. Various modelling techniques can be employed, from simple linear to multilinear and non-linear computational neural networks [1,2]. A detailed study of alpha and beta estrogen receptors (ERα, ERβ) will be presented for 60

different fito estrogenes [3,4]. Models based on artificial neural networks and validated with one of the cross-validation methods will be shown and discussed. The variable selection methods included into the modelling procedure to reduce the number of molecular structure descriptors will be demonstrated and discussed with respect to the impact on the quality of obtained models, prediction ability and robustness.

References:

[1] J. Zupan, J. Gasteiger, *Neural Networks in Chemistry and Drug Design*, Wiley-VCH, Weinheim 1999.
[2] M. Novič, M. Vračko. Artificial neural networks in molecular structures - property studies. In: Leardi, R. (Ed.). *Nature-inspired methods in chemometrics : genetic algorithms and artificial neural networks*, (Data handling in science and technology, 23). 1st ed. Amsterdam [etc.]: Elsevier, 2003, pp 231-256 (Chapter 8).
[3] G.G.J.M. Kuiper, B. Carlsson, K. Grandien, E. Enmark, J. Haggblad, S. Nilsson, J.A. Gustafsson, Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta, *Endocrinology*, 1997, 138, 863-870.
[4] G.G.J.M. Kuiper, J.G. Lemmen, B. Carlsson, J. C. Corton, S.H. Safe, P.T. Van Der Saag, B. Van Der Burg, J. Gustafsson, Interaction of Estrogenic Chemicals and Phytoestrogens with Estrogen receptor beta. *Endocrinology*, 1998, 139, 4252-4263.

[O06] A comparative selectivity analysis on trypsin-like serine protease inhibitors employing CoMFA methodology

B. A. Bhongade, Mali D. S. and A. K. Gadad, Department of Medicinal Chemistry, College of Pharmacy, J. N. Medical College, Belgaum-590 010, Karnataka, India.

A series of indole/benzoimidazole-5-carboxamidines [1,2] are reported to inhibit various trypsin-like serine proteases viz. uPA, tPA, factor Xa, thrombin, plasmin and trypsin. Selective inhibition of these proteases can serve as therapeutic agents in various types of cancer as well as anticoagulant or antithrombosis, thus dual inhibition may results in poor clinical candidate leading to unwanted side effects. 3D-QSAR models were generated for a series of indole/benzoimidazole-5-carboxamidines employing CoMFA [3] technique for the inhibition of uPA [4], tPA, factor Xa, thrombin, plasmin and trypsin in order to understand the selectivity trends of these molecules towards various trypsin-like serine proteases. The molecular superimposition was carried out on template structure by atom-based RMS fit method. The CoMFA models were established from 25-29 molecules, which were validated by predicting the activities of test set of 7-8 compounds. The statistically significant CoMFA models generated using steric and electrostatic fields were used to generate 3D contour maps, which provide the regions in space where appropriate substitutions may alter the selectivity/activity. The present 3D-QSAR studies emphasize the selectivity trends of indole/benzoimidazole-5-carboxamidines, which may be obliging in designing of novel selective serine proteases inhibitors of the interest.

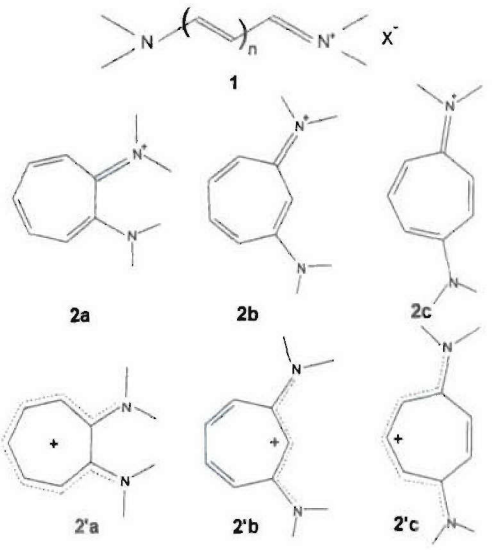
References:

[1] Verner, E. et al., *J. Med. Chem.* **2001**, 44, 2753.
[2] Mackman, R. L. et al., *J. Med. Chem.* **2001**, 44, 3856.
[3] Cramer, R. D. -III.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* 1988, 110, 5959.
[4] Bhongade, B. A.; Gadad, A. K. *Bioorg. Med. Chem.* **2004**, 12, 2797.

[O07] A computational evaluation of bond order and charge distributions in isomeric aminotroponimium ions and their benzo analogues

S. Natarajan Balasubrahmanyam, Jawaharlal Nehru Centre for Advanced Scientific Research, Jakkur, Bangalore; I. N. Narayanan Namboothiri and K. Pius, Department of Organic Chemistry, Indian Institute of Science, Bangalore 560 012.

An evaluation of bond localization and charge redistribution by *ab initio* MO procedures was of interest when the amino(poly)vinyliminium (vinamidinium) structural element 1 is made part of a cyclic system (e.g. aminotropon-2, -3- & -4-iminiums 2a-c; n = 3, 1 & 2 in 1). Differences in energy content among these possibilities can arise because:



- while a "diaminotropylium"-type charge distribution is possible for all three cases (e.g. **2'a** for **2a**), only **2a** can participate as a whole in a vinamidinium-type delocalization. In **2b** and **2c** passage of π-information between the nitrogens for delocalization would be confined to 'odd' sides (**2'b** and **2'c**);
- differences in the degree of destabilization may arise because *syn*-periplanar conformations are forced on **2a** and the vinamidinium part of **2c** (**2'c**) while **1** generally prefer *anti*-periplanar conformations (such a conformation can be maintained in the vinamidinium part of **2b**, however);
- relief of angle strain introduced when the internal angles open out beyond 120° to accommodate a planar 7-membered cyclic structure can be expected to depend on the specific partial bond fixations that come about in **2a-c**. Bond localization could be such as may permit the 7-membered cycles to assume specific puckered conformations, minimizing angle strain;
- The N-C-C-N array may not be uniplanar in the **2a** in order that *syn*-periplanar interaction be relieved, causing it to be different from the 1,3- and 1.4-systems **2b** and **2c** both of which may prefer to be entirely planar.

Modes of electron redistribution in the six isomeric benzo-analogues possible for systems **2** are also examined.

[O08] Computational genomics 1: Identification and characterization of point mutation in coding region of a gene translating Arginine Vasopressin hormone.

Tarun K. Bansal, N. Jaya Prakash, Gayatri Ankem, Ranjith R, Sunil Sharma, Peeyush, Pooja Choudhary, Arun A.K., Akash Jain, Amit K. Goyal, Department of Biological Sciences, BITS-Pilani, 333031, India; Ram Kinkar Roy and S.K. Ray, Department of Chemistry, BITS-Pilani, 333031, India.

Discrimination of a point mutation in the coding region of the gene AVP (NCBI Acc. No. NM_000490) which encodes Arginine Vasopressin hormone, is carried by newly developed parameters namely, Total Modular Information Content (TMIC), Integrated per Class Information Content (IPCIC), Binding Affinity Index (BAI). Parameters originated from application of suitably modified Shannon's Information theory and partial charges of elements of nucleotides. Chem 3D approach is followed to generate the partial charges. Computable discrimination is achieved by incorporating higher order sensitivity operations. The n^{th} order sensitivity value is decoded by freezing first n digits of the characteristic parameter of the gene. Functional weights of codons at different orders of sensitivity vary and it depends on the weight value of the 1^{st} order neighbour codons. It highlights the complexity of the functional characteristics of the genetic elements. Results show that 8^{th} order sensitivity clearly discriminates the point mutation of the studied gene. The computational value of the normal gene is 3.631 and that of mutated gene is

99.580. The methodology has potential for similarity studies of genes, discrimination of the coding sequence, identification of a specific gene and multisite mutational studies. Further Studies are in progress.

[O09] Computer-aided design of topological indices

Vladimir A. Palyulin, Mariya I. Skvortsova, Aleksander Yu.Zotov, Nikolai S. Zefirov, Department of Chemistry, Moscow State University, Moscow, 119992, Russia.

A problem of establishing the quantitative structure-property relationships for chemical compounds is of great importance in modern theoretical chemistry. The methods of molecular structure description play an essential role in these studies. One of the most widely used approaches to this problem is based on the representation of a molecule by weighted graph, usually called "molecular graph". For the quantitative description of molecular structures the invariants of respective molecular graphs ("topological indices") are often used. The topological indices (TIs) are obtained using formal mathematical operations with molecular graphs.

In the present paper the approach for the generation of the sets of topological indices for the modeling structure-property relationships is suggested and the corresponding algorithm is considered. Every topological index is obtained as a result of the sequential fulfillment of a number of basic procedures used in different combinations with various sets of parameters. The basic procedures have been evaluated as a result of analysis of more than a hundred known topological indices of different types.

It has been shown that the suggested scheme allows one to generate a great number of new topological indices as well as basic ones known from the literature. In this process the topological indices could be obtained which unlikely would be constructed manually due to their complexity.

The examples of the application of generated topological indices for building QSPR models are given.

[O10] Computer-aided nanotechnology: Bridging lengthscales and scientific disciplines

Scott D. Kahn, Accelrys Inc, 9685 Scranton Road, San Diego, CA 92121.

Nanotechnology holds tremendous economic and scientific potential, yet it will cost industry a considerable amount of time, money, and resources to research and develop new processes, devices, and synthesis techniques. The use of rational materials discovery software tools in conjunction with experimentation can lower this barrier significantly, and lead to new insights that may not be possible otherwise. Termed computer-aided nanotechnology (CAN), the challenges are equally to increase the applicability of the molecular and mesoscale modeling approaches as they are to the bridging of molecular sciences to the engineering fields.

Technologically important nanomaterials come in all shapes and sizes. They can range from small molecules to complex composites and mixtures. Depending upon the spatial dimensions of the system and properties under investigation, computer modeling of such materials can range from first-principles quantum mechanics, to forcefield-based molecular mechanics, to mesoscale simulation methods, to the prediction of structure-property relationships. All of the above computational techniques are available and integrated within Accelrys' Materials Studio(tm). This talk will illustrate a number of recent applications: (1) carbon nanotubes (CNTs) as nano electromechanical sensors (NEMS); (2) Metal-oxide nanoribbons as chemical sensors; (3) mesoscale modeling of polymer-CNT nanocomposites; and (4) mesoscale diffusion of drug molecules across cell membranes. In each case the molecular science challanges will be broadened to include discussion of the engineering challenges that exist in parallel and that will gate the ability to bring products based upon nanotechnology to market.

[O11] The current scenario of *in silico* drug discovery research

Anil K. Saxena, Medicinal and Process Chemistry Division, Central Drug Research Institute, Lucknow-226001, email: anilsak@hotmail.com

In Silico drug research has its greatest impact on target and lead discovery, which also constitutes the main components of today's drug discovery research. The aim of target discovery is the identification and validation of suitable drug targets for therapeutic intervention, whereas lead discovery identifies novel chemical molecules, effective on these targets or may have potential in the treatment of an ailment. With the near completion of the human genome sequencing, bioinformatics has established itself as an essential tool in target discovery with the *in silico* analysis of gene expression and gene function being its integral part, facilitate the selection of the most relevant targets for a disease under study. Subsequently, cheminformatics involves in the design of targeted, drug-like libraries for optimization of potency, selectivity and ADME/Tox properties through virtual screening which can efficiently manage cost effectiveness and dramatically shorten drug discovery cycle times. However the currently used *in silico* approaches, particularly chemoinformatics suffer from numerous barriers that prevent the realization of its full potential in terms of both effectiveness and efficiency. In our efforts to confront the barriers that plague this technology two strategies: (i) Integration of different complimenting modeling strategies and (ii) a multidisciplinary modeling approach, may have a great potential. In the above context the current state-of-the-art *in silico* drug research technology with respect to their potentials, barriers limiting their use and shall be elaborated on some of our strategies that may confront and overcome the barriers for an effective and efficient use of in silico approaches for drug discovery and development research.

[O12] Designing of less toxic to non target organophosphorus pesticide through multiple correlation analysis (MRA), a part of QSAR

Rini Roy, Aditi Nag Chaudhuri, Department of Microbiology, Lady Brabourne College, P-1/2, Suhrawardy Avenue, Kolkata – 700017, West Bengal, India; and Natasha Das Gupta, Department of Statistics, Lady Brabourne College, P-1/2, Suhrawardy Avenue, Kolkata – 700017, West Bengal, India.

Cost effectiveness and easy degradable nature of Organophosphorus Pesticide (OPs) in biological system and environment have implicated wide use of these types of pesticides in agriculture in India. They all exhibit anti cholinergic effect on non-targets including mammals. The aim of the present study is to find a model structure of an OP having minimal effect on any part of mammalian brain. In our study, we have dealt with the effects of phosphorothionate pesticides like methyl parathion, malathion, phorate, dimethoate, chlorpyrifos, monocrotophos and dichlorvos on acetylcholinesterase (AChE) activity in four different parts of rat brain, viz., hypothalamus (H), striatum (S), cerebellum (CR), cerebrum (C). The AChE activity shows a relationship with the LD50 values and the octanol-water partition coefficient (Pow) of five pesticides. An important conclusion ensues from the above results. Not all the parts of mammalian brain are equally sensitive to the same pesticide as is evident by direct plotting of inverse of AChE activity vs inverse of LD50 and that of AChE activity vs Pow values. Again, that has been supported by statistical calculations. The nature of dependence of pesticide (Pow values) and LD50 on AChE activity is given by MRA. This nature has been supported by multiple correlation coefficient values, which indicate the measure of efficacy for different predicting equations. In most of the cases, the results are satisfactory. The Km values obtained by enzyme kinetics of purified AChE from different parts of rat brain also substantiate the result. From 2-way ANOVA considering five pesticides it was evident that the two factors, i.e., Pow values of the pesticides and LD50 values, both of which affect AChE activity, are interdependent. A multiple regression equation has been formulated from the calculations. Pow and LD50 values of two other pesticides, viz., dichlorvos and monocrotophos, have been applied on the predicted equation to test the fitness of the equation.

[O13] Developing a more efficient system of taxonomy and nomenclature starting from modules derived from the geometrical simplexes of 1, 2 and 3 dimensional space

Seymour B. Elk, Elk Technical Associates, 321 Harris Place, New Milford, New Jersey 07646

Various selected aggregations of atoms, together with the bonds joining them, almost always function as integrated units — rather than being merely the sum of the individual atoms and bonds. Consequently, it is disingenuous to ignore this property when formulating taxonomy and nomenclature. Additionally, many modules which form the foundation upon which chemical structure is predicated can be subdivided into groups corresponding to geometrical simplexes in 1, 2 and 3 dimensions:

(1) The covalent bond, is represented in graph theory as a line segment; i.e., a one dimensional simplex. IUPAC nomenclature, as well as most other systems of chemical nomenclature, restricts its focus to unions of line segments. Moreover, isomorphism between atoms/nodes and bonds/edges is a well established postulate. This isomerism is useful for depicting the organic chemistry nomenclature associated with aliphatic compounds. Furthermore, when more than two atoms are attached to a single atom, this corresponds to "branching" of chains of atoms from a central node. Such description, however, is woefully inadequate for aromatic compounds.

(2) For aromatic compounds, a math model based on the regular hexagon is suggested. However, before applying this idea, attention is focused on the dual graph of the union of hexagons — a set of connected triangles (2-dimensional simplexes).

(3) Just as the benzene module in (2) bears a relation to the 2-simplex, the adamantane module corresponds to the tetrahedron (3-simplex); viz., the four tetrahedral vertices with coordination 3 carbon atoms and the midpoints of the six edges of the tetrahedron with coordination 2 carbon. In other words, let us replace the structural formula $C_{10}H_{16}$ for adamantane with the more meaningful formula: $\underline{C}_4\underline{\underline{C}}_6$, where underscoring indicates a module containing that atom and the indicated number of singly coordinated hydrogen atoms. This modular description in the nomenclature emphasizes the "anchor" property of $\underline{C}$ aggregates vs. the "extender" property of $\underline{\underline{C}}$ groups. This similarity is further attenuated by representing adamanzane, as $N_4\underline{\underline{C}}_6$, in contradistinction to the molecular formula ($C_6H_{12}N_4$), which contains no indication of modular bonding. The universality of this is evident in "traditional" inorganic chemistry, by the oxides of phosphorus, P_2O_3 and P_2O_5. Here both oxides occur as dimers: "trioxide" (actually hexa-oxide) phosphorus atoms function as the anchors of a tetrahedron and oxygen atoms as extenders. Moreover, the pentoxide anchor modules contain a double bonded oxygen atom radially extended from each phosphorus atom. Similar scenarios are expounded for selected boron compounds, etc.

[O14] Effect of geometrical fluctuations on DNA electron transfer rates: A computational study

Satyam Priyadarshy, RKR Group, 13507 Apple Barrel Ct., Herndon, VA 20171, USA.

Electron transfer reactions are crucial to the damage/repair mechanisms of DNA at a long distance. The mechanism of electron transport in DNA has attracted significant attention in last decade, with the central issue around how far and how fast the electron can transfer along the DNA biopolymer. An ability to understand DNA electron transport rests upon the knowledge of how molecular structure controls DNA electronic interactions.

Our previous theoretical studies have focused on studying the effect of the electron tunneling energy and DNA orbital energies on rate of electron transfer. In this paper, effect of geometrical fluctuations in DNA will be discussed. The results are obtained using coupled molecular dynamics and quantum chemical calculations on DNA molecules. Minor geometrical fluctuations in DNA cause large fluctuations in electron transfer rates. Understanding geometrical fluctuation effect is also important for the development of nano-devices that use DNA as their scaffolds.

[O15] Exploring 3D-QSAR of thiazole and thiadiazole derivatives as potent and selective human adenosine A_3 receptor antagonists

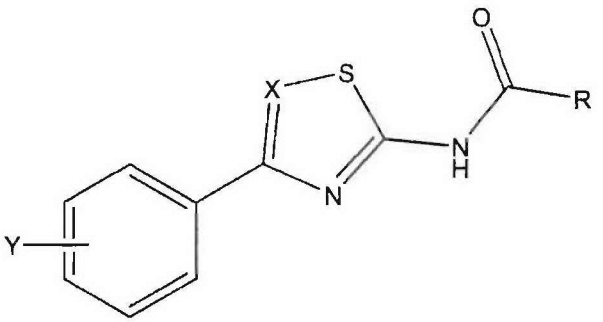
Prosenjit Bhattacharya, J. Thomas Leonard and Kunal Roy, Drug Theoretics and Cheminformatics Lab, Division of Medicinal and Pharmaceutical Chemistry, Department of Pharmaceutical Technology, Jadavpur University, Kolkata 700 032, India, kunalroy_in@yahoo.com.

Binding affinity data [*Bioorg. Med. Chem.* 2004, *12*, 613-623] of thiazole and thiadiazole derivatives (n = 30) for human adenosine A_3 receptor subtype have been subjected to 3D-QSAR analyses by molecular shape analysis (MSA) and molecular field analysis (MFA) techniques using Cerius2 version 4.8 software.

In case of the molecular shape analysis, major steps were (1) generation of conformers and energy minimization; (2) hypothesizing an active conformer (global minimum of the most active compound); (3) selecting a candidate shape reference compound (based on active conformation); (4) performing pairwise molecular superimposition using maximum common subgroup [MCSG] method); (5) measuring molecular shape commonality using MSA descriptors; (6) Determination of other molecular features by calculating spatial, electronic and conformational parameters; (7) selection of conformers; (8) generation of QSAR equations by genetic function algorithm (GFA) or stepwise regression. The best 3D-QSAR equation (MSA) obtained from GFA technique shows 70.0% predicted variance (leave-one-out) and 77.7% explained variance. This equation shows importance of Jurs descriptors (atomic charge weighted positive surface area, relative negative charge and relative positive charge surface area), partial moment of inertia, energy of the most stable conformer and the ratio of common overlap steric volume to volume of individual molecules. In case of stepwise regression, the best relation showed 46.1% predicted variance and 72.3% explained variance.

In case of molecular field analysis, the major steps were (1) generating conformers and energy minimization; (2) matching atoms using maximum common substructure (MCS) search and aligning molecules using default oprion; (3) setting MFA preferences (rectangular grid with 2 Å step size, charges by Gasteiger algorithm, H^+ and CH_3 as probes); (4) creating the field; (5) analysis by G/PLS method. The obtained equation was of excellent statistical qualities: 96.1% explained variance and 71.6% predicted variance.

Statistical reliable 3D-QSAR models obtained from this study suggest that these techniques could be useful to design potent A_3 receptor antagonists.



[O16] Feature extraction of gene expression data using curvilinear component analysis (CCA) and self-organizing map (SOM)

Yogesh P. Badhe, B. D. Kulkarni, Uttam S. Phulwale, S. S. Tambe and Vinay A. Wadekar, Chemical Engineering &Process development Division, National Chemical Laboratory, Pune-411008, India.

Feature extraction is a process of mapping linearly or nonlinearly the original features (measurements) into fewer features, which maximally preserve the information content in the data. In feature extraction, all the available variables are used and data are transformed (using a linear or nonlinear transformation) to a lower dimensional space. Thus, the aim is to replace the original variables by a smaller set of uncorrelated variables. Feature extraction has two motivations: (i) to reduce dimensionality of a multivariable high dimensional data set (*dimensionality*

reduction), and (ii) to produce a low-dimensional representation (ideally 2D or 3D) so that the data may be easily viewed and the relationships and structure in the data are identified (*data projection*).

Dimensionality reduction and multivariate data projection are important issues in pattern recognition and exploratory data analysis. Dimensionality reduction improves the generalization ability of data classifiers and reduces the computational requirements in pattern classification. Data projection enables us to visualize high dimensional data in fewer dimensions to understand better the underlying structure, explore the intrinsic dimensionality, and analyze clusters of multivariate data.

In this study, a gene expression data set consisting of 2329 cDNA fingerprinted (observations) with 139 oligos (variables) is considered for performing feature extraction. The traditionally used the feature extraction paradigms such as the linear principal component analysis (PCA), or non-linear multi-dimensional scaling (MDS), are not efficient in reducing very high dimensional data into two or three dimensions. Thus, we have used two artificial intelligence (AI) based non-linear feature extraction paradigms namely curvilinear components analysis (CCA) and self-organizing map (SOM) to visualize the gene expression data onto a two-dimensional map and to perform their classification. First, CCA is used as a dimensionality reduction technique so that the original 139-dimensional oligo data set is reduced to a 40 dimensional set. Next, SOM is used as a data projection technique for visualization of the 40 dimensional data onto a two-dimensional map and to perform classification. The results obtained thereby indicate that the cDNA data contain 18 clearly visible clusters with well-defined boundaries when viewed in two dimensions. The AI-based techniques presented in this paper are generic in nature and therefore can be effectively utilized for nonlinear feature extraction, dimensionality reduction, classification and low dimensional projection of a wide variety of DNA data.

[O17] Four-color map representation of DNA sequences and their numerical characterization

Milan Randić, National Institute of Chemistry, P.O.B. 3430, 1001 Ljubljana, Slovenia; Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, 5013 Miller Trunk Hwy., Duluth, MN 55811, USA; Nella Lerš, Dejan Plavšić, Ruder Bošković Institute, P.O. Box 180, HR–10002 Zagreb, Croatia; and Alexandru T. Balaban, Texas A&M University at Galveston, Galveston, TX 77553, USA.

We put forward a novel 2-D graphical representation of DNA sequences based on the construction of four-color map obtained from a "spiral" representation of DNA by erasing shared sides between adjacent square cells having the same color. The novel graphical representation is highly compact and it not only allows visual inspection of similarity/dissimilarity between DNA sequences, but also leads to their numerical characterization. The approach is illustrated with the coding sequence of the first exon of β-globin gene of human and a dozen other species.

[O18] Fuzzy set based signal analysis of structured experimental and computational data in chemistry

Lech P. Schulz, Adam Mickiewicz University, Department of Chemistry, Grunwaldzka 6, 60-780 Poznań, Poland

Limited number of essential, quantitative data received in the experiment or as the result of computations is one of the often-encountered peculiarities in chemistry. Therefore, the standard, statistical analysis is difficult to apply in many cases. A more universal tool has preliminarily been tested in works [1] and [2] in the search for parameters responsible for anticancer activity of *diammine platinum* derivatives. Using resemblance *pseudo-metric* [3], an arbitrary set of data can be structured by distinguishing subclasses of points selected due to relationships determined by the *resemblance* between particular data. A feature of the comparisons made is the simultaneous "weighting" of the common and different parts of selected subclasses in question. It is said that the fixed set of data has been structured by these subclasses.

Next, a *fuzzy set* based estimation of the structure received becomes possible by means of certain *criteria of quality* of the *complex systems* involved. The most easily available criteria of this kind are, e.g., *overlapping* and *splitting* [4] though particular problems can require special solutions. The approach proposed assumes that useful properties encoded in molecular systems, like *diammine platinum complexes*, give rise for a signal manifested by a structural feature of a data set.

A procedure for confirmation has been proposed in this work by the search of signals of the same origin in the context of various structures. The investigation is developed both theoretically and in the direction of computer code creations useful in concrete selections of parameters responsible for pharmacological activity.

References:
[1] Schulz, L.P., Chojnacki, H., J. Mol. Struct. 404(1997)87-90.
[2] Schulz, L.P., Chojnacki, H., Int. J. Quantum Chem. 60(1996)1385-1391.
[3] Schulz, L.P., Proceedings of ICCMSE (Kastoria 2003, Simos, T.E., ed.), pp. 571-584, World Scientific, New Jersey 2003.
[4] Schulz, L.P., Preliminary Proceedings of the IIIrd Int. Conference DSTIS (4-6 Sept., 2003), pp. 125-136, Warsaw 2003.

[O19] Graph-theoretically derived extended Hückel rule can explain the aromaticity and unusual properties of condensed polycyclic hydrocarbon networks.

Haruo Hosoya, Ochanomizu University (Emeritus), Bunkyo-ku, Tokyo 112-8610, Japan

Although both experimental and theoretical studies have been performed on the π-electronic stability of condensed nonbenzenoid hydrocarbons, there has been proposed no simple theory for predicting their aromatic or antiarmatic character. Application or extension of the Hückel's ($4n$+2) rule has been unsuccessful. More than thirty years ago the present author proposed the extended Hückel rule by using the aromaticity index, ΔZ, based on his Z-index [1], its application to larger molecules has not been publicized yet. Quite recently it was found that this simple method even without using a computer can predict and explain the π-electronic stability of a number of fairly large condensed nonbenzenoid hydrocarbons. Further, several unusual properties newly found in the mathematical structure of pericondensed tricyclic networks were also deciphered by this analysis. A few examples are shown below.

Acenaphtylene (665), a pericondensed tricyclic hydrocarbon with 12 π-electrons (3R12π), is known to be stable, while a Hafner's hydrocarbon (755) is known to be almost aliphatic. These facts can be explained from the Hückel molecular orbital calculations for the group of isomers also by taking into account their steric strain, although it has already been pointed out that the algebraic structure counting fails in these tricyclic systems [2], and Z-indices of all the isomers were found to be the same. On the other hand, one can quite easily predict and explain these HMO results just by enumerating the ΔZ values with manual addition of all the ring contributions. It was further found that all the 3R12π isomers have five common eigenvalues expressed by $2\cos(2k\pi/11)$ with k=1~5. Other 3R2nπ isomers were also shown to have n–1 common eigenvalues expressed by $2\cos(2k\pi/(2n-1))$ with k=1~n–1. Mathematical structure of the extended Hückel rule and the related unusual properties of polycyclic nonbenzenoid hydrocarbons will be explained in this paper.

References:
[1] H. Hosoya, K. Hosoi, I. Gutman, Theor. Chim. Acta, **38**, 37 (1975).
[2] N. Trinajstic, Chemical Graph Thoery, 2nd. Ed., Boca Raton, FL (1992), pp. 188.

[O20] Graph theory and high-level quantum models for predictive toxicology and graph algorithms for proteomics

K. Balasubramanian, Institute for Data Analysis & Visualization, University of California Davis, Livermore, CA 94550; University of California, Chemistry and Material Science Directorate, Lawrence Livermore National Laboratory, Livermore, California 94550; and Glenn T Seaborg Center, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720.

My talk will deal with integration of graph theory and high-level quantum chemical method that employ large basis sets and electron correlation treatment for mechanistic predictive toxicology. We will show that such techniques are warranted to understand the mechanisms and the metabolic pathways of toxic action. We propose to explore the

cross-fertilization of graph theory and quantum chemistry techniques to derive accurate descriptors that are particularly sensitive to the electronic and stereochemical features of the molecules. We shall consider quantization of graph-theoretical indices and also consider stereo-chemically based and geometrical indices. We shall focus on high-level coupled cluster computations for predictive toxicology of halocarbons. We consider high-level computations of the metabolites including structures of neutral molecules and their anions as well as dissociative electron attachment properties for halomethanes of general formula CX_nY_m; X=H,F; Y=Cl,Br,I; n=0,4; m=4-n. We have considered coupled cluster computations on 55 such species comprising the well-known Crebelli set. These results demonstrate that the dissociative electron attachment is the primary process resulting in hepatotoxicity of these species. The other part of the talk will consider graph-theoretical algorithms for the characterization of the 2d-gel patterns of proteins. We show that complex weighted graphs facilitate characterization of the underlying proteomics patterns.

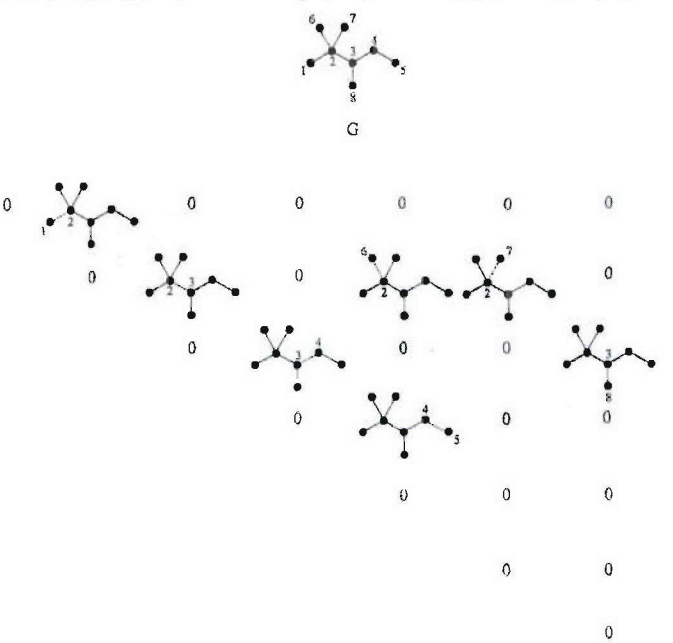
[O21] Graphical matrices as sources of novel invariants for use in QSPR

Sonja Nikolić, Ante Miličević and Nenad Trinajstić, Rugjer Bošković Institute, P.O.B. 180, HR-10002 Zagreb, Croatia.

Graphical matrices are matrices whose elements are subgraphs of the graph rather than numbers. Since the elements of these matrices are (sub)graphs, they are called the *graphical matrices*. Thus far a very little work has been done on these matrices. However, many so-called special matrices, such as the Wiener matrices and the Hosoya matrices, may be regarded to be the numerical realizations of the corresponding graphical matrices. Therefore, the graphical matrices appear to be a promising area of research in chemical graph theory.

The advantage of a graphical matrix lies in the fact that it allows great many possibilities of numerical realizations. In order to obtain a numerical form of a graphical matrix, one needs to select a graph invariant and replace all the graphical elements (subgraphs of some form) by the corresponding numerical values of the selected invariant. In this way, the numerical form of the graphical matrix is established and one can select another or the same type of invariant – this time an invariant of the numerical matrix. Graph invariants generated in this way are *double invariants* in view of the fact that two invariants are used in constructing the targeted molecular descriptor.

We present several graphical matrices that lead to Wiener-Wiener indices, Hosoya-Wiener indices, Hosoya-Balaban indices, Randić-Wiener indices, etc. These indices were considered because the values of the Randić indices, Wiener indices, Balaban indices and Hosoya indices for smaller acyclic fragments (trees) are readily available. As an example, we give the construction of the graphical matrix, based on the concept of the edge-Wiener matrix, that leads to the edge-Wiener-Wiener index ^eWW. It should be noted that the edge-Wiener graphical matrix is constructed by the consecutive removal of the edges from the graph. This is shown below for the hydrogen-depleted graph G representing 2,2,3-trimethylpentane. Since a graphical matrix is a square, $V \times V$, symmetric matrices, it is enough to give for the demonstrative purposes only the upper triangle of the matrix. For graphs without loops, the corresponding graphical matrices have zeros diagonal elements.



The next step is to replace (sub)graphs with their Wiener numbers obtained by summing up the Wiener numbers of acyclic fragments. Below is given the numerical realization of the above edge-Wiener graphical matrix.

0	46	0	0	0	0	0	0
	0	19	0	0	46	46	0
		0	29	0	0	0	46
			0	46	0	0	0
				0	0	0	0
					0	0	0
						0	0
							0

The summation of the matrix-elements in the above matrix-triangle gives the edge-Wiener-Wiener number ${}^{e}WW$ of 2,2,3-trimethylpentane (${}^{e}WW$ =278).

The application of this kind of molecular descriptors to QSPR modeling is described for octanes since the modeling properties of this class alkanes is well-studied in the literature and thus we have a standard against which we tested our models. Our modeling is based on the CROMRsel procedure so devised to give the best possible model for a given number of descriptors.

[O22] Handling tautomerism and geometrical isomerism in drug design and synthesis: The case of phenylpyruvic acid derivatives

Alain J.M. Carpy, LPTC, UMR 5472 CNRS, University of Bordeaux I, 351 Cours de la Libération, 33405 Talence CEDEX, France; Petrus P. Haasbroek, Department of Pharmacy and Pharmacology, University of the Witwatersrand, 7 York Road, Parktown 2193, South Africa; Douglas W. Oliver, Pharmacology, Faculty of Health Sciences, North-West University, Potchefstroom Campus, Private Bag X6001, Potchefstroom 2520, South Africa.

Tautomerism and, to a less extent, geometrical (Z/E) isomerism are often disregarded in drug design and synthesis applications. However things change progressively. The first problem with tautomerism may occur when one wants to retrieve a structure from a chemical database as the different tautomeric states are seldom registered. Further, as registered structures may be used as input for other programs e.g. QSAR applications (molecular similarity, high throughput screening, molecular properties etc) and computer-aided molecular modelling (docking) drug designers and medicinal chemists must be concerned with the implications of tautomerism. Similarity indexes frequently used to identify analogues of active compounds may not be recognised when the compounds are registered as different tautomeric forms with different functional groups.

Tautomeric forms also differ in shape and properties. It has been demonstrated that lipophilicity quantified by log P (octanol/water partition constant) which may be predicted with a number of programs differ from one tautomeric state to another. Obtaining the appropriate tautomeric form improves the quality of physical property predictions. Shape (surface and volume), hydrogen bonding properties of atoms (donor/acceptor), electrostatic interactions used to dock structures in active sites of receptor/enzyme also differ from one tautomeric state to the other. When by chance the crystal structure of the complex between the macromolecular target and a specific ligand is known the docking procedures are facilitated. On the contrary difficulties may occur when the answer to the question: "Which tautomeric form to dock?" is unknown. So when one calculates molecular descriptors and molecular properties he must be highly concerned with the tautomery specially when, in addition, geometrical isomerism occurs on one tautomeric form. On the other hand, tautomerism has been also very often ignored by medicinal chemists in the planning of further synthetic reactions.

Keto/enol tautomerism is important in several areas of biochemistry e.g. the high phosphate-transfer potential of phosphoenolpyruvate results from the fact that the phosphorylated compound is trapped in the less stable enol form, whereas after dephosphorylation it can assume the keto form. A series of phenylpyruvic acid derivatives exhibiting the keto/enol tautomerism has been investigated both experimentally and theoretically. These compounds interact with several enzyme systems. In the active site of the phenylalanine dehydrogenase (a diagnostic enzyme used in the monitoring of phenylketonuria), the phenylpuruvate exists as the keto tautomer potentially conformationally flexible. Similarly we demonstrated that the human carboxypeptidase A might be complexed by the same keto phenylpyruvates. Phenylpyruvate derivatives can also be potential inhibitors of the phenylpyruvate tautomerase activity catalysed by the Macrophage Migration Inhibitory Factor (MIF). Contrary to the complexes with the two mentioned enzymes, phenylpyruvate derivatives exist in the active site of MIF as enol tautomers. Our theoretical

investigation successfully explored the tautomerism and geometrical isomerism of new compounds of the 2-phenylpyruvic acid type synthesised via the azlactone route and the implications in modelling and designing novel bioactive agents interacting specifically with MIF.

[O23] A high dimensional QSAR study on the *Aldose Reductase* inhibitory activity of some flavonoids: Topological descriptors in modeling the activity

Yenamandra S. Prabhakar, Manish K. Gupta, Medicinal and Process Chemistry Division, Central Drug Research Institute, Lucknow-226 001, India; Nobendu Roy, Bioinformatics Division, Satyam Computer Services Ltd, Hyderabad, India; Yenamandra Venkateswarlu, Natural Products Laboratory, Indian Institute of Chemical Technology, Hyderabad-500 007, India, yenpra@yahoo.com.

The inhibition of *Aldose Reductase* (AR) in polyol pathway may prevent and address the cure of the complications such as cataract, neuropathy and retinopathy arising out of the diabetes mellitus. A number of traditional antidiabetic remedies are found to inhibit AR wherein several flavonoids were identified as the active constituents. In this background, Matsuda and coworkers qualitatively examined the structure and AR inhibitory activity of large number of flavonoids and related compounds [1]. With different substitution positions associated with the possibilities for variations, these compounds offer scope for identification and generation of alternative structures for the AR inhibitory activity. As a step in this direction we attempted a CP-MLR (Combinatorial Protocol in Multiple Linear Regression)[2] directed quantitative structure activity relationship (QSAR) study of the AR inhibitory activity of some public domain flavonoids and their derivatives with topological indices from Molconn-Z software [3]. Out of the one hundred and twenty-six molecular connectivity indices formed for these compounds, seventeen indices embedded with the atom type counts, electrotopological states and polarity information contents have been found to model the AR inhibitory activity. Among these seventeen descriptors, the cumulative electrotopological states of methylenes (CH2) and polarity have displayed very high relevance in modeling the inhibitory activity of these compounds. Also, the atom counts and the electrotopological states corresponding to aromatic as well as unsaturated carbons have shown considerable influence on the inhibitory activity. An indicator parameter incorporated in the study suggests that C3 of flavone skeleton is favorable for substituents bulkier than simple hydroxyl group for better inhibitory activity.

References:

[1] Matsuda, M, Morikawa, T, Toguchida, I and Yoshikawa, M. Chem Pharm Ball. **2002** *50*, 788-795.
[2] Prabhakar, Y. S. *QSAR Comb Sci.* **2003**, *22*, 583-595.
[3] MolconnZ programer's toolkit; http://www.edusoft-lc.com/molconn/manuals

[O24] Index of variance of dinucleotide frequencies and evolutionary relationships in gene families

A. Nandy, Environmental Science Programme, Jadavpur Univeristy, Jadavpur, Kolkata 700068, anandy43@yahoo.com

It is the usual practice to define evolutionary relationships between members of a gene family such as the globin genes through an analysis of the amino acid differences in the associated proteins. In some of our previous works we had shown that a similarity index based on the primary DNA sequences had been able to demonstrate broadly the evolutionary relationships between different gene sequences. In this paper we define an index of relationship between two sequences through their dinucleotide frequency spectra and show that within the members of a gene family this new index is able to quantitatively relate the evolutionary relationships that were earlier found through amino acid analysis. We have applied the technique to several gene families and find that the results are quite encouraging.

[O25] Lactone analogues as protein kinase C (PKC) ligands in prostate cancer: An *in silico* approach to drug-design

Shailza Singh, D.K. Sharma, Indian Institute of Technology, Hauz Khas, Delhi-110016, India; B.K. Malik, Institute of Genomics and Integrative Biology, CSIR, Mall Road, Delhi-110007, India.

Prostate cancer metastases are the second most common cause of cancer-related death among men in the United States. Metastasis, the main cause of death for most cancer patients, remains one of the most important but least understood aspects of prostate cancer. Despite recent advances, current diagnostic tests and treatment of prostate cancer have limitations. In the last few years, numerous biomolecules have been investigated with the aim of improving diagnosis, including kallikrein-like proteases, growth factors and neuroendocrine markers. Analysis of susceptibility genes has also been a focus of attention. Extensive research into new therapeutic approaches is also underway, including targeting angiogenesis, immune regulation and stromal-epithelial interactions. Gene therapy, gene chip technology and proteomics have emerged as promising innovations. Defining the mechanisms of action of protein that are required for metastasis to occur may help identify therapeutic agents. Protein Kinase C is one amongst these. Protein kinase C (PKC), phosphorylates target protein that control growth, differentiation and transcription, can inactivate Raf kinase inhibitor protein (RKIP) which functions as a suppressor of metastases, through phosphorylation on serine 153 and alleviate its inhibition of Raf-1.PKC is normally recruited to the plasma membrane and activated by diacylglycerol. Its location near the plasma membrane may place it in close proximity to RKIP, which also binds to phospholipids. As a result, PKC along with RKIP, function as unique selective regulators of the Raf-1/MEK/ERK growth factor signaling cascade. When RKIP is phosphorylated, it releases from Raf-1 and can bind onto G-protein –coupled receptor kinase-2 (GRK-2) preventing GRK-2's ability to inhibit GPCR activity. RKIP activity itself is regulated by PKC. Protein kinase C has been modeled using MODELLER software with 95.6% of the residues in the most favored regions, 3.9% in the additional allowed regions and 0.4% in the generously allowed regions of Ramachandran Plot. The stereochemical quality of PKC predicted through PROCHECK program suggests a fairly good quality model. Molecular Modeling of PKC indicated reduced binding affinity of the representative lactone group and entropy loss due to the relatively flexible hydroxyethyl group. Hence, 4-C-hydroxyethyl 1-2,3-dideoxy-D-glyceropentono-1, 4-lactone is being docked into the predicted active site of PKC as a competitive ligand for further study towards drug-design.

[O26] Mathematical modelling of simple biochemical pathways

C. Suguna, Somdatta Sinha, Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007

Living cells have a wide array of functions, which are performed by biochemical reactions within the cell. These reactions form a complex network of interconnected pathways. As a first step in studying the dynamics of these pathways, we break down the network into smaller functional units. These units are regulated by positive and negative feedback processes. Negative feedback is found most commonly in biological processes as it can stabilize the system to environmental fluctuations and help in conservation of energy. Positive feedback processes, on the other hand, have a destabilizing role but are useful for amplification of signals required in excitable and switching dynamics. We have studied the behaviour of simple biochemical pathways by using mathematical models, which incorporate known features of these reactions. We have studied a minimal model of a biochemical process consisting of a single negative and positive feedback process, regulated by mechanisms commonly found in biological systems and enumerated its behaviour. We have shown that it can exhibit a wide range of behaviours from homeostasis to periodic oscillations, complex oscillations, bifurcations, birhythmicity and chaos. In the birhythmic region we find that the basins of attraction are fractal. We have studied the behaviour of the pathway under different conditions. In the presence of noise we find that the same pathway can show both sensitive and robust behaviour under different conditions. We have also studied the emergent behaviour of aggregates of cells incorporating this pathway in each cell and found that the emergent behaviour is non-intuitive and can be very different from the single cell behaviour.

[O27] Mathematical structural invariants: Developments and applications

Subhash C. Basak, Natural Resources Research institute, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth, MN 55811, USA; sbasak@nrri.umn.edu.

Mathematical approaches, particularly methods of discrete mathematics, have been used for the characterization of structures of chemical and biological systems. Quantitative chemodescriptors are useful not only in deciphering the structural basis of various properties, but also in ordering molecules, measuring similarity/ dissimilarity of molecular systems, and predicting property/ therapeutic activity/ toxicity of chemicals. Whereas chemodescriptors developed for molecules quantify certain aspects of their structure, biodescriptors derived from DNA sequence, genomics and proteomics data are useful in the pattern recognition of such biological macromolecular systems. Our research team at the Natural Resources Research institute (NRR)I has been involved in the development of different classes of novel chemo- and biodescriptors. This presentation will give a brief overview of development of these descriptors and their applications in drug discovery and computational toxicology.

[O28] Modeling and interactions of *Aspergillus fumigatus* lanosterol 14-α demethylase enzymes 'A' and 'B' with azole antifungals

Reena Gollapudy, Sudhir A. Kulkarni, VLife Sciences Technologies Pvt. Ltd. 1 Akshay, 50 Anand Park, Aundh, Pune 411 007, India

Recent identification of the sterol 14-α demethylase genes (CYP51 'A' and 'B') from *Aspergillus fumigatus* and other species by Mellado et al. (J. Clin. Microbiol. 2001, 39(7), 2431–2438), has opened up possibilities of investigating the interactions of azole antifungals with the enzyme(s) from fungi. This study describes for the first time, models of the three-dimensional structures of *A. fumigatus* 14-α demethylases (AF-CYP51 'A' and 'B'), using the crystal structure of *Mycobacterium tuberculosis* 14-a demethylase (PDB code:1EA1) as a template. The paper also describes the various interactions between azole antifungals and the targets from *A. fumigatus*. Quantitative evaluation of these interactions is done using comparative binding energy (COMBINE) analysis to understand contributions of active site residues to ligand activity. It also provides explanation for the activity/inactivity of different ligands for AFCYP51 'A' and 'B'.

[O29] Molecular docking using LigandFit and LibDock

Chandan Raychaudhury, Accelrys K.K., 12th Floor, Discoverer, International Tech Park, Whitefield Road, Bangalore 560066, India.

Docking of chemical compounds to protein active site is one of the fast growing areas of interest in today's drug discovery research. The idea is to identify molecules which could fit to and interact with the protein receptors and could be screened as potential hits. While a number of docking algorithms are available in the literature, it is important that the docking programs be applied to various protein targets to screen different databases of chemical compounds and the usefulness of the programs be understood on the basis of the scientific aspects that the programs are based on. In the present study, we have used two docking programs, LigandFit and LibDock, for docking a number of chemical compounds to different protein receptors. While LigandFit is a shape-based method, LibDock is feature-based. The results have been analyzed from the standpoint of individual algorithm.

[O30] The Molecular Experimental Design (MED) concept: application to the selection and use of molecular descriptors in early ADME prediction

Jacques R. Chretien, MarcoPintore, Nadège Piclin, BioChemics Consulting, Innovation Center, 45074 Orleans Cedex 2, France.

The past fifteen years a wealth of new and powerful techniques such as Combinatorial Chemistry, High Throughput Screening (HTS), Proteomics, Genomics have been largely impulsed in pharmaceutical industry. Unfortunately, the average cost for any blockbuster just arrived on the market has increased drastically and the discovery time remains very high. It prompted us to develop a critical study of the drug discovery process to feel the gap between attractive scientific and/or academic proposal and the industrial requirements derived from Quality Control (QC). The Molecular Experimental Design (MED) concept addresses this challenge.
The MED concept can be expressed under form of "the rule of four", it has been especially designed for the requirements of virtual HTS (v-HTS). We will take chiefly the ADME challenge as a salient example because it remains one of the most crucial points to boost the drug discovery process. Effectively it was a must to go far beyond the limits of the Lipinski rules to avoid most troubles in the clinical phase of drug development. Procedures able to predict ADME properties of any molecule are then required and the proper choice of molecular descriptors remains the core of such a challenge.
Recently, new computational methods based on Adaptive Fuzzy Partitioning (AFP) approach with help of Genetic Algorithms and Fuzzy Logic have been developed by us allowing to develop a number of early ADME predictors [1]. In this contribution, their application to the main pharmacokinetic properties, i.e. oral absorption, bioavailability, volume of distribution and clearance, will be discussed. All models generated were validated by cross-validation, test set and Y-sampling procedures, and most of them were able to predict correctly ADME properties with prediction rates higher than 65-70%. The proposed techniques showed robustness and prediction power higher than those derived from other comparable methods. Moreover, a detailed analysis of different families of descriptors underlined that: i) the average contribution of 3D parameters in increasing the prediction ability as regard to 2D ones is negligible; ii) the fingerprints are not really suitable to model ADME properties, probably as their contribution to the global molecular information, based mostly on structural fragments, is too limited to define so complex properties.

References:
[1] Pintore M, van de Waterbeemd H, Piclin N, Chrétien J., Eur J Med Chem (2003), 38, 427-431.
[2] Ros F., Taboureau O., Pintore M., Chrétien J.R., Chemom. Intell. Lab. Syst. 2003, 67, 29-50.

[O31] Molecular modeling of glutamate and adenosine receptors and design of their new ligands

Nikolai S. Zefirov, Vladimir A. Palyulin, Department of Chemistry, Moscow State University, Moscow 119992, Russia.

Molecular modeling is the powerful instrument for the design of new bioactive compounds. In the present paper the constructed homology-based models for all domains of N-methyl-D-aspartate (NMDA) receptor, for all domains of all known subtypes of metabotropic glutamate receptors mGluR1-8, and four known subtypes of adenosine receptors are discussed.

3D-models of the NMDA receptor glycine binding site have been built using homology modeling on the basis of an X-ray structure of the water-soluble AMPA-sensitive receptor. The docking of agonists and antagonists to these models was used to reveal binding modes of ligands and to explain known structure-activity relationships. Two types of quantitative models, CoMFA and a regression model based on docking energies, were built for a series of antagonists. The CoMFA steric and electrostatic maps were superimposed on the homology-based model, and a close correspondence was marked. The derived computational models have permitted the evaluation of the structural features important for high glycine binding site affinity.

The approach for the analysis of binding selectivity of ligands was suggested. For analyzing the pair-wise selectivity, we propose to use the difference between biological activities (expressed in energy scale) of ligands with respect to different receptor subtypes as a dependent variable for building CoMFA models. The resulting "selectivity fields" indicate the ways of increasing binding selectivity. The applicability of this approach was demonstrated

using Glycine/NMDA and AMPA receptors. The mapping of "selectivity fields" onto the molecular models of the corresponding receptors makes it possible to reveal the reasons of experimentally observed selectivity as well as to suggest additional ways of increasing it.

The models were used for revealing binding modes of agonists and antagonists, as well as for rationalizing known experimental facts concerning structure-activity relationships and for design of new ligands.

[O32] Net-sign identity information index: A novel approach towards numerical characterization of chemical signed graph theory

Prabhat K. Sahu and Shyi –Long Lee, Department of Chemistry and Biochemistry, National Chung Cheng University, Chia-Yi, 621, Taiwan.

In this work, we have made an attempt to derive a novel information theoretic topological index, net-sign identity information index, I_ε from the molecular electronic structure, namely, numerically characterize the net-sign approach, proposed by Lee et al [1] in chemical graph theory. We have recently reported the novel information theoretic topological index and few other properties based information theoretic indices. [2] So far information theory offers quantitative methods for the study of receiving, conserving, processing, transmitting and practical use of information. The relation that exists between structure and properties has been another field of successful application of information-theoretic approach to chemistry. Moreover, the index, I_ε is found to be useful in QSPR modeling.

References:
[1] S. –L. Lee et al. Chem. Phys. Lett. 137 (1987) 279.
[2] P. K. Sahu and S. –L. Lee* Chem. Phys. Lett. (*in press*).

[O33] New approaches to the liquid-solid phase transition of Lennard-Jones fluids

Chizuru Muguruma, Faculty of Liberal Arts, Chukyo University, 101 Tokodachi, Kaizu-cho, Toyota 471-0393, Japan; Yuko Okamoto, Institute for Molecular Science, Okazaki, Aichi 444-8585, Japan and Graduate University for Advanced Studies, Okazaki, Aichi 444-8585, Japan; Masuhiro Mikami, National Institute of Advanced Industrial Science and Technology, Research Institute for Computational Sciences, Chuo-2, 1-1-1 Umezono, Tsukuba, 305-8568, Japan.

The great advancement of computer technology and simulation techniques have made computer simulation an effective tool in many fields of chemistry and physics. However, simulations of complex systems with many degrees of freedom such as spin glasses and biopolymers are still greatly hampered by the multiple-minima problem. This is because conventional canonical simulations at low temperatures tend to get trapped in one of huge number of local-minimum states on the potential energy surface.

The multicanonical (MUCA) and the generalized isobaric-isothermal ensemble algorithms have been introduced as the methods to overcome the multiple-minima problem. The algorithms are based on artificial, non-Boltzmann weight factors. The MUCA algorithm performs a free one-dimensional random walk in potential energy space, while the generalized isobaric-isothermal ensemble algorithms perform a random walk widely in volume space and in potential energy space, which allow the simulation to avoid getting trapped in states of energy local minima. Moreover, one can calculate the expectation values of thermodynamic properties as a function of temperature by applying the histogram-reweighting techniques to the results of one long production run. In the present study, we applied the MUCA Monte Carlo (MC) method and the MC method on isobaric-multithermal ensemble to the liquid-solid phase transition of the Lennard-Jones fluids.

For MUCAMC calculation, we put argon particles in a periodic cubic cell so that the density of the system is 1.65 g/cm^3 and determined the MUCA weight factor under the fixed volume. For the system of 108 argon particles, the probability distribution obtained by the long production run shows that the weight factor is reliable for the temperature range from 60 to 250 K. Averaged thermodynamic properties such as potential energy, pressure, entropy, and free energy tell us that the first-order phase transition takes place at 150 K. The size effects for the

systems of 256 and 500 argon particles are now under consideration. We also carried out an isobaric-multithermal simulation in order to compare the data with the experimental results. We will describe the algorithm itself and present the results of the MC calculations under 1 atm and 3000 atm.

[O34] New developments in variable connectivity indices

Matevž Pompe, University of Ljubljana, Faculty of Chemistry and Chemical Technology, Aškerčeva 5, 1000 Ljubljana, Slovenia; Milan Randić, National Institute of Chemistry, Ljubljana, Slovenia & 3225 Kingman Rd. Ames, IA 50014, USA.

We will present a critical review of not so well known novel *variable* molecular descriptors designed for use in the quantitative structure-property relationship (QSPR) and the quantitative structure-activity relationship (QSAR). In contrast to the traditional molecular descriptors, that is, various topological indices, quantum chemical indices, and indicator variables, which once a molecule is selected can be calculated and remain fixed, the variable molecular descriptors continue to be changed during the regression analysis till their optimal values that result in the minimal standard error are found. Use of optimal variable molecular descriptors not only visibly improves numerous regression results and reduces significantly the number of molecular descriptors to be used in multivariate regression analysis, but also has facilitated interpretation of regressions in some instances. The possibilities of structural interpretation of obtained optimal molecular descriptors will be shown in the series of halogen, oxygen and nitrogen containing aliphatic organic compounds. The study will show the magnitudes of changes of the optimal variable molecular descriptors if different properties or data sets are taken into consideration.

At the end a new development of variable connectivity index to model anti-connectivity properties will be presented. Several properties exists where presence of certain atom or functional group suppresses the signal. None of the classical bond additive or atom additive topological indices has the ability to describe such behavior, this also include current variable topological indices. A new procedure for the calculation of variable connectivity indices will be presented. The modeling of the FID response factors will be shown as an example.

[O35] Novel development of quantitative structure-activity relationship for anti-bacterial activity of newly synthesized Mannich bases using NMR chemical shift as a molecular descriptor

Padmakar V. Khadikar, Department of Chemistry, A.P.S. University, Rewa 486003, India; Dheeraj Mandloi, Institute of Engg. & Technology, Devi Ahilya University, Indore 452017, India; and Sheela Joshi, School of Chemical Science, Devi Ahilya University, Indore 452017, India.

Mathematical models are developed to find quantitative structure-activity relationships (QSARs) that correlate nmr chemical shift and anti--bacterial activity of a series of newly synthesized Mannich bases. In doing so chemical shift in nmr was first established as a molecular descriptor by correlating it with a series of distance-based topological indices. This also provides physical meaning which is probably associated with the nmr chemical shift. Simple as well as multiple linear regressions are performed to find superior models for the prediction of anti-bacterial activity of the Mannich bases. Better results are obtained by the introduction of indicator parameter(s) in the regression analysis. The results are discussed critically using a variety of statistics.

[O36] A novel method of blocking malarial transmission through insect vectors

Ramamurthy Jayalakshmi, Department of Bio-Informatics, Thanthai Hans Roever College, Perambalur, Tamil Nadu, India 621 212; M. Vivekanandan, Department of Biotechnology and Bioinformatics, Bharathidasan University, Tiruchirappalli, India 620 024; Ramanathan Natarajan, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811

Anopheles gambiae is the principal vector of malaria and the analysis of its genome sequence revealed strong evidence for about 14,000 protein-encoding transcripts. Vector and parasite interact at multiple stages and locations, and the nature and effectiveness of this reciprocal interaction determines the success of transmission. Many of the interactions engage the mosquito's innate immunity, a primitive but very effective defense system. In some cases, the mosquito kills the plasmodium parasite, thus blocking the transmission cycle. Prominent expansions in specific families of proteins likely involved in cell adhesion and immunity were noted. An expressed sequence tag analysis (EST) of genes regulated by blood feeding provided insights into the physiological adaptations of a hematophagous insect. Olfaction plays a major role in host preference and blood feeding. It has been identified that four genes encoding candidate odorant receptors from *A. gambiae* that are selectively expressed in olfactory organs, contain approximately seven transmembrane domain. *A. gambiae*'s salivary gland receptor or receptors for sporozoite invasion are the possible target for blocking malarial transmission. In an in vivo bioassay, the monoclonal antibody against the 100-kDa protein inhibited Plasmodium yoelii sporozoite invasion of salivary glands by 70-75%. These results show that *A. gambiae* salivary gland proteins are accessible to monoclonal antibodies that inhibit sporozoite invasion of the salivary glands and suggest alternate targets for blocking the transmission of malaria by this most competent of malaria vectors.

[O37] Partial order: A general tool in ecotoxicological and ecosystems hazard assessment

Rainer Brüggemann, Ute Simon, Leibniz - Institute of Freshwater Ecology and Inland Fisheries, Müggelseedamm 310, 12587 Berlin, Germany; Subhash Basak, Natural Resources Research Institute, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth, Minnesota 55811, USA.

One of the first steps of hazard assessment of chemicals is to formulate a ranking procedure in order to find which chemicals should be tested prior to others. Chemicals however are complicated objects as they affect humans and the environment in different ways. Various aspects of mechanisms by which chemicals affect the environment or humans are unknown. This fact prevents the application of mathematical models, like the exposure models EUSES or E4CHEM (see for example [1]). An additional difficulty is the large number of chemicals in commerce which is increasing (see for example Basak et al. [2]). Thus a theoretical concept is urgently needed to rank if several quantities, characterizing the chemical are to be considered. For example several biological endpoints are to be simultaneously taken into account. The theoretical framework to do this comes from partial order theory which is a subdiscipline of Discrete Mathematics and combines combinatorial and graph theoretical tools. Although partial order theory is explained in the chemical literature [3] we will (1) introduce the methodological concept, and discuss (2) some applications on chemicals [4], and finally (3) show an actual example from water management in Berlin [5].

References:

[1] Heidorn,C.J.A., B.G.Hansen, B.Sokull-Kluettgen, and G.Vollmer, 1997. EUSES - a Practical Tool in Risk Assessment. In: Alef,K., J.Brandt, H.Fiedler, W.Hauthal, O.Hutzinger, D.Mackay, M.Matthies, K.Morgan, L.Newland, H.Robitaille, M.Schlummer, G.Schüürmann, and K.Voigt (eds), **ECO-INFORMA'97** Information and Communication in Environmental and Health Issues, ECO-INFORMA Press, Bayreuth, pp. 143-146; Brüggemann,R., U.Drescher-Kaden, and B.Münzer, 1996. E4CHEM, a simulation program for the fate of chemicals in the environment, GSF-Forschungszentrum für Umwelt und Gesundheit, GSF-Bericht 2/96, Oberschleissheim, p. 229; Brüggemann,R. and U.Drescher-Kaden, 2003. Einführung in die modellgestützte Bewertung von Umweltchemikalien - Datenabschätzung, Ausbreitung, Verhalten, Wirkung und Bewertung, Springer-Verlag, Berlin, p. 519.

[2] Basak,S.C., K.Balasubramanian, B.D.Gute, D.Mills, A.Gorczynska, and S.Roszak, 2003. Prediction of Cellular Toxicity of Halocarbons from computed chemodescriptors: A Hierarchical QSAR Approach. J. Chem. Inf. Comp. Sc. 43: 1103-1109.

[3] Brüggemann R, Halfon E, Welzl G, Voigt K, Steinberg C (2001). J.Chem.Inf. Comp. (J. Chem. Inf. Model.) 41: 918-925.

[4] Brüggemann,R., S.Pudenz, L.Carlsen, P.B.Soerensen, M.Thomsen, and R.K.Mishra, 2001. The Use of Hasse Diagrams as a Potential Approach for Inverse QSAR. SAR and QSAR Environ. Res. 11: 473-487; Carlsen,L., P.B.Soerensen, and M.Thomsen, 2001. Partial Order Ranking - based QSAR's: estimation of solubilities and octanol-water partitioning. Chemosphere 43: 295-302; Carlsen,L., D.B.Lerche, and P.B.Soerensen, 2002.

Improving the Predicting Power of Partial Order Based QSARs through Linear Extensions. J. Chem. Inf. Comp. Sc. 42: 806-811.

[5] Simon,U., R.Brüggemann, and S.Pudenz, 2004. Aspects of decision support in water management - example Berlin and Potsdam (Germany) I - spatially differentiated evaluation. Wat. Res. 38: 1809-1816.

[O38] Physical-chemical properties of PCDD/PCDFs and phthalate esters: A QSPR study

Melek Türker Saçan, Mustafa Özkul, Boğaziçi University, Institute of Environmental Sciences, 34342, Bebek, İstanbul, Turkey; and Safiye Sağ Erdem, Marmara University, Faculty of Arts and Science, Department of Chemistry, Göztepe, İstanbul, Turkey.

PCDDs and PCDFs are compounds with similar chemical properties and they are commonly called as Dioxins. Since the 1930s, there has been a steady increase in environmental levels of dioxins coinciding with the large-scale production and use of chlorinated chemicals. Recent studies have suggested that PCDD/PCDFs can be formed under certain environmental conditions, e.g. composting from the action of microorganisms on chlorinated phenolic compounds. Similarly, PCDD/PCDFs have been reported to be formed during photolysis of highly chlorinated phenol. Due to their persistence and bioaccumulative potential, PCDD/Fs tend to magnify in the food chain and cause adverse effects including neuro, reproductional and immuno toxicities on human and wildlife situated on the top of the food chain.

The physicochemical properties such as octanol/water partition coefficient (K_{ow}), solubility (S), and Henry's law constant (H) are of critical importance for evaluating PCDD/Fs fate and potential exposure to environment and, consequently, for the whole process of environmental risk assessment. In this study a quantitative structure-property relationship model, based on the characteristic root index (CRI) for the prediction of solubility (S), Henry's law constant (H), and octanol/water partition coefficient (K_{ow}) for 32 nonionic organic chemicals is presented. The CRI has been demonstrated correlating with many physicochemical properties including octanol/water partition coefficient (K_{ow}), Henry's law constant (H), solubility (S), soil sorption coefficient (K_{oc}) and Bioconcentration factor (BCF) of other halogenated compounds like PCBs. The CRI was calculated through the use of a series of computer programs including SCIENTIFIC WORKPLACE 3.0. The experimental data of the S, H, and K_{ow} were taken from EPIWIN software. The following three QSPRs were found for S, H, and K_{ow} of PCDD/PCDFs.

$\log S = -0.957(\pm 0.394) - 1.724(\pm 0.085)\ CRI$
$R = 0.978 \quad S.E. = 0.466 \quad F = 411.146 \quad n = 21$

$\log K_{ow} = 2.047(\pm 0.272) + 0.990(\pm 0.059)\ CRI$
$R = 0.960 \quad S.E. = 0.333 \quad F = 285.446 \quad n = 26$

$\log H = -2.576(\pm 0.279) - 0.498(\pm 0.063)\ CRI$
$R = 0.899 \quad S.E. = 0.316 \quad F = 63.387 \quad n = 17$

The CRI was also applied to the prediction of the same properties of phthalate esters which are a key additive in many plastics and hence are an important constituent of many common commercial products and interfere with the reproductive and behavioral health in humans and wildlife, through disturbance of the endocrine systems. The CRI-based models also had high correlations for phthalate esters which were not given in the abstract, but the correlation coefficients decreased significantly in the combined data set of phthalates and PCDD/PCDFs. Therefore, additional semi-empirical descriptors like the energy values of the highest occupied molecular orbital (E_{HOMO}), the energy values of the lowest unoccupied molecular orbital (E_{LUMO}) were used to have good correlations with the used descriptors and the selected properties of the combined data set.

[O39] Prediction of biologic partition coefficients and binding affinities using SAR models

Moiz M. Mumtaz, Hisham A. El-Masri, Computational Toxicology Laboratory, Division of Toxicology, Agency for Toxic Substances and Disease Registry, Atlanta, GA 30333, USA; Douglas M. Hawkins, School of Statistics, University of Minnesota, Minneapolis, MN 55455, USA; Denise Mills, Subash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811, USA.

For some contaminants, toxicological data are not sufficient to evaluate toxic effects and conduct health risk assessments. In such cases, until appropriate data become available, ATSDR and other federal agencies often recommend the use of surrogate values obtained from computational tools such as structure activity relationship (SAR) techniques and physiologically based pharmacokinetic (PBPK) modeling. In an ongoing effort to develop alternative toxicity assessment methods, we have applied SAR to compute: 1) tissue:air partition coefficients, including fat:air, liver:air, and muscle:air, for a group of 46 low molecular weight volatile organic compounds (VOCs); 2) blood:air partition coefficients for a set of 39 VOCs; and 3) aryl hydrocarbon (*Ah*) receptor binding affinity for a set of 34 dibenzofurans using SAR calculations based on theoretical molecular descriptors. These descriptors consisted of four classes based on increased level of complexity and computational demand such as topostructural (TS), topochemical (TC), geometrical (3D) and quantum chemical (QC). Experimental vs. predicted blood:air partition coefficients utilizing only the TC descriptors for some common contaminants were as follows: benzene, 2.88 vs. 2.50; chloroform, 3.04 vs. 3.32; n-hexane, 0.83 vs. 1.28; and toluene, 2.89 vs. 3.38. The results indicate that the cross-validated variance coefficients (R2cv) values when using TS, TC, and 3D descriptors are -0.05, 0.94, and 0.07, respectively. The R2cv for TS+TC+3D combined is 0.94 which shows no improvement over the TC descriptors alone. In conclusion, (1) structure based models using simple descriptors adequately predict toxicological characteristics of some environmental contaminants, and (2) development of extensive or complex descriptors and models might not enhance this capability.

[O40] Predictive transformations between the graphical "databases" of S. A. Shchukarev

Ray Hefferlin, Southern Adventist University, Collegedale, Tennessee 37315, USA, hefferln@southern.edu

Over the course of many decades, scientists of the Chemical faculty of Leningrad State University assembled a huge number of graphical data displays which serve as visual databases for data of molecular properties. The ultimate objective was to formulate mathematical transformations that would make it possible to go from one such database, for instance showing the enthalpies of formation ΔH_f of titanium oxides under normal atmospheric conditions formed from free atoms, to another, for example showing ΔH_a of gaseous fluorides of nitrogen formed from free atoms.[1] Such transformations would allow prediction of additional data. This report proposes a very simple transformation which may serve to fulfill the objective.

Table I shows data from Fig. 155 Ref. [2], i.e., enthalpies of formation of the oxides of Mg and Ba (STP) formed from free atoms. Table I shows species found on the two graphs in the figure; M stands for Mg or Ba; brackets indicate that the species are in the solid phase. The second column shows the number of oxygen atoms per metal atom (the abscissae of the graphs). The third and fourth column vectors give ΔH_f (in kcal/mol). The fifth column gives the ratios for Ba oxides divided by Mg oxides. The average of the ratios is 1.35 with $\sigma = 0.31$. This average makes it possible to go from Mg oxides to Ba oxides.

Species	Oxides	ΔH_f (M=Mg)	ΔH_f (M=Ba)	Ratio
M	0	0	0	1.00
MO	1	-80.3	-116.26	1.45
MO_2	2	-162.14	-229.51	1.42
$(1/2)M_2O$	0.5	-59	-109	1.85
$(1/2)M_2O_2$	1	-113	-177	1.57
[MO]	1	-236.35	-244.64	1.04
$[MO_2]$	2	-303.08	-317.12	1.05
$[MO_4]$	4	-410.66	-437.91	1.07

Fig. 167 of Ref. [2] shows data for the same property of oxides of Li, Na, K, and Rb. Some species represented there are not present in Table I, and visa versa, so comparison of these oxides with those in Table I requires a larger table spanning both sets and some columns will have void entries. An even larger table will be necessary for including other graphs, e.g., the oxides of N, P, As, and Bi as shown in Fig. 157 of Ref. [2].

References:
[1] V. A. Latysheva and R. A. Hefferlin, "Periodic Systems of Molecules as Elements of Shchukarev's "Supermatrix", i.e. the Chemical Element Periodic System," *J. Chem. Inf. Comput. Sci.* vol. **44**, 1202-1209 (2004).
[2] S. A. Shchukarev, Neorganicheskaya Khimia, Vysshaya shkola: Moscow, 1970, vol. 1, 1970.

[O41] Proteomics and bioinformatics to design countermeasures against infectious threat agents

Akbar Khan, Defense Threat Reduction Agency-CBM, 8725 John J. Kingman Road, MS 6201, Fort Belvoir, VA 22060-6201; Alex Walz, Guy Patra, Tim Alefantis, Cesar Mujer, Joe Connolly, and Vito G. DelVecchio, Vital Probes, Inc., 1300 Old Plank Rd, Mayfield, Pa 19433, USA.

Proteomics has been used to define the secretome, membrane composition, and exosporium of *Bacillus anthracis* in comparison to other members of the *B. cereus* group. Proteomic approaches, such as 2-D gel electrophoresis followed by MALDI-TOF and LC/MS-MS, have generated valuable data on genomic and protein biomarkers unique to this important biological warfare agent (BWA). The protein composition of these cellular components has additional importance because they are responsible for most of the offensive mechanisms employed during the infectious process. In addition to the *B. anthracis* toxins such as the edema factor, lethal factor, and protective antigen, other putative offensive proteins that may interact with host cell receptins have been identified *via* probing of gel blots. Thus, such biomarkers will help determine the total virulent capability of a suspected *B. anthracis* isolate by antibody- and nucleic acid-based diagnostic and detection assays. Additionally discovered biomarkers can be used to generate more efficient *B. anthracis* simulants and novel therapeutics. Identification of immunogenic proteins will enable a more directed approach for the selection of protein candidates for vaccine development.

[O42] Protein ontology project

Amandeep S. Sidhu, Tharam S. Dillon, Faculty of Information Technology, University of Technology Sydney, Australia; Baldev S. Sidhu, Punjab State Education Department, India; and Elizabeth Chang, School of Information Systems, Curtin University of Technology, Perth, Australia.

Prediction of protein folding pathway may be evaluated by predicting sub-segments or substructures of proteins. If computational model has right underlying assumptions about what comes first in the pathway, and what comes next, and so on, then blind predictions such as those done as a part of protein structure assessment may validate that model. For correctly defining assumptions and completely understanding processes of Protein Synthesis usually both data and its biological context determines the complete meaning (or semantics) of the protein structure. We define a protein ontology model that describes the concepts of interest in protein complex mechanisms and the protein data source characteristics are

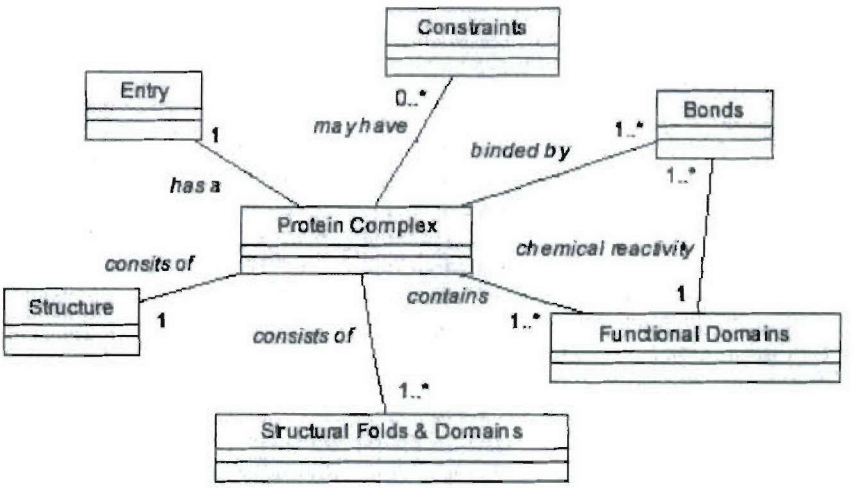


Figure 1: Protein Model

mapped to these concepts. The Protein Ontology Project seeks to provide a set of structured vocabularies for protein domains that can be used to describe cellular products in any organism. The work describes: (1) Protein Sequence and Structure Information, (2) Protein Folding Process, (3) Cellular Functions of Proteins and (4) Molecular Bindings internal and external to Proteins and (5) External Factors affecting Final Protein Conformation (Figure 1). Ontology & Knowledge Base approaches similar to the proposed approach like Gene Ontology [1, 2] and RiboWEB [4, 5] exist for Genes and RNA. The creation of a Protein Ontology [6–8] provides a comprehensive understanding of Protein Complex Mechanisms and will completely map the understanding of Central Dogma. Protein Ontology will facilitate computational processing data, and develop standardized structured data representation model formats for proteomics data. It will make it possible to study relationships among proteins, protein folding, behaviour of protein under various environments, and most importantly cellular function of protein.

References

[1] Harris, M. A., J. Clark, et al. (2004). "The Gene Ontology (GO) database and informatics resource." Nucleic Acids Research 32(Database issue): 258-261.
[2] Yeh, I., P. D. Karp, et al. (2003). "Knowledge acquisition, consistency checking and concurrency control for Gene Ontology (GO)." Bioinformatics 19(2): 241-248.
[3] Abernethy, N. F., J. J. Wu, et al. (1999). "Sophia: A Flexible, Web-Based Knowledge Server." IEEE Intelligent Systems (JULY/AUGUST 1999).
[4] Altman, R. B., M. Bada, et al. (1999). "RiboWEB: An Ontology-Based System for Collaborative Molecular Biology." IEEE Intelligent Systems (September/October 1999): 68-76.
[5] Bada, M. A. and R. B. Altman (1999). Computational Modeling of Structured Experimental Data. Stanford, CA, Stanford Medical Informatics SMI-1999-0764.
[6] Sidhu, A. S., T. S. Dillon, et al. (2004). Protein Knowledge Base: Making of Protein Ontology. HUPO 3rd World Congress 2004, Beijing, China.
[7] Sidhu, A. S., T. S. Dillon, et al. (2004). Making of Protein Ontology. The Second Australian Health & Medical Research (AH&MR) Congress, 21-26th November 2004.
[8] Sidhu, A. S., T. S. Dillon, et al. (2005). The Protein Ontology Project: Structured Vocabularies for Proteins. Data Mining 2005, Greece, Wessex Institute of Technology (WIT) Press.

[O43] QSAR study on 1,4–dihydropyridine-based calcium channel antagonist: Dominating role of molar refractivity

Mamta Thakur, Abhilash Thakur, Softvision College of Biotechnology and Science, Indore India 452001; and Padmakar Khadikar, Research Division, Laxmi Fumigation and Pest control, Pvt. Ltd, 3, Khatipura, Indore 452007, India.

Abstract QSAR study on 1,4- dihydropyridine-based calcium channel antagonists has been carried out using heterogeneous set of molecular descriptors, which includes some distance-based topological indices and Adhoc physicochemical descriptors. The results have shown that molecular refractivity plays a dominating role in this regard. This single parameters in combination with indicator parameters yielded excellent model. The results are critically discussed with a variety of statistical parameters.

[O44] QSAR study on inhibition of *E. coli* by sulfonamides: A case study with no role of positive hydrophobic term

Abhilash Thakur, Mamta Thakur, Department of Chemistry, Softvision College of Biotechnology and Science, Indore India 452001; and Padmakar V. Khadikar, Research Division, Laxmi Fumigation and Pest control, Pvt. Ltd, 3, Khatipura, Indore 452007, India.

The paper describes QSAR study on inhibition of E. Coli by sulfonamides also containing distance based topological indices. The sulfonamide used consists o 39 derivatives with substitution at 2-, 3- and 4- positions as well as containing some disubstituted derivatives. Application of multiple linear regression analysis indicated that combination of distance based topological indices with adhoc molecular descriptors and indicator parameters yielded statistically significant model for modeling inhibitory activity (log1/C) of E. Coli by the sulfonamide used.

Predictive potential of the model is as for trained by cross-validation parameters as well as by using variety of other statistical parameters.

The results have indicated that there is no role of a positive hydrophobic term in the inhibition process, indicating, that the binding of the sulfonamides to the synthetase dose not depend on hydrophobic interactions.

[O45] Quantitative structure-activity relationship (QSAR) modeling of juvenile hormone activity of alkyl (2*E*,4*E*),3,7,11-trimethyl-2,4-dodecadienoates

Subhash C. Basak, Ramanathan Natarajan, Denise Mills, and Brian Gute, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811, USA.

The Hierarchical Quantitative Structure-Activity Relationship (HiQSAR) modeling approach using topostructural, topochemical and geometrical parameters was applied to bioassay data of the alkyl (2E,4E),3,7,11-trimethyl-2,4-dodecadienoates on the yellow-fever mosquito (*Aedes aegypti*), the greater wax moth (*Galleria mellonella*), the yellow mealworm (*Tenebrio molitor*), the house fly (*Musca domestica*), the pea aphid (*Acyrthosiphon pisum*), and the tobacco budworm (*Heliothis virescens*). Comparative regression analyses were performed using ridge regression (RR), partial least square analysis (PLS), and principal component regression (PCR). The class of independent parameters present in the best QSAR model for each of the six insects reveals similarity or dissimilarity of the mechanism of action of the title compounds. The results of the statistical analyses and the HiQSAR modeling will be presented.

[O46] Quantitative structure-activity relationship studies on benzodiazepine hydroxamic acid-based inhibitors of matrix metalloproteinases and tumor necrosis factor-α converting enzyme

S. Kumaran, Department of Pharmacy; and S. P. Gupta, Department of Chemistry, Birla Institute of Technology and Science, Pilani-333031, India.

A quantitative structure-activity relationship (QSAR) study has been made on the inhibitions of some matrix metalloproteinases (MMPs) and tumor necrosis factor-α converting enzyme (TACE) by benzodiazepine hydroxamic acid inhibitors. In our study, we found out that there exit an inverse parabolic relationship between hydrophobicity of the molecule (logP) and inhibition activities of the MMP-1 and TACE, suggesting that inhibition mechanism of these enzymes could be allosteric. The appearance of E-state indices of nitrogen atom (S_N) attached to sulfonyl group in QSAR equations for the MMP-9 and MMP-13 as well as TACE reveals that the mode of interaction of these compounds with these two MMPs and TACE might involve some electronic interactions, where nitrogen play important roles.

[O47] Quantitative structure activity relationships: Model fitting and leave-out-clusters cross-validation

Douglas M. Hawkins, Jessica J. Kraker, University of Minnesota, School of Statistics, 313 Ford Hall, 224 Church Street S.E., Minneapolis, Minnesota, 55455, USA.

QSAR model fitting has two goals. One is to gain a fundamental understanding of the underlying causes of activity. The other is a predictive one: use a collection of compounds with a measured response (chemical or biological activity, or a physical property) and a typically large number of predictors (for example structural or quantum descriptors), to find a function relating the response to these predictors. While both linear and nonlinear models have been used to fit these relationships, surprisingly, linear regression methods such as ridge regression and partial least squares can often predict as well as more flexible models, such as *k* nearest neighbors and neural nets, without overfitting the data.

Available QSAR data often present some difficulties in diagnostics, since p can be larger than n, often drastically so. A widely-used method for model validation is the holdout method. This involves selecting a *holdout* (or *validation*) sample from the entire dataset, fitting the model to the *calibration* data (the remainder of the dataset), and using the holdout sample to assess the fit of the model. While it is intuitively appealing to validate the model using data not involved in the model fitting, it is also certainly arguable that data should not be "wasted", particularly when n is small. This concern is addressed with the use of *leave-one-out (LOO)* cross-validation and n-fold cross-validation; if applied at the proper point in the analysis, these methods use independent validation samples (just as with the holdout method) to assess the fit of the model and have the virtue of using all compounds, both in the fitting and validation steps.

However, conventional cross-validation methods can fail in the situation (seen mainly in drug discovery libraries) that there are clusters of compounds with almost identical structure and almost identical activity. In this setting, both cross-validation methods can be overly optimistic about the quality of the fit.

We present a new validation method in which the omitted groups are neither individuals (as in LOO) or random subsets (as in n-fold), but rather are clusters of compounds with similar structures. We illustrate linear modeling and the proposed "leave-out-clusters" (LOC) approach with a sample QSAR problem.

[O48] Role of hetroatoms in hetrocyclic drug actions: A quantitative aspect

S.P. Gupta, Department of Chemistry, Birla Institute of Technology and Science, Pilani-333031, India

Hetroatoms have significantly different electronic properties than the carbon atoms. They have higher electronegativity, possess lone pair of electron(s), and are capable of forming the hydrogen bonds. Because of these properties, they often play very crucial roles in the binding of drugs with the receptors and thus affect significantly the drug activity. The present article describes through quantitative structure-activity relationships (QSARs) the effect of some important heteroatoms in the activity of drugs like CNS agents, cardiovascular drugs, anticancer agents, anti-HIV drugs, etc. Through QSAR, a quantitative measure of the effect of any heteroatom could be well described.

[O49] Self organizing map and counterpropagation neural networks in structure-property modelling: Examples from environmental science and drug design

Marjan G. Vracko, National institute of chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia.

Self organizing map technique is often used to analyse the data in multi-dimensional space. Basis of this technique is a non-linear projection from multi-dimensional space onto two-dimensional grid of neurons (map). The projection, which is topology preserving but not metric preserving, is achieved via non-linear algorithm known as training. The fundamental property of the trained network is that the similar objects are located close to each other. Counterpropagation neural network is a generalization of self organizing map. Additionally, it takes into account the property (output) values. In the presentation the architecture and the learning strategy of both artificial neural network techniques are discussed (1,2).

Self organizing map and counterpropagation neural network are suitable techniques in molecular structure-property relationship studies (3). They can be applied to build the predictive models, for classification, clustering, determination of outliers, or selection of most relevant descriptors (4). In the presentation we discuss different strategies of application of both methods in structure-property relationship studies. We present several examples where different data sets and different properties, which are interesting in environmental science (aquatic toxicity, mutagenic potency, log P) and in drug design (partitioning coefficients) have been studied. In addition, we discuss how to use the both techniques in analysis of 2-D electrophoreses (proteome) maps.

References:

[1] M. Vračko, M. Novič, J. Zupan. Study of structure-toxicity relationship by a counterpropagation neural network. *Anal. chim. acta*, 1999, 384, 319-332.

[2] M. Vračko, D. Mills, S. C. Basak. Structure-mutagenicity modelling using counter propagation neural networks. *Environ Toxicol Pharmacol.*, 2004, 16, 25-36.
[3] M. Novič, M. Vračko. Artificial neural networks in molecular structures - property studies. In: Leardi, R. (ur.). *Nature-inspired methods in chemometrics : genetic algorithms and artificial neural networks*, (Data handling in science and technology, 23). 1st ed. Amsterdam [etc.]: Elsevier, 2003, pp 231-256 (Chapter 8).
[4] A. Jezierska, M. Vracko, S. C. Basak. Counter-propagation artificial neural network as a tool for the independent variable selection: structure-mutagenicity study on aromatic amines. *Molecular Diversity*, in press.
[5] M. Vracko, S. C. Basak. Similarity study of proteomic maps. Chemometr. Intell. Lab. Syst., 2004, 70, 33-38.

[O50] Software on pair wise sequence alignment: Drug design approach towards *SARS*

Pranav Sharad, Sumeet Jaipuriar; IIIrd Year, B.Tech. Bioinformatics, Dept. of Biotechnology, Vellore Institute of Technology, Vellore-632014 (Tamil Nadu), pranavsharad@yahoo.co.in, sumu_jaipuriar@yahoo.com

Paiwise alignment is the method of aligning two proteins to derive various properties of the sequences such as tandem repeats, conserved domains, phylogeny etc. This software performs pairwise alignment of protein sequences and sequences that code for proteins e.g., DNA. It performs local as well as global alignment using BLOSUM-50 and PAM-250 scoring matrices as per user's choice. The user also views the alignment with gaps and the score of the alignment. It also gives the percentage alignment of the sequences which helps in determining the structural, functional and evolutionary relatedness of the sequences. The software uses Needleman-Wunsch algorithm for global alignment and Smith-Waterman algorithm for local alignment to obtain the dynamic programming matrix. The software is coded in c++ and is compiled under windows.

As a case study for the above mentioned software, we analyze the sequential and thereby structural and functional aspects of the mechanism of action of *sars* virus. The severe acute respiratory syndrome-associated corona virus (SARS-CoV) is the largest positive strand *RNA* virus(100 nm). In the first phase of our study we could arrive at following conclusions:

1.) SARS-CoV is not closely related to other corona viruses.
2.) Still SARS-CoV shares a similar genome organization with other corona viruses.

Entry of corona viruses into target cells is mediated by the viral *S (spike)* protein (150k) .This region gives the virus its antigenic properties and contains the binding site for the cell surface receptor. A *mettalopeptidase* named *angiotensin converting enzyme 2 (ACE-2)* protein is identified as functional receptor for the SARS-CoV. *ACE-2* is present in humans in the lung alveolar epithelial cells and enterocytes of the small intestines which provide possible routes of entry for the SARS-CoV.

In the second phase of our study, with the help of the software we are able to identify the binding sites of the above interaction, on the basis of our results we propose an array of approaches towards drug design for this disease:

1.) inducing mutation into the viral *S-protein* binding site, so that there are conformational changes in the protein and thus resulting in no interaction between the *S-protein* and *ACE-2 receptor protein.*
2.) Addition of a compound structurally similar to the *S-protein,* for ex. the S-protein of *Drosophila melanogaster* introduced into the host, initiates *competitive inhibition, thereby* resulting in no interaction between the host epithelial cells and the virus.

[O51] Three dimensional QSAR based on k-nearest neighbor method

Subhash Ajmani, Sudhir A. Kulkarni, VLife Sciences Technologies Pvt. Ltd. 1 Akshay, 50 Anand Park, Aundh, Pune 411 007, India

A new three dimensional QSAR approach that adopts a k-nearest neighbor principle for generating relationship of molecular fields with the activity (kNN-MFA) is proposed. The kNN-MFA method employs calculated properties (steric, electrostatic, hydrophobic, etc.) at the grid points surrounding aligned molecules as descriptors of chemical

structures. It predicts the biological activity of any compound as the average activity of k most similar compounds. The variable selection is performed using simulated annealing and genetic algorithm to span entire descriptor space. The method is tested for three datasets and compared with CoMFA method in terms of its predictive ability. The resulting model suggest points that are important for explaining activity variation and the interpretation of model is suggested for designing of new compounds with improved activities.

Analysis of results for three test datasets reveal that kNN-MFA method has comparable or better predictive ability compared to CoMFA. However, CoMFA method has simple and intuitive interpretation of models compared to kNN-MFA.

[O52] Topological aspects of aromaticity in inorganic compounds

R. Bruce King, Department of Chemistry, University of Georgia, Athens, Georgia 30602, USA.

Methods derived from graph theory and topology can be used to demonstrate the analogy between the delocalization in the three-dimensional deltahedral boranes $B_n H_n^{2-}$ and carboranes $B_{n-2}C_2H_n$ ($6 \leq n \leq 12$) and that in the two-dimensional planar polygonal hydrocarbons such as cyclopentadienide, benzene, and tropylium. In addition, the concept of Kekulé structures for planar polygonal hydrocarbons can be extended to delocalized three-dimensional polyhedral boranes by using three-center B–B–B bonds instead of carbon-carbon double bonds.

Other interesting types of aromaticity are also found in inorganic compounds. Thus Möbius aromaticity, in which the π-electron network is twisted as in a Möbius strip, can be postulated to occur in the five-center four-electron bond in the equatorial pentagon of pentagonal bipyramidal $RhBi_7(\mu\text{-}Br)_8$ This delocalized bonding in $RhBi_7(\mu\text{-}Br)_8$ can be contrasted with the localized bonding in the octahedral ions $[RhBi_6(\mu\text{-}X)_{12}]^{3-}$ (X = Br, I). Another type of aromaticity is the σ-aromaticity found in cyclopropanes. σ-Aromaticity can also account for the stability of triangular metal carbonyl such as $M_3(CO)_{12}$ (M = Fe, Ru, Os) relative to square metal carbonyls such as $Os_4(CO)_{16}$.

[O53] On topological bond orders of non alternant molecular graphs: A computational approach

S.K. Adhikari, Pailan College of Management & Technology, Amgachia Road, Pailan, 24-Parganas(S), India; and D.K. Sinha, Department of Applied Mathematics, 92 A.P.C Road, Kolkata-700009, India.

A general method for construction as stated below of duplex of a graph involving odd cycles has been developed.This method is utilised to compute the topological bond orders of complex molecules by constructing the sub graphs and considering the conditionality of the graph.

Reference:
Dulal C. Mukherjee, S.K.Adhikari & A.K.Mukherjee, A new method for conservation of non-alternant graphs and calculation of topological bond order. *J. Indian Chem. Soc.* (2vol78,Oct-Dec, pgs 597-599)

[O54] Toxicity QSARs for halogenated aliphatics and melting-point QSPRs for ionic liquids

Steven Trohalaki, The Anteon Corporation, 5100 Springfield Pike, Dayton, OH 45431-1231, USA; Ruth Pachter, Air Force Research Laboratory, Materials & Manufacturing Directorate Wright-Patterson Air Force Base, OH 45433-7702, USA; Kevin T. Geiss and John M. Frazier, Air Force Research Laboratory, Human Effectiveness Directorate, Wright-Patterson Air Force Base, OH 45433, USA.

In order to improve QSARs (Quantitative Structure-Activity Relationships) for halogenated aliphatics (HA) and to better understand the biophysical mechanism of toxic response to these ubiquitous chemicals, we employed

improved quantum-mechanical descriptors to account for HA electrophilicity. We demonstrated that, unlike the lowest unoccupied molecular orbital energy, E_{LUMO}, which was previously used as a descriptor, the electron affinity can be systematically improved by application of higher levels of theory. However, we also showed that employing the reciprocal of E_{LUMO}, which, as explained by frontier molecular orbital (FMO) theory, is more consistent with the biophysical mechanism, improves the correlations with *in vitro* toxicity data of 52 HAs. In addition, FMO theory explains the surprisingly good correlation we observed between the LUMO energies of HA anions with the same *in vitro* toxicity data.

In another study, we used a composite toxicity metric calculated from a set of five *in vitro* hepatotoxicity endpoints determined for 20 HAs. We found that QSARs derived using quantum descriptors calculated from the neutral HA species are statistically similar to QSARs calculated from HA metabolites. In most cases, QSARs derived using descriptors calculated from both neutral HAs and metabolites are statistically superior to those derived using either neutral-HA descriptors or metabolite descriptors. However, in order to properly utilize metabolite descriptors, multiple QSARs, each of which utilizes a set of HAs that form unique metabolites, must be derived and toxicity values calculated therefrom must be averaged. These average toxicity values agree better with experiment than those calculated from the neutral-HA QSARs.

Ionic liquids have recently found application as 'green' alternatives to industrial solvents because of their high solubility and extremely low vapor pressure. However, even basic physical properties, such as density and melting point, are known only for a relatively small number of these materials. Derivation of melting point QSPRs (Quantitative Structure-Property Relationships) for ionic liquids would therefore greatly aid in the molecular design of ionic liquids for specific applications. A series of substituted 4-amino-1,2,4-triazolium bromide, nitrate, and nitroacetamide salts were recently synthesized and their melting points measured. Quantum chemical methods (RHF/6-31G**) were employed to optimize the molecular geometries of the cations of the salts. QSPRs for the melting points were then derived from quantum chemical, thermodynamic, and electrostatic descriptors. Excellent agreement with the experimental data was found. The resulting QSPRs can be interpreted in terms of the melting behavior of ionic liquids.

[O55] Trees (dendrograms and consensus trees) and their topological information

Guillermo Restrepo, Eugenio J. Llanos, José L. Villaveces, Laboratorio de Química Teórica, Universidad de Pamplona, Ciudad Universitaria, Pamplona, Colombia; Observatorio Colombiano de Ciencia y Tecnología, Carrera 7 No. 39 – 08, Bogotá, Colombia.

One of the most versatile technique to study similarity relationships in chemistry is cluster analysis. This methodology starts with the definition of a set of work called Q, which contain all the entities of chemical interest. After, every element in Q is defined by means of its features or properties. This definition allows representing every element as a point in a mathematical space. Next step is the calculus of the similarity relationships among elements in Q, which is made through a similarity function and a grouping methodology. This procedure shows, at the end, groups of elements with common features in a bidimensional graph called dendrogram. However, when we combined several similarity functions and grouping methodologies it is possible to calculate a consensus tree that shows in only one tree the information of several of them. Normally, the application of cluster analysis finishes with the generation of a tree (dendrogram or consensus tree) and the interpretation of the clusters or "branches" shown by the tree. This work is made by the analyst to whom some clusters are relevant and others not so much. To avoid this arbitrariness in the final process of cluster analysis, we developed a mathematical method applying topology to the final tree. This method has been called chemotopological one.

Chemotopological approach starts cutting branches on the tree by means of a mathematical procedure that takes advantage of the selection of an integer $1 \leq k \leq \#Q$, where $\#Q$ means the cardinal or number of elements in Q. Once we cut branches on the tree what we have are neighborhoods for the elements in Q, it means, similarity relationships among elements of Q or similarity neighborhoods to these elements. With these branches we build up a basis for a topology and calculate closures, boundaries, derived sets, interiors and exteriors of several subsets of Q of particular interest. By means of this chemotopological approach we found that the boundary of metals in the set Q of chemical elements are some semimetals. On the other hand we apply the same methodology to other chemical sets (compounds) and we found some results which are in agreement with the orthodox classification of this compounds and other that differ. One important fact of this chemotopological methodology is that once we have a selection of a

value of k we can do an interpretation of the tree without arbitrariness, and we can talk about the mathematical structure of the set Q. But an open question that we resolve in this work is that related to the selection of k.

In chemotopological methodology every k offers, at least, one basis for a topology and at the end we have k different interpretations of a tree. If we have a set Q such that $\#Q$ is a big number, for instance 100, then we have 100 possible interpretations of a tree. Thus, we need to develop a method to select a value of k that shows "good" results or a "good" interpretation of the tree. In this work we develop a mathematical method to find this value of k taking advantage of the mathematical theory of information. This approach starts with the set Q and the definition of its set D_Q of all possible set partitions. On the other hand, taking advantage of the partitions shown by the tree we build up a set D_k that contains all partitions produced by the selection of every k on the tree. After, we compare the partitions of D_k with the ones of D_Q and calculate the probability and the content of information of every D_k. Finally, we have that the "good" results or interpretations of a tree arise when we select the value of k that shows the maximum content of information in D_k.

[O56] The validity domain of QSARs

Romualdo Benigni, Istituto Superiore di Sanita', Viale Regina Elena 299, 00161 Rome, Italy.

In a typical QSAR model: a) the biological endpoint is meaningful internally coherent; 2) the data are good; 3) the chemicals were selected correctly to form a congeneric series supposedly with the same mechanism. The QSAR is empirical, is relative to the chemical space from which it was derived, and is not the expression of "universal" laws. In addition, a chemical series cannot be defined a priori, but only through the combination of chemical information and biological data, thus the existence of a QSAR is a *a posteriori* support to the idea that the chemicals are a series. This empirical character brings an unavoidable load of uncertainty into validation.

In practice, "validation" refers to an array of different procedures, ranging from purely formal (LOO, LMO, training and test sets, etc..., which are basically fitting checks), to external and lateral validation. The external test set molecules have to belong to the QSAR applicability domain: its selection brings also uncertainty, for lack of an "exact" theory on chemical series, and because in the same parameter space it is possible to devise many different substitutions. The new molecules can either fit perfectly the QSAR, or can have different action mechanisms: the latter may occur even with QSARs validated statistically. This because biological activity depends on interactions between molecules and biological structures, of whom we normally have a very limited knowledge.

Thus a primary component of successful QSARs is chemical and biochemical knowledge. A correct use of statistical tools is a necessary, secondary component.

[O57] Virtual high throughput screening in drug design and development

Vellarkad N. Viswanadhan, Molecular Structure and Design, 1, Amgen Center Drive, MS 29-M-B, Thousand Oaks, CA 91320, USA.

Virtual High Through-put Screening (VHTS), a critical technology in drug discovery programs impacts both the early phase of hit to lead development, and the later phase of lead to clinical candidate development. The present talk covers recently developed methods and results in functions ranging from HTS follow-up, scaffold library design for biophysical screening, new core design and *directed diversity* library design based on known active site geometry and modeling brain penetration of drugs.

LATE ADDITIONS

[O58] River water quality fuzzily described: A case study

Ashok Deshpande, Berkeley Initiative in Soft Computing (BISC)-Special Interest Group (SIG)-Environment Management Systems(EMS), India; and Dhananjay Raje, NEERI/CSIR India.

Information on the status of environmental quality is necessary to formulate sound public policies and effective implementation of environmental quality programs. One of the effective ways of communicating such information in general, and river water quality in particular, to policy makers and public at large, is with indices. Would it not be appropriate to straightway define water quality in linguistic terms with some degree of certainty? There exists uncertainty in water quality data (statistical uncertainty) and the imprecision in the expert's perception for defining the water quality parameter (cognitive uncertainty) for the defined usage. We believe that the problem of water quality classification could be approached more effectively using fuzzy logic concepts such as fuzzy set/number, degree of match, fuzzy rule-based systems, approximate reasoning, etc.

The case study relates to the fuzzy description of water quality of river Ganga, India after the of implementation of first pollution abatement measures for bathing purposes at the defined two sampling stations along the river.

[O59] Kinetic of glucose in normal and diabetic persons

V. K. Katiyar, Department of Mathematics, Indian Institute of Technology, Roorkee, Roorkee-247667, India

Body rythms have attracted the interest of many persons in biomedical sciences. These natural rythms range from very fast ones to very slow ones. Some of these processes are induced by the external environment, whereas some other are parts of internal physiology. One of the example of such processes is ADME i.e. absorption, distribution, metabolism and excretion of glucose inside the body. There are basically Hyperglycemic and Hypoglycemic persons, which are broadly related to the effect of glucose.

Our study is confined to the kinetics of glucose in the body and its tolerance in normal and diabetic persons.

[O60] Jahn-Teller distortion in the transition state observed in the deamination of aziridines by difluoroamine: More insights from theory

A. Kalaiselvan, P. Venuvanalingam, Department of Chemistry, Bharathidasan University, Tiruchirappalli-620 024, India; S. Salai Cheettu Ammal, and H. Yamataka, Department of Chemistry, Rikkyo University, Toshima-ku, Tokyo 171-8501, Japan.

cis-2,3-Dimethylaziridine (1) reacts with difluoroamine to give the corresponding alkene and nitrogen, and the reaction has been reported to be stereospecific [1]. Ab-initio and DFT modeling of this reaction reveals that this reaction pass through azamine intermediate (2), as proposed by Bumgardner [2] and not through four membered azetidine intermediate (3). Further, the calculations reveal that azamine cleaves concertedly to give the alkene with same stereochemistry in total agreement with experiment [1]. Ring opening of aziridine takes place in four steps in which a prereactive complex, spiro type TS and an azamine intermediate are formed. The concerted TS formed from 2 is found to be asynchronous and thus it lowers symmetry to become Woodward-Hoffmann allowed reaction [3]. This Jahn-Teller effect on the transition state leads to shifting of lone pair on the exocyclic nitrogen atom of azamine to one of the antibonding C-N σ bonds makes the TS asynchronous and there is no scope for C-C bond rotation. This explains the observed stereochemistry of the reaction. The high exothermicity of the reaction very well explains the release of the ring strain during the ring cleavage. B3LYP barriers are lower and HF barriers are higher when compared to MP2 and QCISD barriers but they all show a similar trend in the energetics of the reaction. It has been found that difluoroamine reacts with *trans*-2,3-dimethylaziridine to give *trans*-2,3-dimethylethylene in a similar way. The more interesting observation here is that azamine undergoes Jahn-Teller distortion to make the reaction Woodward-Hoffmann allowed and thus making the reaction stereospecific.

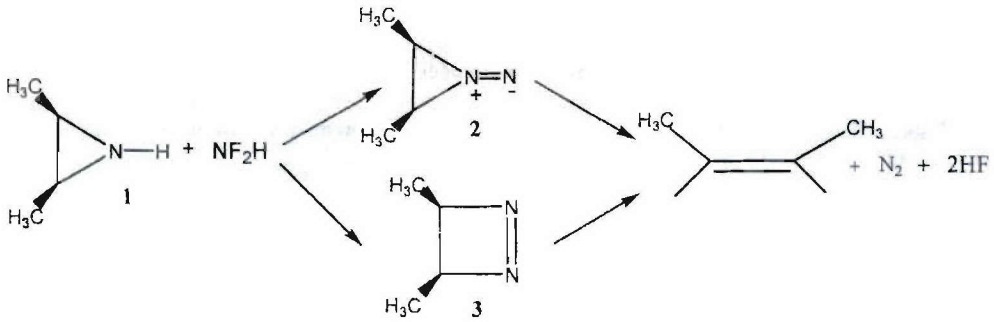


Fig. Scheme for the deamination of aziridine by difluoroamine

References:
[1] Freeman, J. P.; Graham, W. H. *J.Am.Chem.Soc.*, **1967**, *89*, 1761.
[2] Bumgardner, C. L.; Martin, K. J.; Freeman, J. P. *J.Am.Chem.Soc.*,**1963**, *85*, 97.
[3] Yamabe, S.; Minato, T. *J.Phys.Chem.A.*,**2001**, *105*, 7281.

[O61] Tailored similarity: Creation of activity specific molecular similarity spaces

Brian D. Gute, , Natural Resources Research institute, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth, MN 55811, USA; bgute@nrri.umn.edu.

Quantitative molecular similarity analysis (QMSA) methods based on structure spaces using a variety of calculated molecular descriptors and experimental properties have been used in the selection of analogs and estimation of a variety of properties pertaining to chemistry, drug discovery, and environmental toxicology. Traditionally, descriptor sets have been selected either arbitrarily, based on some intuitive notion of the descriptor's association to the property of interest, or by the use of a variety of data reduction techniques. In this study we have developed a method of 'tailoring' the QMSA method to a particular property of interest through the use of indices which are strongly correlated with the property instead of using an arbitrary set of parameters.

Using two very different databases, we have conducted comparative studies to examine tailored *vis-à-vis* two non-tailored (arbitrary) similarity spaces. The spaces are all derived from the same set of topological indices though the methods for selecting indices vary between the arbitrary and tailored approaches. Our results show that the tailored method outperforms the two non-tailored techniques in the estimation of hydrophobicity (logP, octanol/water) for a set of 213 chemicals and Ames' mutagenicity for 95 aromatic and heteroaromatic amines.

POSTERS

[P01] Ab initio study of the addition of bromine to a series of alkenes

Mohammad Shahidul Islam, Raymond A. Poirier, Department of Chemistry, Memorial University of Newfoundland. St. John's, NL A1B 3X7, Canada.

Ab initio calculations have been carried out for the addition of Br_2 to a series of alkenes and mono- and di-substituted ethenes. Although the bromonium ion intermediate has been well studied, to date, no transition state structures along the reaction path has been reported. In this study, a complete reaction path is obtained, including the TS structures. Optimized geometries of the reactants, transition states, intermediates, and products were determined at RHF/6-31G(d), MP2/6-31G(d) and B3LYP/6-31G(d) levels of theory. Once the optimized geometric parameters were obtained, frequencies were calculated using the respective basis sets of the optimized geometries. All of the electronic structure calculations were carried out using the Gaussian03. Thermodynamic properties were computed for all addition reactions, and Intrinsic Reaction Coordinate (IRC) studies were performed to characterize the transition states on the potential energy surface. From the study, it is observed that Bromine (Br_2) adds to alkenes in a two-step mechanism, involving a cyclic bromonium ion intermediate that subsequently undergoes S_N2 attack by the other bromide ion leading to the products. Two transition states were obtained with interesting potential energy surfaces. From the study, it is found that the brominium ion intermediate is very stable for all cases. However, the bromonium ion intermediate lies between two transition states that are also three-membered rings with a nearby counter bromide ion.

[P02] Accelerating QM/MM sampling with pure MM potential: finding stationary points of a peptide-water cluster in an automated way

Pradipta Bandyopadhyay, Department of Biotechnology, Indian Institute of Technology, Guwahati

Though Quantum mechanical/molecular mechanical (QM/MM) method is a major tool to study large molecules, sampling of different structures is a major bottleneck to expand its applicability. Standard molecular dynamics (MD) or Monte Carlo (MC) techniques cannot be used for QM/MM methods because of the high computational cost. One promising way is to use pure MM biasing potential to accelerate QM/MM sampling. Though this technique is rigorous, finding good biasing potential is difficult. In this work, effective fragment potential (EFP), a sophisticated classical potential, is used as the biasing potential. EFP, generated on the fly, is found to be an excellent candidate for the biasing potential. This technique is used to find stationary points of blocked aspartic acid with twelve water molecules using QM/MM potenial is an automated way. Nine stationary points are obtained from only two runs of simulation. This method can be extended to use for a large variety of systems including biomolecules.

[P03] Anti-connectivity: A challenge for QSPR studies

Matevž Pompe, University of Ljubljana, Faculty of Chemistry and Chemical Technology, Aškerčeva 5, 1000 Ljubljana, Slovenia; and Milan Randić, National Institute of Chemistry, Ljubljana, Slovenia & 3225 Kingman Rd. Ames, IA 50014, USA.

Several properties exist where presence of certain atom or functional group suppresses the signal. For instance, the presence of heteroatom in the molecule has negative effect on the response factor of FID detector in gas chromatography. Similar observation could be found in modeling biological activities and reaction rate constants. None of the classical bond additive or atom additive topological indices has the ability to describe such behavior, this also include present state of the art variable topological indices. A new procedure for the calculation of variable connectivity indices will be presented, which enables modeling of the described anti-connectivity properties.

Examples of modeling of the FID response factors and the reaction rate constants will be given. The results obtained by newly defined variable connectivity index will be compared with MLR models containing classical structural descriptors. The results will show that modified variable connectivity index is not only superior to any other single descriptor but enables easier structural interpretation.

[P04] On canonical labeling of proteins of proteomics maps

Milan Randić, National Institute of Chemistry, Ljubljana, Slovenia; Dejan Plavšić, Institute Rudjer Bošković, Zagreb, Croatia; and Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN, USA.

We propose a canonical labeling of protein spots in 2-D proteomics maps which will not only facilitate cataloging of 2-D gel maps but also to facilitate comparisons of gel maps originating from different laboratories. The canonical labels are based on labeling of vertices of an embedded graph derived from partial ordering of the protein spots of a 2-D gel that would result in adjacency matrix associated with the smallest binary number obtained by reading the matrix entries from the left to the right and from the top to the bottom.

[P05] On characterization of 2-D proteome maps using sequential neighborhood graphs

Milan Randić, Marjana Novič, Marjan Vračko and Jure Zupan, National Institute of Chemistry, Ljubljana, Slovenia.

"Sequential Neighborhood Graphs" are constructed by connecting each spot in a proteomics map with NN of its nearest neighbors of higher abundance. In this way one obtains numerical stability of the approach relative to expansion of calculation to cases with increasing number of spots – a featured that was missing in the approach based on the Neighborhood Graphs. Just as was the case with the Neighborhood Graphs we will show that one need not use more than from NN=1 to NN=6 nearest neighbors for satisfactory characterization of proteomics maps. The approach is illustrated on liver cells of rats exposed to various concentration of a particular peroxisome proliferators. The average row sum of constructed distance matrices as well as the leading eigenvalues were used as proteomics map descriptors.

[P06] Comparative 2-D and 3-D QSAR analysis of antifungal griseofulvin analogues.

Dhanshri C. Juvale, Vithal M. Kulkarni, Poona College of Pharmacy, Bharati Vidyapeeth Deemed University, Erandwane, Pune, Maharashtra-411038, India

A variety of quantitative structure-activity relationship (QSAR) methods have been used as possible computational tools to aid drug discovery process. In recent years the QSAR research has tremendously advanced and numerous applications have been reported for biomolecules.

Griseofulvin is a systemic fungicidal for treating superfacial mycoses. Classical QSAR for a series of Griseofulvin analogues was reported in 1971 by C. Hansch and E. J. Lein. They reported that the activity of griseofulvin analogues depend linearly on electronic effect and lipophilicity of substituents α to the carbonyl function. In order to compare and validate the results, we have applied 3-D QSAR CoMFA for the same series. In this method, the superimposition of molecules was carried out by atom based fit (RMS), multi fit and field fit. Along with steric and electronic fields, ClogP was also taken as additional descriptor to account for lipophilicity. Partial least square (PLS) analysis was used in conjugation with the crossvalidation option. Robustness of the CoMFA models was evaluated on test set of 4 molecules. CoMFA models exhibited considerable correlative and predictive properties.

The CoMFA models generated were compared with the results of previously performed 2-D QSAR. The CoMFA results suggest the contribution of electrostatic and steric fields. It also suggests the necessity of

lipophilicity for antifungal activity. A large negative charge favourable polyhedra at the α substituent emphasizes the necessity of electron rich substituents for activity. The yellow contours at α position represent the regions of low steric tolerance. This suggests that the compounds with smaller substituents at this position show higher activity. The 3-D QSAR result emphasizes the role of steric effect in addition to lipophilicity and electronic effect. Thus all the three properties: electronic, steric and lipophilic can be considered for the design of new analogues of griseofulvin.

[P07] Comparison of arbitrary versus tailored similarity spaces in property estimation

Brian D. Gute, Subhash C. Basak, Denise Mills, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN; and Douglas M. Hawkins, School of Statistics, University of Minnesota, Minneapolis, MN.

Quantitative molecular similarity analysis (QMSA) methods using a variety of calculated molecular descriptors and experimental properties have been used to select analogs and estimation a wide variety of properties pertaining to chemistry, drug discovery, and environmental toxicology. Traditionally, descriptor sets have been selected either arbitrarily, intuitively by an expert, or through the use of a variety of data reduction techniques. 'Tailoring' is a new approach to the QMSA method that seeks to use indices that are strongly correlated with the property of interest. Studies have been carried out on two very different databases to examine the effectiveness of tailored *vis-à-vis* arbitrary similarity spaces. The spaces are all derived from the same set of topological indices, only the selection methods vary. Results show that the tailored method outperforms the non-tailored techniques in property estimation.

[P08] Design of *in silico* chemical workbench

Om Prakash Pandey, Ambrish Roy, Ritesh Kumar & Amit Ghate, Bioinformatics Centre, University of Pune, Ganeshkhind, Pune 411 007 India, chemoinfo_msc@yahoo.com

The objective of this project is to develop *in silico* chemical workbench, allowing the computational chemist to perform Ki calculation and data analysis on the hypothetical compounds before they are synthesized.
Most of the drugs bind to their receptor using the non-covalent interactions. Hence the potential strength of interaction between drug and a reasonably matched receptor can be statistically estimated by calculating the strength of noncovalent bonds associated with each functional group, assuming that bound conformation of a drug atleast to its optimal receptor, will be relatively fixed [1].

The Ki, Inhibition Constant, prediction proceeds in two steps. In the first step, an Internet interface is used to draw the structure of compound, which is taken as input from the user and is converted to its unique SMILE [2]. SMILE string were depicted using DEPICTMATCH [2] and then make use of its SMARTS notation to calculate the number of occurrence of functional group in each chemicals.

In the second step, the intrinsic binding energy associated with known functional groups is calculated using stepwise Multiple Linear Regression and Principal Component Analysis and using a training set for which Ki is measured.

The primary application of the result would be to deduce whether a particular molecule represents a good or bad match to its receptor, thus a useful starting point in drug designing. This workbench has been tested using dopamine receptor ligands [3].

References:

[1] P.R.Andrews, D.J.Craik, and J.L.Martin. Functional Group Contributions to Drug-Receptor Interactions.J.Med.Chem. 1984,27,1648-1657.
[2] Daylight Chemical Information System Inc., http://www.daylight.com
[3] S.Shacham et al J.Proteins Str. Fun. And Bioinformatics Vol.57,2004

[P09] Evaluation of LUDI and MOE scoring functions for molecular docking

Philip Prathipati and Anil K. Saxena, Medicinal and Process Chemistry Division, Central Drug Research Institute, Lucknow-226001.

In today's world of high-throughput and in silico screening, the development of virtual screening methodologies based on the molecule's affinity for binding with a target protein is gaining importance to prioritize small molecules as NCEs for synthesis. Among several methods available for the evaluation of molecule's affinity to bind with the target protein, docking is the most effective and widely used. There are three key elements in a typical docking simulation: (1) representation of the 3D-structure of the target protein and ligand (2) conformational space search, and (3), ranking of potential solutions. Despite the major advances in computer technology, the problems arising out of molecular flexibility of the receptor and ligand in the first two steps are being addressed to a considerable extent. Still the problem associated with the ranking of potential solutions in terms of scoring functions remains one of major bottlenecks for molecular recognition and hence it remains a primary focus of current structure-based virtual screening (SVS) technology development. In view of it, three different scoring functions implemented in LUDI along the scoring routine of MOE are evaluated on several thousand protein-ligand complexes (Jacobsson dataset)[1] belonging to four different target proteins viz. estrogen receptor α(ERα), matrix metalloprotease 3 (MMP3), factor Xa (fXa), and acetylcholine esterase (AChE). The results of these studies have been analyzed and compared with the seven other scoring functions for the same sets published earlier.[1]

Reference:

[1] Jacobsson M, Liden P, Stjernschantz E, Bostrom H, Norinder U. Improving structure-based virtual screening by multivariate analysis of scoring data. *J Med Chem.* 2003, 46, 5781-9.

[P10] Hierarchical QSAR (HiQSAR) approach to predicting dermal penetration

Brian D. Gute, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811; Jim E. Riviere, Ronald E. Baynes, Center for Chemical Toxicology Research and Pharmacokinetics (CCTRP), College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27606; and Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811.

Since dermal exposure is one of the primary workplace exposures, predictive models for dermal permeability are invaluable. Such models can aid in our understanding of the chemicals that actually penetrate the stratum corneum and become bioavailable after dermal exposure. Riviere *et al* have developed a new model for skin permeability and the modeling team at NRRI has begun to use their data in developing computationally predictive models for membrane permeability. Exploratory QSAR modeling has been conducted using a series of organochlorines. Membrane permeability is in strong agreement with the lipophilicity of these compounds. Preliminary hierarchical QSAR (HiQSAR) modeling has resulted in adequate models of compound diffusion through the permeable membrane ($Q^2 = 0.81$), but further modeling is being conducted to attempt to refine the models. In addition to various topostructural and topochemical indices, a number of descriptors of molecular shape, size, and electronic character are being examined.

[P11] Highly compact 2-D graphical representation of DNA sequences

Milan Randić and Jure Zupan, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia.

In this contribution we consider a highly compact graphical representation of DNA, which allows visual inspection and numerical characterization of DNA sequences having large number of nucleic acid bases. The approach is illustrated on the DNA sequences of the first exon of human β-globin. The same graphical makes possible graphical representation of protein sequences, which have hitherto evaded 2-D visual representations.

[P12] Hormesis versus traditional dose-response relationship: Cases of algal growth response

Melek Türker Saçan, Bogaziçi University, Institute of Environmental Sciences, 34342, Bebek, Istanbul, Turkey.

Quantitative structure-activity relationships (QSARs) have been used in ecotoxicology as useful tools for hazard assessment and toxicity prediction. Most of the QSAR models are regressions referring to the dose that gives the toxic effect in 50 % of the test species (e.g., IGC_{50} (50% inhibitory growth concentration), LC_{50} (50% lethal concentration)). The traditional toxicity method (i.e., animal/algal models, number of doses, dose spacing, duration and end points measured, etc.) is designed to provide information using a limited number of doses to estimate lowest observed effective level (LOEL) and no-observed effective level (NOEL). There is no doubt that we can predict the acute environmental toxicity of the majority of chemicals reasonably well. However, it is a reality that toxicological dose-response relationships are more complex than previously recognized and that traditional evaluative extrapolation procedures are often no longer viewed as providing accurate estimates of response at low doses. The problem lies in the modeling of the compounds which show biphasic behavior, e.g., stimulatory at low concentration and inhibitory at high concentration (hormesis). The hormetic phenomenon indicates that statistically and biologically significant responses frequently occur below the NOEL. Hormesis has been observed in a wide range of responses in cells, tissues and individuals of a wide range of species. Any response that moves an organism out of its normal homeostatic range or requires expenditure of energy or mobilization of stored resources can be viewed as a negative response. For instance, the low-dose stimulatory effect of industrial effluents/chemicals to algae causes eutrophication in aquatic environment and can eliminate the desirable species (e.g. fishes). Consequently, exposure and response are not always directly related and less is not always better and may in some cases worse. In an investigation of *Dunaliella tertiolecta* (a marine microalga) toxicity data for azo dyes, pulp and paper effluents, and for several selenium compounds besides inhibitory effect at high doses, stimulatory effect was observed at low doses. Thus, the topic will be discussed on chemical classification according to their hormetic behaviors as well as their mode of action (e.g., narcotics, receptor mediated, etc.) and modeling stimulatory effects with QSARs.

[P13] *In silico* design of mineral collectors for froth flotation

Ramanathan Natarajan*, Inderjit Nirdosh, Department of Chemical Engineering, Lakehead University, Thunder Bay, Ontario, ON, Canada P7B 5E1; Ramamurthy Jayalakshmi, Department of Bio-Informatics, Thanthai Hans Roever College, Perambalur, Tamil Nadu, India 621 212.
(*Present Address: Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811)

Quantitative Structure Activity Relationship (QSAR) modeling has been extensively used in drug design and environmental toxicology. The extension of the frontiers of application of this powerful tool to industrially important processes such as mineral processing and solvent extraction was envisaged. Synthesis of new mineral collectors for the froth flotation of ores was set as the first goal, and a few QSAR models were developed for predicting their efficiency of flotation of minerals. The reasonably good predictive ability of these models indicated with certainty the amenability of the flotation efficiency of the chelating collectors to QSAR modeling. This paper explains the subsequent step taken towards the design of a new series of mineral collectors namely, arylhydroxamic acids, for the flotation of base metal ores. Synthesis, laboratory flotation, development of QSAR and validation of the models are the steps involved in the *in silico* design of these collectors. It was found that it was the same compound that gave best yield for sphalerite (ZnS) recovery for three different types of zinc ores, *viz*, copper-zinc ore, copper-lead-zinc ore, lead-zinc ore, tested. The same collector was also found to float nickel ore effectively. The computer aided or the *in silico* methodology was found to be efficient in saving considerable amount of time and money that would be spent on numerous exploratory tests, and it is worth noting that the selection was based on a scientific tool rather than on the traditional trial-and-error method.

[P14] *In silico* screening and structure based drug design for envelope protein of Flaviviruses.

Ragothaman Yennamalli, Naidu Subba Rao, Bioinformatics Centre, Jawaharlal Nehru University, New Delhi - 110067; Thorsten K. Kampmann, Bostjan Kobe, Paul R. Young, Department of Biochemistry, School of Molecular and Microbial Sciences, University of Queensland, St. Lucia, Queensland 4072, Australia.

Flaviviridae are single stranded RNA enveloped viruses. These viruses are transmitted by both ticks and mosquitoes and cause diseases like Dengue Fever and Dengue Hemorrhagic Fever, Tick borne encephalitis, Japanese Encephalitis, West Nile disease, St. Louis Encephalitis. Flaviviruses account for a high mortality after malaria. There are no known drugs or vaccines against these diseases so far. The Envelope or E protein of Flavivirus plays a critical role in infection and is reponsible for the membrane fusion between the virus and the host cell. This protein is known to undergo conformational changes during the fusion mechanism. The pre-fusion and the post-fusion structures, that have been solved, are the starting points for the Structure based Drug Design. Among all the flaviviruses, the *In silico* screeing with dengue viral E protein has been done. We are aiming to inhibit the membrane fusion of dengue virus by targeting the Envelope protein or E protein using Structure based Drug Design.

[P15] On invariants of a 2-D proteome map derived from neighborhood graphs

Milan Randić, National Institute of Chemistry, Ljubljana, Slovenia; Nella Lerš, Dejan Plavšić, Institute Rudjer Bošković, Zagreb, Croatia; and Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN, USA.

"Neighborhood Graphs" for protein spots of proteome maps were used as mathematical object for construction of proteomics map invariants. We will show that one need not used more than from N=1 to N=6 nearest neighbors for satisfactory characterization of proteomics maps. The abundance values of the control group and of liver cells of rats exposed to four peroxisome proliferators are used to construct two component vectors that describe the degree of similarity/dissimilarity between proteome map. The average row sum of constructed distance matrices were used as proteomics map descriptors.

[P16] Invariants of nearest neighbor graphs as biodescriptors of proteomics maps

Varsha Kodali, Subhash C. Basak and Milan Randić, Natural Resources Research Institute, University of Minnesota, Duluth Minnesota, USA.

We have revisited the numerical characterization of proteomics maps based on the approach in which one consider the nearest neighborhood graphs constructed by connecting each spot with NN nearest neighbors. In particular (1) we investigated the effect of drastically increasing the number of spots considered in the construction of neighborhood graphs; and (2) we investigated the effect of visibly increasing the number of the nearest neighbors considered in the construction of neighborhood graphs. Specifically, on one hand we will be comparing numerical characteristics of a selection of proteomics maps based on N = 30 points and based on N = 500 points, and on the other hand we will consider up to 12 nearest neighbors for each spot. As map invariants we will compare the average row sums as well as the leading eigenvalues of associated distance-adjacency matrices.

[P17] Mutagen/non-mutagen classification of congeneric and diverse sets of chemicals using computed molecular descriptors: A hierarchical approach

Subhash C. Basak, Denise Mills, Brian D. Gute, Christian Matson, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811, USA; Douglas M. Hawkins, School of Statistics, University of Minnesota, Minneapolis, MN 55455, USA.

Ridge linear discriminant analysis was used to classify a diverse set of 508 mutagens/ non-mutagens, as well as three structurally homogenous subsets, viz., 260 monocyclic carbocycles and heterocycles, 192 polycyclic carbocycles and heterocycles, and 124 aliphatic alkanes, alkenes, and alkynes. Software programs including POLLY, Triplet, Molconn-Z, Sybyl, and MOPAC were used to calculate a large and diverse set of theoretical molecular descriptors. Subsequently, the descriptors were divided into hierarchical classes based on level of complexity and demand on computational resources. The topostructural (TS) descriptors are at the low end of the hierarchy, encoding information strictly about molecular topology, and they can be calculated very quickly. More complex are the topochemical (TC) descriptors, which encode topological as well as chemical information, such as atom and bond type. The geometrical or 3-dimensional (3D) descriptors make up the third level of the hierarchy, followed by the quantum chemical (QC) descriptors. Most notably, calculating QC descriptors for even a modest-sized data set can be a time-consuming task. The results of this study are reported in terms of the cross-validated correct classification rates. Examining the hierarchical models, there is a significant increase in correct classification rates upon the addition of the TC to the TS descriptors. On the other hand, the addition of 3D and QC descriptors does not result in significant improvement in correct classification rate. In addition, correct classification rates for the relatively homogeneous subsets are comparable to those obtained for the entire set of 508 diverse compounds, indicating that the diverse set of theoretical descriptors is capable of representing the diversity of structural features present in the data set, whereas toxicity of smaller and structurally narrower subsets are predicted by less diverse subsets of descriptors.

[P18] NMR spectral invariants as numerical descriptors for diastereomers

Subhash C Basak and Ramanathan Natarajan, Natural Resources Research Institute, 5013 Miller Trunk Hwy, Duluth, MN 55811.

Topostructural or topochemical invariants derived for a molecule from its molecular graph (hydrogen included or suppressed) are based on the edge count or information content, could differentiate structural isomers and are incapable of differentiating geometrical isomers because 3-D orientation of the atoms in a molecule are not considered in their computation. Though the next generation indices, the geometrical indices such as 3-D Weiner index can account for molecular volume etc., diasteroisomers cannot be distinguished. Though, attempts have been made by Schulz in this line, the indices created by them could not be applied in SAR modeling. NMR, a powerful tool in the hands of chemists, can differentiate diasteroisomers of a compound because it sees the three dimensional disposition, "environment"; of the protons in a molecule that is to say it has a higher dimensional perception of a molecule than that of a chemist who tries to visualize using a molecular graph. This higher dimensional perception is used by us in converting NMR spectra into invariants. The new spectral invariant thus generated can differentiate diastereoisomers. Illustrations of the calculation and future developments of the novel NMR invariant will be presented

[P19] Novel graph theoretical representations of DNA and proteins

Milan Randić and Jure Zupan, National Institute of Chemistry, Ljubljana, Slovenia.

We reports on 3-D, 2-D and 1-D representation of DNA and proteins which has lead to the notion of 2-D "The Table of Proteins." Here to different proteins sequences, as well as to different DNA sequences one can assigned a unique (x, y) coordinates on a 2-D maps. The approach offers alternative numerical and visual

representation of DNA sequence, which are equally suitable for long as well as short sequences of DNA or proteins.

[P20] Ordering the repellency of stereoisomeric topical mosquito repellents by molecular overlay

Ramanathan Natarajan, Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811; Alexandru T. Balaban, Texas A & M University at Galveston, Dept. of Oceanography & Marine Sciences, P. O. Box 1675, Galveston, TX 77553-1675; Jerome A. Klun, and Walter F. Schmidt, USDA Agricultural Research Service, Beltsville Agricultural Research Center, Chemical Affecting Insect Behavior Laboratory, Building 007, Rm 324, 10300 Baltimore Avenue, Beltsville, MD 20705

In spite of some of its deleterious effects, DEET (*N, N* diethyl-3-methylbenzamide) is the most commonly used topical mosquito repellent for nearly five decades. Bayrepel (2-(2-hydroxyethyl)-1-piperidine carboxylic acid-1-methylpropyl ester) is another commercially used compound of this category.

There is a need for more effective and less toxic repellents. To that end, compound like1- (cyclohex-3-ene-1-ylcarbonyl)-2-methylpiperidine (220) have been synthesized and tested. Scientists have also attempted to find more active compounds through the testing of individual configurational isomers instead of the racemic mixtures.

Each Bayrepel and 220 has two chiral centers, and consequently exhibits polychiral diastereoisomerism resulting in 2 pairs of enantiomers. The degree of mosquito repellency of these stereoisomers had been quantified by the USDA. The amide group which is common in these compounds and DEET is considered as the putative pharmacophore. Assuming this is true, one has to find the arrangement of chemical structural features bound to the chiral center(s) in the repellent molecules to tailor the synthesis towards uncovering the most effective repellent compounds possible. Molecular overlay of the optimized geometries of the lowest energy conformers of the configurational isomers was investigated as a method to elucidate the role of chiral centers in 220 and Bayrepel. It was found that the presence of a chiral carbon alpha to the nitrogen with *S* configuration in the piperidine ring is essential to the 3-D arrangement of the atoms of the pharmacophore for effective repellant activity. It is noteworthy that the most active stereoisomer of 220 has the closest resemblance to DEET.

[P21] Overall path connectivity–Two non-degenerate indices for alkanes

Ramanathan Natarajan, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811; T. C. Murali, T. M. Anbazhagan, 58 Heber Road, Beema Nagar, Trichy, Tamil Nadu, India 620 001.

It was necessary for a specific purpose to create a database of all alkane-isomers from carbon number 1 to 15 with their SMILES notation and commonly used Topological Indices (TIs). A Non-degenerate index is essential to differentiate the isomers possible for a given carbon number. Two new indices, Overall Path Multiplicity index (OPM) and Mean Path Connectivity index (MPC) based on overall connectivity are proposed. The new indices are calculated according to the equations:

$$OPM = \sum {}^{i}\chi \times K_i$$

$$MPC = \sum \frac{{}^{i}\chi}{K_i}$$

where ${}^{i}\chi$ is the path connectivity index of order i and K_i is the number of paths of length i (path number). The new indices where found to be non-degenerate and all possible alkanes for a given carbon number could be enumerated up to carbon number 15 using either of them. The computer program developed uses carbon number as the input to enumerate the isomers and calculate several TIs. The new indices are not mutually related (correlation coefficient r = -0.4045) and MPC is very poorly correlated ($r < 0.2$) to several commonly

used topological indices. However, it is closely related to Balaban J index (r = 0.8501). Ordering of alkanes and QSAR modeling of properties of alkanes will be presented. Calculation of these two indices is extended to cycloalkanes, benzenes and compounds with hetero atoms.

[P22] π-electron ring partition for Clar structures

Milan Randić, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia and
3225 Kingman Rd., Ames, IA 50014; and Alexandru T. Balaban, Texas A& M at Galveston, TX, USA.

According to Clar's Aromatic Sextet model only selection of Kekulé valence structures play dominant role for representation of benzenoid hydrocarbons. We have examined those Kekulé valence structures and report on the partition of π-electrons to individual benzene rings for these Kekulé structures. The resulting distribution of π-electrons to individual rings clearly shows significant variations in local aromatic character (measured by the magnitude of partition of π-electrons) between individual benzene rings in polycyclic benzenod hydrocarbons which parallels intuitive and empirical approach of Clar.

[P23] Potential site mapping for intramolecular hydrogen bonding arrangement in different ionic forms of glyphosate: An analysis using DFT based MESP maps

M. Mohamed Naseer Ali and P. Kaliannan, Department of Physics, Bharathidasan University, Tiruchirappalli – 620024, India.

Glyphosate [N-(phosphonomethyl) glycine] is the active ingredient in the formulation of the broad spectrum herbicide Roundup. It is an effective broad spectrum herbicide exhibiting many of the characteristics desired in environmentally friendly herbicides. Over the last two decades a number of physiological, biochemical and genetic experiments have demonstrated that glyphosate acts as a herbicide by inhibiting the enzyme 5-enolpyruvoylshikimate-3-phosphate synthase (EPSPS) which is involved in the synthesis of plant growth amino acids Tyrosine, Phenylalanine and Tryptophan. To function optimally glyphosate should overlap as completely as possible with the natural substrate, phospho-enol-pyruvate (PEP), of EPSPS. However, optimum overlap in the extended form of glyphosate with PEP at the critical anionic recognition centers is somewhat precluded due to the extra methelene group present in glyphosate. In order to accommodate the configurational restrictions in PEP, the phosphonate and carboxylate functionalities in glyphosate should be bound in a more pinched form. This type of pinched form is possible only in the arrangement wherein, the intramolecular hydrogen bonding network of phosphonate and carboxyl groups with amino nitrogen exists. Thus, in this context, it is interesting to analyse the formation of potential intramolecular hydrogen bond interaction in glyphosate and its different ionic forms thus, enabling the possible prediction of potent ionic form for tight binding with its target enzyme. Hence, in the present paper efforts are made to construct Molecular Electrostatic Potential maps (MESP) for various ionic forms of glyphosate and to predict potential hydrogen bonding interaction sites using MESPs. The results of the analysis on the characteristics of MESPs, the calculated hydrogen bonding parameters and the resulting potent form of glyphosate will be presented in detail.

[P24] Prediction of binding affinity for oestrogen receptors with artificial neural networks: alpha and beta receptor selectivity study

Elena Boriani, Morena Spreafico, On leave from the Institute Mario Negri, Milano, Italy; Marjana Novič, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia.

Kohonen Artificial Neural Networks are used in a preliminary analysis to predict experimental binding affinity of a range of substances towards Oestrogen Receptor Alpha and Beta. Several compounds as oestrogenic chemicals, phytoestrogens and natural and synthetic estrogens are described through the information encoded in

molecular descriptors, used to train and test the network. In addition, a virtual screening with docking techniques is provided. Artificial Neural Networks and Virtual Docking are useful in silico tools to analyze binding affinity of potential oestrogen-like compounds. The selectivity capabilities of the neural network models designed for prediction of binding affinities to alpha and beta oestrogen receptors are studied. The predictions of the external validation set of compounds, initially eliminated from the investigated dataset of 60 chemicals[1,2], are considered as the key results to asses the model capabilities.

References:

[1] G.G.J.M. Kuiper, B. Carlsson, K. Grandien, E. Enmark, J. Haggblad, S. Nilsson, J.A. Gustafsson, Comparison of the ligand binding specificity and transcript tissue distribution of estrogen receptors alpha and beta, *Endocrinology*, 1997, **138**, 863-870.

[2] G.G.J.M. Kuiper, J.G. Lemmen, B. Carlsson, J. C. Corton, S.H. Safe, P.T. Van Der Saag, B. Van Der Burg, J. Gustafsson, Interaction of Estrogenic Chemicals and Phytoestrogens with Estrogen receptor beta. *Endocrinology*, 1998, **139**, 4252-4263.

[P25] Prediction of blood: brain penetration of chemicals using computed molecular descriptors

Christian T. Matson, Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, 5013 Miller Trunk Highway, Duluth MN 55811, USA; and Lester R. Drewes, Department of Biochemistry and Molecular Biology, School of Medicine, University of Minnesota Duluth, Duluth, MN 55812, USA.

Prediction of blood: brain entry of chemicals is important both for drug discovery and environmental protection. A drug designer would like to know whether candidate chemicals for the design of psychoactive drugs would be sufficiently available at the specific receptor site. This will be guided by the BBB entry of the chemicals. In environmental protection, people in USA are collectively exposed to over 75,000 chemicals. The regulator would like to know how many of them will be entering the brain from environmental exposure. But the exhaustive experimental testing of all chemicals for their BBB entry is not possible for the enormous cost and the large number of animals necessary for such testing. So, there is a need for the development of computational models for the prediction of BBB entry from molecular structure directly without the input of any experimental data.

We have developed computational models for the prediction of BBB penetration of chemicals from their computed molecular descriptors. The descriptors used include topological and geometrical indices. The usefulness of this approach in BBB research will be discussed.

[P26] Prediction of thrombin and trypsin inhibition constants using self organizing map and counterpropagation neural networks

Špela Župerl, Gregor Mlinšek, Marjana Novič, Tomaž Šolmajer, Jure Zupan, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia; Gregor Mlinšek, University Medical Center, Zaloška 2, 1000 Ljubljana, Slovenia; Tomaž Šolmajer, Lek, d.d., Drug Discovery, Verovškova 57, 1526 Ljubljana, Slovenia.

In this work we present models for prediction of the Ki values for binding of an inhibitor into the active site of thrombin and trypsin receptor. It is supposed that the binding mechanism is similar for both receptors, which enables us to apply the same modeling methods in both cases. The two data sets used to train and validate the models consist of 65 and 60 inhibitor-enzyme complexes for thrombin and trypsin receptors, respectively. The molecular electrostatic potential at the van der Waals surfaces of atoms in the protein's active site was calculated with QM/MM procedure [1]. Two data sets were used to construct models based on the counterpropagation artificial neural network method. Self organizing maps were used to divide the set in training and test set. In addition, 12 complexes containing the same ligands in both data sets were chosen for final validation of the models. The two models determined as optimal were validated by the corresponding validation set. The average error of predicted Ki values (RMS) from our models were 1,06 for thrombin and

0,49 for trypsin complexes. For 5 complexes from both data sets the Ki values are not known. The values are predicted by our models and the accuracy of these predictions will be obtained from further experiments.

On the basis of obtained results we hypothesize that the binding mechanism for thrombin complexes differs from trypsin and additional variables besides MEP values should be included into the model.

Reference:

[1] G. Mlinšek, M. Novič, M. Hodošček, T. Solmajer, Prediction of Enzyme Binding: Human Thrombin Inhibition Study by Quantum Chemical and Artificial Intelligence Methods Based on X-ray Structures, *J. chem. inf. comput. sci.*, 2001, **41**, 1286-1294.

[P27] QSAR and molecular modeling studies on substituted tropane analogs as dopamine transporter ligands

Aparajita Chattopadhyay, Anshuman Dixit and Anil K. Saxena, Medicinal and Process Chemistry Division, Central drug Research Institute, Lucknow, India-226001.

Cocaine is a widely abused drug and is associated with dopamine transporter. Dopamine transporter ligands that serve as cocaine antagonists or replacements are of current interest as a remedy for cocaine addiction. It has been demonstrated that there are at least two binding sites for (R) Cocaine one with high and other with low affinity, but the site responsible for the addicting properties of cocaine is unknown. The binding of the ligands to the specific site on dopamine transporter blocks the reuptake of dopamine in the presynaptic neuron, with a resultant increase in accumulation of excess dopamine in the synaptic cleft. In order to identify the essential structural and physicochemical requirements for dopamine transporter ligands for their binding on the dopamine transporter, the 2D and 3D QSAR studies have been carried out on some tropane analogs[1-3] with widely varying substitution patterns.

The statistically significant 2D QSAR equations (r>0.7) were developed which show the positive influence of electronic parameters *viz.* sigma(□), HOMO energy, dielectric energy and dipole moment on overall binding affinity. The application of 3D QSAR methods *viz.* CoMFA[4], advance CoMFA and CoMSIA[5] alongwith CoMFA region focusing resulted in highly significant 3D QSAR CoMFA (q^2=0.511) and CoMSIA (q^2=0.547) models with good predictive ability for the training set of 52 molecules. These models were also validated for their prediction to external test set of 16 compounds. Where the predicted and observed activities were found to be in good accordance and thus the models may be useful in designing potent dopamine trans.

References:

[1] Meltzer, P.C.; Blundell, P.; Yong, Y.F.;Chen, Z.; George, C.; Gonzalez, M.D.; Madras, B.K. *J. Med. Chem.* **2000**, 43, 2982-2991
[2] Meltzer, P.C.; Wang,B.; Chen, Z.; Blundell, P.; Jayaraman, M.; Gonzalez, M.D.; George, C.; Madras, B.K. *J. Med. Chem.* **2001**, 44, 2619-2635.
[3] Blough, B.E.; Keverline, K.I.; Nie, Z.; Navarro, H.; Kuhar, M.J.; Carroll, F.I. *J. Med. Chem.* **2002**, 45, 4029-4037.
[4] Cramer, R.D.; Patterson, D.E.; Bunce, J. D.; *J. Am. Chem. Soc.* **1988**, 110, 5959-5967.
[5] Klebe, G.; Abraham, U.; Mietzner, T. *J. Med. Chem.* **1994**, 37(24), 4130-4146.

[P28] QSAR studies of insect repellents and design of better insect repellents

Javendra B. Bhonsle, Apurba K. Bhattacharjee, Walter Reed Army Institute of Research, Department of Medicinal Chemistry, Division of Experimental Therapeutics, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA; and Raj K. Gupta, Office of Director, Research, Plans & Programs, U.S. Army Medical Research & Material Command, 504 Scott Street, Ft. Detrick, MD 21702, USA.

The Quantitative Structure Activity Relationship (QSAR) study involves several reported insect repellents from various classes of organic compounds. The compounds include carbamates derived from N,N-disubstituted diamines, N-substituted ureas & cyclic ureas, various mono & di-substituted benzamidines, benzylamides &

cyclohexanamide, several monocarboxylic esters of aliphatic diols, cyclic analogs of lactic acids, 2-oxazolidines, quinolines, isoquinolines, quinoxalines, alkyl triethylene glycol monoethers, benzyl ethers, sulfonamides, alicyclic, bicyclic, & unsaturated acetals, aminoacetals, and carboxamide acetals. The descriptors used in this study include molecular surface areas, logP, AlogP, AlogP98, hydrophilicity, lipophilicity, heat of formation, vapor pressure, boiling point, melting point, number of H-bond acceptors & donors, molecular orbital energies, superpolarizability, partial charges mapped on molecular surface areas (Jurs descriptors), geometrical descriptors that characterize molecular shapes (shadow indices), and principal moment of inertias. Several QSAR models have been generated employing multiple linear regression analysis, principal components regression analysis, partial least squares analysis, genetic function approximation analysis, and genetic partial least squares analysis. Our previously reported chemical feature based pharmacophore for insect repellency is found to be consistent with the QSAR results. These results should allow us to design novel insect repellents with more potent activity.

[P29] QSAR studies of quinolone antibacterials against *M. fortuitum* and *M. smegmatis* using theoretical molecular descriptors

Manish C. Bagchi, Bhim C. Maiti, Indian Institute of Chemical Biology, 4 Raja S.C. Mullick Road, Jadavpur, Calcutta 700032, India; Denise Mills, Subhash C. Basak, Natural Resources Research Institute, University of Minnesota-Duluth, 5013 Miller Trunk Highway, Duluth, MN 55811, USA

The re-emergence of tuberculosis infections which are resistant to conventional drug therapy has steadily risen in the last decade. Several of the quinolone antibacterials have been examined as inhibitors of *M. tuberculosis* infection as well as other mycobacterial infections. But there is hardly any study to examine specific structure-activity relationships of the quinolone antibacterials against mycobacteria. The present paper is an attempt to establish a quantitative structure-activity relationship modelling for a series of anti mycobacterial compounds. Most of the anti mycobacterial compounds do not have sufficient physicochemical data and predictive methods based on experimental data are of limited use in this situation. Hence the need for the development of QSARs utilising theoretical molecular descriptors that can be calculated directly from molecular structures. Descriptors associated with chemical structures of N-1 and C-7 substituted quinolone derivatives as well as 8-substituted quinolone derivatives with good anti mycobacterial activities against *M. fortuitum* and *M. smegmatis* have been evaluated. Ridge regression models using cross validated R^2, based on significant topostructural, topochemical and 3-D geometrical indices are developed for the prediction of molecular descriptors against anti bacterial activities of the above compounds. The comparison of the relative effectiveness of the molecular descriptors in the Ridge regression model shows some interesting results.

[P30] QSAR study on modeling of the rate of glycine conjugation of some benzoic acid derivatives: A topological approach.

Padmakar V. Khadikar, Jyoti Singh, Vijay Agrawal, Department of Chemistry, A.P.S.University, Rewa – 486008, India; Anand G. Fadnis, Department of Pharmaceutical Chemistry, Govt. Model Autonomous Holkar Science College, Indore-452017, India.

The rate of glycine conjugation(M,n mol/mg of protein) of series of benzoic acid derivatives in rat liver ($\log M_{liv}$) and kidney ($\log M_{kid}$) was modeled using a large series of molecular descriptors including distance based topological indices.

Simple as well as multiple linear regression were performed to separate the equations (models) that relates the structural features(topological indices) to the biological activities: $\log M_{liv}$ and $\log M_{kid}$. The regression analysis has indicated that the best model for the prediction of activities is obtained in multi parametric regression which included the indicator parameter. The results are critically discussed using variety of statistics including cross validated parameters.

[P31] QSTR models of juvenile hormone mimetic compounds for *Culex pipiens* larvae

Jessica J. Kraker, Douglas M. Hawkins, School of Statistics, University of Minnesota, 224 Church Street SE, Minneapolis, MN 55455, USA; Denise Mills, Ramanathan Natarajan, Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811.

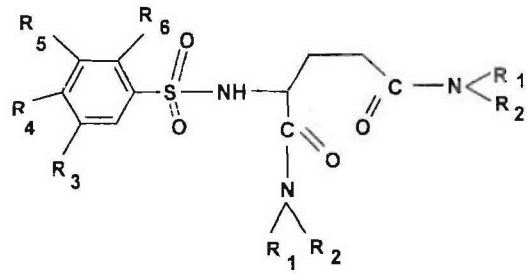
The goal of this study was to develop quantitative structure-toxicity relationship (QSTR) models for predicting insect juvenile hormone (JH) activity of 304 JH mimetic compounds against Culex pipiens. The activities (pI50) of the compounds were modeled using various calculated predictors to find a function relating the activities to the predictors. Linear models can often predict as well as more flexible models (such as k nearest neighbors or neural nets) without overfitting the data. However, in a regression setting, using highly correlated predictors to estimate a response can lead to difficulties in estimation, due to uncertainty in estimating the coefficients of the predictors. Many methods have been suggested to alleviate this problem, including ridge regression (RR), principal component regression (PCR), and partial least squares (PLS).

Our analysis of 304 compounds was to predict measured toxicity to mosquitoes, using various chemodescriptors. These were: 920 atom pair descriptors (denoted AP), 268 topological indices (denotes DES). In addition to separate modeling using each type, we combined them into a 1188-predictor run (BTH). The models were compared using the cross-validated squared multiple correlation q^2 of the regressions. RR performed best, (BTH data gave $q^2 = 0.621$, AP data gave $q^2 = 0.600$). PLS also performed reasonably well with the BTH data ($q^2 = 0.549$) or AP data ($q^2 = 0.515$), but PCR performed poorly.

To both lighten computation and improve interpretability, thinning of the descriptors was investigated. This involved stepwise selection of 100, 150 or 200 of the predictors. Whereas conventional predictor thinning by such methods as stepwise regression has been largely discredited, a moderate pruning of the predictor list did nothing to diminish the fit of the models. The best fits came with the AP data, thinned to 150 variables, using ridge regression ($q^2 = 0.680$) and PLS ($q^2 = 0.633$) methods.

[P32] Quest for new anticancer agents: A new approach of QSAR study of glutamamide analogs

Tarun Jha, Bikash Debnath, Shovanlal Gayen, Soma Samanta, Anindya Basu and Balaram Ghosh, Department of Pharmaceutical Technology, P.O. Box 17020, Jadavpur University, Kolkata 700 032, India, e-mail: tjjupharm@yahoo.com



(1)

Chemical approach to treat the biological disorder of cancer needs new frontiers. Quest of new anticancer agents lead to new approaches of chemical structure and biological activity relationship quantitatively. As a part of our continuous search for new anticancer agents[1] of the composite programme of rational drug design[1-2], a new approach of Quantitative Structure Activity Relationship (QSAR) study used to perform with 56 (fifty six) synthesized 1,5, N, N'-substituted-2-(substituted benzenesulphonyl) glutamamide (1) analogs and biological activity of the synthesizes compounds. Appreciable correlations were obtained in the QSAR study using refractotopological state atom index (RTSA) and electrotopological state atom index (ETSA). The study

showed the possible pharmacophore requirement for anticancer activity for such types of compounds. It will help for further tailoring of the molecules to get a new lead.

References:

[1] Samanta, S., Srikanth, K., Banerjee, S., Debnath, B. Gayen, S. and Jha, T. (2004) *Bioorg. Med. Chem.*, **12**, 1413-1423.
[2] Debnath, B., Gayen, S., Basu, A., Srikanth, K. and Jha, T. (2004). *J. Mol. Mod.*, **In press**.

[P33] Similarity of proteomic maps: Using of similarity index and self organizing maps

Marjan G. Vracko, National institute of chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia; Subhash C. Basak, Natural Resources Research Institute, Duluth, MN 55811, USA; Kevin Geiss, Human Effectiveness Directorate, Air Force Research Laboratory, Dayton, OH 45433, USA; Frank Witzmann, Indiana University School of Medicine, Indianapolis IN 46202, USA.

This study was conducted to assess the utility of computational tools for analyzing 2-dimensional protein gel expression patterns. The data under investigation consist on the input side from measured two-dimensional electrophoresis maps (proteomic maps) and on the output side from six measured toxicities ($EC50_{MTT}$, $EC50_{LDH}$, $EC20_{SH}$, LEC_{LP} , LEC_{ROS} , LEC_{CAT}). The maps have been recorded for non-treated cells (control) and the cells treated with 14 halocarbons. Two strategies are presented to study the similarity among proteomic maps: taking the similarity index as biodescriptor and using of self organizing map technique.

The study with the similarity index as descriptor was done in following steps.

1. we selected a sub-map of 188 spots with the lowest standard deviation (SD)
2. selected spots have been classified according their intensity
3. for each class we calculated similarity index between treated sample and control
4. similarity indices were correlated to toxicity values

One selection of spots correlates with toxicity with correlation coefficient r = 0.73. The spots of this selections were taken as biodescriptors to build the 'quantitative map-toxicity relationship model'. Counterpropagation neural network technique was applied for modelling. The correlation coefficient for regression line after the leave-one-out test was 0.83. Regarding the second strategy we present the self organizing map where proteomic maps are organized in two dimensional array of neurons according the similarity relationship among them.

[P34] Similarity-based chemical clustering techniques

Brian D. Gute and Subhash C. Basak, Natural Resources Research Institute, University of Minnesota-Duluth, Duluth, Minnesota 55811.

This study proposes the use of clustering techniques within a chemical similarity space derived using calculated molecular descriptors to identify related groups of chemicals and to choose representative chemicals from these groups for testing. These techniques are as applicable to drug discovery or industrial chemical development, i.e. data mining, as they are to toxicity estimation and address possible issues of mixture toxicity. Through the use of these techniques it is possible to use a minimalistic approach to testing rather than full factorial design analysis. The results reported here will focus on clustering a large library of psoralen derivatives, a family of chemicals used to treat both psoriasis and leukemia, and clustering a set of 228 chemicals that have been identified as components of JP-8 (jet propellant formulation #8), a fuel mixture that is in widespread use by the United States military and is known to cause immunosuppression.

[P35] A simple algorithm to unique representation of chemical structure - cyclic/ acyclic functionalized non-chiral hydrocarbons

Yenamandra S. Prabhakar, Medicinal and Process Chemistry Division, Central Drug Research Institute, Lucknow-226 001, India; Email: yenpra@yahoo.com

An algorithm, based on the 'vertex priority values' has been proposed to uniquely sequence and represent connectivity matrix of chemical structure of cyclic/ acyclic fictionalized non-chiral hydrocarbons. In this the 'vertex priority values' have been assigned in terms of atomic weights, subgraph lengths, loops, heteroatom content. In this the terminal vertices have been considered separately after completing the sequencing of the core vertices. This approach provides a multi-layered connectivity graph, which can be put to use in comparing two or more structures or parts thereof for any given purpose. Also, the basic vertex connection tables generated here come to use in the computation of characteristic matrices/ topological indices and in storing, sorting and retrieving of chemical structures from databases.

[P36] Soft-HLA: Software for HLA laboratory and Asian Indian Donor Marrow Registry

Rajeev Kalia, Uma Kanga, N K Mehra, Department of Transplant Immunology and Immunogenetics, All India Institute of Medical Sciences , New Delhi- 110 029.

The Asian Indian Donor Marrow Registry (AIDMR) in India is the first registry of voluntary bone marrow donors in India for patients of Asian Indian origin requiring bone marrow transplantation. The registry was established in1994 and has by now recruited more than 3000 voluntary donors, most of whom are tested for both HLA class I as well as class II genes. Patients who do not have an HLA identical sibling in the family have been greatly benefited from it. The maintenance and search of donor and patient's data was done manually. Since the retrieval and analysis is laborious and time consuming a need was felt for developing an efficient computer software.

HLA specialists in collaboration with data management experts and statisticians have been involved in designing the software. The ultimate goal of the exercise is to develop a user friendly system, by which patients requiring stem cell transplantation can be helped in their search for an unrelated but optimally matched donor.

We have employed the Visual Basic 6.0 as front-end (client) and Oracle 8 as back-end (server) system. Crystal Reports 7.0 was used as the report generation tool.

This Application was found to be useful for the Research and Development of the Transplantation Immunology department. All data of the patients and the donors has been maintained in the databases. This database can be routinely used for producing reports and the same can be analyzed for publishing research papers etc. The data bank of the AIDMR has been regularly utilized for fulfilling 'search requests' for an unrelated donor. With the automated system, reports can be generated and the data can also be utilized for analysis.

[P37] Stereochemical structure-activity relationship studies of insect repellents

Ramanathan Natarajan, Subhash C. Basak, Natural Resources Research Institute, University of Minnesota Duluth, Duluth, MN 55811; Alexandru T. Balaban, Texas A & M University at Galveston, Dept. of Oceanography & Marine Sciences, P. O. Box 1675, Galveston, TX 77553-1675; Jerome A. Klun, USDA Agricultural Research Service, Beltsville Agricultural Research Center, Chemical Affecting Insect Behavior Laboratory, Building 007, Rm 324, 10300 Baltimore Avenue, Beltsville, MD 20705.

A contemporary interest in medical entomology is the design of chemicals that will be effective repellents against blood-sucking arthropods. Major impetus for this trend of research arises out of the need to discover new generation of repellents to supplement or replace the standard repellent worldwide, *N,N*-diethyl-3-methylbenzamide (DEET). The development of resistance of arthropod vectors against existing chemical control agents in certain parts of the world is also a contributing factor in this trend of research.

Insect vectors (or their biological receptors) recognize repellents based on their chemical characteristics. For repellents such as (1S, 2'S)-methylpiperidinyl-3-cyclohexene-1-carboxamide (SS220) and 2-(2-hydroxyethyl)-1-piperidine carboxylic acid-1-methylpropyl ester (Bayrepel), which contain chiral centers, some mosquitoes recognize differently stereoisomers of the same repellent. Therefore, we need to study the relationship between chirality measures and the degree of repellency. With this end in view we carried out stereochemical structure activity relationship (SSAR) studies of SS220 and Bayrepel.

In the first approach we calculated a quantitative measure of chirality. The plot of decrease in biting (repellency) by mosquitoes (*Aedes aegypti*) versus the calculated chirality measure shows that the chirality descriptor is capable of discriminating the strong repellents from the weak ones. The second approach consisted of studying the degree of crowding of the putative pharmacophore (N−C=O group) by non-bonded interaction by the neighboring groups. This was accomplished by molecular mechanics (MM2) calculations which showed that less active repellents have more crowding by neighboring atoms.

[P38] Study of double proton transfer in formic acid dimer using *ab initio* molecular dynamics simulation

Sharan Shetty, Sourav Pal, Theoretical Chemistry Group, Physical Chemistry Division, National Chemical Laboratory, Pune-411008, India ; D. G. Kanhere, Department of Physics, University of Pune, Pune-411007, India.

Proton transfer processes such as multiple proton transfer, which involves more than one proton transfer has been one of the most fundamental and important reactions in chemistry and biology. Tautomeric proton transfer in the hydrogen bonded base pair of DNA and proton relay systems in many enzymes are typical examples of proton transfer in biological systems. It has been recently suggested that formation of 'short-strong' hydrogen bonds or 'low-barrier' hydrogen bonds play a crucial role in stabilizing transition states of enzymatic reactions.

Formic acid dimer has been used as the prototypical molecule to study the double proton transfer reactions [1-3]. In the present study [4], we have used *ab initio* molecular dynamics simulation to study the nature of double proton transfer in formic acid dimer at constant temperature (canonical ensemble) for a period of 25 pico seconds. We have shown that there is a synchronous motion of the protons from one formic acid monomer to the other. During the double proton transfer, both the monomer units of the formic acid dimer approach each other forming short strong hydrogen bonds at the transition state. We have also shown that, at the transition state there is resonance in both the O-C-O structures of the formic acid dimer, which can be attributed to the synchronous motion of both the protons. We have used electron localization function to study the evolution of bonding during the double proton transfer.

In this presentation, we wish to discuss on the issues such as, the synchronous nature of the double proton transfer, the formation of short-strong hydrogen bonds and the resonance in formic acid dimer. We also discuss how the above concept can be extended to study the proton transfer reactions in DNA base pairs.

References:

[1] Chang, Y. T.; Yamaguchi, Y; Miller, W. H.; Schaefer H. F. *J. Am Chem. Soc.* **1987**, 109, 7245-7253.
[2] Kim, Y. *J. Am. Chem. Soc.* **1996**, 101, 9755.
[3] Tautermann, C. S.; Loferer, M.; Voegele, A. F.; Liedl, K. R. *J. Chem. Phys.* **2004**, 120, 116501-1657.
[4] Shetty, S, Kanhere, D.G., Pal, S. to be *communicated*.

[P39] 3D pharmacophore model for insect repellent activity and discovery of new repellent candidates

Apurba K. Bhattacharjee, Department of Medicinal Chemistry, Division of Experimental Therapeutics, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA; Watanapom Dheranetra, Division of Entomology, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA; Daniel A. Nichols, Department of Medicinal Chemistry, Division of Experimental Therapeutics, Walter Reed Army Institute of Research, 503 Robert Grant Avenue, Silver Spring, MD 20910, USA; and Raj K. Gupta, Office of Director, Research, Plans & Programs, U.S. Army Medical Research & Material Command, 504 Scott Street, Ft. Detrick, MD 21702, USA.

To better understand the mechanism of insect repellency and to identify new repellent candidates, we performed a three-dimensional quantitative structure-activity (QSAR) study and developed a pharmacophore model for potent repellent activity from several known diverse insect repellent compounds by using the CATALYST 3D-QSAR methodology. The protection time for repellent activity of the compounds was taken from earlier published studies. The pharmacophore was found to have three hydrophobic sites (two aliphatic and one aromatic) and a hydrogen-bond acceptor site in specific locations in the three dimensional geometry of the molecules that are critical for potent repellent activity. The pharmacophore showed an excellent correlation (correlation=0.9) between the experimental protection time afforded by the compounds in the training set and their predicted protection time. The validity of the pharmacophore model goes beyond the list in the training set and is found to map well on a variety of other insect repellents. By mapping this model on one of the more potent analogue we have generated a three-dimensional shape based template which allowed a search of our in-house compound database and discover four new potential insect repellent candidates.

[P40] Topological descriptors in modeling the antimalarial activity of 4-(3',5'-disubstitutedanilino) quinolines

Manish K. Gupta and Yenamandra S. Prabhakar, Medicinal and Process Chemistry Division, Central Drug Research Institute, Lucknow-226 001, India, yenpra@yahoo.com.

Establishing a correlation between the conjoined systems will come to aid in furthering the understanding of the participating systems under investigation and in tuning them for optimum response. In drug research, the chemical structure, biological response, toxicity, physiological concentrations etc. are conjoined systems and development of model(s) for optimizing biological response is always a cherished goal. One of the contemporary chemotherapeutic demands is the development of new and alternative antimalarial agents to address the proliferation of multidrug-resistant strains of Plasmodium parasites. Research interest have been renewed in quinoline based antimalaril agents due to the existence of large information base and the low cost involved in the preparation of alternative/ new drugs or drug like molecules in this chemical class. In this scenario, amodiaquine a 4-aminoquinoline analogue has emerged as potential lead for modification since the chloroquine-resistant *Plasmodium* parasites are not automatically cross-resistant to it. This has prompted for the synthesis and evaluation of some 7-chloro-4-(3',5'-disubstitutedanilino)quinolines as alternative antimalarial agents [1]. In this background, we have undertaken a CP-MLR (Combinatorial Protocol in Multiple Linear Regression) [2] based quantitative structure-activity relationship (QSAR) study to rationalize the antimalarial activity profile of these compounds. More than four hundred eighty descriptors from Empirical, Constitutional, Topological, Molecular Walk Counts, Modified Burden Eigenvalues, Galvez Topological Charge Indices, Atom Centered Fragments and 2D Auto Correlations descriptor classes have been considered from DRAGON software [3] to characterize the varying substituent groups of these compounds. The study has revealed that in these compounds the antimalarial activity is correlated with the number of double bonds (nDB), Randic's shape index (path/walk-2 count; PW2), different information content (IC) indices (Bond, Structural and Complementary of orders two to five), and eigenvalues sum from the electronegativity weighed distance matrix. In the models, nDB is associated with negative regression coefficient suggesting a favorable situation for saturated templates in the substituent groups for better activity. The positive regression coefficient PW2 suggests favorable nature of the increased branching in the substituent groups for the activity. The regression coefficients of ICs are in agreement with each other as well as with other descriptors. In addition to these, an indicator parameter defined to account for the nitrogen content of 5'-position substituent has strongly correlated with the activity and suggests the importance of amines compared to other functional groups.

References:

[1] Sandrine, D.; Girault, S.; Maes, L.; Debreu-Fontaine, M.-A.; Labaeid, M.; Grellier, P.; Sergheraert, C.; *J. Med. Chem.* **2001**, *44*, 2827-2833.
[2] Prabhakar, Y. S. *QSAR Comb Sci.* **2003**, *22*, 583-595.
[3] Dragon software (version 3.0-2003) by Todeschini, R.; Consonni, V.; Mauri, M.; Pavan, M.; Milano, Italy. http://disat.unimib.it/chm/Dragon.htm

[P41] Topological estimation of binding affinities for HEPT and Nevirapin analogues with HIV-1 reverse transcriptase

Vijay K Agrawal, Krishna C. Mishra, R. Sharma and P.V. Khadikar, QSAR and Computer Chemical Laboratories, Department of Chemistry, APS University, Rewa-486 003, vijay-agrawal@lycos.com.

The topological estimation of binding affinities for HEPT and Nevirapin analogues with HIV-1 reverse Transcriptase has been carried out using initially a large set of distance based topological indices. The data under investigation consists of 20 HEPT and 20 Nevirapin analogues. The regression analysis has shown that out of the large pool of topological indices used the first-order connectivity index play dominant role, which gave excellent results in multiparametric correlations. The results are discussed critically using cross-validation method.

Author Index

(Paper Presented by the Author in **Bold**)

LATE LATE ADDITIONS

[O 62] In search of new scaffold from data mining of Antituberculosis compounds.

Indira Ghosh & Om Prakash, Bioinformatics Center, IBB, University of Pune, Pune 411107

As a part of a program to look for new antitubercular agents and targets we have tried a novel way to find a Minimum Common Bioactive Sub-structure (MCBS) present, both across the diverse chemical class and within the particular chemical class, in the various known drugs as well as compounds with known MICs [Minimum Inhibitory Concentration].
For this study, we have built a small Database of Compounds having known M.tb MICs. All the compounds are taken from published literature. While building a database we tried to make our database as diverse as possible in both structure and bioactivity, at least with respect to Mycobacterium. The database searching has been able retrived most of the pharmacophore reported [1]
On the assumption that the activity information is encoded in the fingerprint of the structure, we performed an HQSAR (hologram QSAR) [2] study on this diverse class of anti-Mycobacterial compounds in order to derive the 2D motif(s) responsible for biological activity. In the field of lead designing by fragment approach the identification of active motif for anti-tuberculosis will help to develop novel anti-Mycobacterial compounds. Our motif has been successful in identifying some of the known anti-Mycobacterial drugs, eg, Rifampicin.

References:
[1] Use of Genomics and Combinatorial Chemistry in the Development of New Antimycobacterial Drugs; C.E. Barry, III et.al, Biochemical Pharmacology, Volume 59, Issue 3, Pages pp 221-231 (1 February 2000)
[2] HQSAR: A new, Highly Predictive QSAR Technique; Tripos Technical Notes, Tripos Tech. Notes Vol 1, No.5. pp 1-7, Oct. 1997

[O 63] Gene to Drug – A Computational Pathway.

B. Jayaram, Department of Chemistry & Supercomputing Facility for Bioinformatics & Computational Biology, IIT, Delhi, India – 110016. Email: bjayaram@chemistry.iitd.ac.in; Website: www.scfbio-iitd.org

A novel chemical model (ChemGene1.0) for genome analysis has been developed. The model could distinguish genes from non-genes in 120 bacterial genomes. Efforts to extend the methodology to eukaryotic genomes with eventual application to human genome are in progress. A computational pathway has been designed combining Bioinformatics tools with ab initio methods for protein structure prediction. The pathway could successfully bracket native-like structures in the 100 lowest energy structures for 10 small alpha helical globular proteins. Efforts are on to make the methodology comparable to the accuracies and speeds attained by homology model building softwares but without database dependence. A comprehensive, active site-directed lead design protocol (christened Sanjeevini 1.0) has been developed. The protocols could successfully sort drugs from non-drugs for COX-2 target. Validation on other drug targets is in progress.
Progresses recorded at the Supercomputing Facility in the areas of Genome Analysis, Protein Structure Prediction and Active-Site Directed Drug Design will be reported.

[O 64] Interaction of DNA with clusters of amino acids in proteins.

Sathyapriya.R and Saraswathi Vishveshwara. Molecular Biophysics Unit, Indian Institute of Science, Bangalore. Karnataka. spriya@mbu.iisc.ernet.in

Protein-DNA recognition is guided by plenteous non-covalent interactions like hydrogen bonds, van der Waals, electrostatic, Cation-π and π-π interactions at the protein-DNA interface. These interactions occur between the protein residue and the base of DNA in a pairwise manner. However, the recognition motifs may constitute a cluster of spatially interacting residues in a protein. In the present study, we have detected such clusters in proteins from the protein-DNA complexes, which interact with the bases of DNA. A Graph Spectral algorithm

is used to detect side chain clusters comprising Arg, Lys, Asn, Gln and aromatic residues from the proteins interacting with the DNA.
We have considered the protein data set involved in Cation-π stair motif interaction. Our analysis shows that the interaction of proteins with DNA is through clusters in about 50% of proteins and through individual residues in the other proteins from the data set. Furthermore, an inspection of the clusters has revealed additional interactions in a few cases, which have not been reported earlier. The geometry of interaction of the DNA base and the protein residue is quantified in relation to the centroid and the normal to the base by the distance (d) and the angle (θ) parameters. Interestingly, the Cation-π and the hydrogen-bonded interactions fall into distinct set of (d, θ) values.
Analysis of the cluster composition reveals the presence of two different types of clusters, Type I and Type II, based on the presence or absence of aromatic residues (Phe, Tyr) in them. Arg residue is found to be predominant in the clusters compared to Lys, Asn and Gln. Some of the residues in the clusters also interact directly with the DNA base and these are the newly identified interactions from our study.
Residue conservation in the clusters is examined. Apart from the residues identified previously, a few more residues mainly (Phe, Tyr) have also been identified as conserved and interacting with the DNA. Interestingly, a few residues, which are a part of the interacting clusters but not directly interacting with the bases of the DNA, have also been conserved. This emphasizes the importance of recognizing the protein side chain cluster motifs interacting with the DNA, which could serve as signatures of protein-DNA recognition in the families of DNA binding proteins.

[O 65] Significance of High Performance Computing (HPC) in Bioinformatics

Janaki Chintalapati; Uddhavesh Sonavane; Dattatraya Bhat; Santosh Atanur; Sameer Ingle; and Rajendra R. Joshi * email: rajendra@cdacindia.com; uddhaveshs@cdacindia.com
Bioinformatics Team, Scientific and Engineering Computing Group, Ganeshkhind, Centre for Development of Advanced Computing, Pune – 411007 INDIA

High-throughput techniques like DNA sequencing, Microarrays etc. are revolutionizing life sciences. The numerous genome-sequencing projects are generating huge sequence data, flooding the various databases. High Performance Computing (HPC) is must to store and analyze these databases. Scientifically and computationally challenging problems like protein folding, comparative genomics, systems biology etc, demand for big number crunching machines of many tera-flops or even petaflops.
At CDAC, attempts have been made to address various scientific and computational problems using HPC. For example, High Performance Computing has been used to carry out a 100ns molecular dynamics simulation to study early stages of protein folding. Similarly, a study was carried out to detect motifs in the upstream sequences of arobidopsis genes and further study was done to understand the complex phenomena of gene regulation by analyzing the gene expression data. Efforts have also been directed towards creating a user-friendly web-computing portal for various sequence analysis and molecular modeling codes, which could help biologists to use HPC for their research without knowing the intricacies of parallel computing.

[P 42] Conformations of Polar Molecules in Hydrophobic Protein Cavities – Molecular Dynamics Study

Przemysław Miszta, Wiesław Nowak, Theoretical Molecular Biophysics Group,Instytut Fizyki, Uniwersytet M. Kopernika, ul. Grudziądzka 5, 87 – 100 Toruń, Poland

The electric dipole moment of a molecule is a very useful notion in studies of chemical and biological processes. Many reaction paths do depend on a polarity of the environment and substrates. Electrical properties are often used in QSAR studies [1].
The search for active inhibitors of enzymes requires good estimates of polarities of protein active sites. To this end fluorescent probes are often used [2,3]. However, internal molecular processes may affect the interpretation of the experimental result. It would be useful to have multiple sources of information on the distribution of the probe within the protein cavity, for example from computer modeling. Such a distribution may depend on the electronic state of the probe if the polarity of the molecule changes upon the excitation. Charge transfer systems

often have such property. We used as a model system 6-propionyl-2-(N,N)-dimethylaminonaphtalene (PRODAN) molecule buried in the apomyoglobin H93G heme pocket [3]. Numerous molecular dynamics simulations (>3ns) were performed in order to collect information on the distributions of the probe inside the protein cavity. Different electronic states of PRODAN were sampled. Simulations, done within the MOIL force field model (MOIL code [4]), have shown that PRODAN finds several docking sites in the protein pocket. There is no single simple way for a mathematical representation of this structural information. In the poster we discuss possible descriptions of data obtained from those particular computer experiments. The same data will be presented as (i) distributions of arbitrary chosen vectors in R3 coordinate space, (ii) time evolutions of selected tensors of inertia, (iii) trajectory of a point in a multidimensional subspace of the protein + PRODAN conformational space. We would like to find out a more elegant mathematical description of this problem and we hope that discussions with mathematical chemists may be helpful in reaching this goal.
Acknowledgment: This research was supported in part by UMK grant (446-F).

[1] S. C. Basak, D. Mills, B. D. Gute, "Predicting Bioactivity and Toxicity of Chemicals from Mathematical Descriptors: A Chemical-cum-Biochemical Approach", to appear in Advances in Quantum Chemistry, D.J. Klein and E. Brandas, Eds., Elsevier–Academic Press, (2005).
[2] Z.R. Grabowski, K. Rotkiewicz, W. Rettig, Chem. Rev. 103 (2003) 3899-4032.
[3] A.B.J. Parusel, W. Nowak et al., "Comparative Theoretical Study on Charge-Transfer Fluorescence Probes: 6-Propionyl-2-(N,N-dimethylamino)naphthalene and Derivatives." J. Phys. Chem. A, 102 (1998) 7149-7156.
[4] R. Elber, A. Roitberg, et al., Comput. Phys. Comm. 91 (1994) 159.

Balasubramanian Krishnan
University of California & Lawrence Livermore National Laboratory
USA
kbala@ucdavis.edu

Basak Kanika
Natural Resources Research Institute
USA

Basak Subhash
Natural Resources Research Institute
USA
sbasak@nrri.umn.edu

Bhattacharjee Apurba
Walter Reed Army Institute of Research
USA
apurba.bhattacharjee@na.amedd.army.mil

Bhonsle Jayendra
Walter Reed Army Institute of Research
USA
jbhonsle@comcast.net,
jayendra.bhonsle@amedd.army.mil

Bruggemann Rainer
Leibniz-Institute of Freshwater Ecology and Inland Fish
Germany
brg_home@t-online.de

Carpy Alain
University of Bordeaux I
France
a.carpy@lptc.u-bordeaux1.fr

Chretien Jacques
BioChemics Consulting
France
jacques.chretien@biochemics-consulting.com

Gute Brian
Natural Resources Research Institute
USA
bgute@nrri.umn.edu

Hefferlin Ray
Southern Adventist University
USA
hefferln@southern.edu

Hosoya Haruo
Ochanomizu University
Japan
hosoya@cc.ocha.ac.jp

Islam Mohammad Shahidul
Memorial University of Newfoundland
Canada
v33msi@mun.ca

Kahn Scott
Accelrys Inc.
USA
scott@accelrys.com

Kevin Geiss
Air Force Research Laboratory
USA

Khan Akbar
Defence Threat Reduction Agency
USA
Khan@dtra.mil

King R. Bruce
University of Georgia
USA
rbking@sunchem.chem.uga.edu

Kodali Varsha
Natural Resources Research Institute
USA
koda0012@d.umn.edu

Kraker Jessica
University of Minnesota
USA
stoerijj@stat.umn.edu

Kulkarni Amit
Accelrys Inc.
USA
amit@accelrys.com

Matson Christian
Natural Resources Research Institute
USA
mats0126@d.umn.edu

Miszta Przemyslaw
Faculty of Physics, Astronomy and Informatics, Nicolaus Copernicus University in Torun,ul. Grudziadzka 5/7, 87-100 Torun
POLAND
miszta@phys.uni.torun.pl

Muguruma Chizuru
Chukyo University
Japan
muguruma@lets.chukyo-u.ac.jp

Mumtaz Mohammad, Moiz
ATSDR
USA
mgm4@cdc.gov

Natarajan Ramanathan
Natural Resources Research Institute
USA
rnataraj@nrri.umn.edu

Niemi Jerry
Natural Resources Research Institute
USA
gniemi@nrri.umn.edu

Novic Marjana
National Institute of Chemistry
Slovenia
marjana.novic@ki.si

Palyulin Vladimir
Moscow State University
Russia
vap@org.chem.msu.su

Pompe Matevz
University of Ljubljana
Slovenia
matevz.pompe@uni-lj.si

Priyadarshy Satyam
RKR Group
USA
satyam@priyadarshy.com

Randic Milan
National Institute of Chemistry
Slovenia
mrandic@msn.com

Restrepo Guillermo
Universidad de Pamplona
Colombia
grestrepo@unipamplona.edu.co

Sahu Prabhat
National Chung Cheng University
Taiwan
d9124002@ccu.edu.tw

Schulz Lech
Adam Mickiewicz University
Poland
lsch@artemida.amu.edu.pl

Sem Daniel
Marquette University
USA
daniel.sem@marquette.edu, daniel.sem@mu.edu

Sidhu Amandeep S.
Faculty of Information Technology
University of Technology, Sydney
PO Box 123
Broadway, NSW 2007,
Australia
asidhu@it.uts.edu.au

Trohalaki Steven
Air Force Research Laboratory
USA
Steven.Trohalaki@wpafb.af.mil

Turker Sacan Melek
Bogaziçi University
Turkey
msacan@boun.edu.tr

Viswanadhan Vellarkad
Amgen, Inc.
USA
visv@amgen.com

Vracko Marjan
National Institute of Chemistry
Slovenia
marjan.vracko@ki.si

Agrawal Vijay
A. P. S University
India
vijay-agrawal@lycos.com

Ali Mohamed Naseer
Bharathidasan University
India
naseer_1977in@yahoo.co.in

Anbazhagan T.M
Crux Fusion Computer Center, Bangalore
India
anbuarvind@msn.com

Avti Pramod Kumar
Postgraduate Institute of Medical Education
and Research (PGIMER),
Chandigarh, India.
pramod_pgi@hotmail.com

Badhe Yogesh
National Chemical Laboratory
Pune, India
yphadhe@che.ncl.res.in

Bagchi Manish
Indian Institute of Chemical Biology
Calcutta, India
mcbagchi@iicb.res.in

Bajaj Sanjay
SGGS College of Pharmacy
India
sanjaybajaj02@yahoo.co.in

Balasubrahmanyam S. Natarajan
Jawaharlal Nehru Centre for Advanced
Scientific Research
India
snobchem@hotmail.com

Bansal Tarun
BITS-Pilani
India
skray@bits-pilani.ac.in

Barbhuiya Bahrul Islam
Assam University
Silchar, Assam
India

Barthwal Ritu
IIT, Roorkee
India
ritubfbs@iitr.ernet.in

Bhat Mohsin
University of Kashmir
India
mohsin_skb@yahoo.co.in

Bhat Parvaiz
University of Kashmir
Kashmir, India
parvaizahmmeddin@yahoo.com

Bhongade Bhoomendra
KLES's College of Pharmacy
Belgaum, Karnataka, India
babhongade@rediffmail.com

Biswal Devendra Kumar
Orissa University Agriculture & Technology
Orissa, India
dev_bioinfo@yahoo.com

Chakraborty Uday Shankar
Assam University
Silchar, Assam
India

Chandan Badapanda
Orissa University Agriculture & Technology
Orissa, India
cbp_bioinfo@yahoo.com

Chaudhuri Aditi Roy
Lady Brabourne College
Calcutta, India
odditybio@yahoo.com

Das Sukesh
Assam University
Silchar, Assam
India

Devi Siyamala A.
India

Dhakshinamoorthy A
Madurai Kamaraj University
Madurai, Tamil Nadu, India
amardm02@rediffmail.com

Dharma P
University of Madras
Chennai, India
lathap5@yahoo.co.in

Dixit Anshuman
Central Drug Research Institute
India
anilsak@hotmail.com, anshudx@indiatimes.co.in

Dudeja Sarvjit
University of Delhi
Delhi, India
sarvjitdudeja@hotmail.com

Fadnis Anand
Govt. Model Autonomous Holkar Science College
India
fadnisanand1@rediffmail.com

Ganesan Aravindhan
Annamalai University
Chidambaram, Tamil Nadu, India
garveend@yahoo.co.in

Ghosh Raja
Calcutta University
Calcutta, India
sanku_74@rediffmail.com

Gollapudy Reena
VLife Sciences Technologies Pvt. Ltd.
India
reenag@vlifesciences.com

Gulati Palak
NIPER,
Mohali, Punjab, India
palak_niper@yahoo.co.in

Gupta Manish Kumar
Central Drug Research Institute
India
manishgt@rediffmail.com

Gupta S.P.
Birla Institute of Technology and Science (BITS)
India
spg@bits-pilani.ac.in

Gupta Subhendu
Calcutta University
Calcutta, India
sg1942@rediffmail.com

Jain Manish
Department of Biotechnology, College of Agriculture
Rajasthan Agricultural University,
Bikaner, Rajasthan, India.
jainmanish80in@yahoo.com

Jaipuriar Sumeet
Vellore Institute Of Technology
India
sumu_jaipuriar@yahoo.com

Jayram B
IIT(Delhi)
Delhi, India
bjayaram@chemistry.iitd.ernet.in

Jayalakshmi Ramamurthy
Thanthai Hans Roever College
Perambalur, Tamil Nadu, India
aarjaya@hotmail.com

Juvale Dhanshri
Poona College of Pharmacy
Pune, Maharashtra, India
dhanshri_j@yahoo.com

Kalia Rajeev
All India Institute of Medical Sciences
New Delhi, India
rajeevkalia@hotmail.com

Kalyanamoorthy Subha
Annamalai University
Chidambaram, Tamil Nadu, India
subha_2410@yahoo.co.in

Karmakar Amrit
Assam University
Silchar, Assam
India

Katiyar V.K
IIT Roorkee
India
vktmafma@iitr.ernet.in

Kousar A.Tazien
Dept of Crystallography and Biophysics,
University of Madras
Chennai, India
kousar_taz@yahoo.co.in

Kulkarni Sudhir
VLife Sciences Technologies Pvt. Ltd.
India
sudhirk@vlifesciences.com

Kumar R.Praveen
Dept of Crystallography and Biophysics,
University of Madras Chennai, India
r_praveenkumar1982@rediffmail.com

Kumaraguru Sathish
Annamalai University
Chidambaram, Tamil Nadu, India
rassnaa@yahoo.co.in

Kumaran Shanmuga Sundram
Birla Institute of Technology and Science
India
kumarany2k_2000@yahoo.co.uk

Kumari Priti
India
prityonlineah@yahoo.co.in

Kumari Manisha
India
manisha_151200@yahoo.co.in

Louis Bruno
College of Pharmaceutical Sciences, Medical College,
Trivandrum
India
louisb4425@yahoo.com

Majhi Manash Chandra
Orissa University Agriculture & Technolog
Orissa, India
manash_rs_bio@yahoo.com

Mali Deepak
KLES's College of Pharmacy
Belgaum, Karnataka, India
mali_ds@rediffmail.com

Mandloi Dheeraj
Devi Ahilya University
India
dheeraj_m@sancharnet.in

Mapari Jayashree
15, Manisha Society, Pune
India
dr_jaya@rediffmail.com

Mathura S
Dept. of Crystallography and Biophysics,
University of Madras
Chennai, India
mathurashri@yahoo.co.in

Mirsamadi Neda
JNU
Delhi, India
nedamirsamadi@yahoo.com

Murali T.C
Department of Bioinformatics and Biotechnology,
Bharathidasan University, Trichirappalli,
Tamil Nadu, India

Nandy Ashesh
Jadavpur University
Calcutta, India
anandy43@yahoo.com

Prabhakar Yenamandra
Central Drug Research Institute
India
yenpra@yahoo.com

Prabuanand A. G.
India

Prathipati Philip
Central Drug Research Institute
India
anilsak@hotmail.com, phil_d_@yahoo.com

Priya G. S.
India

Ragothaman Yennamalli
Jawaharlal Nehru University
Delhi, India
ragothaman@gmail.com,
ymr5823@students.jnu.ac.in

Rajagopal Satyapriya
MBU, IISc.
Bangalore, India
spriya@mbu.iisc.ernet.in

Rajarajenderan A.
India
rajan_baski2003@yahoo.com

Rajasekaran Rajalakshmi
Annamalai University
Chidambaram, Tamil Nadu , India
saran_907@yahoo.co.in

Raskar Pravinkumar
RSIRT
Dr. C. V. Raman Scientific Institute of R & T
India
rsirt@rediffmail.com

Ray S.K.
BITS-Pilani
India
skray@bits-pilani.ac.in

Raychaudhury Chandan
Accelrys K.K.
India
chandan@accelrys.com

Roy Ambrish
University of Pune
Pune, Maharashtra, India
chemoinfo_msc@yahoo.com

Roy Kunal
Jadavpur University
Calcutta, India
kunalroy_in@yahoo.com

Roy Nobendu
Strand Genomics Pvt. Ltd.
Bangalore, India
nobendu@strandgenomics.com

Roy Rini
Lady Brabourne College
Calcutta, India
rini_roy@rediffmail.com

Saini Jagmohan Singh
NIPER
Mohali, Punjab, India
jags_niper@yahoo.co.in

Saxena Anil
Central Drug Research Institute
India
anilsak@hotmail.com

Selvadurai Jayashankar
Dept of Crystallography and Biophysics, Madras University
Chennai, India
jaidna@hotmail.com

Sharad Pranav
Vellore Institute of Technology
India
pranavsharad@yahoo.co.in

Sharma Manishika
NIPER
Mohali, Punjab, India
manishikas@yahoo.com

Shetty Sharan
National Chemical Laboratory
Pune, Maharashtra, India
sharan@comp.ncl.res.in

Singh Jyoti
Department of Chemistry, A. P. S University
Rewa, M.P, India
vijay-agrawal@lycos.com

Singh Shailza
Indian Institute of Technology (Delhi)
Delhi, India
shailza_iitd@yahoo.com

Srikrishna Arvind M.
Annamalai University, Chidambaram, Tamil Nadu
India
arvindsrikrishna@yahoo.co.in

Subramanian Jyothi
Nicholas Piramal India ltd.
India
jyothis@nicholaspiramal.co.in

Suguna C.
Centre for Cellular and Molecular Biology
Hyderabad, Andhra Pradesh, India
suguna@ccmb.res.in

Thakur Abhilash
Softvision College of Biotechnology and Science
India
abhilashthakur@yahoo.com

Thakur Mamta
Softvision College of Biotechnology and Science
India
mamtaathakur@yahoo.co.in

Thomas Leonard Joseph
C/o Dr. K. Roy, Division of Medicinal
& Pharmaceutical Chemistry,
Department of Pharmaceutical Technology,
Jadavpur University
Calcutta, India
alabavae@yahoo.com

Tripahti Shailesh Kumar
IIT(Guwahati)
Assam, India
shailesh@iitg.ernet.in

Verma Jitender
NIPER
Mohali, Punjab, India
jeetu270@rediffmail.com

Vishveshwara Saraswathi
MBU, IISc.
Bangalore, India
sv@mbu.iisc.ernet.in

Wadekar Vinay
National Chemical Laboratory
Pune, India
vinaywadekar@che.ncl.res.in

Chattopadhyay Aparajita
Reddy's Lab
appu_bits2000@yahoo.co.in

Debnath Sitra
Assam University,
Silchar, Assam
India

Kanthinath Rahul
Assam University,
Silchar, Assam
India

Mujundar Mohammad Afjal Hussain
Assam University,
Silchar, Assam
India

Enabling Scientific Innovations!

Seascape is a US-India Computational Science company offering software, research consulting, education and software services for Drug Discovery and Nanotechnology

Sunil@seascapelearning.com
www.seascapelearning.com
USA India

Seascape

Seascape Software Categories

- Molecular Modelling & Analysis
- Combinatorial Chemistry And Diversity
- Bioinformatics & Structure Based Design
- Virtual Screening
- Chemical Informatics
- Pharmacophore Perception
- QSAR & ADMET
- Reaction Kinetics

Seascape

Seascape R&D Services

Drug Discovery

- Library Design
- Pharmacophore Perception
- Structure Based Drug Design
- Protein Structure Elucidation
- ADME/Tox Properties
- Prediction of Physical Properties
- Simulations
- Data Mining
- Chemometrics
- Bioinformatics

Nanotechnology

- Property Prediction
- Simulations
- Structure Elucidation

Methodologies

- Molecular Dynamics / Molecular Mechanics
- Semi-empirical and Ab Initio Calculations
- Density Functional Theory
- QSAR/QSPR (Including 3D and 4D QSAR)
- Combinatorial Chemistry and Molecular Diversity Analysis
- Virtual Screening
- Docking
- Conformational Analysis

Seascape

Seascape Software Services

Application Expertise

- Molecular Modelling
- Chemical registration/data base systems
- Lab information systems/Notebooks
- Knowledge base development
- Data curation
- Text-based data mining
- eLearning

Software Expertise

- Object Oriented Design
- Structured Programming Environments
- GUI development
- Graphics and Visualization
- Design of computer based learning solutions

Seascape

Seascape R&D Network

Network of world-class Scientific experts from India

Drug Discovery and Research Services offered in USA

Seascape provides market access and project management

SYBYL Base

Industry Standard for Drug Discovery

- Provides molecular modelling environment
  - GUI for all SYBYL modules and Unity
  - visualization of molecular structures
- Provides basic modelling tools
  - sketch, modify and visualize structures
  - calculate energies and minimize molecules
    - variety of different force fields
  - Data organization
    - molecular databases, molecular spreadsheet
  - Molecular descriptor calculation
    - basic descriptors, mol weight, dipole moment, ...
  - Scripting - automate tasks, prototype new ideas
    - SYBYL Programming language (SPL)

ADF

Universal Density Functional Package for Chemists

- What does it do?
  - First principles electronic structure calculations, based on Density Functional Theory.
- Why is it useful?
  - DFT provides superior accuracy to Hartree-Fock theory and semi-empirical approaches. In contrast to conventional ab initio methods (MP2, CI, CC), it enables accurate treatment of systems with several hundreds of atoms (or several thousands with QM/MM).
- Why is it cool? [What is its truly unique value?]
  - ADF is efficient due to a combination of linear scaling and parallelization techniques.
  - unique chemical analysis methods (energy + charge decomposition) not available in other programs
  - main focus on inorganic chemistry / transition metal compounds
  - Reliable and robust for any element in periodic system
  - Spectroscopy is its differentiating use

Seascape

TranSIESTA

Modelling of electrical properties of nanoscale devices

- What does it do?
  - TranSIESTA-C, is a first principles electronic structure program capable of modelling electrical properties of nanostructured systems coupled to semi-infinite electrodes. The two electrodes, for instance, could be a nanotube and a metal, and the nanostructure could be the interface region between the two systems. Other typical systems include molecules between metal surfaces and interfaces between materials.
- Why is it useful?
  - This product that can perform calculations on integrated nano systems such as metal-molecule-metal. In addition, it can calculate electron structure and transport across interfaces and junctions, which is invaluable in nano electronics and ultra small-scale semiconductors.
  - One can do calculations on isolated molecules as well as periodic systems such as metals, crystals and carbon nanotubes.

Seascape

4DQSAR

Quantative 3D Pharmacophore Models of Biopotency

- What does it do?
  - Build 3D-QSAR models incorporating the conformational, alignment, and pharmacophore degrees of freedom
- Why is it useful?
  - It is used to create and screen against 3D-pharmacophore QSAR models and can be used in receptor-independent or receptor-dependent modes.
  - Alignment problem is taken care of. This is particularly acute when dealing with diverse compounds or when multiple conformations are involved.
- Why is it cool? [What is its truly unique value?]
- Things that you can get from 4D-QSSR, but not from CoMFA which are usually critical in lead optimization and identification are:
  - The loss in biological activity due to ligand conformational entropy.
  - A prediction of the active conformation of each ligand using a given 4D-QSAR model.
  - All 4D-QSAR models, as opposed to one CoMFA model, consistent with a given training set. That is, the best model for the training set may NOT be the best model for the series of ligands being explored.
  - An explicit QSAR equation where the relative importance of each pharmacophore site [descriptor term] can be directly compared to all pharmacophore sites. CoMFA does not provide an explicit, simple QSAR equation.
  - The ability to apply the method to both receptor-independent and receptor-dependent (structure-based) QSAR design problems.

Seascape

Reaction Kinetics....

- Khimera is for atomistic simulations in combustion, plasma chemistry, and materials science for quantum electronics and using quantum-chemical program packages. It estimates thermodynamic, transport and kinetic parameters from fundamental molecular data and performs reactor level simulations.
- CARAT can be used for CAlculation of RATe constants and mechanism building in the fields of gas-phase chemical technology, combustion, plasma chemistry, and materials science.
- COSMOS is a for calculation of leakage current and capacitance of the MOSFET structures.
- Chemical WorkBench is designed for kinetic simulation and conceptual design of physical and chemical processes, reactors, and technologies.

Seascape

Why work with Seascape?

Best-in-class scientific software

Best Price/Value

Seascape R&D Network

Sunil@seascapelearning.com
www.seascapelearning.com
USA India

Seascape

Enabling Scientific Innovations

Drug Discovery & Nanotechnology Research

Computational Science & Molecular Modeling Software

World class Collaborative Research & Informatics

Training & Education Services

Seascape

www.Seascapelearning.com
info@Seascapelearning.com
USA INDIA

BIOINFORMATICS AT C-DAC

About C-DAC

Established in March 1988, as a Scientific Society of the Ministry of Communications and Information Technology, Government of India, the Centre for Development of Advanced Computing (C-DAC), is primarily an R&D institution involved in the design, development and deployment of electronics and advanced Information Technology (IT) products and solutions. Some of the technologies that C-DAC has dealt with are High Performance Computing (HPC), Natural Language Processing (NLP), Artificial Intelligence (AI), e-Learning, Multilingual multimedia computing, Geomatics, Cyber security, Real time systems & software, Data warehousing, Data mining, Digital/ Broadband wireless networks, Scientific modeling & visualization. The key sectors addressed by C-DAC are Finance, Healthcare, Power, Steel, Defence, Telecom, Agriculture, Industrial Controls, Broadcasting, Education and e-Governance.

Bioinformatics Activities @ C-DAC

The bioinformatics team at C-DAC deals with the development, porting and optimizing of codes on the PARAM series of parallel supercomputers, in the various areas like molecular modeling, sequence analysis, comparative genomics, quantum chemistry etc. The widely used molecular modeling codes like AMBER, CHARMM, GROMACS, FTDock, DARWIN and sequence analysis codes like BLAST, FASTA, ClustalW, MEME, Glimmer are available on the PARAM Padma. The bioinformatics team is also carrying out various domain activities in the area of sequence analysis, comparative genomics, systems biology and molecular modeling. Some of the major activities include reconstruction of metabolic pathways for detecting drug targets, ortholog detection, microarray data analysis and correlation with genome sequence data, gene annotation, protein folding studies etc. The team has also developed a web-computing portal (GIPSY) with a user-friendly environment for the above bioinformatics applications available on the PARAM Padma. Such an environment can relieve the scientist from the burdens associated with the inessential and often arcane details of the hardware and software systems. The recently inaugurated Bioinformatics Resource and Applications Facility (BRAF) of C-DAC provides a high-end supercomputing facility to the bioinformatics research community with remote access to the existing bioinformatics software available on the PARAM Padma. The Bioinformatics Team of C-DAC is collaborating with National Center for Cell Sciences (NCCS, Pune) for a project on EST sequence assembly and annotation, and with Indian Institute of Chemical Technology (IICT, Hyderabad) for conducting an Advanced Course in Bioinformatics (ACB). Some of the other collaborators include Jawaharlal Nehru University (JNU, Delhi), National Chemical Laboratory (NCL, Pune), IBM Life sciences, University of Pune etc.

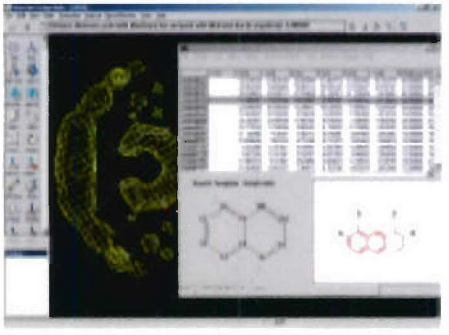
Contact:
Dr. Rajendra R. Joshi
Coordinator, Bioinformatics,
Scientific and Engineering Computing Group (SECG)
C-DAC, Pune 411007
Email: rajendra@cdacindia.com
http://bioinfo-portal.cdacindia.com

VLife Molecular Design Suite

enhancing

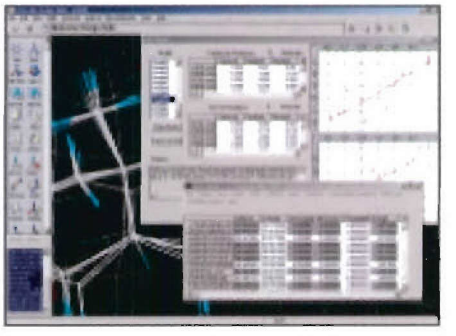
Computer Aided Discovery Research

vLife sciences
adding value to life



VLife MDS 2.0

- Build, analyze & visualize small molecules and bio-molecules
- Perform conformational analysis systematically as well as randomly
- Evaluate & visualize three dimensional molecular properties
- Create your own molecule database customized for focused research
- Database searches based on 2D & 3D templates, fingerprints, descriptors, pharmacophores & receptor information
- Homology modeling, active site ID and docking

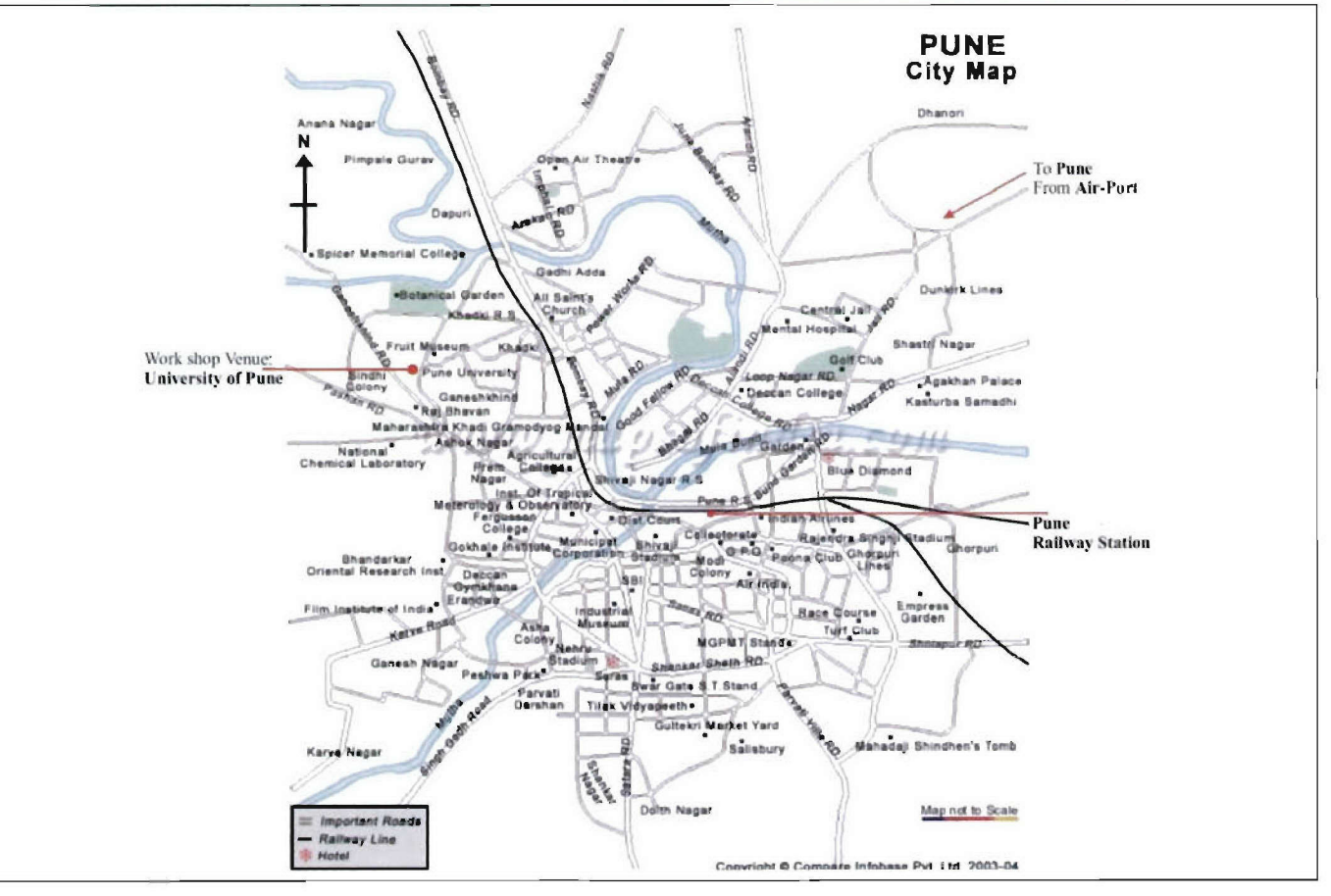


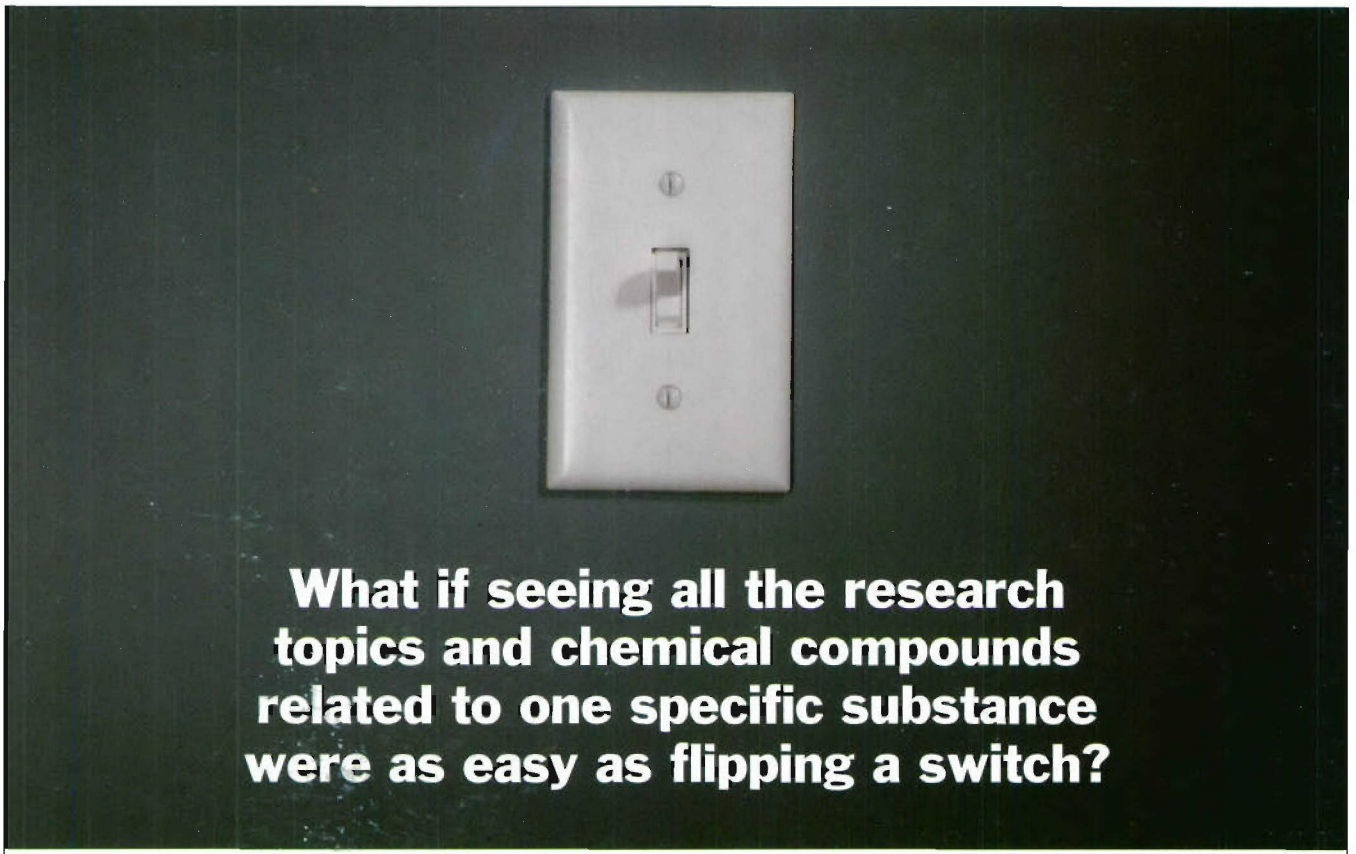
QSARPlus and 3D QSARPlus

- Evaluation of 500 descriptors from wide range of categories
- Use our powerful data manager along with mathematical functions
- Avail host of regression methods & neural networks for range of analysis
- Perform 3D QSAR to establish relationship between molecular fields
- Create combinatorial libraries right from template & substituents
- Predict activities of generated libraries using saved regressions

System Requirements: Windows OS, Pentium IV or equivalent, graphics accelerator card, 256 MB or more RAM

For an evaluation copy visit http://www.vlifesciences.com or contact us at VLife Sciences Technologies Pvt Ltd; Tel: +91 20 2588 6737





What if seeing all the research topics and chemical compounds related to one specific substance were as easy as flipping a switch?

It is.

SciFinder provides a single point of access to the world's largest collection of organic and inorganic substance information, seamlessly linked to relevant patent and journal literature.

Start with what you have—draw a generic or specific structure, substructure, or reaction, or type in a formula or trade name—and SciFinder will lead you to a wealth of chemical information.

You can view the substance in the CAS Registry,℠ the world's largest chemical substance database, and click to find chemical property information, commercial suppliers, regulatory information, and much more.

Another click takes you to cited references and the electronic full-text documents in which the substance appears, such as journal articles and patents from all over the world, from the beginning of the 20th century to yesterday.

To find out how the substance is prepared or used in synthetic processes, you can click to reaction information, which leads you to associated reaction types, starting materials, catalysts, and solvents and reaction conditions used in preparation.

Comprehensive, intuitive, seamless—SciFinder enlightens you. It's part of the process. To find out more, contact our sales representatives, Sci-Edge Information at 91 20 2545 3620 (India), 614-447-3700 (worldwide), mwaikar@sciedgeinfo.com, or visit www.cas.org/SCIFINDER.

SciFinder®

Part of the process.™

With STN, you can find exactly the science and technology information your business needs to make important decisions. A change in molecular stereochemistry can radically alter the strategic direction of your entire research project and possibly your entire company. That's why STN gives you so many ways to research substances, your competitors, and your industry. Our tightly integrated system ties together published research, journal literature, patents, structures, stereochemistry, trade names, physical properties, sequences, and other data from the world's largest collection of chemical substance information—CAS Registry.℠ So when you're dealing with issues vital to your business, use STN and get the results you want. Precisely.

STN®

To find out more, contact our sales representatives in India, Sci-Edge Information, at 91 20 2545 3620 or info@sciedgeinfo.com

JAPAN

STN International℠ c/o Japan Association for International Chemical Information (JAICI)
Phone: 81 3-5978-3621
Internet: www.jaici.or.jp

STN International c/o Japan Science and Technology Agency (JST)
Phone: 81 3-5214-8402
Internet: pr.jst.go.jp/db/STN/index.html

EUROPE

STN International c/o FIZ Karlsruhe
Phone: 49 7247/808-555
Internet: www.stn-international.de

NORTH AMERICA

STN International c/o Chemical Abstracts Service
Phone: 800-753-4227 or 614-447-3700
Internet: www.cas.org/stn.html