Fourth International Symposium

Computational Methods in Toxicology and Pharmacology

Integrating Internet Resources

(CMTPI-2007)

Abstract Book

Moscow, Russia

September 1-5, 2007

Co-Chairmen
Nikolai Zefirov, Full Member of Russian Academy of Sciences
Alexander Archakov, Full Member of Russian Academy of Medical Sciences

Local Organizing Committee
Chairman
Vladimir Poroikov (Institute of Biomedical Chemistry of RAMS, Russia)
Co-Chairman
Oleg Mikheev (NC Group/NVK "VIST")
Secretary
Paulina Gabusu (NC Group/NVK "VIST")
Members
E. Beyer (Institute of Biomedical Chemistry of RAMS, Russia)
T. Fedoronchuk (Institute of Biomedical Chemistry of RAMS, Russia)
A. Kabankin (Center of Theoretical Problems of Physical-Chemical Pharmacology of RAS, Russia)
A. Lagunin (Institute of Biomedical Chemistry of RAMS, Russia)
A. Lisitsa (Institute of Biomedical Chemistry of RAMS, Russia)
S. Novikov (Institute of Human Ecology and Hygiene of Environment of RAMS, Russia)
N. Pshenichnikova (NC Group/NVK "VIST")
V. Palyulin (Moscow State University, Russia)
E. Shilova (National Research Center for Biologically Active Compounds, Russia)

International Organizing Committee
A. Carpy (University of Bordeaux I, France)
J. Devillers (CTIS, France)
Bo Tao Fan (University of Paris 7, France)
A. Geronikaki (Aristotelian University of Thessaloniki, Greece)

International Scientific Advisory Board
S. Basak (University of Minnesota, USA)
H. Chuman (University of Tokushima, Japan)
A. Kel (Biologische Datenbanken GmbH, Germany)
D. Livingstone (Chemquest, UK)
K.T. No (BMDRC, Yonsei University, Korea)
M. Nicklaus (National Cancer Institute, NIH, USA)
O. Raevsky (Institute of Physiologically Active Compounds of RAS, Russia)
A. Richard (Environmental Protection Agency, USA)
A. Saxena (Central Drug Research Institute, India)
A. Tropsha (University of North Carolina in Chapel Hill, USA)
A. Worth (EC Joint Research Center, Italy)
J. Yao (Shanghai Institute of Organic Chemistry, China)

Program

01.09.2007		

12:00 - 18:00	**Registration**	
18:00 - 18:40	***Opening Ceremony***	
18:40 - 19:00	***To the memory of Prof. Bo Tao Fan***	
19:00 - 19:45	Plenary lecture	**Ann Richard**. *National Center for Computational Toxicology, EPA, NC, USA.* NEW PUBLIC DATA & INTERNET RESOURCES IMPACTING PREDICTIVE TOXICOLOGY
20:00 - 22:00	***Welcome Party***	

02.09.2007		

Internet Tools & Databases (I)
Chairs: Alexander Kel and Alexander Tropsha

09:00 - 09:40	Plenary Lecture	**Marc Nicklaus**. *National Cancer Institute, National Institute of Health. NCI-Frederick, MD, USA*. INTERNET RESOURCES INTEGRATING MANY SMALL-MOLECULE DATABASES
09:40 - 10:00	Major Talk	**Wolf-Dietrich Ihlenfeldt**. *Xemistry GmbH, Germany.* A VIRTUAL FILE SYSTEM FOR THE PUBCHEM CHEMICAL STRUCTURE AND BIOASSAY DATABASE
10:00 - 10:20	Major Talk	**Andrey Lisitsa**. *Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., Moscow, Russia.* CYTOCHROME P450 KNOWLEDGEBASE FOR DRUG DISCOVERY
10:20 - 10:40	Major Talk	**Alexander Kos**. *AKos GmbH, Steinen, Germany.* WORKFLOW ENVIRONMENT FOR ESTIMATING BIOLOGICAL EFFECTS
10:40 - 11:10	***Coffee Break***	

Internet Tools & Databases (II)
Chairs: Roman Efremov and Marc Nicklaus

11:10 - 11:30	Major Talk	**Anil Saxena**. Central Drug Research Institute (CDRI), Lucknow, India. INTERNET RESOURCES IN GPCR MODELLING
11:30 - 11:50	Major Talk	**Athina Geronikaki**. Aristotle University of Thessaloniki, Thessaloniki, Greece. PREDICTION OF BIOLOGICAL ACTIVITY VIA INTERNET. MEDICINAL CHEMIST'S POINT OF VIEW
11:50 - 12:05	Oral Communication	**Greg Pearl**. ACD/Labs, Toronto, Ontario, Canada. EFFECT OF DESCRIPTOR SELECTION: COMPARING LOGD TO LOGP IN DRUG-LIKENESS PROFILING
12:05 - 12:20	Oral Communication	**Alexey Lagunin**. *Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., Moscow, Russia.* PHARMAEXPERT: ESTIMATING DRUG-DRUG INTERACTIONS AND FINDING COMPOUNDS WITH MULTIPLE MECHANISMS OF ACTION
12:20 - 12:35	Oral Communication	**Russel Naven**. *Lhasa Limited, Leeds, UK.* IMPROVING THE PREDICTIVE PERFORMANCE OF AROMATIC AMINE MUTAGENICITY IN DEREK FOR WINDOWS THROUGH THE ANALYSIS OF PROPRIETARY DATA
12:35 - 14:30	***Lunch***	
15:00 - 19:00	***Sightseeing Tour***	

03.09.2007		

"OMIC"-Sciences and Bioinformatics (I)
Chairs: Johann Gasteiger and Andrey Lisitsa

09:00 - 09:40	Plenary Lecture	**Alexander Kel**. *BIOBASE GmbH, Wolfenbuttel, Germany.* EXPLAIN™: FINDING MOLECULAR MECHANISM OF DISEASE FROM PROMOTER MODELS TO SIGNALING PATHWAYS
09:40 - 10:00	Major Talk	**Ivan Rusyn**. *Department of Environmental Sciences and*

		Engineering; University of North Carolina at Chapel Hill, USA. BRIDGING INVESTIGATIVE TOXICOLOGY AND DISEASE-ORIENTED RESEARCH BY BUILDING A MOUSE-TO-HUMAN PARADIGM
10:00 - 10:20	Major Talk	**Razif Gabdoulline**. *EML Research GmbH, Molecular and Cellular Modeling Group, Heidelberg, Germany.* ENZYME KINETIC PARAMETER ESTIMATION IN SYSTEMS BIOLOGY USING PROTEIN STRUCTURES
10:20 - 10:35	Oral Communication	**Hanjun Zou**. *Intertek Consumer Goods North & East China, Shanghai, China.* MOLECULAR INSIGHT INTO THE INTERACTION BETWEEN IFABP AND PA BY USING MM-PBSA AND ALANINE SCANNING METHODS
10:35 - 10:50	Oral Communication	**Sergey Moshkovskii**. *Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., Moscow, Russia.* ANALYSIS OF PEAK INTENSITY CORRELATIONS TO MAXIMIZE BIOLOGICAL MEANING OF PROTEOME DIAGNOSTIC MASS-SPECTRA
10:50 - 11:20	***Coffee Break***	

"OMIC"-Sciences and Bioinformatics (II)
Chairs: Izet Kapetanovic and Greg Pearl

11:20 - 12:00	Plenary Lecture	**Johann Gasteiger**. *Computer-Chemie-Centrum and Institute of Organic Chemistry, University of Erlangen-Nuremberg; Molecular Networks GmbH, Erlangen, German .* EXPLORATION INTO BIOCHEMICAL PATHWAYS
11:55 - 12:15	Major Talk	**Milan Randic**. *National Institute of Chemistry, Ljubljana, Slovenia.* ON GRAPHICAL REPRESENTATION OF PROTEINS AND PROTEIN ALIGNMENT
12:15 - 12:30	Oral Communication	**Adel Golovin**. *EMBL-EBI, Hinxton Hall, Genome Campus, Cambridge, UK.* A DATABASE SEARCH AND RETRIEVAL SYSTEM FOR THE ANALYSIS AND VIEWING OF BOUND LIGANDS, ACTIVE SITES, SEQUENCE MOTIFS AND 3D STRUCTURAL MOTIFS
12:30 - 12:45	Oral Communication	**Carlos da Silva**. *School of Pharmaceutical Sciences of Ribeirao Preto, University of Sao Paulo, Ribeirao Preto - SP, Brazil.* HOMOLOGY MODELING AND COMPUTER-AIDED DESIGN OF HNRNP K MARKER AND LIGANDS IN HEAD AND NECK CANCER
12:45 - 14:00	***Lunch***	

QSAR and Environmental Toxicity (I)
Chairs: James Devillers and Ann Richard

14:00 - 14:40	Plenary Lecture	**Kyoung Tai No**. *Bioinformatics and Molecular Design Research Center, Seoul, Korea.* PREDICTION OF ENVIRONMENTAL TOXICITIES
14:40 - 15:00	Major Talk	**Jacques Chretien**. *BioChemics Consulting SAS, Orleans, France.* NEW IN SILICO STRATEGIES FOR ASSESSING ENVIRONMENTAL CHEMISTRY AND REACH CHALLENGES
15:00 - 15:20	Major Talk	**Daniel Zuaboni**. *MerckSerono Geneva Research Center, University of Lausanne, Geneva, Switzerland.* CONSENSUS MODELS AND META-MODELS FOR THE PREDICTION OF LOGP USING NEURAL NETWORKS
15:20 - 15:35	Oral Communication	**Yuriy Abramov**. *Pfizer Global Research and Development, Groton, USA.* REACTIVE METABOLISM – QSAR AND QUANTUM CHEMISTRY MODELING
15:35 - 15:50	Oral Communication	**Mariana Novic**. *National Institute of Chemistry, Ljubljana, Slovenia.* MODELLING, OPTIMIZATION, AND VARIABLES SELECTION IN QSA(P)R STUDIES. THE ROLE OF COUNTER-PROPAGATION NEURAL NETWORK
15:50 - 16:20	***Coffee Break***	
16:20 - 18:20	**Poster Session**	

QSAR and Environmental Toxicity (II)
Chairs: Jacques Chretien and Kyoung Tai No

09:00 - 09:40	Plenary Lecture	**James Devillers**. *CTIS, Rillieux La Pape, France.* STRUCTURE-ACTIVITY MODELING OF A DIVERSE SET OF ANDROGEN RECEPTOR LIGANDS
09:40 - 10:00	Major Talk	**Paola Gramatica**. *QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Structural and Functional Biology, University of Insubria, Varese, Italy.* IN SILICO SCREENING OF ESTROGEN-LIKE CHEMICALS BASED ON QSAR MODELS
10:00 - 10:20	Major Talk	**Nathalie Marchand-Geneste**. *University of Bordeaux, Talence, France.* e-ENDOCRINE DISRUPTING CHEMICAL DATABASES FOR DERIVING QSAR MODELS
10:20 - 10:35	Oral Communication	**Natalja Fjodorova**. *Laboratory of Chemometrics, National Institute of Chemistry, Ljubljana, Slovenia.* QSAR MODELLING OF CARCINOGENICITY FOR REGULATOTY USE IN EUROPE
10:35 - 10:50	Oral Communication	**Vyacheslav Nikolaev**. *Arbuzov Institute of Organic and Physical Chemistry, Russian Academy of Sciences, Russia.* QUASICRYSTALLINITY OF A LIQUID STATE AS A BASIS FOR CREATING QSPR-MODELS AND DESCRIBING SYNERGIC EFFECTS IN BINARY SYSTEMS
10:50 - 11:20	*Coffee Break*	

Dedicated to the Memory of Martyn Ford

Computer-Aided Drug Discovery (I)
Chairs: Athina Geronikaki and Vladimir Poroikov

11:20 - 12:00	Plenary Lecture	**David Livingstone**. *ChemQuest, Sandown; Centre for Molecular Design, University of Portsmouth, UK.* QSAR STUDIES USING THE PARASHIFT SYSTEM
11:55 - 12:15	Major Talk	**Florent Barbault**. *ITODYS, University Paris, France. CYCLIN DEPENDENT KINASE (CDK2/CDK4) SELECTIVITY ELUCIDATED BY MOLECULAR DYNAMICS AND QUANTUM CHEMISTRY STUDIES*
12:15 - 12:30	Oral Communication	**Peter Fedichev**. *Quantum Pharmaceuticals, Moscow, Russia.* BIOLOGICAL SPECTRA ANALYSIS: LINKING BIOLOGICAL ACTIVITY PROFILES TO MOLECULAR TOXICITY
12:30 - 12:45	Oral Communication	**Catia Teixeira**. *ITODYS, University Paris, France .* 2D/3D QSAR AND MOLECULAR MODELLING STUDIES OF NEW CLASS OF HIV-1 GP41 INHIBITORS
12:45 - 14:00	*Lunch*	

Computer-Aided Drug Discovery (II)
Chairs: Daniel Domine and Douglas Oliver

14:00 - 14:40	Plenary Lecture	**Oleg Raevsky**. *Institute of Physiologically Active Compounds, Russian Academy of Sciences, Chernogolovka, Russia.* HYDROGEN BONDING PARAMETRIZATION IN QSAR AND DRUG DESIGN
14:40 - 15:00	Major Talk	**Hiroshi Chuman**. *Institute of Health Biosciences, The University of Tokushima, Japan.* TOWARD BASIC UNDERSTANDING OF THE PARTITION COEFFICIENT LOG P AND ITS APPLICATIONS IN QSAR
15:00 - 15:20	Major Talk	**Jianhua Yao**. *Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, China.* NEW PREDICTION SYSTEM FOR MUTAGENICITY: CISOC-PSMT AND ITS APPLICATIONS
15:20 - 15:35	Oral Communication	**Irene Kouskoumvekaki**. *Center for Biological Sequence Analysis, BioCentrum, Technical University of Denmark, Lyngby, Denmark.* PREDICTION OF PH-DEPENDENT AQUEOUS SOLUBILITY

		OF DRUG-LIKE MOLECULES AND DRUG CANDIDATES WITH CHEMOINFORMATICS TOOLS
15:35 - 15:50	Oral Communication	**Ilkay Yildiz**. *Ankara University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, Tandogan-Ankara, Turkey.* THREE-DIMENSIONAL COMMON-FEATURE HYPOTHESES AND 2D-QSAR STUDIES ON SOME ANTIMICROBIOLOGICALLY ACTIVE AMIDES AGAINST DRUG-RESISTANCE *STAPHYLOCOCCUS AUREUS*
15:50 - 16:20	***Coffee Break***	

Computer-Aided Drug Discovery (III)
Chairs: David Livingstone and Vladimir Palyulin

16:20 - 17:00	Plenary lecture	**Alexander Tropsha**. *School of Pharmacy, University of North Carolina at Chapel Hill, USA.* COMBINATIONAL QSAR MODELING OF CHEMICAL TOXICANTS TESTED AGAINST *TETRAHYMENA PYRIFORMIS*
17:00 - 17:20	Major Talk	**Douglas Oliver**. *Unit for Drug Research and Development, School of Pharmacy, North-West University Potchefstroom Campus, Potchefstroom, South Africa.* MOLECULAR MODELLING AND MEDICINAL CHEMISTY OF BIOACTIVE POLYCYCLIC CAGE COMPOUNDS
17:20 - 17:40	Major Talk	**Esin Aki**. *Pharmaceutical Chemistry Dept., Faculty of Pharmacy, Ankara University, Ankara, Turkey.* PHARMACOPHORE ANALYSIS OF TOPOISOMERASE II INHIBITORY ACTIVE BENZAZOLES
17:40 - 17:55	Oral Communication	**Roman Efremov**. *Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia.* MOLECULAR MODELING OF MEMBRANE PROTEINS – PERSPECTIVE DRUG TARGETS
17:55 - 18:10	Oral Communication	**Kimito Funatsu**. *The University of Tokyo, Department of Chemical System Engineering, Japan.* APPLICATION OF THE NOVEL MOLECULAR ALIGNMENT METHOD USING THE HOPFIELD NEURAL NETWORK TO 3D-QSAR
18:10 - 18:25	Oral Communication	**Yuriy Vorobjev**. *Institute of Chemical Biology and Fundamental Medicine of Siberian Brunch of Russian Academy of Sciences, Novosibirsk, Russia. AB INITIO* DOCKING METHOD COMBINING CAVITY SEARCH WITH MOLECULAR DYNAMICS
18:25 - 18:45	***Presentation of CMTPI-2009***	
19:30	***Gala Dinner***	

05.09.2007

Computer-Aided Drug Discovery (IV)
Chairs: Hiroshi Chuman and Jean-Pierre Doucet

09:00 - 09:20	Major Talk	**Vladimir Palyulin**. *Department of Chemistry, Moscow State University, Moscow, Russia.* MFTA-BASED DESIGN OF ACTIVE STRUCTURES
09:20 - 09:40	Major Talk	**Daniel Domine**. *Merck Serono Geneva Research Center, Geneva, Switzerland.* OPTIMIZING THE USE OF TOXICITY PREDICTION TOOLS ACROSS DRUG DISCOVERY STAGES
09:40 - 10:00	Major Talk	**Petko Petkov**. *Laboratory of Mathematical Chemistry, Bourgas As. Zlatarov University, Bourgas, Bulgaria.* MECHANISM BASED COMMON REACTIVITY PATTERN (COREPA) MODELING OF AHR BINDING AFFINITY
10:00 - 10:15	Oral Communication	**Tuomo Kalliokoski**. *Department of Pharmaceutical Chemistry, University of Kuopio, Finland.* FIELDCHOPPER, MOLECULAR FIELD-BASED VIRTUAL SCREENING METHOD. DESCRIPTION AND EVALUATION
10:15 - 10:30	Oral Communication	**Marjan Vracko**. *National Institute of Chemistry, Ljubljana, Slovenia.* QSAR STUDY ON A SET OF QUINOXALINE

		COMPOUNDS AS ANTITUBERCULOSIS DRUGS USING THE COUNTER PROPAGATION NEURAL NETWORKS
10:30 - 10:45	Oral Communication	**Banafsheh Javahery**. *Department of Chemistry, Imam Hossein University, Tehran, IRAN.* STRUCTURAL-ACTIVITY RELATIONSHIP (SAR), ATOMIC ELECTRON DENSITY AND CONFORMATIONAL INVESTIGATION OF FENTANYL ANALOGUES (FA)
10:45 - 11:15	***Coffee Break***	

SCI-MIX
Chairs: Paola Gramatica and Ismail Yalcin

11:15 - 11:35	Major Talk	**Takashi Okada**. *School of Science and Technology, Kwansei Gakuin University, Japan.* PHARMACOPHORE IDENTIFICATION BY DATA MINING
11:35 - 11:55	Major Talk	**Konstantin Balakin**. *Chemical Diversity Research Institute, Khimki, Moscow Region, Russia.* NEW INSIGHTS FOR HERG INHIBITION USING MAPPING TECHNIQUES
11:55 - 12:10	Oral Communication	**Marcia Ferreira**. *Instituto de Quimica, Universidade Estadual de Campinas, Brazil.* CHEMOMETRIC INVESTIGATIONS OF MULTIDRUG RESISTANCE IN STRAINS OF THE PHYTOPATHOGENIC FUNGUS *PENICILLIUM DIGITATUM*
12:10 - 12:25	Oral Communication	**Melek Turker Sacan**. *Bogazici University, Inst?tute of Environmental Sciences, Istanbul, Turkey.* ASSESSMENT AND MODELLING OF THE TOXICITY OF SUBSTITUTED AROMATIC COMPOUNDS TO FIVE AQUATIC SPECIES
12:25 - 12:40	Oral Communication	**Sonja Nikolic**. *The Rugjer Boskovic Institute, Zagreb, Croatia.* THE ZAGREB APPROACH TO THE STRUCTURE-PROPERTY-ACTIVITY MODELING
12:40 - 14:00	***Lunch***	
14:00 - 14:45	Plenary Lecture	**Nikolay Zefirov**. *Department of Chemistry, Moscow State University, Moscow, Russia.* MEDICINAL CHEMISTRY AND MATHEMATICS
14:45 - 16:00	***Closing of the Symposium***	

List of Posters

1. <u>Roustam M. Abdeev</u>, Sergey A. Brouskin, Tatiana A. Nikolskaya, Eleonora S. Piruzian. *Center for Theoretical Problems of Physical-Chemical Pharmacology Russian Academy of Sciences.* SURVEY OF PSORIASIS CANDIDATE GENES BY USING BIOINFORMATICS.

2. <u>Vladimir Agadjanyan</u>, Eduard Oganesyan. *Pyatigorsk State Pharmaceutical Academy, Russia.* APPLICATION OF THE QUANTUM-CHEMICAL ANALYTICAL METHODS TO SUBSTANTIATE THE ANTIRADICAL ACTIVITY IN THE FLAVONE, CHALCONE AND CINNAMIC ACID HYDROXYDERIVATIVE SERIES.

3. <u>Ruslan Aliev</u>, Svetlana D.Demukhamedova, Irada N.Alieva, Niftali M.Godjaev. *Institute for Physical Problems, Baku State University, Baku, Azerbaijan.* STUDY OF PEPTIDE BOND DEFORMATION IN MODEL DIPEPTIDES BY THE SEMIEMPIRICAL QUANTUM CHEMISTRY METHODS.

4. <u>Sabiha Alper-Hayta</u>, Ilkay Yildiz, Ozlem Temiz-Arpaci, Esin Aki-Sener, Ismail Yalcin. *Faculty of Pharmacy, Ankara University, Ankara, Turkey.* FREE-WILSON STUDY ON SOME MICROBIOLOGICALLY ACTIVE BENZAZOLES.

5. <u>Elena Andreeva</u>, Oleg A. Raevsky. *Institute of Physiologically Active Compounds, Russian Academy of Sciences, Russia.* A DEVELOPMENT OF NEAREST NEIGHBOR METHOD FOR LIPOPHILICITY CALCULATION.

6. <u>Galina N. Apryshko</u>, Irina S. Golubeva, Olga V. Gorunova, Nadia P. Yavorskaya, Dmitry A. Filimonov, and Vladimir V. Poroikov. Russian *Cancer Research Center of Russian Academy of Medical Sciences.* COMPUTER-BASED PREDICTION OF ACTIVITY AND EXPERIMENTAL TESTING OF N-GLYCOSIDES OF SUBSTITUTED INDOLO[2,3-a]CARBAZOLES IN SEARCH FOR NEW ANTITUMOR SUBSTANCES.

7. <u>Galina N. Apryshko</u>, Dmitry A. Filimonov and Vladimir V. Poroikov. Russian *Cancer Research Center of Russian Academy of Medical Sciences.* COMPUTER-BASED SEARCH FOR NEW ANTITUMOR DRUGS USING THE RCRC RAMS DATABASE ON ANTITUMOR SUBSTANCES.

8. <u>Anatoly G. Artemenko</u>, Victor E. Kuz'min, Eugene N. Muratov, Leonid G. Gorb, Mohammad Qasim, and Jerzy Leszczynski. *A.V. Bogatsky Physico-Chemical Institute of the National Academy of Sciences of Ukraine.* 2D QSAR ANALYSIS OF NITROAROMATIC TOXICITY OF THE *TETRAHYMENA PYRIFORMIS.*

9. <u>Natalia V. Artemenko</u>, Tim James, Darren Fayne, David G. Lloyd. *School of Biochemistry & Immunology, Trinity College Dublin, Ireland.* A SIMPLE MODEL FOR THE PREDICTION OF BRAIN-BLOOD BARRIER PENETRATION BASED ON 2D AND SUBSTRUCTURAL DESCRIPTORS USING PLS STATISTICS.

10. <u>Konstantin Balakin</u>, Andrey A. Ivashchenko, Yan A. Ivanenkov. *Chemical Diversity Research Institute, Moscow, Russia.* SMART MINING: A NOVEL TECHNOLOGY FOR ADDRESSING ADME/TOX ISSUES.

11. <u>Anna Balandina</u>, Sergey S. Karamzin. *National Research Center for Hematology, Lomonosov Moscow State University, Moscow, Russia.* SPATIAL DYNAMICS OF THROMBIN GENERATION IN PLASMA.

12. <u>Florent Barbault</u>, Hu Rongjing, Delamar Michel, Doucet Jean-Pierre, Ruisheng Zhang. *ITODYS, Universite Paris, France.* QSAR STUDIES ON A SET OF PYRIMIDINE NUCLEOSIDES AS ANTI HIV-1 AGENTS.

13. <u>Elena V.Shilova</u>, Violetta L. Kovaleva, Athina Geronikaki, Dmitry Blinov, Vladimir V. Poroikov, Oxana V. Proskurina. *Mordovian Ogarev State university, Saransk, Russia.* ESTIMATION OF BIOLOGICAL ACTIVITY SPECTRA FOR 2-AMINOTHIAZOLES AND 2-AMINOBENZOTHIAZOLES.

14. <u>Anna Boiko</u>, Alexander V. Veselovsky, Vladlen S. Skvortsov, Oleg A. Raevsky. *Moscow State University, Russia.* INTERACTION OF THE ESTROGEN RECEPTOR WITH LIGAND FROM DIFFERENT CHEMICAL CLASSES: 3D-QSAR WITH COMSIA ANALYSIS.

15. <u>Oya Bozdag-Dundar</u>, Ilkay Yildiz. *Ankara University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, Ankara, TURKEY.* MOLECULAR STRUCTURE OF 3-[2-(4-CHLORO-PHENYL)-2-OXO-ETHYL]-5-(2,4-DICHLORO-THIAZOLE-5-YL-METHYLENYL)- THIAZOLIDINE-2,4-DIONE AS ANTIFUNGAL AGENT.

16. <u>Sergey A. Brouskin</u>, Roustam M. Abdeev, Tatiana A. Nikolskaya, Eleonora S. Piruzian. *Vavilov Institute of General Genetics, Russian Academy of Sciences.* COMPARISON OF PSORIASIS AND CROHN'S DISEASE PATHOLOGICAL PROCESSES AT THE LEVEL OF GENE NETWORK INTERACTIONS BY BIOINFORMATICS METHODS.

17. Alex Bunker. *Faculty of Pharmacy, University of Helsinki, Finland.* DETAILED MOLECULAR DYNAMICS SIMULATION OF PROTEIN-LIGAND INTERACTIONS FOR TWO IMPORTANT DRUG TARGETS: CATECHOL-O-METHYL TRANSFERASE (COMT) AND PROLYL OLIGOPEPTIDASE (POP).

18. Alexander Chernorudskiy, Alejandro Garcia, Eugene Eremin, Anastasia Shorina-Zhabereva, Ekaterina Kondratieva and Murat Gainullin. *Institute of Applied and Fundamental Medicine, Nizhny Novgorod State Medical Academy, Russia,* UBIPROT DATABASE AND ANALYSIS OF PROTEIN UBIQUITYLATION FOR APPLICATION IN PHARMACOLOGICAL RESEARCH.

19. Anton Chugunov, Valery N. Novoseletsky, Roman G. Efremov. *Institute of Bioorganic Chemistry, Russian Academy of Sciences, Russia.* QUALITY OF COMPUTER-BUILT MODELS OF MEMBRANE PROTEINS ASSESSED VIA KNOWLEDGE-BASED APPROACH.

20. James Devillers, Pascal Pandard, Eric Thybaud, and Anne Merle. *CTIS, France.* INTERSPECIES RELATIONSHIPS IN THE ACUTE TOXICITY OF XENOBIOTICS.

21. Alexander Dmitriev, Alexey A. Lagunin, Alexey V. Zakharov, Olga A Filz, Dmitry A Filimonov and Vladimir V Poroikov. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Moscow, Russia.* COMPUTER PREDICTION OF SUBSTRATES, INHIBITORS AND INDUCERS OF CYTOCHROME P450.

22. Jean-Pierre Doucet, Ana G. Maldonado, Michel Petitjean and Bo-Tao Fan. *ITODYS, Universite Paris, France.* USER CUSTOMIZED MOLECULAR DIVERSITY ANALYSIS USING THE MOLDIA SOFTWARE.

23. Pavel Dyachkov, Nina V. Kharchevnikova, Alexandr V. Dmitriev, Alexei V. Kuznetzov, Vladimir V. Poroikov. *N.S. Kurnakov Institute of General and Inorganic Chemistry, Russian Academy of Sciences, Moscow, Russia.* QUANTUM CHEMICAL SIMULATION OF CYTOCHROME P450 CATALYZED AROMATIC OXIDATION: METABOLISM, TOXICITY, AND BIODEGRADATION OFBENZENE DERIVATIVES.

24. Doga Erturk, Melek Turker Sacan. *Bogaziçi University, Instıtute of Environmental Sciences, Turkey.* MODELING THE TOXICITY OF BENZENE DERIVATIVES TO THE TADPOLE (*Rana japonica*) WITH THE CHARACTERISTIC ROOT INDEX AND SEMIEMPIRICAL MOLECULAR DESCRIPTORS.

25. Marcia Ferreira, Mirian C. S. Pereira, Rudolf Kiralj. *Instituto de Quimica, Universidade Estadual de Campinas, Brazil.* COMPUTATIONAL STUDY OF ARTEMISININ INTERACTION WITH HEME AND ITS POSTERIOR DECOMPOSITION.

26. Olga Filz, Alexey Lagunin, Dmitriy Filimonov, Vladimir Poroikov. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* COMPUTER-AIDED PREDICTION OF HERG CHANNELS REGULATION.

27. Eduard S. Fomin, Nikolay A. Alemasov, Artem S. Chirtsov, Arseniy E. Fomin. *Institute of Cytology and Genetics, Siberian Branch of Russian Academy of Sciences, Russia.* MOLKERN AS NEW EFFECTIVE ENGINE FOR DRUG DISCOVERY SOFTWARE.

28. Alexander Golbraikh, Hao Zhu, Kun Wang, Mei Wang Bell and Alexander Tropsha. *School of Pharmacy, University of North Carolina, USA* QSAR ANALYSIS OF IMBALANCED TOXICITY DATASETS USING K-NEAREST NEIGHBORS APPROACH.

29. Veniamin Grigor'ev, Eric E. Weber, Oleg A. Raevsky. *Institute of Physiologically Active Compounds, Russian Academy of Sciences, Russia.* CLASSIFICATION SAR MODELS OF CHEMICALS TOXICITY TO GUPPY, FATHEAD MINNOW AND RAINBOW TROUT ON THE BASIS OF STRUCTURE SIMILARITY.

30. Paola Gramatica and Ester Papa. *Department of Structural and Functional Biology, University of Insubria, Varese, Italy.* PBTs SCREENING BY MULTIVARIATE ANALYSIS AND QSAR MODELING.

31. Paola Gramatica, Pamela Pilutti and Ester Papa. *Department of Structural and Functional Biology, University of Insubria, Varese, Italy.* PREDICTION OF POLYCYCLIC AROMATIC COMPOUNDS' MUTAGENICITY AND GENOTOXICITY BY QSAR CLASSIFICATION MODELS.

32. Paola Gramatica and Huanxiang Liu. *Department of Structural and Functional Biology, University of Insubria, Varese, Italy.* QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDY OF SELECTIVE LIGANDS FOR THE THYROID HORMONE RECEPTOR.

33. Tad Hurst, *ChemNavigator, Inc, San Diego, USA.* 3DPL – AN ULTRA-FAST, INTERNET ENABLED PROTEIN-BASEDSEARCHING/DOCKING SYSTEM FOR BIOLOGISTS AND CHEMISTS.

34. <u>Svava Jonsdottir</u>, Irene Kouskoumvekaki, Olivier Taboureau, Niclas T. Hansen, Jens E. P. Larsen, Preeti Singh, Flemming S. Jorgensen and Soren Brunak. *Center for Biological Sequence Analysis (CBS), BioCentrum, Technical University of Denmark, Kongens Lyngby, Denmark.* CHEMOINFORMATICS TOOLS FOR PREDICTION OF VARIOUS PROPERTIES OF RELEVANCE TO DRUG DISCOVERY.

35. <u>David Karlinsky</u>, Michael E. Popov, Alexander P. Kaplun. *Moscow State University, Russia.* FINDING BINDING SITES FOR LOW-MOLECULAR LIGANDS ON THE SURFACE OF C1Q WITH MOLECULAR DOCKING.

36. <u>Veronika Khairullina,</u> Svetlana Kirlan, Anatolyi Gerchikov, Lidiya Tyurina, Garifa Garifullina, Alexander Kolbin. *Bashkir State University, Ufa, Russia.* MOLECULAR DESIGN Of ANTIOXIDANTS BY SUPERPOSITION OF THEIR EFFICIENCY AND LOW LEVEL OF A TOXICITY.

37. <u>Veronika Khayrullina</u>, Azat Mukhametov, Anatoliy Gerchikov, Garifa Garifullina, Lidija Turina, Felix Zarudiy. *Bashkir State University, 32, Frunze Str., Ufa, Russia.* THEORY OF PATTERN RECOGNITION IN THE DEVELOPMENT OF THE LOW-TOXIC ANTIINFLAMMATORY DRUGS.

38. <u>Olga V. Kharitonova</u>, Svetlana V. Kurbatova. *Samara State University, Samara, Russia.* QSAR STUDY OF NITROGEN-CONTAINING BIOLOGICAL ACTIVE HETEROCYCLES.

39. <u>Olga Koborova</u>, Alexey V. Zakharov, Alexey A. Lagunin, Dmitrii A. Filimonov, Alexander Kel, Fedor Kolpakov, Ruslan Sharipov and Vladimir V. Poroikov. *Moscow State University, Russia.* COMPUTER-AIDED PREDICTION OF PROMISING ANTI-TUMOR TARGETS TAKING INTO ACCOUNT INFORMATION ABOUT PROBABLE SIDE EFFECTS.

40. <u>Ludmila Kochetova</u>, Natalya Scheglova, Tatyana Kustova and Natalya Kalinina. *SEI HPE Ivanovo State University, Ivanovo, Russia.* ARYLSULFONILATIONOF α-AMINOACIDS: QSAR AND COMPUTER MODELING OF REACTION'S MECHANISM.

41. <u>Fedor Kolpakov</u>, Vladimir Poroikov, Ruslan Sharipov, Luciano Milanesi, and Alexander Kel. Institute *of Systems Biology, Design Technological Institute of Digital Techniques SB RAS, Russia.* CYCLONET - AN INTEGRATED DATABASE AND COMPREHENSIVE TOOL FOR COMPUTATIONAL PHARMACOLOGY.

42. <u>Nadezhda Kovalevskaya</u>, Yegor Smurnyy, Berry Birdsall, James Feeney, Vladimir Polshakov. *Center for Drug Chemistry, Moscow, Russia.* SOLUTION STRUCTURES AND DYNAMICS OF THE DIHYDROFOLATE REDUCTASE COMPLEXES WITH ANTIBACTERIAL DRUG TRIMETHOPRIM .

43. <u>Sunil Kulkarni</u>, Joel Paterson and Jiping Zhu. *Safe Environments Programme, Health Canada, Ottawa, Canada.* INTEGRATED APPROACH TO ASSESS THE DOMAIN OF APPLICABILITY OF SOME COMMERCIAL (Q)SAR MODELS .

44. <u>Mikhail Kumskov</u>, Nikolai S. Zefirov, Igor V. Svitanko, Alexei M. Zaharov, Ludmila A. Ponomareva. *Zelinsky Institute of Organic Chemistry, Moscow, Russia* UNIFIED REPOSITORY OF QSAR MODELS AND MOLECULES - THE SYSTEM OF PROPERTY PREDICTION OF THE CHEMICAL COMPOUND.

45. <u>Nikita Kuznetsov</u>, Vladimir Koval, Olga Fedorova and Yury Vorobjev. *Institute of Chemical Biology and Fundamental Medicine of SB RAN, Novosibirsk, Russia.* ANALYSIS OF PROCESSING MECHANISM OF HUMAN 8-OXOGUANINE-DNA GLYCOSYLASE ENZYME OF DNA REPAIR SYSTEM.

46. <u>Jens Larsen</u>, Olivier Taboureau, Flemming S. Jorgensen, Soren Brunak and Svava O. Jonsdottir. *Center for Biological Sequence Analysis (CBS), Technical University of Denmark, Denmark.* A MULTI-MUTAGENICITY ENDPOINT MODEL BUILT ON TOXICOPHORES.

47. <u>Olga Levtsova</u>, Konstantin V. Shaitan. *Moscow State University, Russi.* MOLECULAR DYNAMICS OF ZERVAMICIN II AND ITS MUTANTS IN DIFFERENT SOLVENTS.

48. <u>Nathalie Marchand-Geneste</u>, Alain Carpy, James Devillers and Jean-Marc Porcher. *Universite Bordeaux 1, Talence, France.* e-EXPERIMENTAL ENDOCRINE DISRUPTOR BINDING ASSAY (e^{3}dba) DATABASE.

49. <u>Nathalie Marchand-Geneste</u>, Thibault Laborde, James Devillers and Jean-Marc Porcher. *Universite Bordeaux 1, Talence, France.* E-ENDOCRINE DISRUPTING CHEMICAL DATABASES FOR DERIVING QSAR MODELS.

50. <u>Marica Medic-Saric</u> and Vesna Rastija. *Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia.* QSAR STUDY OF ANTHOCYANINS, ANTHOCYANIDINS AND CATECHIN AS INHIBITORS OF LIPID PEROXIDATION.

51. <u>Marica Medic-Saric</u>, Ivona Jasprica, Zeljko Debeljak and Ana Mornar. *Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia.*QSAR STUDY OF FLAVONOIDS' IMMUNOMODULATORY ACTIVITY.

52. <u>Svetlana Mikhaylichenko</u>, Vladimir V. Dounin, Vladimir N. Zaplishny, Alexey A. Chesnyuk. *University of Toronto Scarborough, Toronto, Canada.* SYNTHESIS AND POTENTIAL BIOACTIVITY OF NEW PYRAZOLYL-SYM-TRIAZINE DERIVATIVES.

53. <u>Svetlana Mikhaylichenko</u>, Vladimir V. Dounin, Vladimir N. Zaplishny. *University of Toronto Scarborough, Toronto, Canada.* EFFECTIVITY OF THE PASS PROGRAM IN PREDICTING OF BIOLOGICAL ACTIVITY FOR SYM-TRIAZINE DERIVATIVES.

54. <u>Oleg A. Raevsky</u>, Dmitry Moiseev, Olga E. Raevskaya, Elena V. Bovina, Klaus-Jorden Schaper. *Institute of Physiologically Active Compounds, Russian Academy of Sciences, Russia.* A GENERAL PHYSICOCHEMICAL QSPR MODELS OF CHEMICALS PARTITIONING IN SYSTEM SOLVENT–WATER–AIR.

55. <u>Umpei Nagashima</u>, Sumie Tajima, Makoto Haraguchi. *Research Institute for Computational Science, National Institute of Advanced Industrial Science and Technology, Japan.* ENERATION OF NOVEL EQUATION FOR THE LOGALITHMIC RATIO OF THE OCTANOL-WATER PARTITION FUNCTION: LOG P ESTIMATION USING PERCEPTRON TYPE NEURAL NETWORK - DEVELOPMENT OF COMPUTER AIDED MOLECULAR DESIGN SYSTEM: MolWorks.

56. <u>Ludmila Naumenko</u>, Pavel M. Vassiliev, Alexander A. Spasov. *Volgograd State Medical University, Volgograd, Russia.* THE «MICROCOSM» INFORMATION TECHNOLOGY SYSTEM FOR BIOLOGICAL ACTIVITY PREDICTION OF ORGANIC COMPOUNDS: THE PREDICTION OF HEMORHEOLOGICAL ACTIVITY.

57. <u>Alexey Neznanov</u>, Victor A. Kokhov, Sergey V. Tkachenko. *Moscow Power Engineering Institute (Technical University), Moscow, Russia.* NEW MATHEMATICAL MODELS AND SOFTWARE TOOLS FOR COMPLEXITY AND SIMILARITY ANALYSIS OF MOLECULAR GRAPHS.

58. <u>Lyudmila Ognichenko</u>, Victor E. Kuz'min, Anatoly G. Artemenko. *Physico-Chemical Institute of the National Academy of Sciences of Ukraine, Odessa, Ukraine.* NEW STRUCTURAL DESCRIPTORS OF MOLECULES ON THE BASIS OF SYMBIOSIS OF THE INFORMATIONAL FIELD MODEL AND SIMPLEX REPRESENTATION OF MOLECULAR STRUCTURE .

59. <u>Igor Orshanskiy</u>, Ksenia B. Tereshkina, Yegor V. Tourleigh, Konstantin V. Shaitan. *Moscow State University, Russia.* SECONDARY STRUCTURE OF ISGNAL PEPTIDES OF NS2 PROTEIN OF HCV AND ITS STABILITY.

60. <u>Evgenija Osipova</u>, Vladimir Kuz'min, Natalija Kuz'mina. *The State Scientific Research Institute of Organic Chemistry and Technology, Russia.* THE GEOMETRICAL APPROACH TO FORECASTING EFFECTIVE NARCOTIC DEPENDENCE REMEDIES.

61. <u>Ruslan Ostapchuk</u>, Vladimir Zatsepin, Vladimir Ivanchenko. *All-Russian Institute of Scientific and Technical Information, Moscow, Russia.* DEVELOPMENT OF THE INFORMATION SUBSYSTEM ON MATHEMATICAL MODELS IN TOXICOLOGY AND PHARMACOLOGY.

62. <u>Annick Panaye</u>, James DEVILLERS, Jean-Pierre DOUCET, Nathalie MARCHAND- GENESTE and Jean-Marc PORCHER. *ITODYS, Universite Paris, France.* DECISION TREES FOR CLASSIFICATION OF ENDOCRINE DISRUPTORS.

63. <u>G.L. Perlovich</u>, N.N. Strakhova, V.P. Kazachenko, K.-J. Schaper, O.A. Raevsky. *Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds, Russian Academy of Sciences, Chernogolovka, Russia.* SULFONAMIDES AS A SUBJECT TO STUDY MOLECULAR INTERACTIONS IN CRYSTALS AND SOLUTIONS: SUBLIMATION, SOLUBILITY, SOLVATION, DISTRIBUTION.

64. <u>Pavel Polishchuk</u>, Victor E. Kuz'min, Anatoly G. Artemenko, Svetlana Yu. Makan, Sergey A. Andronati. *A.V. Bogatsky Physico-Chemical Institute of the National Academy of Sciences of Ukraine.* NEW APPROACH OF QUANTITATIVE INTERPRETATION OF DECISION TREE MODELS AND ITS APPLICATION FOR SOLUTION OF QSAR TASKS.

65. <u>Elena Ponomarenko</u>, Andrey V. Lisitsa, Alexander I.Archakov. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* TEXT MINING TOOLS IN ANALYSIS OF HIGH-THROUGHPUT DATA.

66. <u>Oxana Proskurina</u>, Violetta L. Kovaleva, Athina Geronikaki, Elena V.Shilova, Vladimir V. Poroikov, Dmitry S. Blinov. *National Research Center on Biologically Active Compounds, Staraya Kupavna, Russia.* ANTIINFLAMMATORY AND ANALGESIC ACTIVITY OF NEW DERIVATIVES OF 2-AMINOBENZOTHIAZOLE.

67. <u>Mikhail Pyatnitskiy</u>, Andrey V Lisitsa, Alexander I Archakov. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* PREDICTION OF PROTEIN-PROTEIN INTERACTIONS: PHYLOGENETIC PROFILES AND CLUSTER ANALYSIS.

68. <u>Timothy Pyrkov</u>, Yuri A. Kosinsky, Roman G. Efremov. *Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow,, Russia.* LIGAND-SPECIFIC SCORING FUNCTIONS: IMPROVED RANKING OF DOCKING SOLUTIONS.

69. <u>Sergei Rahmanov</u> and Vsevolod J. Makeev. *GosNIIGenetika, Moscow, Russia.* KNOWLEDGE-BASED POTENTIALS ALLOW PLACEMENT OF WATER AND IONS IN MACROMOLECULAR INTERFACES.

70. <u>Vasily Ramensky</u>, Natalia I. Zaitseva, Anatoly R. Rubinov, Victor V. Zosimov. *Institute of Molecular Biology of Russian Academy of Sciences, Applied Acoustics Research Institute, Russia.* COMPUTATIONAL DRUG DESIGN BASED ON A NOVEL APPROACH TO LOCAL SIMILARITY OF PROTEIN BINDING SITES.

71. <u>Alexander Rasdolsky</u>, Gerasimenko V.A., Filimonov D.A., Poroikov V.V., Weber E.J., Raevsky O.A. *Institute of Physiologically Active Compounds, Russian Academy of Sciences, Russia.* CLASSIFICATION OF ORGANIC COMPOUNDS BY THEIR MODE OF TOXIC ACTION ON *DAPHNIA MAGNA* BY THREE DIFFERENT METHODS.

72. <u>Joseph Rebehmed</u>; Florent Barbault; Catia Teixeira; Francois Maurel; BoTao Fan. *ITODYS, Universite Paris, Paris, FRANCE.* COMFA/COMSIA AND 2D QSAR STUDIES OF DIARYLPYRIMIDINE HIV-1 REVERSE TRANSCRIPTASE INHIBITORS.

73. <u>Irine Sarvilina</u>, Anna Krishtopa, Yuri Maklyakov, Daria Gordienko. *South Scientific Centre of Russian Academy of Sciences, Russia.* AIDA: SOFTWARE FOR MODELING OF GLUCOSE-INSULIN INTERACTIONS IN PATIENTS WITH DIABETES MELLITUS TYPE 1.

74. <u>Anastasia Sergeiko</u>, Alla Stepanchikova, Boris Sobolev, Sergey Zotchev, Dmitry Filimonov, Alexey Lagunin, Vladimir Poroikov. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* COMPUTER-AIDED DESIGN OF POLYKETIDES WITH THE REQUIRED PROPERTIES.

75. <u>Maria Shadrina</u>, Bella Grigorenko, Alexander Nemukhin. *Moscow State University, Russia.* MODELING CATALYTIC MECHANISMS OF HYDROLYSIS OF NUCLEOSIDE TRIPHOSPHATES (NTP).

76. <u>Ruslan Sharipov</u>, Yuriy Kondrakhin, Fedor Kolpakov. *Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia.* EXPRESSION ANALYSIS OF NF-κB-REGULATED GENES IN BREAST CANCER. META-ANALYSIS OF MICROARRAY DATA.

77. <u>Dmitry Shcherbinin</u>, Alexander V. Veselovsky. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* COMPUTER MODELING OF INTERACTIONS BETWEEN HIV-1 INTEGRASE AND STYRYLQUINOLINES.

78. <u>Tianxiang Shen</u>, Feng Li, Jianhua Yao. *Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, China.* CISOC-CHIMS: CHEMICAL INFORMATION MANAGEMENT SYSTEM.

79. <u>Dmitry Shulga</u>, Alexander Oliferenko, Sergey Pisarev, Vladimir Palyulin, Nikolay Zefirov. *Institute of Physiologically Active Compounds, Russian Academy of Sciences, Russia.* AST TOOLS FFOR CALCULATION OF ATOMIC CHARGES WELL SUITED FOR DRUG DESIGN.

80. <u>Irina Shutova</u>, Olga V. Mel'nik, Vladimir P. Golubovich. *The Institute of Bioorganic Chemistry of the National Academy of Sciences of Belarus.* COMPUTER DESIGN OF LOW-MOLECULAR COMPOUNDS OF A PEPTIDE KIND – ACTIVATORS OF DIRECTIONAL DIFFERENTIATION OF STEM CELLS.

81. <u>Natalya Skvortsova</u>, Sergey M. Novikov. *A.N. Sysin Research Institute of Human Ecology and Environmental Healt, Russian Academy of Medical Sciences, Russia.* DATABASE OF ENVIRONMENT POLLUTANTS FOR HEALTH RISK ASSESSMENT AND PROGNOSIS OF THE NEW CHEMICAL SUBSTANCE UNKNOWN CHARACTERISTICS.

82. <u>Vladlen Skvortsov</u>, Alexei A. Lagunin, Dmitrij A. Filimonov, Oleg A. Raevsky. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* INTEGRATED SYSTEM FOR PREDICTION OF THE BIOLOGICAL ACTIVITY, TOXICITY AND PHYSICOCHEMICAL PROPERTIES.

83. <u>Egor Sobolev</u>, Oleg V. Sobolev and Dmitry A. Tikhonov. *Institute of Mathematical Problems of Biology, Russia.* ONLINE RESOURCE FOR THEORETICAL STUDY OF HYDRATION OF BIOPOLYMERS.

84. <u>Mariya Solodilova</u>, Alexey V. Polonikov, Vladimir P. Ivanov. *Kursk State Medical University, Kursk, Russia.* APPLICATION OF SET-ASSOCIATION APPROACH TO SEARCHING FOR SINGLE NUCLEOTIDE POLYMORPHISMS OF "ENVIRONMENTAL GENES" RESPONSIBLE FOR COMPLEX DISEASE SUSCEPTIBILITY.

85. <u>Milan Soskic</u>. *Faculty of Agriculture, University of Zagreb, Zagreb, Croatia.* INHIBITION OF THE HILL REACTION BY ORGANOPHOSPHORUS INSECTICIDES AND SIMPLE ORGANIC MOLECULES: A QSAR STUDY.

86. <u>Mikhail Stryapan</u>, Mikhail Islyaikin. *Ivanovo State University of Chemistry and Technology, Ivanovo, Russia.* ELECTRON AND SPATIAL STRUCTURE PECULIARITIES OF MACROHETEROCYCLIC COMPOUDS.

87. <u>Boris Sukhov</u>, Aidar T. Gubaidullin*, Igor A. Litvinov*, Svetlana F. Malysheva, Alexander V. Vashchenko, Vladimir I. Smirnov, Yulia A. Grigor'eva, Nina K. Gusarova, Boris A. Trofimov. *Institute of Organic and Physical Chemistry of RAS, Russia.* "PROPELLER" CHIRALITY AND SPONTANEOUS RESOLUTION OF TRISTYRYLPHOSPHINE CHALCOGENIDES INTO ENANTIOMERS.

88. <u>Alexey Sulimov</u>, Vladimir B Sulimov, Alexey N Romanov, Fedor V Grigoriev, Olga A Kondakova, Peter A Bryzgalov and Denis A Ostapenko. *Reseach Computer Center at M.V. Lomonosov Moscow State University, Moscow, Russia.* WEB-ORIENTED SYSTEM KEENBASE FOR NEW DRUGS DESIGN.

89. <u>Catia Teixeira</u>, Florent Barbault, Joseph Rebhemed, Karen Urgin, Nawal Serradji, Francoise Heymans, Francois Maurel. *ITODYS, Universite Paris, France.* SYSTEMATIC CONFORMATIONAL STUDIES OF NEW POTENTIAL HIV-1 ENTRY INHIBITORS TARGETING CD4-G.

90. <u>Ozlem Temiz-Arpaci</u>, Betul Tekıner-Gulbas, Ilkay Yıldız, Esin Akı-Sener, Ismail Yalcın. *Faculty of Pharmacy, Ankara University, Ankara, Turkey.* 2D-QSAR STUDIES OF SOME BENZOXAZOLES AGAINST *S. aureus.*

91. <u>Ksenia Tereshkina</u>, Konstantin V. Shaitan and Eduard V. Tereshkin. *Moscow State University, Russia.* PROGRAM-APPARATUS SOLUTIONS AT ION CHANNEL COMPUTER MODELLING BY THE EXAMPLE OF GABA-A CHANNEL.

92. <u>Olga Tikhonova</u>, Vladlen S. Skvortsov, Veniamin Yu. Grigor'ev and O.A.Raevsky. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* 3D-QSAR MODELS OF IRREVERSIBLE AND REVERSIBLE INHIBITORS OF ACETYLCHOLINESTERASE.

93. <u>Lidija Tyurina</u>, Liana Sementeeva, Alexander Kolbin, Rail Valitov. *State Petroleum Technological University, Ufa, Russia.* BANKS of MODELS AND KNOWLEDGE BASE OF RESULTS SAR - BASIS OF MOLECULAR DESIGN AND PREDICTION.

94. <u>Francisco Torrens</u> and Gloria Castellano. *Institut Universitari de Ciencia Molecular, Universitat de Valencia, Edifici d'Instituts de Paterna, Spain.* CHARGE-TRANSFER INDICES FOR MULTIFUNCTIONAL AMINO ACIDS.

95. <u>Daria Tsareva</u>, Alexander A. Oliferenko, Vladimir A. Palyulin, Nikolai S. Zefirov. *Moscow State University, Russia.* COMPUTATIONAL DESIGN OF ACETYLCHOLINESTERASE INHIBITORS BY FRAGMENT LINKING.

96. <u>Irina Tsygankova</u>. *Institute of Experimental and Theoretical Biophysics RAS.* FRAGMENTAL DESCRIPTORS IN QSAR MODELS: SELECTION AND INTERPRETATION.

97. <u>Nikolay Tverdokhlebov</u>, Alexey Zhuchkov, Boris Alperovich, Alexander Kravchenko. *Telecommunication Centre "UMOS", Russia.* GRID-SERVICES FOR EVIDENCE-BASED PHARMACOLOGY.

98. <u>Andrew Voronkov</u>, Igor I. Baskin, Vladimir A. Palyulin, Nikolai S. Zefirov. *Moscow State University, Russia.* COMPUTER BASED DRUG DESIGN OF LIGANDS FOR CRD DOMAIN OF FRIZZLED RECEPTOR AS PROSPECTIVE ANTICANCER COMPOUNDS.

99. <u>Dmitry Yakovlev</u>, Pavel M. Vassiliev, Vladimir V. Poroikov, Dmitry A. Filimonov, Alexander A. Spasov, Maxim V. Chernikov and Alexandra N. Zheltukhina. *Volgograd Science Centre, Russian Academy of Medical Sciences, Russia.* THE 3D-MODELING, COMPUTER PREDICTION AND EXPERIMENTAL TESTING OF 5-HT$_3$-ANTISEROTONIN ACTIVITY OF NEW CHEMICAL COMPOUNDS.

100. <u>Alexey Zakharov</u>, Alexey Lagunin, Dmitry Filimonov and Vladimir Poroikov. Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia. QSAR MODELLING OF ACUTE TOXICITY IN THE FATHEAD MINNOW.

101. <u>Spela Zuper</u>, Sabina Passamonti, Marjana Novic. *Laboratory of Chemometrics, National Institute of Chemistry, Ljubljana, Slovenia.* QSAR STUDY OF BILITRANSLOCASE TRANSPORT ACTIVITY.

102. <u>Sehan Lee</u>. *Yonsei University, Seoul, Korea.* IN SILICO PREDICTION OF CYP450-MEDIATED METABOLISM PROFILE.

103. <u>Youngyong In</u>, Sung Kwang Lee, Pil Je Kim, Kyoung Tai No. *Bioinformatics and Molecular Design Research Center, Seoul, Korea.* A QSAR STUDY OF ACUTE TOXICITY TO FATHEAD MINNOW (*PIMEPHALES PROMELAS*).

104. <u>Liana Sementeeva</u>, Lidija Tyurina, Alexander Kolbin, Olga Tyurina, Tatjana Solominova, Vladislav Kirlan. *Ufa State Petroleum Technological University, Ufa, Russia.* THE INTEGRATED COMPUTER SYSTEM «SARD-TOX–OVOLMEP» FOR DESIGN AND PREDICTION OF CHEMICAL COMPOUNDS WITH AN ESTIMATION of VARIOUS ASPECTS Of THEIR ACTION.

105. <u>Dmitry S. Druzhilovskiy</u>, Dmitry F. Filimonov, Alexey A. Lagunin, Tatyana A. Gloriozova, Vladimir V. Poroikov. *Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Russia.* COMPUTER-ASSISTED PREDICTION OF CYTOTOXICITY FOR HIV-1 INTEGRASE INHIBITORS.

NEW PUBLIC DATA & INTERNET RESOURCES IMPACTING PREDICTIVE TOXICOLOGY

Ann Richard[1], Maritja Wolf[2], Clar-Lynda Williams-Devane[3], Richard Judson[1]

[1]National Center for Computational Toxicology, US EPA, RTP, NC; [2]Lockheed Martin, Contractor to the US EPA, RTP, NC; [3]Bioinformatics Graduate Program, NC State University, Raleigh, NC 27599; E-mail: richard.ann@epa.gov

High-throughput screening (HTS) technologies, along with efforts to improve public access to chemical toxicity information resources and to systematize older toxicity studies, have the potential to significantly improve predictive capabilities in toxicology. Important developments include: 1) large and growing public resources that link chemical structures to biological activity and toxicity data in searchable format, and that offer more nuanced and varied representations of activity; 2) standardized relational data models that capture relevant details of chemical treatment and effects of published in vivo experiments; and 3) the generation of large amounts of new data from public efforts that are employing HTS technologies to probe a wide range of bioactivity and cellular processes across large swaths of chemical space. By annotating toxicity data with associated chemical structure information, these efforts link data across diverse study domains (e.g., 'omics', HTS, traditional toxicity studies), toxicity domains (carcinogenicity, developmental toxicity, etc) and database sources (EPA, FDA, NCI, DSSTox, PubChem, GEO, ArrayExpress, etc.). Public initiatives (such as ToxML) are developing systematized data models of toxicity study areas and introducing standardized templates, controlled vocabularies, hierarchical organization, and powerful relational searching capability across captured data. Cheminformatics and data models, in turn, are providing the underpinning for the large public HTS efforts of the NIH Molecular Libraries Initiative, as well as new toxicity-targeted HTS programs within the EPA and the NIEHS National Toxicology Program. These initiatives are turning the structure-activity paradigm on its head, using chemicals to probe biological space and generating "biological profiles" of chemicals that, along with chemical structure considerations, offer the promise of providing richer, and more relevant and predictive associations to in vivo responses. *This work was reviewed by EPA and approved for publication, but does not necessarily reflect EPA policy.*

INTERNET RESOURCES INTEGRATING MANY SMALL-MOLECULE DATABASES

Igor V. Filippov, Markus Sitzmann, <u>Marc C. Nicklaus</u>

Computer-Aided Drug Design (CADD) Group, Lab. of Medicinal Chemistry, Center for Cancer Research, National Cancer Institutes, National Institutes of Health, DHHS. NCI-Frederick, Bldg. 376, 376 Boyles St., Frederick, MD 21702, USA; E-mail: mn1@helix.nih.gov

We present new tools and services developed by the CADD Group, NCI, NIH, in the context of our chemoinformatics and drug development work, made available on the CADD Group's web site http://cactus.nci.nih.gov. These tools are designed for searching for structures in very large databases of small molecules. One of them is a web service for very rapid structure lookup in an aggregated collection of currently more than 80 databases comprising more than 27 million unique structures. This Chemical Structure Lookup Service (CSLS) contains toxicology-related databases, catalogs of commercially available samples, drugs, assay results data sets, and databases in several other categories. CSLS allows the user to find out very rapidly in which one(s) of all these databases a given structure occurs independent of the representation of the input structure, by making use of InChIs as well as CACTVS hashcode-based identifiers. These new, calculable, identifiers are designed to take into account tautomerism, different resonance structures drawn for charged species, and presence of additional fragments. They make possible fine-tunable yet rapid compound identification and database overlap analyses. We also present a powerful substructure search tool, implemented in the form of a web service. Using a CACTVS-based search engine operating in distributed mode across a small Linux cluster, it allows for powerful substructure searches in "web time" in databases of millions of compounds.

A VIRTUAL FILE SYSTEM FOR THE PUBCHEM CHEMICAL STRUCTURE AND BIOASSAY DATABASE

Wolf-D. Ihlenfeldt

Xemistry GmbH, Auf den Stieden 8, D-35094 Lahntal, Germany; E-mail: wdi@xemistry.com

The PubChem chemical structure and bioassay database (pubchem.ncbi.nlm.nih.gov) has established itself as one of the premier information sources for chemical structures and assay data accessible via the Internet. PubChem provides a convenient interactive Web interface for the execution and result display of standard structure and text-based queries. However, the capabilities for formulating complex queries, and to access and download specific data sets, are limited. Because of the necessity to integrate PubChem into the existing Entrez framework, instead of designing a new streamlined interface for handling its peculiar data content, in many cases queries are awkward to set up and execute if they become more complex.

Recently, PubChem has published a specification for its Power User Gateway (PUG). This interface allows enterprising users to retrieve selected PubChem data via an XML-based Web service. The text-oriented cluster of Entrez databases has been accessible for a long time via similarly structured access modules termed Eutils. Eutils also support access to the text components of the PubChem deposition structure and standardized compound databases. Finally, individual record downloads and similar operations are possible via creatively crafted Web URLs, mimicking those encountered in interactive sessions.

By virtue of these mechanisms, PubChem does in principle provide a larger degree of programmatical accessibility than other chemistry Web databases. Nevertheless, they are extremely difficult to use in their native, raw form.

In order to address this difficulty, we have implemented a virtual file I/O module for the Cactvs Chemoinformatics Toolkit. It provides access to the PubChem compound database as a virtual file. The supported feature set starts with simple record-based I/O and extends to the execution of structure queries of higher complexity than possible via the PUG. Users of the toolkit may now script the same toolkit commands for the PubChem database as they can for a local read-only structure file. Behind the scenes, the I/O module leverages the described three access mechanisms in an optimized fashion, re-routing as many of the operations needed to perform the command to the PubChem servers as possible, but also transparently utilizing downloaded records for local processing in case operations are requested which exceed the capabilities of the systems operating on the PubChem site. A strength of our implementation is that it fully understands the native ASN.1 data specification for the database contents and is therefore not restricted to working on the sometimes coarse approximations of structure configuration which are available as SD-file records.

With this work, we think we are presenting the first usable solution for scripted access to the PubChem database. Given its emerging importance for drug-related research, we hope that this software will be generally useful for a broad audience.

CYTOCHROME P450 KNOWLEDGEBASE FOR DRUG DISCOVERY

Andrey V. Lisitsa, Alexander I. Archakov

Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., Pogodinskaya str. 10, Moscow, 119121, Russia; E-mail: fox@ibmh.msk.su

Side effects discovered in many cases withdraw drug candidates from the late phases of discovery process. Some of the side effects can be foreseen at early stages by analysis of the interactions between putative lead and cytochromes P450. As is known these enzymes mediates the biotransformation of over 60% of drugs. While application of SAR, QSAR and other predictive techniques the quality of raw data is of much importance. The statistical analysis case be easily hampered by spontaneous fluctuation of experimentally collected results, especially is these were collected by different research laboratories. In order to prepare the high-quality database for investigation of interaction between compounds and cytochromes P450 since 1999 we maintain the cytochrome P450 knowledgebase. The knowledge-based approach has been taken to provide the automated means for data retrieval. Medline abstracts are used as the primary source for selection of new information. The relatedness of the newly coming document to the knowledgebase topic is assessed based on the frequency of discriminating terms. Controlled vocabularies and linguistic patterns are applied to the relevant document to extract information and to insert into the database fields. Each knowledgebase record is focused upon the published experimental study of interaction between the specified form of cytochrome P450 and a low molecular weight effectors. The effectors are classifies into substrates, inducers and inhibitors. Records are equipped with appropriate references. Automatically created records are manually verified by the experts in the field of cytochromes P450 before storage into the knowledgebase. To the current moment about three thousands facts were thoroughly collected and made publicly available on Web (http://cpd.ibmh.msk.su). Possibilities of application of this data in the field of drug discovery are discussed.

WORKFLOW ENVIRONMENT FOR ESTIMATING BIOLOGICAL EFFECTS

Alexander J. Kos

AKos GmbH, Austr. 26, D-79585, Steinen, Germany; E-mail: akos@akosgmbh.de

In silico and experimental results are more reliable if generated with different methods. If the results converge, the reliability increases. For instance, carcinogenicity is tested against a variety of cell lines. Predicting active compounds works better when considering a set of biological effects.

Similar, we want to develop a program that predicts reliable biological effects of compounds by using a set of prediction tools and databases. If all the methods converge, we can trust the results. The answer is a spreadsheet giving an overview of the results of all the methods.

Integrating all available resources that predict or contain biological effects is a monumental task way beyond what we can do. We will illustrate our idea on practical examples of the overall workflow. We will use Microsoft's Window Workflow Foundation using as an object the program PASS. In another example, we will integrate different data sources using Elsevier MDL's Isentris.

Our vision is to build a container for a workflow of estimating toxicity on existing software like Windows Workflow Foundation, Pipeline Pilot, InforSense KDE, Isentris, and an Open Source Project software that allows the user to attach without any programming licensed software packages and databases to predict with a one button approach toxicity.

INTERNET RESOURCES IN GPCR MODELLING

Anil K. Saxena, Intekhab Alam, Anshuman Dixit, Mridula Saxena

Division of Medicinal and Process Chemistry, Central Drug Research Institute (CDRI), Chattar Manzil Palace, P.O.Box 173, Lucknow 226001, India; E-mail: anilsak@gmail.com

G-Protein coupled receptors (GPCRs) belong to the superfamily of integral membrane proteins characterized by seven transmembrane helix (TM1-TM7).They are encoded by about 5% of the total human genome and comprises 45% of the available drug targets. The three dimensional (3-D) structure of protein is important for drug design. However the progress in solving the GPCR structures has been slow. Out of the 907 human GPCRs reported till date, the 3-D structure has not been solved for any of them using standard experimental techniques viz. NMR, X-ray crystallography. In view of this, in-silico approaches in solving the protein structure from amino acid sequence have been used. The algorithms available for structure prediction are homology modeling/comparative modeling, threading, and ab initio folding. Among these methods, the homology modeling which builds the 3-D models of the proteins by aligning the target sequence to an evolutionary related sequence (template) has been widely used. The only GPCR whose 3-D structure from X-ray crystallography is available and is commonly used in homology modeling of GPCR is bovine rhodopsin. The derived homology models of GPCRs are validated and refined on the basis of biochemical, pharmacological as well as mutation data. Several web servers and computer programs are available that automate the homology modeling process. Some of these include Modeler, Swiss-Model server, Homer etc. These tools have been used to generate reliable homology models of human histamine H1 receptor (HRH1) and thrombin receptor (PAR-1) which explain the binding mode of the standard antagonists of these receptors and may be useful in designing their novel antagonists.

PREDICTION OF BIOLOGICAL ACTIVITY VIA INTERNET. MEDICINAL CHEMIST'S POINT OF VIEW

Athina Geronikaki[1], Dmitry Druzhilovsky[2], Alexey Zakharov[2], Vladimir Poroikov[2]

[1]School of Pharmacy, Department of Pharmaceutical Chemistry, Aristotle University of Thessaloniki, Thessaloniki, Greece; [2]Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: geronik@pharm.auth.gr

Chemical compounds synthesized in the framework of academic research are considered as a valuable source for discovery of new leads (C&EN, Apr. 16, 2007, p.42). However, being limited in resources, the medicinal chemist from University should apply a rational design in the studies, to obtain molecules with the appropriate pharmacodynamics and pharmacokinetics. Rational design of drug-like compounds is based on computer-aided estimation of different characteristics, including biological activity, physical-chemical properties, etc. In addition to the numerous commercially-available computer programs, there exist some computational tools freely available via Internet. Among them, the programs for prediction of melting point, boiling point, critical temperature, critical pressure and other physical-chemical properties (http://www.molecularknowledge.com/Online/Estimation/ online1.htm), solubility, lipophylicity and pKa (http://vcclab.org/lab/alogps/start.html), some kinds of biological activity (GPCR ligands, ion channel modulators, kinase inhibitors, nuclear receptor ligands – http://www.molinspiration.com/cgi-bin/properties), irritant, mutagenic, tumorigenic and reproductive effects (http://www.organic-chemistry.org/prog/peo/ index.html), drug-likeness (http://www.molsoft.com/mprop/), etc. About 2500 pharmacotherapeutic effects, molecular mechanisms of action, adverse & toxic effects, metabolic terms can be predicted with PASS INet (http://www.ibmc.msk.ru/PASS). We have analyzed the applicability of computational tools available via Internet for solution of different tasks in the drug discovery process, compared the results of prediction of the same property obtained with different tools, considered the advantages and drawbacks of different approaches. Some examples of new biologically active compounds found on the basis of computer predictions will be presented, including local anesthetics, NSAIDs, anxiolytics, cognition enhancers, antituberculosic agents.

EFFECT OF DESCRIPTOR SELECTION: COMPARING LOGD TO LOGP IN DRUG-LIKENESS PROFILING

Greg Pearl, Sanji Bhal, Ian Peirson, Karim Kassam

ACD/Labs, 110 Yonge Street, 14th Floor, Toronto, Ontario Canada M5C 1T4

The much-publicised "Rule of 5" has been widely adopted amongst the pharmaceutical industry as the first step in the virtual screening of compound libraries, in an effort to eliminate lead candidates that are deemed to have poor physicochemical properties. Although LogP is a useful descriptor, it fails to take into account any variation in lipophilicity of a drug due to the potential ionisation at a key biological pH. Given that >95% of commercial pharmaceuticals contain an ionisable moiety, we propose that LogD could be used as a alternate descriptor for lipophilicity in the Rule of 5 in order to reduce the number of potential false-positives that are eliminated in screening. With the advances in *in-silico* prediction capability, accurate values of pKa, LogP and LogD can be rapidly computed and assessed. To this end, the adapted Rule of 5 was applied to a series of commercial compound libraries and notable improvements in pass rate were attained.

PHARMAEXPERT: ESTIMATING DRUG-DRUG INTERACTIONS AND FINDING COMPOUNDS WITH MULTIPLE MECHANISMS OF ACTION

Alexey Lagunin, Dmitry Filimonov, Tatyana Gloriozova, Vladimir Poroikov

Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: alexey.lagunin@ibmc.msk.ru

On the basis of structural formula of chemical compound, computer program PASS predicts ~2800 kinds of biological activity including ~370 pharmacotherapeutic effects, ~2300 mechanisms of action, ~40 toxic/side effects and ~110 metabolic terms with average accuracy ~93% (http://www.ibmc.msk.ru/PASS). PharmaExpert is the program, which interprets PASS predictions taking into consideration known mechanism-effect(s) and effect-mechanism(s) relationships, and provides a flexible mechanism for selection of compounds with desirable but without unwanted kinds of biological activity in libraries of chemical compounds. Knowledgebase of the current version of PharmaExpert (March 2007) contains information about 6003 names of biological activity, 9101 their synonyms, 5108 mechanisms of action, 634 pharmacotherapeutic effects, 61 adverse effects/toxicities, 227 metabolic terms, and 4998 relationships between these terms. Since PASS predictions contain a plethora of information about probable biological actions of chemical compounds, using PharmaExpert it is possible to select compounds with the required multiple mechanisms of action. Also, analyzing PASS predictions it is possible to estimate the probable drug-drug interactions, when a combination of compounds might lead to additive, synergistic or antagonistic effects. The last option may be particularly important in application to herbal medicines contained the mixtures of natural compounds. With PASS and PharmaExpert, most probable specific effects of separate substances can be analyzed as well as interactions between them. Application of PASS and PharmaExpert will be presented on example of finding of antihypertensive substances with combined mechanisms of action and dual COX/LOX inhibitors. Possibilities to find agents with multiple mechanisms of action in databases of commercially available samples will be analyzed.

IMPROVING THE PREDICTIVE PERFORMANCE OF AROMATIC AMINE MUTAGENICITY IN DEREK FOR WINDOWS THROUGH THE ANALYSIS OF PROPRIETARY DATA

Russell T. Naven, Laura Gibson, Richard V. Williams

Lhasa Limited, 22-23 Blenheim Terrace, Woodhouse Lane, Leeds, LS2 9HD, UK; E-mail: russell.naven@lhasalimited.org

Aromatic amines as a chemical class are widely associated with mutagenicity. Mutagenic potential is often attenuated, however, by structural or physicochemical factors which in practice lead to inactivity in the Ames test. The mutagenicity of aromatic amines is currently described by four alerts in the Derek for Windows (DfW) knowledge base. These alerts have been derived largely on the basis of data from the public domain and have been reported to result in a number of false positive predictions. A study was recently undertaken to further increase the specificity of these alerts using proprietary mutagenicity data for over 350 aromatic amines and amides. Two strategies were developed to understand how the mutagenic activity of an aromatic amine is influenced by its environment. Firstly, an analysis of those inactive compounds predicted to be Ames mutagens was undertaken to identify structural motifs associated with inactivity. This found, for example, that a hydroxyl substituent at the ortho position adjacent to the aromatic amine significantly reduces mutagenic activity. Secondly, a study of the relationship between the physicochemical properties and absolute mutagenic activity was carried out to establish thresholds which could be used to distinguish between active and inactive compounds. This highlighted, for example, the importance of molecular size and the energy of the highest occupied molecular orbital (HOMO) as key factors for determining whether an aromatic amine will be active or inactive in the Ames test. Many of the conclusions of the study can be rationalised in terms of the first step in the mutagenic mechanism, which involves cytochrome P450 induced oxidation of the aromatic amine to the hydroxylamine. The results of this exercise have been used to refine the DfW coverage of aromatic amine mutagenicity, illustrating the process whereby non-confidential knowledge can be derived from confidential, proprietary data leading to a significant improvement in predictive performance.

EXPLAIN™: FINDING MOLECULAR MECHANISM OF DISEASE FROM PROMOTER MODELS TO SIGNALING PATHWAYS

Alexander Kel[1], Nico Voss[1], Tatyana Konovalova[2], Philip Stegmaier[1], Olga Kel-Margoulis[1], Tagir Waleev[2], Gerd Schmitz[3], Edgar Wingender[1,4]

[1]BIOBASE GmbH, Halchtersche Str. 33, D-38304 Wolfenbuettel, Germany; [2]A.P. Ershov's Institute of Informatics Systems, 6, Lavrentiev ave., 630090 Novosibirsk, Russia; [3]Institute for Clinical Chemistry and Laboratory Medicine University Hospital Regensburg Franz-Josef-Strauss-Allee 11, D-93053 Regensburg, Germany; [4]Dept. Bioinformatics, UKG/Univ. Goettingen, Goldschmidtstr. 1, 37077 Goettingen, Germany

Different signal transduction pathways leading to the activation of transcription factors converge at key molecules that master the regulation of many cellular processes. Such crossroads of signaling networks often appear as "Achilles Heels" causing a disease when not functioning properly.

We developed a novel computational tool, ExPlain™ for causal interpretation of gene expression data. It performs a rather unusual way of analysis through considering the earlier causes that have led to the observed gene expression changes rather then analysing the later effects of those changes. First of all, promoters of differentially expressed genes are analyzed and specific combinations of transcription factors (Composite Modules) regulating these genes are hypothesized. Next, analysis of signal transduction network upstream of these transcription factors allows us to reveal key signaling molecules that can master the observed gene expression profile. The method utilizes data from three databases (TRANSFAC®, TRANSPATH® and HumanPSD http://www.biobase-international.com/).

Affymetrix microarray data have been taken from clinical studies of a genetic disorder of elastic fiber system, Pseudoxanthoma elasticum (PXE), which is exhibited as a specific skin disease. A set of 150 promoters of differentially expressed has been compared to the set of 300 promoters of genes that did not showed any significant change of expression. Analysis of composite modules has revealed a highly significant combination of transcription factor binding sites for such factors as NF-kB, IRF, SOX-9, AML and OCT and factor pairs: SP-1/ERG-1 and AP-1/OCT. This composite promoter model was able to discriminate more then 80% of the differentially expressed promoters from the background promoters. Finally, the analysis of the signal transduction pathways upstream of these transcription factors helps to identify several potential key molecules such as the ActR-II, which is an important factor of the Atrophin-1 (DRPLA) pathway. This analysis helps to generate hypothesis on novel perspective therapeutically targets for PXE disease.

BRIDGING INVESTIGATIVE TOXICOLOGY AND DISEASE-ORIENTED RESEARCH BY BUILDING A MOUSE-TO-HUMAN PARADIGM

Alison I. Hege[1,2], Daniel Gatti[2], Paul B. Watkins[3], David W. Threadgill[1,4], Ivan Rusyn[1,2]

[1]Curriculum in Toxicology; [2]Department of Environmental Sciences and Engineering; [3]Division of Gastroenterology and Hepatology; [4]Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599, USA; E-mail: iir@unc.edu

Classical toxicology focus is on the dose-response relationships and mode-of-action paradigms. Yet, the effects of many drug and environmental chemical exposures are context dependent, with genetic diversity a major variable that is usually overlooked in safety testing. Recent advances in our basic understanding of the genetic diversity of the human genome, as well as genomes of model organisms used in toxicity and safety assessment testing, not only provide an improved basis for species comparisons, but also serve as a lead into designing new approaches for understanding and estimating the risks in a genetically diverse population.

The novel uses of gene expression and SNP genotyping data in combination with traditional toxicology experiments will be shown using two examples. First, by using exposures to different chemicals on a panel of mouse inbred strains that model genetic diversity in a population, we show that underlying genetic variation has a significant impact on the molecular and phenotypic endpoints and adds an important dimension to toxicology. Specifically, we propose and validate a strategy using a Mouse Model of the Human Population (MMHP) to identify genetic polymorphisms and novel mechanisms contributing to xenobiotic-induced liver injury in humans. Second, we dissected genetic networks that control liver gene expression by combining large-scale quantitative mRNA expression analysis with genetic mapping in a reference population of BXD recombinant inbred mouse strains for which extensive SNP, haplotype and phenotypic data is publicly available. We describe several genetic loci that control expression of large numbers of genes. By using expression (e)QTL mapping, we identified the *Serpina3* family of genes as potential novel master regulators of transcription in the liver. In conclusion, a combination of knowledge of toxicology, gene expression profiling and the latest information on the genetic diversity in the mouse and human assists in understanding a "toxicity susceptibility state" in liver in response to toxicants. The data obtained from these studies will enable us to determine what genetic variants correlate with susceptibility or resistance to liver disease, thus potentially helping to establish whether certain genetic polymorphisms in humans could identify a susceptible population for each exposure.

ENZYME KINETIC PARAMETER ESTIMATION IN SYSTEMS BIOLOGY USING PROTEIN STRUCTURES

R.R. Gabdoulline, M. Stein, R.C. Wade

EML Research gGmbH, Molecular and Cellular Modeling Group, Schloss-Wolfsbrunnenweg 33, 69118 Heidelberg, Germany

The successful modelling of metabolic and signalling pathways in systems biology requires a consistent set of enzymatic kinetic parameters. Sometimes, these parameters are not available from the literature or were obtained under different experimental conditions. Therefore computational methods to estimate these parameters are needed. We are developing simple and computationally efficient procedure to relate protein structural information to enzymatic kinetic parameters.

The qPIPSA, quantitative Protein Interaction Property Similarity Analysis, approach relates differences in molecular interaction fields between enzymes to the ratios of their kinetic parameters. This procedure can be used to determine unknown kinetic parameters when enzyme structural information is available and kinetic parameters have been measured for related enzymes, e.g. orthologues from other species, or under different conditions, e.g. a different pH.

The protein structure modeling protocol is paid much attention to, and it ensures that differences between structural models reflect the differences between the protein sequences, rather than random fluctuations in protein structure. The simple physical meaning of correlations allows detection of outliers, arising due to varying importance of contributions like protein stability and conformational changes to the kinetic parameters.

It is shown that when the measurement conditions and the protein structural models are consistent correlations between interaction fields and kinetic parameters can be established for sets of related enzymes or for an enzyme under a range of environmental conditions.

MOLECULAR INSIGHT INTO THE INTERACTION BETWEEN IFABP AND PA BY USING MM-PBSA AND ALANINE SCANNING METHODS

Hanjun Zou[1,2], Hualiang Jiang[2]

[1]Chemicals Registration Services, Toys, Food & Hardlines Division, Intertek Consumer Goods North & East China, 9/F, No.2 Building, Shanghai Comalong Industrial Park 889 Yishan Rd., Shanghai 200233, China; [2]Drug Discovery and Design Center, State Key Laboratory of Drug Research, Shanghai Institute of Materia Medica, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, and Graduate School of the Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China; E-mail: steven.zou@intertek.com

The Molecular Mechanics–Poisson–Boltzmann surface area (MM-PBSA) method combined with alanine-scanning mutagenesis is a very important process for rational drug design. In this study, molecular dynamics (MD) and MM-PBSA were applied to calculate the binding free energy between the rat intestinal fatty acid binding protein (IFABP) and palmitic acid (PA) to gain insight to the interaction details. Equally spaced snapshots along the trajectory were chosen to compute the binding free energy, which yields a result highly consistent with experimental value with a deviation of 0.4 kcal/mol. Computational alanine scanning was performed on the same set of snapshots by mutating the residues in IFABP to alanine and recomputing the $\Delta\Delta G_{binding}$. By postprocessing a single trajectory of the wild-type complex, the average unsigned error of our calculated $\Delta\Delta G_{binding}$ is below 1.5 kcal/mol for most of the alanine mutations of the non-charged residues (67% in total). To further investigate some particular mutants, three additional dynamical simulations of IFABP Arg126Ala, Arg106Ala and Arg106Gln mutants were conducted. Recalculated binding free energies are well consistent with the experimental data. Moreover, the ambiguous role of Arg106 caused by the free energy change of the opposite sign when it is mutated to alanine and glutamine respectively is clarified both structurally and energetically. Typically, this can be attributed to the partial electrostatic compensation mainly from Arg56 and the obvious entropy gain in Arg106Ala mutant while not in Arg016Gln mutant. The presented structural model of IFAPB-PA complex could be used to guide future studies such as designing ligands with high affinities.

ANALYSIS OF PEAK INTENSITY CORRELATIONS TO MAXIMIZE BIOLOGICAL MEANING OF PROTEOME DIAGNOSTIC MASS-SPECTRA

Sergei A. Moshkovskii, Mikhail A. Pyatnitsky, Alexander I. Archakov

Institute of Biomedical Chemistry, 10 Pogodinskaya Str., Moscow, 119121, Russia; E-mail: smosh@mail.ru

Since 2001, many works have been publish that suggest to employ matrix-assisted lazer desorption-ionization time-of-flight (MALDI-TOF) mass spectra of easily accessible human biosamples for classification between disease and normal subjects. This field of diagnostic proteomics encompasses disorders, such as cancers and some other chronic life-threatening diseases. In such method, mass-spectrometry peak intensities from each spectrum are used to teach classifier, e.g., genetic algorithm, support vector machine, etc, which should make diagnostic decision. Unsupervised approach often leads to the inclusion to the classifier artifact and unknown peaks. We propose the analysis of peak intensity correlations between samples to avoid artifacts and to provide diagnostic mass-spectra with biological meaning. Such analysis which was performed using our own mass-spectra for ovarian cancer diagnostics outlines the following advantages of the method. (i) This analysis may be used to assess spectra quality. (ii) It detects groups of correlating peaks allowing identification of the proteins that form the signal. (iii) The use of variables accounting the correlation groups improves the performance of diagnostic classifier based on support vector machine.

EXPLORATION INTO BIOCHEMICAL PATHWAYS

Johann Gasteiger[1,2], Martin Reitz[1], Oliver Sacher[2]

[1]Computer-Chemie-Centrum and Institute of Organic Chemistry, University of Erlangen-Nuremberg, Naegelsbachstr. 25, 91052 Erlangen, Germany; http://www2.chemie.uni-erlangen.de; [2]Molecular Networks GmbH, Naegelsbachstr. 25, 91052 Erlangen, Germany; http://www.mol-net.de; E-mail: Gasteiger@chemie.uni-erlangen.de

Biochemical processes in living organisms are often represented by complicated two-dimensional networks. Finding the desired information, and, in particular, perceiving relationships between individual reactions in such networks can be quite difficult. We have stored the contents of the poster "Biochemical Pathways" originally distributed by Boehringer Mannheim (now Roche) in a reaction database and have enriched it with additional information. The database contains 1,500 structures and 2,200 reactions. Small as this database is, it nevertheless stores information on the most important reactions, those that keep us alive.

A retrieval system, C@ROL, has been developed to access this rich source of information: Searches can now be performed for names, full structures and substructures, reaction partners, enzymes and coenzymes, organisms, reaction centers, etc. By using a standard structure format, other chemical databases and computer programs can be connected to this database. Furthermore, connection to bioinformatics databases can be made through enzyme names and the enzyme EC codes [1].

As an application of this database, we have investigated the geometric and electronic requirements of enzyme reactions. Three-dimensional models were automatically built by the 3D structure generator CORINA for all molecules involved in biochemical pathways. This then allowed us to test the transition state hypothesis, stating that the role of an enzyme is primarily in stabilizing the transition state of a reaction.

In order to investigate the electronic requirements of enzyme reactions, various physicochemical effects such as charge distribution as well as inductive, resonance, and polarizability effects were calculated for the atoms and bonds of the reaction center, i.e., those bonds directly participating in the reaction. These values were then used to train a self-organizing (Kohonen) neural network, clustering these reactions. These clusters by and large correspond to the classification of enzymes by the EC code. However, sometimes differences are observed indicating deficiencies of the EC classification and pointing out that the physicochemical descriptors show finer details of enzyme reactions.

Thus, this database provides deeper insights into the mechanisms of biochemical pathways and can also be used for making inferences on the metabolism of compounds. The BioPath database has been made accessible to the public on the internet at http://www.mol-net.de/biopath/index.html/.

1. M. Reitz, O. Sacher, A. Tarkhov, D. Trumbach, J. Gasteiger, *Org. Biomol. Chem.,* **2004**, 2, 3226-3237.

ON GRAPHICAL RERPESENTATION OF PROTEINS AND PROTEIN ALIGNMENT

Milan Randic[1,2]

[1]National Institute of Chemistry, Ljubljana, Slovenia; [2]Drake University, Dept. of Mathematics & Computer Sci., Des Moines, Iowa; E-mail: mrandic@msn.com

While graphical representations of DNA were initiated in mid 1985 by Hamori, which was soon followed by Gates, Jeffrey, Nandy and others, the initiation of graphical representations of proteins awaited for another 20 years to finally emerge in 2005. We will outline recent developments on graphical representations of proteins, which were initiated by modifications of the "Chaos Game" approach of Jeffrey for graphical representation of DNA, which also lead to numerical characterizations of the accompanying graphical representations of proteins. One route to graphical representation of proteins was based on the notion of the "Virtual Genetic Code." An alternative approach, in which the algorithm of Jeffrey was restricted to triplets of amino acids, has lead to the 8x8 Tables of Codons, which serve as a template for protein representations. Very recently numerical and graphical representation of proteins that are mostly devoid of arbitrary conventions of ordering of amino acids, based on star graphs, has been proposed. Finally, we will describe the "spectrum-like" representations of proteins, which offers a route to graphical comparisons of proteins and has been modified for graphical visualization of the proton alignments (to be illustrated). We will end the presentation with an illustration of the latest approach to the graphical alignment of proteins in which proteins are represented bys zigzag curves over a 2D map of the 20 natural amino acids. The 2D amino acid maps are obtained by ordering amino acids with respect to a pair of their physico-chemical properties.

A DATABASE SEARCH AND RETRIEVAL SYSTEM FOR THE ANALYSIS AND VIEWING OF BOUND LIGANDS, ACTIVE SITES, SEQUENCE MOTIFS AND 3D STRUCTURAL MOTIFS

Adel Golovin, Kim Henrick

EMBL-EBI, Hinxton Hall, Genome Campus, Cambridge, UK; E-mail: golovin@ebi.ac.uk

The three-dimensional environments of ligand binding sites have been derived from the parsing and loading of the PDB entries into a relational Macromolecular Structure Database (1). For each bound molecule the biological assembly of the quaternary structure has been used to determine all contact residues and a fast interactive search and retrieval system has been developed. The database was extended with small 3D structural motifs, PROSITE (2) patterns and profiles, Catalytic Sites. Novel algorithms for chemical substructure search, for dihedral angles sequences search, for sequence patterns search, for super-secondary structure motifs matches and for small 3D structural motif groups searching are incorporated. The interface provides functionality for visualization and creating a search criteria. It provides sequence and 3D multiple alignment options along with a number of statistical charts. The service is available at http://www.ebi.ac.uk/msd-srv/msmotif/new.

HOMOLOGY MODELING AND COMPUTER-AIDED DESIGN OF HNRNP K MARKER AND LIGANDS IN HEAD AND NECK CANCER

Carlos H. T .P. Silva, Vinicius B. da Silva, Andreia M. Leopoldino

School of Pharmaceutical Sciences of Ribeirao Preto, University of Sao Paulo, Ribeirao Preto - SP, Brazil; E-mail: tomich@fcfrp.usp.br

Head and neck cancers (HNC) are a diverse group of diseases, each with its own distinct features, and treatment considerations. Despite improvements in diagnosis, high recurrence and local management, long-term survival rates for patients have not increased significantly over the last years. In this work, we have selected some target proteins in HNC to molecular and structural investigations including hnRNPK protein, which is involved in splicing and cell cycle progression.

In this work we have performed virtual screening with the recently solved third KH domain [1], using a database of 8.8 million of drug-like compounds. The best ligands have been selected for purchase and enzymatic assay using structural and pharmacophoric constraints. Results reinforce that Arg59 plays an important role for binding, pointing out the necessary structural modifications which must be done in the novel ligands.

1. P.H. Backe, A.C. Messias, R.B. Ravelli, M. Sattler, *Structure,* **2005**, 13, 1055.

PREDICTION OF ENVIRONMENTAL TOXICITIES

Young Yong In[1], Sung Kwang Lee[1], Pil Jae Kim[2], Kyoung Tai No[1,3]

[1]Bioinformatics and Molecular Design Research Center, Seoul 120-749, Korea; [2]National Institute of Environmental Research, Incheon 404-170, Korea; [3]Department of Biotechnology, Yonsei Univ., Seoul 120-749, Korea; E-mail: ktno@yonsei.ac.kr

Prediction of environmental and xenobiotic toxicity is crucial for systematic management of millions of chemicals and for the reduction of drug developing cost by introduction of early stage toxicity prediction. For this purpose, we start to develop an integrated toxicity prediction system that includes both environmental and xebobiotic toxicities. Though the system is in construction, we already have finished some toxicity prediction engines. In the presentation, i) some newly developed descriptors that contain both structural and propertirs information, ii) calcinogenicity and, mutagenecity predictoin method that is mainly based on topological descriptor and QSTR, iii) as an environtal toxicity, an acute fish toxicity model, and iv) a phase I metabolite prediction model will be introduced.

In the prediction of toxicity, the major limitation and difficulty is how to introduce the mode of action (MOA) information of each toxicity end point. To increase the predictability of each toxicity end point, it is necessary to classify chemicals according to their MOA and to develop QSTR model for each class. In order to introduce the MOA to the QSTR prediction engines, we introduce gene markers that are related to a MOA of a certain toxicity end point. In our research, using the markers (signatures), the chemicals will be classified into MOAs of a certain end point and the QSTR model will be developed by introducing both molecular properties and signatures as descriptors.

Finally, we will introduce the strategy how to integrate developed engines, to standardize data and I/O, and to service to end users.

NEW IN SILICO STRATEGIES FOR ASSESSING ENVIRONMENTAL CHEMISTRY AND REACH CHALLENGES

Jacques R. Chretien, Marco Pintore

BioChemics Consulting SAS, 16 rue L. de Vinci, 45074, Orleans, France; E-mail: jacques.chretien@biochemics-consulting.com

The introduction of REACH system could potentially involve the use of 3.9 million animals for experimental toxicity studies. Simultaneously European regulation to be applied in 2009 implies to stop and/or to reduce drastically animal testing. It prompted us to work within the previous EC DEMETRA Project focused on acute toxicity by using alternative procedure based on [QSAR] [1].

The general aim of CAESAR, a EC financed project under FP6, is then to produce (Q)SAR models for predicting the toxicity of chemical substances. These *in silico* models are designed to be used for regulatory purposes, more specifically to assist the implementation of the proposed REACH system. The CAESAR work involve a robust and detailed characterization of the most relevant endpoints as defined in EC documents relating to the requirements for REACH. Utmost attention is paid to obtain a solid experimental basis by the use of toxicity data that have been checked for quality at all stages.

In this work, several (Q)SAR methods and set of descriptors are used and compared on the first priority regulatory endpoints selected inside CAESAR, i.e. bioaccumulation and skin sensitization, associating 473 and 200 chemicals, respectively. These methods cover linear and non linear algorithms based on PLS, Fuzzy Logic and Artificial Neural Networks. Good preliminary results are obtained for each endpoint, by different (Q)SAR approaches, and the best cross-validation and test set scores reach accuracy values up to 80%. Moreover, the presence of false negative predictions is carefully evaluated to reduce the risk of predicting potentially harmful compounds to be non-toxic. The applicability domain of each (Q)SAR model is also described, both in terms of its boundaries and density in chemical and biological space. The best (Q)SAR models developed in this project, and their protocols, will be placed on the project's Internet site to allow the widest possible free access.

1. Quantitative Structure –Activity Relationships [QSAR] for Pesticide Regulatory Purposes. Emilio Benfenati Ed., Elsevier, Amsterdam, **2007**.

CONSENSUS MODELS AND META-MODELS FOR THE PREDICTION OF LOGP USING NEURAL NETWORKS

Daniel Zuaboni[1,2], Christophe Cleva[1], Pierre-Alain Carrupt[2]

[1]MerckSerono Geneva Research Center, 9 Chemin des Mines, 1211 Geneva, Switzerland; [2]LCT-Pharmacochemistry, School of Pharmaceutical Sciences University of Geneva, University of Lausanne, 30 Quai Ernest-Ansermet, CH-1211 Geneva 4, Switzerland

Numerous methods have been developed for predicting the logP of compounds from their chemical structure. Most of these methods are based on different approaches and perform differently on different series of compounds. The aim of this project is to combine these different prediction methods in the form of consensus models or meta-models in order to improve the accuracy of prediction compared to a single method.

A consensus model takes as input the different results obtained by different prediction models and combines them to produce a single output value. For this study we used several well known prediction models (KowWin, Vlogp, ClogP, Marvin logP, and several versions of AlogP) as input and we trained feed-forward neural networks on datasets extracted from the PhysProp database to obtain consensus models that outperform each and every single method.

In a following step, we identified for each input model the main substructures leading to a significant worsening of the predictions. By adding this structural information in the form of binary descriptors to the input of the neural network, we obtained meta-models that showed improved performance compared to the consensus models.

REACTIVE METABOLISM – QSAR AND QUANTUM CHEMISTRY MODELING

Yuriy Abramov

Pfizer Global Research and Development, Eastern Point Road, Groton, CT 06340, USA; E-mail: Yuriy.A.Abramov@pfizer.com

The cytochrome P450 enzymes may be involved in biotransformation of relatively inert drugs into highly reactive metabolites, commonly referred to as bioactivation. It is recognized that the reactive metabolites can form conjugates with GSH, causing hepatotoxicity. Two independent approaches were used for modeling the bioactivation. A C5.0 QSAR model was implemented for screening purposes. An overall accuracy of the model is reasonably high, 84% for the test set. The quantum chemistry approach was also utilized for getting important insights into specific mechanisms of the bioactivation.

MODELLING, OPTIMIZATION, AND VARIABLES SELECTION IN QSA(P)R STUDIES. THE ROLE OF COUNTER-PROPAGATION NEURAL NETWORK

Marjana Novic

National Institute of Chemistry, Hajdrihova 19, POB 660, 1001 Ljubljana, Slovenia; E-mail: marjana.novic@ki.si

Counterpropagation artificial neural network (CP ANN) has been often successfully applied in structure-based drug design studies [1-2] as well as in modelling of adverse molecular properties, i.e. structure-toxicity studies [3-5]. The characteristics of CP-ANN learning process allows unsupervised clustering of objects (studied molecules) followed by the supervised construction of the response surface. In the first type of applications we developed models for use in prediction of thrombin-inhibitor binding constants on the basis of the complex crystal structures. Thrombin belongs to serine protease family and inhibitor specificity, in particular towards analogous mammalian protease trypsin is required, and so we have analogously modelled the inhibitor binding affinity towards trypsin. Contrary to structure-based drug design, in which the aim is to find new potent compounds, the structure-toxicity studies are focused to the prediction of an adverse property of existing chemicals. With the same modelling method (CP ANN) several toxicity studies have been performed. An example is given on of endocrine disrupters (estrogen binders) study.

1. Gregor Mlinsek, Marjana Novic, Miha Kotnik, Tomaz Solmajer. Enzyme binding selectivity prediction : [alpha]-thrombin *vs* trypsin inhibition. *J. chem. inf. comput. sci.*, **2004**, 44, no. 5, 1872-1882.
2. Marjana Novic, Zaneta Nikolovska-Coleska, Tomaz Solmajer. Quantitative structure-activity relationship of flavonoid p56[sup](lck) protein tyrosine kinase inhibitors. A neural network approach. *J. chem. inf. comput. Sci.,* **1997**, 37, no. 6, 990-998.
3. Marjan Vracko, Marjana Novic, Jure Zupan. Study of structure-toxicity relationship by a counterpropagation neural network. *Anal. chim. acta*, **1999**, 384, 319-332.
4. Alessandra Roncaglioni, Marjana Novic, Marjan Vracko, Emilio Benfenati. Classification of potential endocrine disrupters on the basis of molecular structure using a nonlinear modeling method. *J. chem. inf. comput. sci.*, **2004**, 44, no. 2, 300-309.
5. Federico Marini, Alessandra Roncaglioni, Marjana Novic. Variable selection and interpretation in structure-affinity correlation modeling of estrogen receptor binders. *J. chem. inf. mod.*, **2005**, vol. 45, no. 6, 1507-1519.

STRUCTURE-ACTIVITY MODELING OF A DIVERSE SET OF ANDROGEN RECEPTOR LIGANDS

J. Devillers[1], J.P. Doucet[2], A. Panaye[2], N. Marchand-Geneste[3], J.M. Porcher[4]

[1]CTIS, 3 Chemin de la Graviere, 69140 Rillieux La Pape, France; [2] Universite Paris 7 Denis Diderot, ITODYS, 1 rue Guy de la Brosse, 75005 Paris, France; [3]Universite de Bordeaux I, LPTC, ISM – UMR 5255 CNRS, 351 Cours de la Liberation, 33405 Talence CEDEX, France; [4]INERIS, Parc Technologique ALATA, BP n° 2, 60550 Verneuil en Halatte, France

Numerous chemicals released into the environment can interfere with normal, hormonally regulated biological processes to adversely affect development and/or reproductive function in wildlife and humans. Due to the ability of these chemicals to interfere with the endocrine systems, they have been labeled as endocrine disruptors (EDs). SARs and QSARs are powerful screening tools to detect potential EDs and to prioritize them for more intensive and costly evaluations based on *in vitro* and *in vivo* assays.

In this context, androgen-receptor binding data (active/inactive) for a large set of about 200 structurally diverse chemicals, described by CODESSA descriptors encoding topological and physicochemical properties, were used for deriving structure-activity models. Different types of artificial neural networks and support vector machines with different kernel functions were tested as statistical tools. The performance of a classical discriminant analysis was also estimated. The comparison exercise was performed on the basis of the same learning and testing sets as well as from the same set of selected descriptors. The modeling performances as well as the technical advantages and limitations of each statistical method have been critically analyzed.

IN SILICO SCREENING OF ESTROGEN-LIKE CHEMICALS BASED ON QSAR MODELS

Huanxiang Liu, Ester Papa, <u>Paola Gramatica</u>

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Structural and Functional Biology, University of Insubria, via Dunant 3, 21100 Varese (Italy); E-mail: paola.gramatica@uninsubria.it

Increasing concern is shown by the scientific community, regulators, and the public about endocrine-disrupting chemicals (EDCs) that, in the environment, are adversely affecting human and wildlife health through a variety of mechanisms, mainly estrogen receptor-mediated mechanisms of toxicity. Because of the large number of EDCs in the environment, there is a great need for an effective tool of rapidly assessing ED activity in the toxicology assessment process and in the context of the new European REACH policy. Here, classification and regression QSAR models were developed to predict the estrogen receptor binding affinity based on a large data set of heterogeneous chemicals and theoretical molecular descriptors from DRAGON. The built OLS regression model, based on eight descriptors, was validated comprehensively (internal and external validation, Y-randomization test) and all the validations indicate that the proposed QSAR model is robust and satisfactory (Q^2 = 0.75-0.85). Comparison with similar studies revealed that our model seems to outperform all the others on the whole, although it is impossible to have an absolute measure of comparison. For the classification models, three nonlinear classification methodologies: Least Square Support Vector Machine (LS-SVM), Counter Propagation Artificial Neural Network (CP-ANN), and k-nearest Neighbor (kNN) were applied, by using four molecular structural descriptors as inputs. All three methods can give satisfactory prediction results both for training and prediction sets, and the most accurate model was obtained by the LS-SVM approach. In addition, our models were also applied to about 58 000 discrete organic chemicals; about 76% were predicted not to bind to an estrogen receptor. Thus, the proposed models are very satisfactory and externally predictive, can provide a practical tool for the rapid screening of the estrogen activity of organic compounds, allowing the quick identification of possible environmental estrogens, based only on chemical structure.

E-ENDOCRINE DISRUPTING CHEMICAL DATABASES FOR DERIVING QSAR MODELS

Nathalie Marchand-Geneste[1], James Devillers[2], Jean-Christophe Dore[3], Jean-Marc Porcher[4]

[1]Universite Bordeaux 1 ; CNRS UMR 5255 ISM-LPTC, 351 cours de la Liberation, 33405 Talence, France; [2]CTIS, 3 Chemin de la Graviere, 69140 Rillieux La Pape, France; [3]Departement Regulations, Developpement et Diversite Moleculaire, USM 0502, UMR 8041 CNRS, Museum National d'Histoire Naturelle, 63 rue de Buffon, 75005 Paris, France; [4]INERIS, BP2, 60550 Verneuil-en-Halatte, France; E-mail: n.geneste@ism.u-bordeaux1.fr

There is increasing evidence that numerous chemicals released into the environment by human activities have the potential to alter the normal functions of the endocrine system in wildlife. These xenobiotics, mimicking natural hormones, are called endocrine disrupting chemicals (EDCs). The aim of this study was to catalogue the different EDC database (biological data, chemical descriptor data, etc.) resources available on the Internet for deriving structure-activity models. *In vitro* and *in vivo* experimental data on nuclear receptor binding assays, *in vitro* test methods reliability and variability assessment for detecting EDCs and comparison between biological assays will be presented and critically analysed.

QSAR MODELLING OF CARCINOGENICITY FOR REGULATOTY USE IN EUROPE

Natalja Fjodorova, Marjana Novic, Marjan Vracko

Laboratory of Chemometrics, National Institute of Chemistry, Hajdrihova 19, SI- 1000 Ljubljana

In the context of EU legislation, such as REACH and the Cosmetics Directive (Council Directive 2003/15/EC), it is anticipated that (Q)SARs will be used more extensively, in the interests of time- and cost-effectiveness and animal welfare.

Many different (Q)SAR models for prediction of properties relevant for chemical management exist, including models for prediction of carcinogenic potency. The models are described in the literature, however, most of them are poorly described in terms of Five principles for validation of (Q)SAR models, which have been adopted by OECD (Organization for Economical Cooperation and Development).

CAESAR is one of European Projects aimed at development the models for five properties relevant for chemical management and to provide the users all details, which are necessary to fulfill the requirements of Five principles. The predictions of properties together with all modeling details can be easily used in chemical regulation.

In this presentation we will show the example of QSAR modelling of chemical carcinogens according with principles of validation adopted by OECD in scope of European project CAESAR.

QUASICRYSTALLINITY OF A LIQUID STATE AS A BASIS FOR CREATING QSPR-MODELS AND DESCRIBING SYNERGIC EFFECTS IN BINARY SYSTEMS

Vyacheslav F. Nikolaev

Arbuzov Institute of Organic and Physical Chemistry, Russian Academy of Sciences, Arbuzov str. 8, Kazan, 420088 Russia; E-mail: mobin7@yandex.ru

Quasicrystallinity approach to the analysis of pure and binary liquids gives two models.

Description of physicochemical constants of pure liquids [1]. The first model based on such molecular characteristics as molecular refraction (MR), dipole moment (μ), and molar volume (M/d) is suggested for quantitatively describing molar physicochemical properties *F* (surface tension, vaporization enthalpy, boiling point, viscosity, etc.) of pure molecular liquids. Experimental data on vaporization enthalpies $\Delta_{VAP}H$, boiling points t_b, molar surface tension σ, and viscosities η of pure non-electrolytes were treated using the general correlation equation

$$F = k_0 + k_1\, \mu^2\, (d/M) + k_2\, MR^2\, (d/M) \quad (1)$$

Description of physicochemical properties of binary mixtures [2]:

$$F_{mix} = F_1 \cdot f_1 + F_2 \cdot f_2 + A \cdot f_2 \cdot \exp(B \cdot f_2) + C \cdot f_1 \cdot f_2 + D \cdot f_1 \cdot \exp(E \cdot f_1), \quad (2)$$

with F_{mix}, F_1, F_2 – molar physicochemical properties of the component mixture (1+2) and those of pure components, respectively; f_1, f_2 – molar fractions $(f_1 + f_2 = 1)$, $A \cdot f_2 \cdot \exp(B \cdot f_2)$, –structural asymmetric electrostatic and specific contribution in the region of the first component excess, $D \cdot f_1 \cdot \exp(E \cdot f_1)$ – contribution upon the excess of the second component. The approach is applicable for the analysis of the properties of the solute (e.g. solvatation enthalpy, solubility etc.) depending on the composition of the binary solvent, as well as for the analysis of the integral interaction effects of any systems with different character and structural organization (e.g. activity of binary medicine, symbiosis of plants or strains, productivity of mixed two-component seeding, catastrophe theory, coexistence of ethnic groups, interference of cultures etc.). The factors determining the structure and individuality of social group, might be common mentality, culture, language (the importance of common language for the results of joint activity is clearly demonstrated in the Bible parable about the Babylon tower).

1. V.F. Nikolaev, *Russ. J. Phys. Chem.*, **2006**, Vol. 80, Suppl. 1, pp. S21–S25.
2. V.F. Nikolaev, I.V. Nikolaev, V.E. Kataev, *Russ. J. Phys.Chem.*, **2006**, Vol. 80, Suppl. 1, pp. S26–S30.

QSAR STUDIES USING THE PARASHIFT SYSTEM

David J. Livingstone[1,2], Timothy Clark[3], Brian D. Hudson[2], Martyn G. Ford[2]

[1]ChemQuest, Sandown, UK; [2]Centre for Molecular Design, University of Portsmouth, UK; [3]Computer-Chemie-Centrum, Universitat Erlangen-Nurnberg, Germany; E-mail: Davel@chemquest.uk.com

Following the pioneering work of Corwin Hansch, most QSAR studies consisted of regression models based on substituent constants or topological indices. This began to change with the application of computational chemistry programs to provide molecular descriptors [1] and the use of multivariate statistical methods [2] to build mathematical models. The situation today is that there is a bewildering variety of descriptors that may be used to characterise molecules and a host of mathematical techniques that can be used to build QSAR models. It is, however, perhaps time to re-evaluate the way in which we approach this task. When compounds interact with a biochemical system what the components "see" is the corresponding molecular surfaces. Thus, it is proposed that molecules are characterised by properties calculated on a molecular surface. This, of course, requires the generation of an appropriate molecular surface and relevant surface properties. Preliminary work has shown that such surface properties have utility in the estimation of some fundamental molecular properties [3]. This presentation will examine the use of this approach in the construction of QSAR models.

1. R.C. Glen and V.S. Rose. Computer program suite for the calculation, storage and manipulation of molecular property and activity descriptors. *J. Mol. Graph.*, **1987**, 5, 79-86.
2. R.M. Hyde and D.J. Livingstone. Perspectives in QSAR: Computer Chemistry and Pattern Recognition. *J. Comp.-Aid. Mol. Design*, **1988**, 2, 145-157.
3. B. Ehresmann, M. J. de Groot and T. Clark. Surface-Integral QSPR Models: Local Energy Properties. *J. Chem. Inf. Model*, **2005**, 45, 1053-1060.

CYCLIN DEPENDENT KINASE (CDK2/CDK4) SELECTIVITY ELUCIDATED BY MOLECULAR DYNAMICS AND QUANTUM CHEMISTRY STUDIES

Aixiao Li[1,2], Florent Barbault[1], Francois Maurel[1], Baoshan Wang[2], BoTao Fan[1†], Michel Delamar[1]

[†]This communication is dedicated to the memory of Prof. BoTao Fan

[1]ITODYS, University Paris 7 – CNRS UMR 7086 ; 1 rue Guy de la Brosse 75005 Paris, France; [2]Molecular Simulation and Chemical Information Laboratory, Department of Chemistry and Molecular Science, University of Wuhan, Hubei 430072, China; E-mail: florent.barbault@paris7.jussieu.fr

Cyclin-dependent kinases (CDKs) are a family of Ser/Thr kinases that play a central role in eukaryotic cell cycle regulation [1, 2]. CDKs have being identified as important targets for therapeutic intervention in cancer [3, 4] and small molecules inhibitors have been identified and reported [5-7]. Most of them are targeting the CDK4 enzyme but, unfortunately, they also inhibit CDK2 and may cause serious adverse effects. Indeed, CDK4 presents 46% and 72%, respectively, of sequence identity and similarity in comparison to CDK2. This is quite large but logical since they share the same biological function. Therefore, it is desirable to design more selective inhibitors with a broader range of biological profiles, that would present less adverse effects. It is also highly desirable to obtain more information about the origin of the selectivity that is observed with some inhibitors.

To elucidate this it is necessary to understand the differences in the interaction mechanism of CDK2 and CDK4 with an inhibitor showing selectivity. An interesting compound, called 2PU, inhibits CDK4 and presents a real selectivity for CDK4 compared to CDK2 [8]. This work presents a complete study of the interaction process differences between CDK2/CDK4 by two different methods.

First, both systems (CDK4-2PU and CDK2-2PU) were extensively studied by molecular dynamics simulations in explicit solvent. To study the influence of target flexibility, both CDKs were studied in the free and complexed states leading to a total of 6 simulations of 10ns trajectory length. A detailed study of H bonding network was also performed. Both CDK2 and CDK4 binding free energies with 2PU were finally calculated using the well-known MMPBSA method and normal modes analysis for entropy contributions.

We then studied CDK2 and CDK4 complexes using quantum chemistry methods. We were thus able to identify the protein residues most interacting with 2PU.

This structural and interaction mode analysis give new insight into the interpretation of the selectivity process. Computational methods are also compared and discussed in terms of complementarity and reliability.

1. T. Hunter, J. Pines, *Cell*, **1994**, 79, 573-582.
2. C.J. Sherr, *Science*, **1996**, 274, 1672-1677.
3. M.D. Garrett, A. Fattaey, *Curr Opin Genet Dev*, **1999**, 9, 104-111.
4. K.R. Webster, *Expert Opin Investig Drugs*, **1998**, 7, 865-887.
5. A. Huwe, R. Mazitschek, A. Giannis, *Angew Chem Int Ed Engl*, **2003**, 42, 2122-2138.
6. T.M. Sielecki, J.F. Boylan, P.A. Benfield, G.L. Trainor, *J Med Chem*, **2000**, 43, 1-18.
7. P.L. Toogood, *Med Res Rev*, **2001**, 21, 487-498.
8. M. Ikuta, K. Kamata, K. Fukasawa, T. Honma et al., *J. Biol. Chem.*, **2001**, 276, 27548-27554.

BIOLOGICAL SPECTRA ANALYSIS: LINKING BIOLOGICAL ACTIVITY PROFILES TO MOLECULAR TOXICITY

Peter O. Fedichev, Andrei A. Vinnik

Quantum Pharmaceuticals; E-mail: peter.fedichev@q-pharm.com

Starting from the premise that biological activity including toxicity results from the capacity of small organic molecules to modulate the activity of the proteome, we set out to investigate whether calculated inhibition values could be used for measuring and quantifying toxicity of the molecules. Using a 1,148-small molecule compound database and a diversified set of 476 human protein, we show that inhibition values, determined by in-silico methods, provide a precise molecular property estimation. When broad biological activity of molecules is presented in spectra form, organic molecules can be sorted by quantifying differences between biological spectra, which define molecular properties, including toxicity. This methodology, in which we have applied biological spectra analysis, provides the capability to estimate a number of measures (LD50, MRDD and TD50) of different forms of toxicity (acute toxicity, chronic toxicity, carcinogenicity, neurotoxicity, etc.). We believe that our methodology derived from biological activity profiles provides a more specific estimate of the toxic dose threshold of chemicals in humans compared to current risk assessment models which extrapolate from animals to humans employing multiple uncertainty safety factors.

2D/3D QSAR AND MOLECULAR MODELLING STUDIES OF NEW CLASS OF HIV-1 GP41 INHIBITORS

Catia Teixeira, Florent Barbault, Joseph Rebehmed, Francois Maurel, BoTao Fan[†]

[†] This communication is dedicated to the memory of Prof. BoTao Fan

ITODYS, University Paris 7 – CNRS UMR 7086 ; 1 rue Guy de la Brosse 75005 Paris, France; E-mail: catia.teixeira@jussieu.paris7.fr

Currently, anti-HIV drugs combinations therapies only target HIV-1 protease and reverse transcriptase [1]. Unfortunately, most of these molecules presents numerous shortcomings such as viral resistances and adverse effects. In addition, these drugs are involved to later stages of infection [2]. Therefore, it is necessary to develop new drugs which are able to block the first steps of viral cycle life. A compound, Enfuvirtide [3], prevents the fusion between the virus and cellular membranes. This 36-mer peptide is derived from gp41 CHR region but suffers from several limitations. The purpose of this project is to design N-substituted pyrrole small molecules that will mimic the Enfurvirtide/gp41 interactions and inhibit HIV viral entry.

To do that, 2D and 3D QSAR models were derived from a first set of 23 HIV-1 gp41 inhibitors. The 2D-QSAR studies of the derivatives were performed using CODESSA software package. A linear correlation is obtain with five-descriptors. The 3D-QSAR was done with the comparative molecular fields analysis (CoMFA) method. This model presents a eight components regression equation with a good correlative and predictive capabilities.

At the same time, a detailed molecular docking study was performed for the 23 compounds with the autodock3 software. The goal of this work was to obtain the bioactive conformations of these compounds. The descriptors involved in the 2D-QSAR equations are related to the main factors that influence the inhibitory activity of the HIV-1 gp41. Moreover, the results obtained from 3D-QSAR study were superimposed on the gp41 active site and the main interactions were studied. These findings provide us very good advices for future structural modifications of this new class of entry inhibitors for better HIV-1 activity.

1. A.S. Fauci, *Nat. Med,.* **2003**, 9, 839-843.
2. A.K. Debnath et al; *J. Med. Chem.,* **1999**, 42, 3203-3209.
3. E. De Clercq; *J. Clin. Virol.,* **2004**, 30, 115-133.

HYDROGEN BONDING PARAMETRIZATION IN QSAR AND DRUG DESIGN

Oleg A. Raevsky

Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds, Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia; E-mail: raevsky@ipac.ac.ru

In spite of all attempts to arrive at a better understanding of the role of water and of hydrogen bonds in biological systems we are far from a satisfactory situation. Analysis of different approaches to characterize hydrogen bonding and its contribution in chemical properties and activity is presented in the report. Different types of H-bond descriptors including indirect H-bond descriptors, indicator variables, two dimensional thermodynamics descriptors, surface H-bond descriptors, distance H-bond potentials are considered. Examples of successful application of H-bond descriptors in QSAR and Drug Design including solubility and partitioning of chemicals in water-solvent-gas systems, permeability and absorption in humans, classification of pharmacokinetic properties, biological target inhibition, aquatic toxicity are presented. The ubiquitous application of indirect and/or indicator parameters of H-bonding processes in QSAR and Drug Design to cast a false color real H-bonding role. Studies based on direct thermodynamic parameters of H-bonding and exact three-dimensional structures of H-bonding complexes have essentially improved our understanding of complex processes of solvation and specific intermolecular interactions. These studies consider the structure of liquid water, new X-ray data for specific H-bonding complexes, quantitative estimation of contribution of H-bond acceptor and donor factors and volume-related terms in chemicals solvation processes partitioning in water/solvent/air systems, a refinement in the PSA approach, improvement of Grid potentials, calculation schemes of optimum H-bonding potential values for any concrete H-bonding atoms in any complexes which consider the nature of interacting atoms and the influence of substituents. These developments ensure real quantitative description of H-bonding and the successful application of direct H-bonding descriptors in QSAR and Drug Design.

TOWARD BASIC UNDERSTANDING OF THE PARTITION COEFFICIENT LOG P AND ITS APPLICATIONS IN QSAR

Hiroshi Chuman

Institute of Health Biosciences, The University of Tokushima, Graduate School, Shomachi 1-78, Tokushima, 770-8505, Japan; E-mail: hchuman@ph.tokushima-u.ac.jp

The log P value has been the first choice for the molecular hydrophobicity descriptor and in fact it has been widely used in QSAR studies. However, it is still now difficult to understand the partitioning phenomenon in terms of physical chemistry and the physicochemical meaning of log P appeared in enzymatic QSAR equations.

In my first topic, an attempt to understand and predict log P is addressed. We formulated a simple model that expressed by two terms, the solvent accessible surface area and the solvation energy difference (ΔE_{sol}) between aqueous and organic solvent phases. The E_{sol} value was calculated using the ab initio SCRF-MO method. The log P value except for those of hydrogen bond donors was shown to be analyzable reasonably well by this model.

The second topic is a practical application of log P in the field of clinical pharmacy. It is important to evaluate the risk potential of drugs excreted from mothers to their babies. The log M/P value (concentration ratio between milk and plasma) of structurally diverse compounds was nicely predicable with log P, molecular size and degree of ionization.

The last topic is a reinterpretation of classical QSAR equations using the ab initio MO calculation of a whole protein-ligand complex. The results can help us to understand how classical QSAR descriptors such as σ and log P work effectively in comparison with the change of electronic structure induced by complex formation.

NEW PREDICTION SYSTEM FOR MUTAGENICITY: CISOC-PSMT AND ITS APPLICATIONS

Jianhua Yao, Quan Liao, Tianxiang Shen, Shengang Yuan

Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, 354, Fenglin Road, Shanghai, 200032, China; E-mail: yaojh@mail.sioc.ac.cn

The study of prediction of toxicity is very important and necessary because measurement of toxicity is typically time-consuming and expensive. In this work, four kinds of fragments: Atom, Star, Path and Ring, which were generated by the methods described in our former publications [1, 2] acted as descriptors. Recursive Partitioning (RP) method was used to select descriptors. RP and Support Vector Machines (SVM) were used to construct structure-toxicity relationship models, RP model and SVM model, respectively. The performances of the two models are different. The prediction accuracies of the RP model are 80.2% for mutagenic compounds in MDL's toxicity database, 83.4% for compounds in CMC and 84.9% for agrochemicals in in-house database respectively. Those of SVM model are 81.4%, 87.0% and 87.3% respectively.

Based on the study, we developed prediction system for mutagenicity: CISOC-PSMT (fig. 1) which was authorized in China [3]. It is being applied in TCM study, pesticide discovery, drug discovery and innovation of new chemical entities.

(a)   (b)
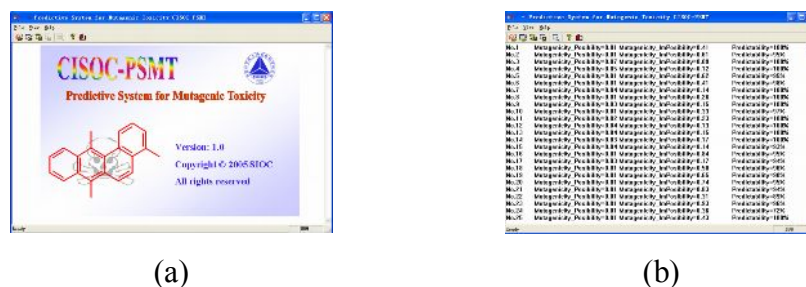
Fig.1 CISOC-PSMT, (a) Interface of the system, (b) Prediction results

1. Q. Liao, J.H. Yao, F. Li, S.G. Yuan, J.P. Doucet, A. Panaye, B.T. Fan. CISOC-PSCT: A predictive system for carcinogenic toxicity. *SAR QSAR Environ. Res.,* **2004**, 15, 217-235.
2. Q. Liao, J.H. Yao, S.G. Yuan. SVM approach for predicting LogP. *Molecular Diversity*, **2006**, 10, 301-309.
3. CISOC-PSMT: Prediction System for Mutagenic Toxicity, Reg. No. 042160, (China).

PREDICTION OF PH-DEPENDENT AQUEOUS SOLUBILITY OF DRUG-LIKE MOLECULES AND DRUG CANDIDATES WITH CHEMOINFORMATICS TOOLS

Irene Kouskoumvekaki[1], Niclas Tue Hansen[1], Fredrik Bjorkling[2], Flemming Steen Jorgensen[3], Soren Brunak[1], Svava Osk Jonsdottir[1]

[1]Center for Biological Sequence Analysis, BioCentrum, Technical University of Denmark, DK-2800 Lyngby, Denmark; [2]Topotarget A/S, Symbion Science Park, Fruebjergvej 3, DK-2100 Copenhagen, Denmark; [3]Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen, Denmark; E-mail: irene@cbs.dtu.dk

Aqueous solubility and its dependence on the pH of the gastrointestinal tract are key factors that determine the bioavailability of drugs. Thus, reliable computational methods to estimate pH-dependent aqueous solubility of drug-like molecules and drug candidates from their molecular structure would be desirable in the drug design process.

In this work, the Henderson-Hasselbalch equation [1] has been employed, which describes pH-dependent solubility as a function of the intrinsic solubility (S_o) and the acid / base dissociation coefficients (pK_a / pK_b). A predictive model for the intrinsic solubility has been developed based on artificial neural networks (ANN) that have been trained on a druglike PHYSPROP subset of 4548 compounds. For the prediction of the acid / base dissociation coefficients, a commercial tool (MARVIN, www.chemaxon.com) has been used, after being validated on a dataset of 467 molecules from the PHYSPROP database. The best performing ANN has an RMSE of 0.71 using three-part validation, while the MARVIN module has an RMSE of 0.71 pH-units.

A dataset of 27 drugs with experimentally determined pH dependent solubility curves has been assembled from the literature for the validation of the combined model. An average RMSE of 0.79 logS-units suggests that pH-dependent solubility can be modelled with almost the same accuracy as the intrinsic solubility and can be used for evaluating existing commercial and in-house libraries [2]. Furthermore, the combined model has been tested in a series of Histone Deacetylase (HDAC) inhibitors from the drug optimization pipeline of Topotarget A/S and efforts have been made for its refinement for this particular family of compounds [3].

1. C.A.S. Bergstrom, K. Luthman, P. Artursson. Accuracy of calculated pH-dependent Aqueous Drug Solubility. *Eur. J. Pharm. Sci.,* **2004**, 22, 387-398.
2. N.T. Hansen, I. Kouskoumvekaki, F.S. Jorgensen, S. Brunak, S.O. Jonsdottir. 'Prediction of pH-Dependent Aqueous Solubility of Drug-like Molecules'. *J. Chem.Inf. Model.,* **2006,** 46(6), 2601-2609.
3. I. Kouskoumvekaki, S.M. Vadlamundi, N.T. Hansen, F. Bjorkling, S.O. Jonsdottir, 'Prediction of pH-Dependent Solubility of Histone Deacetylase (HDAC) Inhibitors', manuscript in preparation.

THREE-DIMENSIONAL COMMON-FEATURE HYPOTHESES AND 2D-QSAR STUDIES ON SOME ANTIMICROBIOLOGICALLY ACTIVE AMIDES AGAINST DRUG-RESISTANCE *STAPHYLOCOCCUS AUREUS*

Ilkay Yıldız, Tugba Ertan, Ozlem Temiz-Arpaci, Kayhan Bolelli, Esin Aki-Sener, Ismail Yalcin

Ankara University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, 06100 Tandogan-Ankara, Turkey; E-mail: iyildiz@pharmacy.ankara.edu.tr

The dramatically rising prevalence of multi-drug resistant microbial infections in the past few decades has become a serious health care problem. In particular, the emergence of multi-drug resistant strains of Gram-positive bacterial pathogens such as methicillin-resistant *Staphylococcus aureus* and *S. epidermis* and vancomycin-resistant *Enterococcus* is a problem of ever-increasing significance [1-5]. In order to prevent this serious medical problem, the elaboration of the new types of the previously known drugs is a very actual task.

Inhibitor effects against drug-resistance *Staphylococcus aureus* of a new series benzamide and phenylacetamide derivatives which have a nitro or amine group attached on position 4 or 5 of *N*-(2-hydroxyphenyl) binding them were investigated [6].

In this study, common-features hypotheses are generated by using Catalyst 4.9 [7] for finding the chemical features among a set of some amide derivatives given in Figure and quantitative structure-activity relationships was performed in order to determine the lead optimization by using the Hansch analysis method. The analysis was carried out on 27 analogues of which 17 were used in the training set and the rest considered for the test set. Physicochemical and indicator parameters were used in QSAR study.

The best equation obtained from the QSAR analysis is given below;

$$\log 1/C = 3{,}40\ (\pm 0{,}09) - 0{,}656\ (\pm 0{,}04)\mathbf{I_x} + 0{,}664\ (\pm 0{,}06)\ \mathbf{B_{1R_1}} + 0{,}249\ (\pm 0{,}05)\sigma\mathbf{R_2}$$

$(n = 17; r = 0.965; s = 0.0875; F = 118.86\ (p < 0.05); Q^2 = 0.9297; s\text{-PRESS} = 0.19889)$

The QSAR results reveal that a phenylacetamide system decreased the activity while the minimum width of position R_1 and having an electron withdrawing group on position R_2 increased the activity.
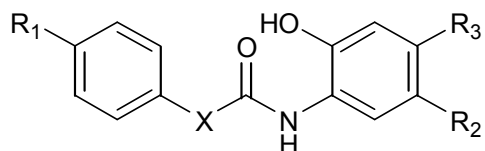
$X = -, CH_2;$
$R_1 = H, F, Br, Cl, CH_3, C_2H_5, C(CH_3)_3;$
$R_2 = R_3 = H, NO_2, NH_2$

1. A. Dalhoff, *Infection,* **1994**, 22, 111.
2. V. Lee, S. Hecker, *J. Med. Chem.*, **1999**, 19, 521.
3. D. Livermore, *Int. J. Antimicrob. Agents,* **2000**, 16, S3.
4. K. Poole, *Curr. Opin. Microbiol.*, **2001**, 4, 500.
5. D. Abbanat, M. Macielag, K. Bush, *Expert Opin. Investig. Drugs,* **2003**, 12, 379.
6. T. Ertan, I. Yildiz, S. Ozkan, OT. Arpacı, F. Kaynak, E. Aki-Sener, I. Yalcin, U. Abbasoglu, *Bioorg. & Med. Chem.*, **2007**, 15, 2032.
7. Accelrys Inc. Catalyst 4.9, **2004.**

COMBINATIONAL QSAR MODELING OF CHEMICAL TOXICANTS TESTED AGAINST *TETRAHYMENA PYRIFORMIS*

Hao Zhu[1], Denis Fourches[2], Alexandre Varnek[2], Ester Papa[3], Paola Gramatica[3], Tomas Oberg[4], Igor Tetko[5], Alexander Tropsha[1]

[1]CB # 7360, Beard Hall, School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA; [2]Faculte de Chimie, Universite Louis Pasteur, 4, rue B. Pascal, Strasbourg 67000, France; [3]Department of Structural and Functional Biology, University of Insubria, Via J.H. Dunant 3 - 21100 Varese, Italy; [4]University of Kalmar, Department of Biology and Environmental Science, SE-391 82 Kalmar, Sweden; [5]Institute for Bioinformatics, Munich Information Center for Protein Sequences (MIPS), Ingolstaedter Landstrasse 1, 85764 Neuherberg (bei Munich)

Selecting suitable quantitative structure-activity relationships (QSAR) approaches for a specific toxicity endpoint and accurate toxicity assessment for both drug candidates and environmental chemicals is one of the critical issues in computational toxicology. We have compiled an aqueous toxicity dataset containing 1,093 unique compounds tested in the same laboratory over several years against *tetrahymenapyriformis*. A modeling set consisting of 644 compounds was randomly selected from the original set and distributed to five cheminfomatic groups of co-authors. The remaining 449 compounds in the original set were used as an evaluation set to test the predictive power of individual models. Each group employed their favorite QSAR approaches and descriptors for model development. Jointly, our virtual collaboratory generated 11 different types of validated QSAR toxicity models for the training set. The best models had the Leave One Out (LOO) cross-validation correlation coefficient $R^2(q^2)$ of 0.93 for the training set and the correlation coefficient R^2 for the external evaluation sets as high as 0.83. The results demonstrated that the evaluation of the models based on the statistical parameters obtained for the modeling set only may mislead the selection of the externally predictive models. We have developed a consensus model based on averaging the aqueous toxicity values predicted by 11 individual models. The consensus model resulted in the best prediction accuracy for both training and external evaluation sets with q^2 as high as 0.95 and R^2 of 0.86, respectively. The model applicability domain could be used to balance the prediction accuracy vs. the chemistry space coverage based on the user tolerance with respect to the error level. This study presents an example of a fruitful international collaboration between researchers that use different QSAR techniques but share general principles of model development and validation.

MOLECULAR MODELLING AND MEDICINAL CHEMISTY OF BIOACTIVE POLYCYCLIC CAGE COMPOUNDS

Douglas W. Oliver, Sarel F. Malan, Alain J. M. Carpy

Unit for Drug Research and Development, School of Pharmacy, North-West University Potchefstroom Campus, Potchefstroom, South Africa

Saturated polycyclic hydrocarbon structures such as the monocyclic octane, bicylic norbornane and tricyclic adamantane has drawn the attention of several research groups since the 1930s. In the 1950s the synthesis of the so called "bird-cage", pentacyclo[$5.4.0.0^{2,6}.0^{3,10}.0^{5,9}$]undecane-8,11-dione **1** (Scheme 1), also known as Cookson's diketone was reported. This pentacyclic cage diketone is the product of the intramolecular photocyclized Diels Alder adduct of *p*-bensoquinone and cyclopentadiene.

Scheme 1
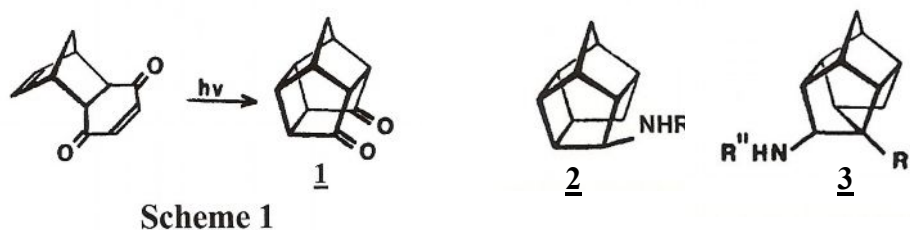
The conversion of this diketone to its monoketone analog formed the basis of a variety of mono substituted derivatives of type **2**. Furthermore, acid based rearrangement reactions of hydoxyl substituted compounds lead to amongst others, the unique D_3-trishomocubane symmetrical compounds, of type **3**, which consists of only five membered carbon rings. The D_3 stereoisomerism of the trishomobubane affords unique chemical challenges with potential medicinal implications. The medicinal chemistry of these cage compounds gained momentum in the 1980s with the discovery of the calcium channel modulating effects and anti-viral activity thereof. Our research efforts identified biological activites for these structures in various pharmacological areas i.e., dopaminergic, catecholaminergic, NMDA, focussing on disorders, in particular that of the central nervous system, such as neurodegeneration (Parkinson's Disease). Molecular modelling studies of these cage structures furthermore revealed unique interactions with HIV protease and nitric oxide synthase enzymes. These polycyclic structures have proofed to be very useful in drug discovery research in particular during the past 25 years.

PHARMACOPHORE ANALYSIS OF TOPOISOMERASE II INHIBITORY ACTIVE BENZAZOLES

Ismail Yalcin, Esin Aki-Sener, Ilkay Yildiz, Ozlem Temiz-Arpaci, Berk Zafer, Betul Tekiner-Gulbas, Sabiha Alper

Pharmaceutical Chemistry Dept., Faculty of Pharmacy, Ankara University, Tandogan 06100 Ankara, Turkey

Since the activity of topoisomerases is essential for several cellular processes such as replication, transcription and chromosome condensation, investigation of inhibitory activities of eukaryotic topoisomerases is widely used in anticancer drug development.

Their antitumor activity is related to the formation of protein-concealed DNA strand breaks, resulting in the stabilization by the drug of an intermediary complex of the Topo II reaction [1]. These drug-induced cleavable complexes have been proposed to be the primary action responsible for the antitumor activity.

In this study, a new series of 2,5-disubstituted-benzoxazole and benzimidazole, benzothiazole and oxazolopyridine derivatives, has been investigated for their inhibitory activity on eukaryotic DNA topoisomerases II in cell free system. The goal of this research was that predictions from the structure activity relationships of these tested compounds possibly will lead to design more active new DNA topoisomerases II inhibitors. Their activities were found to be between 11.4 and 433.2 IC_{50} µM. Etoposide was used as a reference drug for the inhibitory effect.

For the pharmacophore analysis using Catalyst [3] for finding the chemical features among a set of some Topo II inhibitor compounds as given below.
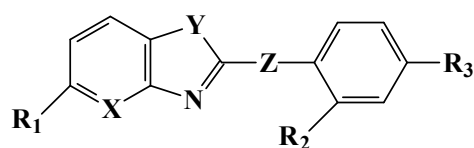
X= CH, N
Y= NH, S, O
Z= -, CH_2, CH_2S, CH_2O, CH_2CH_2

R_1= CH_3, $COOCH_3$, NO_2, H, Cl, NH_2
R_2= H, F, CH_3
R_3= H, NH_2, Cl, NO_2, CH_3, C_2H_5, OC_2H_5, OCH_3

1. P. D'Arpa, L.F. Liu. *Biochim Biophys Acta*, **1989**, 989, 163-177.
2. A. Pinar, P. Yurdakul, I. Yildiz, O. Temiz-Arpaci, N.L. Acan, E. Aki-Sener, I. Yalcin. *Biochim Biophys Res Comm*, **2004**, 317, 670-674.
3. Accelrys Inc. Catalyst 4.9, **2004**.

MOLECULAR MODELING OF MEMBRANE PROTEINS – PERSPECTIVE DRUG TARGETS

Roman G. Efremov

M.M. Shemyakin & Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Ul. Miklukho-Maklaya, 16/10, 117997 GSP, Moscow V-437, Russia; E-mail: efremov@nmr.ru

Membrane and membrane-active peptides and proteins (MPs) play crucial role in numerous cell processes acting as highly specific and efficient drugs or drug targets. MPs therefore attract growing interest in medicine and biotechnology. Because of experimental difficulties with their structural characterization, essential attention is given now to molecular modeling. Here we review our recent results on computer simulations of different types of MPs (antimicrobial peptides, transmembrane (TM) segments of MPs and their oligomers, G-protein coupled receptors, and others) in various membrane-mimetic environments (lipid bilayers, detergent micelles) [1-2]. To gain insight into structural and functional properties of the objects under study, a combined *in silico* approach was elaborated. It combines molecular dynamics and Monte Carlo simulations of MPs in implicit and explicit membranes, detailed analysis of their hydrophobic properties, and assessment of the quality of TM domains in proteins using original empirical scoring function. Testing against experimental data showed that the calculations permit correct assessment of a number of phenomena which accompany binding of MPs to membranes, determine their spatial structure in the membrane-bound state, and drive peptide-peptide interactions in lipid bilayers. The modeling results were shown to be useful for rational design of new biologically active molecules – putative prototypes of drugs. Among them are antimicrobial peptides with reduced hemolytic activity, peptides capable of binding to TM domains of receptor tyrosine kinases and others. The proposed approach extends considerably the class of studied biological problems. Perspectives of the modeling techniques in drug design are discussed.

1. R.G. Efremov, D.E. Nolde, A.G. Konshina, N.P. Syrtcev, A.S. Arseniev, *Curr. Med. Chem.*, **2004**, 11:2421.
2. R.G. Efremov, A.O. Chugunov, T.V. Pyrkov, J.P. Priestle, A.S. Arseniev, E. Jacoby, *Curr. Med. Chem.*, **2007**, 14:393.

APPLICATION OF THE NOVEL MOLECULAR ALIGNMENT METHOD USING THE HOPFIELD NEURAL NETWORK TO 3D-QSAR

Kimito Funatsu

The University of Tokyo, Department of Chemical System Engineering, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan; E-mail: funatsu@chemsys.t.u-tokyo.ac.jp

Comparative Molecular Field Analysis (CoMFA) has been widely used as a powerful 3D-QSAR (Quantitative Structure-Activity Relationship) tool in the field of medicinal chemistry. CoMFA is frequently used as standard QSAR technique, but some problems still remain. The molecular alignment is one of the key problems in QSAR study. In the CoMFA and most other 3D-QSAR techniques, a proper alignment between molecules is necessary. Recently, we invested and proposed the novel molecular alignment method with Hopfield Neural Network (HNN). This alignment method is based on methodology which solves the pattern-matching problem. The molecules are represented by four kinds of chemical properties (hydrophobic group, hydrogen-bonding acceptor, hydrogen-bonding donor, and hydrogen-bonding donor/acceptor), and then those properties between two molecules are corresponded each other using HNN. In this paper, we apply the molecular alignment method to three-dimensional quantitative structure-activity relationship (3D-QSAR) analysis. Two data sets (Human epidermal growth factor receptor-2 inhibitors and cyclooxygenase-2 inhibitors) were investigated to validate our method. The robust and predictive CoMFA models could be successfully obtained in both data sets.

AB INITIO DOCKING METHOD COMBINING CAVITY SEARCH WITH MOLECULAR DYNAMICS

Yury N. Vorobjev

Institute of Chemical Biology and Fundamental Medicine of Siberian Brunch of Russian Academy of Sciences, 8 Ac.Lavrentiev Ave., Novosibirsk 630090, Russia

A hierarchical algorithm of *ab initio* docking of a flexible ligand is developed and implemented. The algorithm combines exhaustive cavity search on protein surface and calculation of docking surface grid with global optimization of ligand positions via molecular dynamics: 1) calculation of a probe accessible surface (PAS) of protein; 2) analysis of the PAS and calculation of a docking surface grid as a virtual positions of accommodation of a probe chemical groups; 3) estimation of the quality score for docking surface grid points; 4) making of PAS grid compatible ligand image; 5) search of ligand image positions on the docking surface grid and preliminary ranking of a set of binding sites; 6) global rotational/translational optimization of the ligand for the list of preliminary binding sites via full atom molecular dynamic simulating annealing with variable force field for flexible ligand and protein. Docking method is tested on a variety of protein-ligand complexes for a small and large ligands with size up to eight residue-peptide. The docking method successively finds a set of binding sites and the major binding mode as a mode with highest binding energy estimated at the final stage of MD optimization. With high probability, more than 90% , the major binding mode are in agreement with X-ray data, within RMSD < 3 E for ligand atoms, for tested protein-ligand complexes with variety of proteins and ligands: benzamidine, biotin, Ile-Val, argatroban, and complexes of HIV-1 protease with XK263, VAC, KNI272, etc.

MFTA-BASED DESIGN OF ACTIVE STRUCTURES

<u>Vladimir A. Palyulin</u>, Eugene V. Radchenko, Andrey A. Melnikov, Nikolay S. Zefirov

Department of Chemistry, Moscow State University, Moscow 119992 Russia; E-mail: vap@org.chem.msu.su

Molecular Field Topology Analysis (MFTA) [1] is a useful approach to QSAR modelling and molecular design of novel promising structures. MFTA model is based on a so-called molecular supergraph – a simple graph such that the molecular graphs of all training set structures can be represented as its subgraphs. Uniform descriptor set for statistical analysis is obtained by superimposing the structure onto the molecular supergraph. Each supergraph vertex is assigned the values of effective atomic charge, van der Waals radius and/or other local parameters for the corresponding atom. For unoccupied vertices the neutral descriptor values are used.

Especially interesting is the joint application of QSAR models and the structure generator designed for QSAR studies [2] in search for novel bioactive compounds. However, existing general-purpose generators do not take into account the features of MFTA-based QSAR models. On one hand, it is difficult to obtain all the structures of interest while avoiding the combinatorial explosion. On the other hand, many generated structures fall outside the applicability domain of an MFTA model.

We have developed a method of molecular graph generation in conjunction with MFTA models based on the consideration of fragmental supergraph and structural constraints (e.g. preventing the generation of unstable, toxic or over-decorated structures). In deterministic mode, it builds all possible connected molecular graphs that can be represented as subgraphs of MFTA supergraph and satisfy the required constraints. In stochastic mode, a representative subset of this structure set is constructed. The proposed approach is implemented in the convenient Java-based software. The tests show that it facilitates the search for compounds having better predicted activity [3].

1. V.A. Palyulin, E.V. Radchenko, N.S. Zefirov, *J. Chem. Inf. Comp. Sci.*, **2000**, 40 (3), 659-667.
2. A.A. Melnikov, V.A. Palyulin, N.S. Zefirov, *Doklady Chemistry*, **2005**, 402 (1), 81-85.
3. A.A. Melnikov, E.V. Radchenko, V.A. Palyulin, N.S. Zefirov, *Doklady Chemistry*, **2007**, in press.

OPTIMIZING THE USE OF TOXICITY PREDICTION TOOLS ACROSS DRUG DISCOVERY STAGES

D. Domine, C. Cleva, E. Sebille, W. Sauer, V. Barbie, C. Merlot

Merck Serono Geneva Research Center (Geneva, CH)

It has now been recognized for long that in silico ADMET (absorption, distribution, metabolism, excretion, and toxicity) models are important tools in combating late-stage attrition in the drug discovery process. As a result, a plethora of models has been designed and is now available to monitor and optimize these properties along the development path. In this presentation, some external and internal in silico toxicity models are presented and their optimal use is discussed with respect to the stage at which they are applied. We show that successful safety monitoring relies on the close integration between in-silico and experimental approaches and that different combinations of in-silico methods are needed for different stages in the drug discovery process. Methods for automated extraction and combination of software outputs are discussed. Practical case studies dealing with library design and lead optimization are presented.

MECHANISM BASED COMMON REACTIVITY PATTERN (COREPA) MODELING OF AHR BINDING AFFINITY

Petko I. Petkov[1], R. Serafimova[1], J. C. Rowlands[2], R. Budinsky[2], M. Denison[3], O. Mekenyan[1]

[1]Laboratory of Mathematical Chemistry, Bourgas As. Zlatarov University, 8010 Bourgas, Bulgaria; [2]Toxicology and Environmental Research & Consulting, 1803 Building, The Dow Chemical Company, Midland, Michigan 48674; [3]Department of Environmental Toxicology, Meyer Hall, One Shields Avenue, University of California, Davis, CA 95616 USA; E-mail: p_petkov@btu.bg

The aryl hydrocarbon receptor (AHR) is a ligand-activated transcription factor responsive to both natural and synthetic environmental compounds, with the most potent agonist being 2,3,7,8-tetrachlotrodibenzo-p-dioxin (TCDD or dioxin). Dioxins are a group of environmental contaminants that raise concern because of their widespread presence and persistence. AHR mediates a wide variety of biochemical and toxic effects of these chemicals. While the specific mechanism for these toxic effects is not known, there is evidence suggesting that the AHR is involved in metabolic transformation of natural aromatic hydrocarbons, as a part of more complicated mechanism that may involve alteration of cell cycle control. The aim of this work was to develop a COmmon REactivity PAttern (COREPA) based model of AHR ligands within different AHR reporter gene inducing activity ranges thus building a categorical SAR model. The training set of AHR ligands included dioxins, dibenzofuranes, biphenyls, PAHs, flavonoids, indirubins and aromatic hydrocarbons. The synthetic ligands in the training set were ranked into three reporter gene inducing activity classes: (1) strong inducers with $EC_{50} < 10$ nM (80 chemicals), (2) weak inducers with EC_{50} ranged between 10 and 100 nM (57 chemicals), and (3) non-inducers with $EC_{50} > 100$ nM (87 chemicals). The conformational distributions of chemicals were analyzed and compared to define the commonality between biologically similar chemicals within 2D and 3D structural space. The COREPA analysis suggested two different binding mechanisms called dioxin- and biphenyl-like, respectively. The dioxin-like model predicts a mechanism that requires a favorable interaction with a receptor nucleophilic site in the central part of the ligand and with electrophilic sites at both sides of the principal molecular axis. Dioxin-like ligands in this class included dioxins, dibenzofuranes, flavonoids and indirubins. The biphenyl-like model predicted a stacking type interaction with the AHR allowing electron charge transfer from the receptor to the ligand; which was confirmed by the appearance of the energies of frontier molecular orbitals (E_{LUMO}) as a discriminating parameter. The total concordance of the derived categorical model was 84% whereas the Pearson's coefficient was - 0.90. Additional COREPA models were based on descriptors such as planarity, energy gap and partially charged surface area and supported the hypothesis for a strong H-bonding interaction between reactive and planar aromatic molecules and the receptor.

FIELDCHOPPER, MOLECULAR FIELD-BASED VIRTUAL SCREENING METHOD. DESCRIPTION AND EVALUATION

Tuomo Kalliokoski, Toni Ronkko, Antti Poso

Department of Pharmaceutical Chemistry, University of Kuopio, P.O.Box 1627, FIN-70211, Finland

Virtual Screening (VS) is a routine part of modern drug discovery. Ligand-based VS has been shown to be effective. There are plethora of methods available. Others use single molecule structure as a query while others use a group of molecules for screening. Here we present a 3D-method called FieldChopper (FC) that uses multiple ligands to build a model based on molecular fields which then can be used to screen a molecular database. FC uses electrostatic and Van der Waals molecular fields represented as grid boxes with spacing of 1E. FC requires a fast molecular superimpositioning method like BRUTUS or ROCS.

FC models are built from several active compounds. Model building algorithm creates a histogram of field values. From this histogram, peaks are detected and are used in the scoring algorithm. Scoring algorithm requires molecules to be superimposed to the template used in the model building process. A molecule is compared to model point-by-point and total

score consists of electrostatic and Van der Waals scores.

To evaluate performance of FC, we conducted retrospective virtual screening using recently published DUD dataset. 12 targets were selected based on number of ligands. We compared the effect of FC scoring to BRUTUS and ROCS based superimpositions with Receiver Operating Characteristic (ROC) curve and ROC Area Under the Curve (ROC AUC). Templates for superimpositioning were taken from Protein Databank protein-ligand complexes. 15 ligands per target were selected for the models using GRIND-descriptors.

FC yields higher ROC AUCs than basic BRUTUS or ROCS scoring in most cases and seems to be independent from superimpositioning algorithm used. These results suggest that this method may be useful in VS when several active compounds are known.

QSAR STUDY ON A SET OF QUINOXALINE COMPOUNDS AS ANTITUBERCULOSIS DRUGS USING THE COUNTER PROPAGATION NEURAL NETWORKS

Marjan Vracko[1], Manish C. Bagchi[2], Sisir Nandy[2], Payel Gosh[2]

[1]National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia; [2]Indian Institute of Chemical Biology, 4 Raja S. C. Mullick Road, Jadavpur, Calcutta 700032, India; E-mail: marjan.vracko@ki.si

A set of 47 quinoxaline compounds and their activities against *Mycobacterium tuberculosis* expressed as MIC (minimum inhibitory concentration) have been used for QSAR modelling. The structures have been represented with descriptors, which can be classified as constitutional, electrostatic, geometrical, physicochemical and topological ones. As the modelling technique the self organizing maps (SOM) and counter propagation neural networks (CP NN) have been selected, both of the methods have been often used for QSAR modelling. Several models were built up and tested considering different classes of descriptors. Generaly, the models built with constitutional, topological, and electrostatic descriptors seem to be superior to the models built with geometrical and physicochemical ones. In further steps we have applied the counter propagation neural networks to analyze clusters in the set of compounds and for classification of compounds. Four classes have been defined accordingly to the activity (MIC) of compounds. The classification results for leave one out test are: 49% of compunds are classified corectly with constitutional descriptors, 40% with electrostatic ones, 36% with geometrical ones, 40% with physicochemical ones, and 30% with topological descriptors.

STRUCTURAL-ACTIVITY RELATIONSHIP (SAR), ATOMIC ELECTRON DENSITY AND CONFORMATIONAL INVESTIGATION OF FENTANYL ANALOGUES (FA)

Hossein Fakhraian, Teimour Nezamoleslam, Monireh Babaei Panbehriseh, Banafsheh Javahery

Department of Chemistry, Imam Hossein University, Tehran, IRAN; E-mail: Fakhraian@yahoo.com

Fentanyl analogues (FA) are highly potent and clinically widely used narcotic analgesic and represent a particular class of μ agonist characterized by highly potency, rapid onset and short duration of action. A very large number of FA have been synthesized, with the aim of establishing SAR and to find clinically more useful drugs with better pharmaceutical profile such as high potency and less side effects (e.g. respiratory depression and lower addiction potential). The most well-known FA are fentanyl, 3-methyl fentanyl, ohmfentanyl, lofentanyl, carfentanyl, alfentanyl and sufentanyl (Scheme 1).
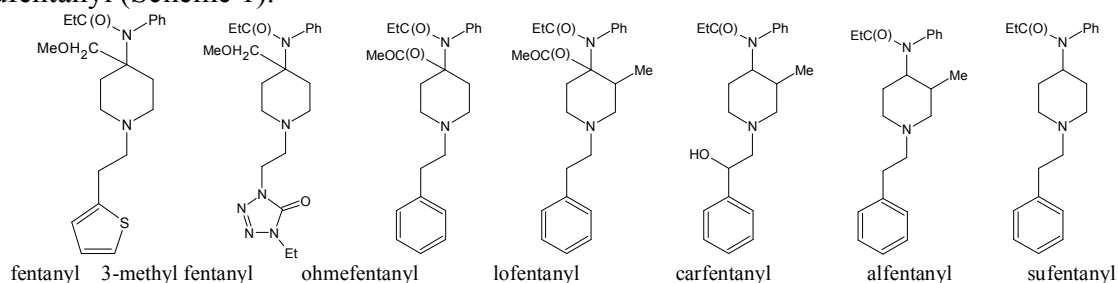
Scheme (1)

ED_{50} of different FA are reported but these data are not all considered together to deduce structural-activity relationship. In this contribution, the effect of 5 groups contributing to the analgesic activity of FA (70 compounds) are reinvestigated. The best groups are determined and more effective fentanyl analogue is suggested (Scheme 2).
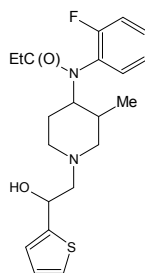
Scheme (2)

We have tried to determine the relative weight of effective groups on ED_{50} of FA, but comparison of different data indicates that the actions of different groups on ED_{50} are not independent. In another part, atomic electron density and conformational studies of FA by Abinitio methods were correlated with their respective ED50.

1. N. Lalined, J. Moliterni, D. Wright, H. Spencer, M. Ossipov, *J. Med. Chem.,* **1990**, 33, 2876.
2. L.D. Micovic, *J. Serb. Chem. Soc.,* **2004**, 69(11), 843 (Review paper).
3. F. Janssen, J. Torremans, P.A.J. Janssen, *J. Med. Chem.,* **1986**, 29, 2290.

PHARMACOPHORE IDENTIFICATION BY DATA MINING

Takashi Okada, Masumi Yamakawa, Satoshi Fujishima, Norito Ohmori

School of Science and Technology, Kwansei Gakuin University, Gakuen 2-1, Sanda, Hyogo, 669-1337, Japan; E-mail: okada-office@ksc.kwansei.ac.jp

Pharmacophore identification is a pending issue, when we have to extract it from a database of drugs with a variety of skeletons. Our current project is to create a knowledge base of pharmacophore for various activities. This presentation shows the methods employed in the pharmacophore identification using an example of dopamine D1 agonist activity.

Structural formulae with the agonist activity to D1, D2 and Dauto receptors were used as the learning data. We first generated linear fragments from the compounds. Each fragment is described by two terminal atoms and the connecting bonds using SMILES like notation. The number of adjacent atoms and the existence of a hydrogen atom are also included into the atom information. This description gives a higher support to important fragments with functional groups at the terminals, and chemists can understand their meanings.

The cascade model was used to generate rules that describe the essential conditions of descriptors to the D1 agonist activity. The method uses the BSS value to show the strength of a rule, which indicates a higher value when the rule condition largely changes the activity distribution and the supporting instances are numerous. For example, the strongest rule employed a catechol structure (O2H-c3:c3-O2H) as the main condition. It selected 52 from 369 compounds, whereby the D1 activity ratio increased from 17% to 96%. The application of this condition changes the appearance ratio of other fragments, and we could find that the ratio of ethylamine connected to an aromatic carbon also increased from 18% to 81%.

Finally, a knowledge refinement process in the structure space was applied to grasp the whole pharmacophore. It accepts the fragment in the main condition, and expands the structure by attaching various types of atoms and bonds. The substructure with the highest BSS value is selected and a greedy optimization is repeated while the BSS value does not decrease. Chemists can browse the supporting compound structures where the active and inactive compounds are depicted separately. We could recover a dopamine structure itself, starting from the catechol structure in the rule. Two other types of pharmacophore were also discovered which were judged to be reasonable by medicinal chemists.

NEW INSIGHTS FOR HERG INHIBITION USING MAPPING TECHNIQUES

Konstantin V. Balakin[1], Sean Ekins[2,3], Nikolay P. Savchuk[1], Andrey A. Ivashchenko[1], Yan A. Ivanenkov[1]

[1]Chemical Diversity Research Institute, Rabochaya 2a, Khimki, Moscow reg. 114401, Russia; [2]ACT LLC, 601 Runnymede Ave, Jenkintown, PA 19046; [3]Department of Pharmaceutical Sciences, University of Maryland, 20 Penn Street, Baltimore, MD 21201, USA; E-mail: kvb@iihr.ru

The human ether-ago-go related gene (hERG) can be inhibited by drugs and this may lead to QT prolongation and possibly fatal cardiac arrhythmia. We have collated literature data for 99 diverse hERG inhibitors to generate Kohonen self-organizing maps and Sammon nonlinear maps. These advanced mapping algorithms allow the generation of a two-dimensional image of a multidimensional property space for the analyzed objects, and thus represent powerful tools for visual analysis of the structure-activity relationships. Our aim was to investigate whether these models could be used to predict the binding of diverse molecules and offer novel insights into hERG inhibition. The mapping approaches used molecular descriptors required for hERG inhibition that were not reported previously, highlighting the importance of molecular shape. The Sammon map model provided the best qualitative classification of the test set (95% correct) compared with Kohonen map (81% correct). The consensus approach provided equivalent levels of correct predictions and may be more desirable in a pharmaceutical setting to avoid this undesirable property in new molecules by using the strengths of different approaches. This study illustrates that data from various sources can be combined to generate statistically valid models of hERG inhibition.

CHEMOMETRIC INVESTIGATIONS OF MULTIDRUG RESISTANCE IN STRAINS OF THE PHYTOPATHOGENIC FUNGUS *PENICILLIUM DIGITATUM*

Marcia M. C. Ferreira, Rudolf Kiralj

Instituto de Quimica, Universidade Estadual de Campinas, Campinas 13083-862, SP, Brazil; E-mail: marcia@iqm.unicamp.br

Demethylation inhibitor (DMI) resistance by pathogenic fungi is a serious problem in agriculture and medicine. *P. digitatum* (the green mold) causes important postharvest diseases of citrus fruits. The present work studies *P. digitatum* strains (DMI-resistant, moderately resistant and sensitive) and 4 DMIs and 3 non-DMIs by means of Principal Component Analysis (PCA) and Hierarchical Cluster Analysis (HCA) and Partial Least Squares (PLS). Novel types of relationships between toxicant structure and fungal resistance, and between the resistance and fungal genome, were established.

Biological activity datasets for *P. digitatum* strains with respect to DMIs (triflumizole, fenarimol, bitertanol, pyrifenox) and non-DMIs (cycloheximide, 4-nitroquinoline-*N*-oxide, acriflavine) were generated from experimental EC_{50} (effective inhibitory concentration for 50% radial growth inhibition) and MIC (minimal inhibitory concentration) from literature [1]. The datasets contained: pEC_{50}=-log(EC_{50}/mol dm^{-3}); $pECr_{50}$=pEC_{50}(standard)–pEC_{50}; descriptors (*a*, *b*, *c*, |a|, |c|) from regressions pMIC=a+b pEC_{50} and c=a/b for each toxicant; 8 and 16 morphological descriptors (radii, circumferences and areas of fungal cultures) for the growth of 39 strains without and with toxicant, respectively; pEC_{50} from multiple measurements for DMIs, non-DMIs and all 7 substances. Genome structure descriptors related to constitutive and toxicant-induced expression levels of *CYP51* and *PMR1* genes in diverse *P. digitatum* strains were generated, as well as their products with molecular descriptors for 4DMIs. The descriptors were correlated with the corresponding pEC_{50} activity at PLS level, applying leave-one-out crossvalidation and external validation.

Novel Activity-Structure Relationships (ASRs) resulted from exploratory analyses of the activity datasets. Relationships between toxicants structure and strain features (baseline resistance, morphology, origin/target) are visible, can be rationalized and used in predictions. Quantitative Genome-Activity Relationship (QGAR) and Quantitative Genome/Structure-Activity Relationship (QGSAR) are novel relationships between pEC_{50} and genome/toxicant descriptors, with satisfactory PLS statistics (QGAR: 3PCs, R=0.90, Q=0.89, SEV=0.34; QGSAR: 5PCs, R=0.93, Q=0.92, SEV=0.286). The primary contribution to DMI resistance comes from the *CYP51* gene, secondary from the *PMR1* gene and its interaction with toxicants. QGSAR and QGSAR can aid in detecting resistance strains of *P. digitatum* and development of novel antifungals.

1. R. Nakaune et al., *Mol. Genet. Genom.*, **2002**, 267, 179 and references cited therein.

ASSESSMENT AND MODELLING OF THE TOXICITY OF SUBSTITUTED AROMATIC COMPOUNDS TO FIVE AQUATIC SPECIES

Doga Erturk[1], Melek Turker Sacan[1], Safiye Sag Erdem[2]

[1]Bogazici University, Institute of Environmental Sciences, 34342, Bebek/Istanbul/Turkiye; [2]Marmara University, Faculty of Arts and Sciences, Chemistry Department, Goztepe Istanbul/Turkiye; E-mail: msacan@boun.edu.tr, doga@turcek.org.tr

A previous QSPR study with the toxicity of 39 aromatic compounds to *Scenedesmus obliquus* having strong relationship with the Characteristic Root Index (CRI) and the energy of the lowest unoccupied molecular orbital (E_{LUMO}) [1] encouraged us to search a similar relationship between the same descriptors and the toxicity of substituted aromatic compounds to other five aquatic species: protozoa (*Tetrahymena pyroformis*), bacteria (*Vibrio fischeri*), algae (*Chlorella vulgaris*), daphnid (*Daphnia magna*) and fish (*Poecilia reticulata*). The toxicity database in this study contains a number of polar narcotics such as aniline, phenol and other benzene derivatives such as monohalogenated and mononitro-substituted benzenes, anilines, and phenols, and compounds including dinitroanilines, mononitropolyhalogenated anilines and phenols. 48 h-LC_{50} values of 68 compounds were used for the model developed for protozoa while 30 min-EC_{50} values of 67 compounds for bacteria; 15 min-LC_{50} values of 17 compounds for algae; 48 h-LC_{50} values of 41 compounds for daphnid and 96 h-LC_{50} values of 31 compounds for fish were used to develop relevant QSPRs. QSPRs have been developed using the CRI and E_{LUMO} as descriptors for five aquatic species using multiple linear regression (MLR). Suitable QSPR models ($0.75 < r < 0.91$) to predict acute toxicity of substituted aromatic compounds were obtained. The CRI was a sufficient descriptor for all cases. However, inclusion of E_{LUMO} increased the predictive ability of the models developed for algae and protozoa. Two-descriptor models were developed for protozoa and algae whereas one-descriptor models were developed for bacteria, daphnid and fish. The predictive accuracy of the proposed models was compared with the commonly used K_{ow} model and recently published studies in which toxicity models were developed. Particular emphasis has been made to clearly define the boundaries for the application of the alternative developed models as well as the quality of estimates.

1. M.T. Sacan, M. Ozkul, S.S. Erdem, *Chemosphere*, **2007**, 68, 695-702.

THE ZAGREB APPROACH TO THE STRUCTURE-PROPERTY-ACTIVITY MODELING

Sonja Nikolic, Bono Lucic, Nenad Trinajstic

The Rugjer Boskovic Institute, P.O.B. 1016, HR-10002 Zagreb, Croatia; E-mail: sonja@irb.hr

The Zagreb approach is based on the CROMRsel procedure which is a multivariate procedure that has been designed to select the best possible model among the set of models obtained for a given number of descriptors, the criterion being the standard error of estimate [1]. The nonlinearities also introduced in the set of descriptors by enlarging it with squares and cross-products of initial descriptors. A wide range of descriptors, in their orthogonalized form, is used, including topological indices. The application of the Zagreb approach produced a number of successful QSPR/QSAR models [*e.g.*2-6].

1. B. Lucic, N. Trinajstic, *J. Chem. Inf. Comput. Sci.*, **1999**, 39, 121-132.
2. B. Lucic, N. Trinajstic, S. Sild, M. Karelson, A.R. Katritzky, *J. Chem. Inf. Comput. Sci.*, **1999**, 39, 121-132.
3. A.R. Katritzky, K. Chen, Y. Wang, M. Karelson, B. Lucic, N. Trinajstic, T. Suzuki, G. Schuurmann, *J. Phys. Org. Chem.*, **2000**, 13, 80-86.
4. B. Lucic, D. Amic, N. Trinajstic, *J. Chem. Inf. Comput. Sci.*, **2000**, 40, 403-413.
5. A. Milicevic, S. Nikolic, N. Trinajstic, A. Juric, *Molecules,* **2004**, 9, 1208-1221.
6. D. Amic, D. Davidovic-Amic, D. Beslo, V. Rastija, B. Lucic and N. Trinajstic, *Curr. Med. Chem.,* **2007**, 14, 827-845.

MEDICINAL CHEMISTRY AND MATHEMATICS

Nilolay S. Zefirov
Full Member of Rus. Acad. Sci., Full member of International Academy of Mathematical Chemistry, Distinguished Professor of MSU; E-mail: zefirov@org.chem.msu.ru

Lecture will consider basic principles of application of discrete mathematics to organic & medicinal chemistry problems.

The points will be considered are the following:

   A. Structural generations and manipulations.

      1. Structural generators.

      2. Enumerations.

      3. Diversity and similarity. Descriptors.

      4. Search for novel reactions (Formal-logical approach).

   B. Structure – activity (property) relationships.

      1. QSAR and Neural Networks.

      2. Molecular modeling and drug design (exemplified by successful search of anti-Alzheimer drug).

SURVEY OF PSORIASIS CANDIDATE GENES BY USING BIOINFORMATICS

Roustam M. Abdeev[1], Sergey A. Brouskin[2], Tatiana A. Nikolskaya[2,3], Eleonora S. Piruzian[2]

[1]Center for Theoretical Problems of Physical-Chemical Pharmacology Russian Academy of Sciences, Kosigin str. 4, Moscow, Russia, 119991; [2]Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin str. 3, Moscow, Russia, 119991; [3]«Metalogik» Ltd, Gubkin str. 3, Moscow, Russia, 119991; E-mail: art-arm@yandex.ru

Psoriasis is a complex genetically determined skin disorder that occurs as a consequence of interactions between wide groups (networks) of genes/proteins. The main criterion for searching of genes candidates for pathological process in specific organs and tissues is a significant change in their expression level. The strategy also includes the selection of transcription factors that triggered the change of expression in large gene groups. Up to date, microarrays is a highly throughput and informative tool for studying the expression of large gene groups. In our bioinformatics researches we used GEO DataSets (http://www.ncbi.nlm.nih.gov/geo/) database, contained results of microarray experiments estimating the level of genes' expression. To analyze and systematize huge amount of incoming genetic information we've used software "MetaCore" produced by GeneGO Inc (USA). The level of expression of about 12000 genes estimated in microarrays experiments was analyzed. The number of genes with more then two fold change in the level of its expression under psoriasis condition was shown to be 1118. To visualize genetic processes at the level of gene network interactions we used maps of network gene interactions with superposed changes in the levels of genes expression. The selected maps of gene interactions described the main alternated processes in the cell. Thus, it is remarkable that many processes where expression pattern has been changed lead to expression increase of important transcriptional factors (Jun, Fos, STAT families etc) participating in regulation of essential for cell and organism processes such as cell cycle and apoptosis, proliferation and cell adhesion. We hypothesize that for expression change of large co-working gene groups involved in some cellular processes key transcriptional factors triggering signaling cascades should change its expression. That is why, genes that have significantly changed its expression level during psoriasis and belonged to transcriptional factors are potential genes-candidates for psoriasis. Further our researches will be focused on experimental support of this hypothesis.

APPLICATION OF THE QUANTUM-CHEMICAL ANALYTICAL METHODS TO SUBSTANTIATE THE ANTIRADICAL ACTIVITY IN THE FLAVONE, CHALCONE AND CINNAMIC ACID HYDROXYDERIVATIVE SERIES

Vladimir Agadjanyan, Eduard Oganesyan

Pyatigorsk State Pharmaceutical Academy, Pr. Kalinina 11, Pyatigorsk, 357532, Russia; E-mail: avsvova@mail.ru

Reactivity indexes in the series of the cinnamic acid – chalcone – flavone hydroxy derivatives in respect of active oxygen forms have been investigated by semi-empirical quantum – chemical methods (PM3, AM1). The new index – summary unsaturation index was proposed according to the values of the bond orders, bonding numbers and valencies of the appropriate carbon atoms in the structures analyzed. It quantitatively equals to the sum of differences between the valency and the bonding numbers for each carbon atom in the molecule and, to our opinion, it soundly characterizes the compound capability to play the role of free radical traps. Taking into account the enthalpy formation of the intermediate adducts of the reactions with the hydroxyl radical the electronic density distribution as well as the p_z – orbital contribution into the top occupied molecular orbital for the compounds studied we came to the conclusion that the primary radical attachment centre localizes on the C_3 atom of the propenone fragment. Calculation of the H – O bond order and enthalpy formation of the intermediate radicals in the homolytic bond splitting allowed to conclude that the compounds with the hydroxygroup at flavone nucleus C_3 atom as well as at $C_{3'}$, $C_{4'}$ of the B ring are likely to show the highest antioxidant activity. On the basis of the quantum – chemical analysis data and pharmacological activity prognostic results of still virtual structures (PASS - program) the synthesis of some cinnamic acid hydroxyderivatives has been carried out. The data obtained on the basis of the preliminary pharmacological screening are sure to correlate with those of the quantum – chemical analysis and the prognostic results of the compound activity according to PASS program. The cinnamic acid hydroxyderivatives proved to possess the capability to act not only as free radical traps but also as substances preventing the enzyme inactivation of the antioxidant defence, in particular superoxide dismutase, that is considered to be of great importance.

STUDY OF PEPTIDE BOND DEFORMATION IN MODEL DIPEPTIDES BY THE SEMIEMPIRICAL QUANTUM CHEMISTRY METHODS

Ruslan E.Aliev, Svetlana D.Demukhamedova, Irada N.Alieva, Niftali M.Godjaev

Institute for Physical Problems, Baku State University, Z.Khalilov str., 23, Baku, AZ 1148, Azerbaijan; E-mail: svetlanabest@mail.ru

Quantum chemical calculations for the model dipeptide molecules, composed of L- and D-stereoisomers of arginine and proline amino acids were carried out by the CNDO/2, CNDO/M and MINDO/3 semiempirical methods on the base of the program LEV. The program worked out at the Institute after name of Vernadsky of the Russian Academy of Sciences [1].

The main aim of this work is investigation of the relationship between electronic structure of the L-Arg-L-Pro, D-Arg-L-Pro, L-Arg-D-Pro, D-Arg-D-Pro and spatial structures of the model dipeptide molecules. The effects of the peptide bond deformation at the scanning along C-N bond were investigated. In order to reduce the effect of terminal polar groups the N-end of the dipeptides was modified by acetyl and the C-end was modified by N-methylamide. As a result two additional peptide bonds are constructed.

At first, theoretical conformational analysis was used to study the spatial structure and conformational properties of dipeptide molecules. The low-energy conformations of these molecules were found, the values of dihedral angles of the backbone and side chains of the amino acid residues constituting these peptides were determined, and the energies of intra- and interresidual interactions were estimated. Calculation models were constructed on the base of coordinates of atoms in accordance with results of theoretical conformational analysis. The main electronic parameters such as electron density distribution, electron and nuclear forces, total dipole moments and dipole moments of individual bonds in dependence of low-energy conformational state were analyzed. The calculation results have shown that the D-izomerization of the arginine residue in the positions 1 and 2 accompanied by the large changes of the scanning energy along C-N peptide bond. In result the following conclusion was made: C-N peptide bond in L-Arg-L-Pro dipeptide is more stable in comparison with other dipeptide molecules.

1. Gribov L.A. Light elucidation of vibrations. The description of the program and instruction to the user // OOO "INLAN", **1995**.

FREE-WILSON STUDY ON SOME MICROBIOLOGICALLY ACTIVE BENZAZOLES

Sabiha Alper-Hayta, Ilkay Yildiz, Ozlem Temiz-Arpaci, Esin Aki-Sener, Ismail Yalcin

Ankara University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, 06100 Tandogan, Ankara - Turkey

The Free-Wilson approach is an application of multiple regression analysis of QSAR methodology. The basic assumption of Free-Wilson analysis is that within a homologous series of drugs, individual segments of molecules make additive and constant contributions to biological activity. If such contributions are known, biological activity can be estimated by simple addition for all the compounds obtainable by any new combination of segments involved. As Hansch analysis, Free-Wilson analysis can be applied to homologous series where only substituents are varied in a constant molecule [1-3].

Because of the need of new and different antibacterial agents and the many effective antimicrobial drugs possess heterocyclic systems, in this study, QSAR analysis of some novel isosteric heterocyclic compounds including benzoxazole, benzimidazole, benzothiazole, and oxazolo(4,5-b)pyridine derivatives [4] shown in figure were determined for the biological activity against Gram-positive bacterium, *Staphylococcus aureus*. Predictions for the lead optimization have been described by the results obtained from Free-Wilson analysis with the determination of the activity contributions of the structural descriptors performed by using IBM-computer working with MINITAB statistic package.
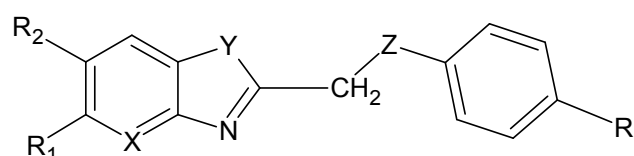
R = H, Cl
R_1 = H, NO_2, Cl, CH_3, $COOCH_3$
R_2 = H, NO_2, CH_3
Y = O, S, NH
Z = O, S, NH, CH_2
X = CH, N

Figure

1. S.M. Free and J. M. Wilson, *J. Med. Chem.*, **1964**, 7, 395-399.
2. M. E. Wolff, Burger's Med. Chem. John Wiley & Sons Ltd. New York, Vol: I, **1980**, 407.
3. R. Franke, Theo. Drug Design Methods, Elsevier Science Pub., Amsterdam, Vol:VII, **1984** 160, 256.
4. I. Yildiz-Oren, I. Yalcin, E. Akı-Sener, N. Ucarturk, *Eur. J. Med. Chem.*, **2004**, 39, 291-298.

A DEVELOPMENT OF NEAREST NEIGHBOR METHOD FOR LIPOPHILICITY CALCULATION

Elena P. Andreeva, Oleg A. Raevsky

Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds, Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia, E-mail: andreeva@ipac.ac.ru

8 subsets of chemicals containing diverse compounds with its at least 20 structure relatives (Tanimoto coefficient ≥ 0.5) were chosen (software MOLDIVS) from the main database of 11514 compounds inclucive experimental logP values (software SLIPPER). The dataset composed of all these subsets (245 compounds) has been formed. In each datasets molecular were calculated (software HYBOT).

Linear models for logP calculation with 4, 3 and 2 descriptors (polarizability (α), $\sum C_a$, $\sum C_d$, $\sum Q^-$) have been built for 8 subsets, for the whole database (11514 compounds) and for composed dataset (245 compounds). The best criteria were chosen for each datasets on the base of statistical data. Coefficients of linear equations (b_i) were used to construct calculation models for prediction logP by nearest neighbor method:

$$logP = logP_n + b_1*(\alpha - \alpha_n) + b_2*(\Sigma C_d - \Sigma C_{dn}) + b_3*(\Sigma Q^- - \Sigma Q^-_n) + b_4*(\Sigma C_a(o) - \Sigma C_a(o)_n).$$

Lipophilicity calculation for each subset based on the whole database (11514), composed dataset (245) and itself was made. Calculated logP values have been compared with experimental logP values and statistical criteria were obtained for each case of prediction.

A conclusion that the most stable is a model of prediction based on two molecular descriptors α and $\sum C_a$ (o) was made on the base of all calculations.

Three modes of calculation logP values for different training sets based on abovementioned equation was curried out using as a main database for prediction the whole database 11514, the proper dataset 245 and different 8 subsets. The comparison of statistical criteria for of calculation logP values and experimental logP values was provided.

The results demonstrate that logP calculation models on the base of structure similarity and physicochemical descriptors such as molecular polarizability and hydrogen-bond acceptor factor for selected subsets of chemicals are stable independently of used modes of calculations.

COMPUTER-BASED PREDICTION OF ACTIVITY AND EXPERIMENTAL TESTING OF N-GLYCOSIDES OF SUBSTITUTED INDOLO[2,3-a]CARBAZOLES IN SEARCH FOR NEW ANTITUMOR SUBSTANCES

Galina N. Apryshko[1], Irina S. Golubeva[1], Olga V. Gorunova[1], Nadia P. Yavorskaya[1], Dmitry A. Filimonov[2], Vladimir V. Poroikov[2]

[1]Russian Cancer Research Center of Russian Academy of Medical Sciences, 24, Kashyrskoe Shosse, Moskow, 115478, Russia; [2]Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: g48@mail.ru

Computer system PASS version 1.917 (Prediction of Activity Spectra for Substances, http://www.ibmc.msk.ru/PASS) was used for predicting cytostatic and antineoplastic activities (CA and AA, reference activities in antitumor drug discovery) of several types of N-glycosides of substituted indolo[2,3-a]carbazoles obtained in the chemical laboratory of Russian Cancer Research Center of Russian Academy of Medical Sciences (RCRC RAMS). All substances were evaluated *in vitro* for effect on tumor cells CaOv, SKOV3, MCF-7. In accordance with the chemotherapeutic criteria the substances with $IC_{50} \geq 10^{-5}$ M were considered as unpromising for testing *in vivo*. With the help of PASS there were obtained probability values 0.3-0.5 for CA and 0.7-0.8 for AA for substances with $IC_{50} \leq 10^{-5}$ M (high cytostatic effect). Predicted probabilities of AA for some glycosides derivatives of indolo[2,3-a]carbazoles with $IC_{50} \geq 10^{-5}$ M which were considered as not cytostatic *in vitro* and had low predicted probability of CA ($\approx 0,2$) were 0.6-0.7. One of them (LHS-1156) in spite of not predicted and not experimentally found CA on the basis of high predicted AA (0.71) was studied *in vivo* on several experimental tumor models (breast carcinoma Ca-755, cervix tumor RShM-5, Fisher leukemia, lung carcinoma LLC, melanoma B-16). LHS-1156 produced 70-85% of tumor growth inhibition and revealed a low total toxicity. In some cases significant antineoplastic effect of LHS-1156 (62% of tumor growth inhibition) lasts for 25 days. The coincidence of predicted and experimental data suggests that computer system PASS can be used for preliminary evaluation of activity of low toxic antitumor indolo[2,3-a]carbazoles.

COMPUTER-BASED SEARCH FOR NEW ANTITUMOR DRUGS USING THE RCRC RAMS DATABASE ON ANTITUMOR SUBSTANCES

Galina N. Apryshko[1], Dmitry A. Filimonov[2], Vladimir V. Poroikov[2]

[1]Russian Cancer Research Center of Russian Academy of Medical Sciences, 24, Kashyrskoe Shosse, Moskow, 115478, Russia; [2]Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: g48@mail.ru

The computer-based by system PASS (http://www.ibmc.msk.ru/PASS) prediction of antineoplastic activity (AA) was provided for 1694 substances of different chemical classes got from the Database on antitumor substances that was created in Russian Cancer Research Center of Russian Academy of Medical Sciences (RCRC RAMS). For all these substances in the Database there are data on positive or negative results of AA experimental testing obtained in RCRC RAMS laboratories. It was shown that in the case of experimental testing of substances based on the positive AA prediction only 46% of initially proposing substances are to be entitled for testing. The preliminary computer prediction of AA permits to increase the number of substances revealing AA from 42,7% to 60% and to decrease the number of substances not revealing AA from 57,3% to 18,5%. It leads to reducing of wasteful spending for testing of verily not AA substances by a factor of three. The results of testing of substances effect on different experimental tumors taken from the Database of RCRC RAMS were used for creating training set for computer system PASS in order to predict the action on about 14 experimental tumor models for new substances. The quantitative data on AA and cytostatic activity from the RCRC RAMS Database on antitumor substances permit to form training set for semiquantitative prediction of AA. PASS is used in RCRC RAMS for preliminary evaluation of substances as candidates for drug development by predicting AA and cytostatic activities, mechanisms of action known for anticancer drugs, interactions with cell and molecular targets, and possible toxic effects. Results of computer prediction of biological activities by PASS and training sets formed on the basis of the Databases of RCRC RAMS are used for planning experimental testing of new chemical substances.

2D QSAR ANALYSIS OF NITROAROMATIC TOXICITY OF THE TETRAHYMENA PYRIFORMIS

Anatoly G. Artemenko[1], Victor E. Kuz'min[1], Eugene N. Muratov[1,2], Leonid G. Gorb[2,3], Mohammad Qasim[3], Jerzy Leszczynski[2]

[1]A.V. Bogatsky Physical-Chemical Institute National Academy of Sciences of Ukraine, Lustdorfskaya Doroga 86, Odessa 65080, Ukraine; [2]Computational Center for Molecular Structure and Interactions, Jackson State University, 1400 J.R. Lynch Str., Jackson 39217, Mississippi, USA; [3]US Army ERDC (SpecPro), Vicksburg, Mississippi, 39180, USA; E-mail: artanat@ukr.net

The present study applies the Hierarchical Technology for Quantitative Structure - Activity Relationships (HiT QSAR) i) to evaluate the influence of the structure of 95 various nitroaromatic compounds (including some widely known explosives) on their toxicity to the ciliate *Tetrahymena pyriformis;* ii) for virtual screening of toxicity of new nitroaromatic derivatives; iii) analysis of the characteristics of the substituents in nitroaromatic compounds as to their influence on toxicity. The 50% inhibition growth concentration (IGC_{50}) was used to develop QSAR models based on Simplex representation of molecular structure. During the second part of the work the whole initial set of compounds was divided into 3 interpenetrating sets depending on the possible mechanism of action. Obtained 2D QSAR PLS models are quite satisfactory (R^2=0.86–0.98; Q^2=0.71–0.95). The predictive ability of QSAR models was confirmed through usage of three different test sets (maximal similarity with training set, also minimal and random choice, taking into account toxicity range only) for any obtained model (R^2_{test}=0.54–0.97). Molecular fragments that promote and interfere with toxicity were defined. The initial division into different sets was confirmed by developed QSAR analysis, i.e. the models developed for structures with one mechanism (e.g. redox cyclers) cannot satisfactory predict the others (e.g. participating in nucleophilic attack, etc.). However, the reliable predictive model can be obtained for all the compounds, regardless of mechanism, when structures of different action modes are sufficiently represented in the training set. Virtual screening of toxicity for new nitroaromatics has been carried out. It was shown that substituent interference in the benzene ring plays the determining role for toxicity.

A SIMPLE MODEL FOR THE PREDICTION OF BRAIN-BLOOD BARRIER PENETRATION BASED ON 2D AND SUBSTRUCTURAL DESCRIPTORS USING PLS STATISTICS

Natalia V. Artemenko, Tim James, Darren Fayne, David G. Lloyd

Molecular Design Group, School of Biochemistry & Immunology, Trinity College Dublin, Dublin 2, Ireland; E-mail address: artemenn@tcd.ie

Prediction of various properties tightly connected with ADME/Toxicity plays an increasingly important role in the research projects of many pharmaceutical companies. A knowledge and deep comprehension of complex physiological characteristics is currently very demanding due to the high desirability of novel, and more progressive, medicines for targets which are located in the central nervous system (CNS). In accordance with this view, such a property as blood-brain barrier (BBB) partitioning is determined by a number of different metabolic effects maintaining the homeostasis of CNS and plays a leading role in drug design. A strong focus on the modelling of BBB penetration for a diverse set of compounds, including potential drug candidates, was the main goal of our current study due to the fundamental importance of this characteristic for drug design and modelling.

The major focus of recent computational QSAR approaches to predict BBB penetration was based on different physicochemical descriptors which are often linked somehow to log P or some surface/volume characteristics. Our goal was to develop an accurate model of BBB penetration that is reproducible and simply implemented in practice.

Our current study is based on a diverse set consisting of 165 compounds in total that were collected from literature and put into 2 subsets: training and external test set consisting of 148 and 17 compounds correspondingly. All questionable compounds characterised by ambiguous information from different sources or very specific chemical structure were deleted. We report our results obtained using a combination of various of 2D descriptors, including such properties log P and topological polar surface area, substructural information on structure, and PLS statistics implemented in the MOE software package. For calculating substructural descriptors we utilised our in-house developed code for decomposition of molecular graphs into sets of fragments and computation of occurrence numbers for all of them which was integrated into the Pipeline Pilot package. Our best model ($n = 148$, $N_F = 3$; $R^2 = 0.799$; $RMSE_{train} = 0.287$; $Q^2 = 0.673$) could be characterised by a high level of predictability providing an RMSE value for an external test set of 0.160.

SMART MINING: A NOVEL TECHNOLOGY FOR ADDRESSING ADME/TOX ISSUES

Konstantin V. Balakin, Andrey A. Ivashchenko, Yan A. Ivanenkov

Chemical Diversity Research Institute, Rabochaya 2a, Khimki, Moscow reg. 114401, Russia; E-mail: kvb@iihr.ru

Multivariate data mining techniques can serve as the basis for advanced ADME filters. The scientists at CDRI have developed a software program, Smart Mining, for early evaluation of several important pharmacokinetic parameters including ADME/Tox-related and physico-chemical parameters. This PC-based program can calculate more than 100 different molecular descriptors and model complex molecular properties using non-linear mapping algorithms. Generated models demonstrated good predictive power in the internal and external validation experiments with up to 80-90% compounds classified accurately. The accuracy level achieved can be used as a guide in modifying and optimizing these pharmacokinetic properties in chemical libraries before synthesis. To-date, these algorithms were validated on human intestinal absorption, blood–brain-barrier, plasma half-life, volume of distribution, plasma protein binding, CYP450 substrate/non-substrate potential, as well as cytochrome P450 binding affinity models. The Smart Mining-based collection of predictive classification tools is both extensive and well validated in multiple library design projects. These methods are particularly suited for the rapid evaluation of large compound libraries in connection with early ADME/Tox profiling.

SPATIAL DYNAMICS OF THROMBIN GENERATION IN PLASMA

Anna N. Balandina[1,2], Sergey S. Karamzin[1]

[1]National Research Center for Hematology, 4a, Novozikivskii pr., Moscow, 125167, Russia; [2]Lomonosov Moscow State University, GSP-2, Leninskie Gory, Moscow, 119992, Russia; E-mail: a_balandina@inbox.ru, karamzin@yandex.ru

The function of the blood clotting system is the prevention of blood loss during the vascular system integrity damage. Many factors including pharmacological substances influence the clotting system. Even a small change in clotting balance can lead to either bleeding or thrombosis. Thrombin plays a central role in blood coagulation. Investigation of the spatio-temporal thrombin distribution during the clotting is an important pharmacological problem. It will help to obtain significant information about the influence of the new medical products on the hemostasis. The developed method is based on the registration of the fluorescent product cleaved by thrombin. Clotting was initiated in a thin layer of plasma by bringing it in contact with the fibroblast monolayer surface. Aminomethylcoumarin (AMC) fluorescence was stimulated by ultraviolet light and registered by the CCD camera. The AMC diffusion from the production area and the substrate diffusion into the consumption region were taken in account when performing calculation. The result of thrombin signal restoration strongly depends on experimental noises, because of the presence of double differentiation operation in the restoration procedure, which leads to strong amplification of high-frequency noise. Thus, selection of an effective method for experimental AMC signal filtration is a problem of great importance for thrombin distribution restoration. The comparison of following methods of noise reduction was made: low-pass filtering by high-order FIR filters, wavelet smoothing, high-order polynomial regression, piecewise approximation by b-splines were used. The best results in thrombin signal restoration were achieved by spline approximation, which involves AMC signal division into pieces and their smoothing by 3-degree b-splines with preservation of function smoothness in the knots of joining before every differentiation operation. In the first stage of coagulation there was a peak of thrombin near the activator. Then, one can observe wave propagation of thrombin from the fibroblast layer deep into the plasma. The propagation rate of the thrombin front was about 0.8-1.0 mm in 30 minutes. The introduced method is a new highly sensitive physiological tool, which has no analogues and gives unique information about mechanisms of blood coagulation.

QSAR STUDIES ON A SET OF PYRIMIDINE NUCLEOSIDES AS ANTI HIV-1 AGENTS

Rongjing Hu[1,2], Florent Barbault[1], Michel Delamar[1], Jean-Pierre Doucet[1], Ruisheng Zhang[2]

[1]Interfaces, Traitements, Organisation et Dynamique des Systemes (ITODYS), Universite Paris 7 - Denis Diderot and CNRS, UMR 7086, 1 rue Guy de la Brosse, 75005 Paris, France; [2]Department of Chemistry, Lanzhou University, Lanzhou, 730000, China

Quantitative structure-activity relationship (QSAR) studies on a set of pyrimidine nucleosides [1] [2] (Fig.1) as antiviral agents for HIV-1 were performed using a comprehensive set of geometrical, electrostatic and quantum-chemical descriptors.

Fig.1.

$R_5 = H, CH_3, C_2H_5, C_3H_7, CH=CH-Br, F, Br, I, CF_3, NH_2, NHCH_3, N(CH_3)_2, OH, OCH_3, OCH_2CH_3, OCH_2C\equiv CH, OCH_2CH=CH_2, OCH_2CN, SCN, SCH_3$

$R_{5'} = OH, N_3$ $R_{3'} = N_3, NH_2, I, H, OH$ $X = O, NH$

$Y = H, F, CH_3$
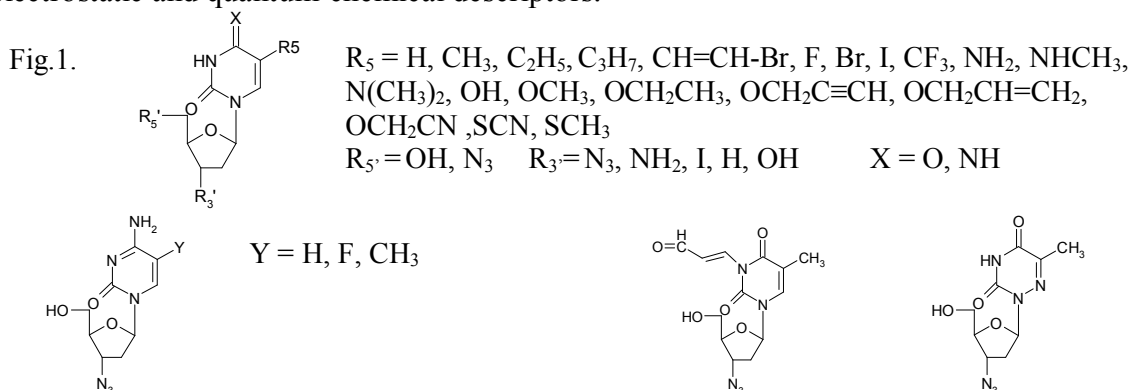
3D molecular structures were built using Hyperchem and geometry optimization calculations were performed using first Molecular Mechanics then PM3 in MOPAC. There are 45 pyrimidine nucleosides in all. The 34 molecules for which precise bioactivity values are known were used the training set to build the model. Multiple linear regression (MLR) was utilized to select the molecular descriptors (one electrostatic and 5 quantum descriptors) and construct a linear model with R^2 and mean square error (MSE) of 0.729 and 0.36 for the training set. However the predicted values for the whole set of molecules showed significant deviations from the experimental data. The support vector machine (SVM) and projection pursuit regression (PPR) methods were then utilized to construct non-linear prediction models, leading to good correlation coefficient (R^2) of 0.850 and 0.841 and mean square errors (MSE) of 0.22 and 0.21 for the training set, respectively, in the case of SVM (Fig.2). The prediction results of the SVM and PPR models are better than those of MLR and are compared to previous results from the literature [3]. Our results might also provide some insight into the structural features favouring the bioactivity of pyrimidine nucleosides.
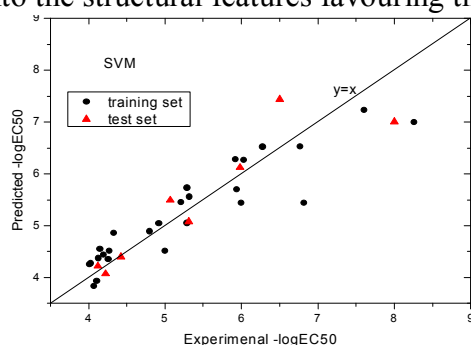
Fig.2. SVM Predicted activity values vs. experimental values for the training set and test set.

1. T.S. Lin, J.Y. Guo, R.F. Schinazi, C.K. Chu, J.N.Xiang, W.H. Prusoff, *J.Med.Chem.*, **1988**, 31, 336.
2. C.K. Chu, R.F. Schinazi, M.K. Ahn, G.V. Ullas, Z.P. Gu., *J. Med.Chem.,* **1989**, 32, 612.
3. M. Mahmoudian, *Pharm. Res.,* **1991**, 8, 43.

ESTIMATION OF BIOLOGICAL ACTIVITY SPECTRA FOR 2-AMINOTHIAZOLES AND 2-AMINOBENZOTHIAZOLES

Elena V. Shilova[1], Violetta L. Kovaleva[1], Athina Geronikaki[2], Dmitry S. Blinov[3], Vladimir V. Poroikov[4], Oxana V. Proskurina[1]

[1]National Research Center on Biologically Active Compounds, Kirov st. 23, Staraya Kupavna, 142450, Moscow region, Russia, E-mail: vnc@pc-club.ru; [2]School of Pharmacy, Aristotelian University of Thessaloniki, University Campus, 54006 Thessaloniki, Greece; [3]Mordovian Ogarev State University, Bol'shevistskaya st. 68, Saransk, 43000, Russia; [4]Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, Pogodinskaya st. 10, Moscow, 119121, Russia

Development of new potent and safety drugs with antiasthmatic, anti-inflammatory (non-steroidal), antiallergic and analgesic activity is important. It is known QSAR methods are useful both in search for a new lead compounds and in development of more active analogs without adverse effects. The spectrum of biological activity of 2-aminothiazoles and 2-aminobenzothiazoles bearing various substituents at 2-amino group and at 4 and 5 positions of cycles was estimated using computer program PASS (http://www.ibmc.msk.ru/PASS) among 369 possible pharmacological effects, 2055 molecular mechanisms, 39 side effects and toxicity and 66 metabolism. Prediction of biological activity permits to suppose presence of such activities as analgesic, anti-inflammatory, mucomembranous protector, antihistaminic, antiobesity, mediator release inhibitor, antiasthmatic. Introduction of adamantyl radical at 4 position was predicted to enhance anti-inflammatory effect (maximal Pa 0.655) through cycllooxygenase inhibition (maximal Pa 0.640). Substances with benzoizothiazolyl group show antiasthmatic activity (maximal Pa 0.553). Derivatives of morpholinylpropionic acid at 2-amino group have high antiobesity (maximal Pa 0.768) and mucomembrane protector (maximal Pa 0.902) effects. Presence of substituted benzylidenes at 5 position of thiazole ring leads to appearance of non-opioid analgesic effect (maximal Pa 0.743). All tested structures were predicted to have 0-6 possible side effects (maximal Pa 0.348) and low toxicity. Analgesic and anti-inflammatory activity of some substituted 2-aminothiazoles bearing benzyliden group was confirmed in experiments in vitro and in vivo.

INTERACTION OF THE ESTROGEN RECEPTOR WITH LIGAND FROM DIFFERENT CHEMICAL CLASSES: 3D-QSAR WITH COMSIA ANALYSIS

Anna V. Boiko[1] , Alexander V. Veselovsky[2],Vladlen S. Skvortsov[2], Oleg A. Raevsky[3]

[1]Lomonosov Moscow State University, Leninskie gory, Moscow, 119992, Russia; [2]Institute of Biomedical Chemistry of RAMS, 10, Pogodinskaya Str., Moscow, 119121, Russia; [3]Institute of physiological active compounds RAS, Chernogolovka, Moscow region, Russia; E-mail: anna.boiko@mail.ru

One of the molecular target of the toxic compounds are the nuclear receptors, in particular estrogen, androgen receptors etc. The toxicity based on the interaction of pollutants with these receptors called endocrine toxicity. The detail mechanisms of it are still unknown that limit the adequate prediction of toxicity of endocrine-disrupting chemicals. Using molecular modeling methods we investigate of ability to predict interaction of estrogen binding compounds from different chemical classes with estrogen receptors. The database with 515 compounds was used in this investigation. These compounds were docked in ligand-binding domain of rat estrogen receptor using docking program Dock 6.0. It allows to predict the position of the most compounds. Analysis of predicted complexes showed that it is possible to select two main positions of ligands in binding site. The first position is equivalent to position of ligands from known crystal structures, the second one – perpendicular to it, accommodating the entrance of binding site. Ligands from the first position in conformation and alignment obtained by docking procedure were used for further design of 3D-QSAR with CoMSIA model. The steric, electrostatic, hydrophobic and donor-acceptor fields were taken into consideration. The cross-validation analysis showed that QSAR model can be designed with eliminated 6 compounds from the training set. The designed model have the good statistic values ($Q^2 = 0,646$; $s_{cv} = 0,78$) Thus, the obtained result showed the ability to design the 3D-QSAR with CoMSIA model able to predict the binding affinity of compounds from different chemical classes to ligand binding domain of estrogen receptor.

MOLECULAR STRUCTURE OF 3-[2-(4-CHLORO-PHENYL)-2-OXO-ETHYL]-5-(2,4-DICHLORO-THIAZOLE-5-YL-METHYLENYL)- THIAZOLIDINE-2,4-DIONE AS ANTIFUNGAL AGENT

Ilkay Yildiz, Oya Bozdag-Dundar

Ankara University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, 06100, Tandogan, Ankara-Turkey

In the past 10 years there has been a major expansion in the development of antifungal drugs, but there are still weaknesses in the range and scope of current antifungal chemotherapy [1]. New developments have included the modification of existing drug molecules to eliminate toxicity and improve activity. Therefore, we are also trying to screen the antifungal activity beside antibacterial to be able to luckily discover a new lead compounds.

In this study, the molecular geometry in the ground state of 3-[2-(4-chloro-phenyl)-2-oxo-ethyl]-5-(2,4-dichloro-thiazole-5-yl-methylenyl)-thiazolidine-2,4-dione (Figure), which had to be found two-fold dilution better antifungal activity against *Candida albicans* ATCC 10145 than standard drug, miconazole [2], has been calculated using Becke`s three-parameter hybrid method with the Lee, Yang, and Parr correlation functional (B3LYP) methods with 3-21G (d,p) basis set [3,4]. Besides, its data has been compared with X-ray crystal structure of it.
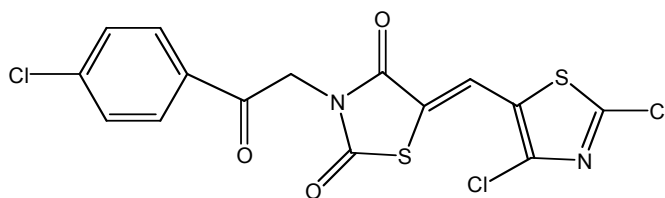
Figure

1. J.R. Graybill, *Clin. Infect. Dis.*, **1996**, 22(2), 166.
2. O. Bozdağ-Dündar, O. Ozgen, A. Mentese, N. Altanlar, E. Kendi, R. Ertan, *Bioorg. Med. Chem.* (submitted)
3. A.D. Becke, *J. Chem. Phys.*, **1993**, *98*, 5648-5652.
4. C. Lee, W. Yang, R.G. Parr, *Phys. Rev. B,* **1988**, *37*, 785-788.

COMPARISON OF PSORIASIS AND CROHN'S DISEASE PATHOLOGICAL PROCESSES AT THE LEVEL OF GENE NETWORK INTERACTIONS BY BIOINFORMATICS METHODS

Sergey A. Brouskin[1], Roustam M. Abdeev[2], Tatiana A. Nikolskaya[1,3], Eleonora S. Piruzian[1]

[1]Vavilov Institute of General Genetics, Russian Academy of Sciences, Gubkin str. 3, Moscow, Russia, 119991; [2]Center for Theoretical Problems of Physical-Chemical Pharmacology, Russian Academy of Sciences, Kosigin str. 4, Moscow, Russia, 119991; [3]Metalogik» Ltd, Gubkin str. 3, Moscow, Russia, 119991; E-mail: brouskin@vigg.ru

Psoriasis that is a skin disorder, and Crohn's disease that is a damage of bowels epithelium, are complex diseases, which are the consequences of mis-regulation of genetic program. Because of fast development of genomic and postgenomic technologies, and as consequence creating of vast databases of microarrays experiments and scientific publications it is possible to identify now the genes associated with these pathological processes. The survey was performed with bioinformatics tools, by way of creating maps and networks, contained candidate genes. It allows identifying sub processes that could be important in pathology development. Then it is required to verify it with standard molecular methods. In our bioinformatics researches we used GEO DataSets (http://www.ncbi.nlm.nih.gov/geo/) database, contained results of microarray experiments estimating the level of genes' expression. The level of expression of about 12000 genes estimated in microarrays experiments was analyzed with "MetaCore" software (GeneGo Inc., USA). The number of genes with more then two fold change in the level of its expression under psoriasis condition was shown to be 1118 and 884 during Crohn's disease. We have analyzed microarrays experiments data concerned studying pathogenesis and determined that expression change under psoriasis and Crohn's disease affect such processes as signal transduction, immune response, regulation of cell cycle and others. On the basement of database and publication analysis we have identified some genes include Bax, Bcl-2, Caspase-3 and genes encoding transcriptional factor AP-1 (c-jun, JunB, JunD, c-fos, FosB, Fra-1 and Fra-2 and others). Our results obtained by analysis of genes candidates expression during psoriasis and Crohn's disease can contribute to identification of new targets for medical preparations used for sickness treatment.

DETAILED MOLECULAR DYNAMICS SIMULATION OF PROTEIN-LIGAND INTERACTIONS FOR TWO IMPORTANT DRUG TARGETS: CATECHOL-O-METHYL TRANSFERASE (COMT) AND PROLYL OLIGOPEPTIDASE (POP)

Alex Bunker

DDTC, Faculty of Pharmacy, University of Helsinki; E-mail: alex.bunker@helsinki.fi

COMT and POP are both important and mature drug targets, for which several potential inhibitor ligands have been found. These proteins have been successfully crystallized containing a variety of ligands. For the first time long time precise atomistic molecular dynamics calculations have been made. For COMT it has been found that the protein structure present in the crystallized form is partially dependent on the presence of the ADOMET co-enzyme and Mg^{++} ion.

Simulations have been performed using the 4 existing COMT crystal structures (1H1D, 1VID, 1JR4, and 2CL5 in the PDB database). POP has been simulated with and without covalently bound inhibitor, and with the inhibitor with covalent bond cut, to investigate the path taken in leaving the active site.

UBIPROT DATABASE AND ANALYSIS OF PROTEIN UBIQUITYLATION FOR APPLICATION IN PHARMACOLOGICAL RESEARCH

Alexander Chernorudskiy, Alejandro Garcia, Eugene Eremin, Anastasia Shorina-Zhabereva, Ekaterina Kondratieva, Murat Gainullin

Institute of Applied and Fundamental Medicine, Nizhny Novgorod State Medical Academy, 10/1, Minin Sq., Nizhny Novgorod, 603005, Russia; E-mail: biochem@gma.nnov.ru

Post-translational protein modification with ubiquitin, or ubiquitylation, has a dramatic impact on diverse cellular processes and is involved in pathogenesis of severe human diseases. Many constituents of ubiquitin-dependent pathway (e.g. enzymes forming ubiquitin conjugating cascade and 26S proteasome) are rated as attractive potential drug targets. In contrast, a wide range of proteins targeted for ubiquitylation is currently out of pharmacologists' interest. This may be due to the lack of systematization of data about particular ubiquitylated proteins. We try to solve this problem by elaborating a specialized web-based resource and virtual tools, which enable to accumulate and systematize experimental data about various protein substrates of ubiquitylation and to use it for further computational analysis. This resource termed UbiProt Database is available at http://ubiprot.org.ru for free. Currently the database contains information about more than 400 individual proteins from various organisms. Each database entry contains retrievable information about overall characteristics of a particular protein, ubiquitylation features, related ubiquitylation and de-ubiquitylation machinery and literature references reflecting experimental evidence of ubiquitylation. Pharmacological utility of the collected dataset can be illustrated by XIAP (X-linked inhibitor of apoptosis) ubiquitylation. A computational analysis of XIAP 3D structure (PDB 1tft) reveals a steric interference of an ubiquitin acceptor site (Lys 322) and a binding site for potential proapoptotic/anticancer compound. Hence, we hypothesize that the attachment of rather big modifier like ubiquitin can impair drug interaction with the target protein and in such a way decrease or modify drug activity. The utility of UbiProt is also being extended for usage in proteomics tasks. To fill a general need for MS-based identification of ubiquitylated proteins, UbiProt will be supplemented with an algorithm for virtual fragmentation of target proteins (included in UbiProt dataset as well as supplied by user) and calculation of peptide masses, taking into account all possible variants of ubiquitylation. In summary, UbiProt is considered to be a useful resource for multidisciplinary research concerned with protein ubiquitylation.

QUALITY OF COMPUTER-BUILT MODELS OF MEMBRANE PROTEINS ASSESSED VIA KNOWLEDGE-BASED APPROACH

Anton O. Chugunov, Valery N. Novoseletsky, Roman G. Efremov

Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, ul. Miklukho-Maklaya 16/10, GSP Moscow 117997, Russia; E-mail: batch2k@yandex.ru

Integral membrane proteins (MP) are pharmaceutical targets of exceptional importance, given e.g. the fact that very significant MP subfamily, G–protein-coupled receptors (GPCR), are targets for >50% of currently marketed drugs. Although availability of spatial structures of MPs is a strong prerequisite for drug design applications, it is currently limited due to imperfection of spatial structure determination techniques. That is why, computational structure prediction methods gain special importance. Here we present a novel approach for packing "quality" assessment of computer-built models of transmembrane (TM) domains of α-helical MPs, built upon the analysis of the spatial organization of proteins with known structure. The method is based on the concept of membrane–protein environment classes, whereby each amino acid residue is described in terms of its environment polarity and accessibility to the membrane. Residues' preferences to five predefined environment classes were calculated from the analysis of non-redundant set of high-resolution X-ray structures of α-helical MPs. A number of tests reveal that the proposed "membrane score" approach is useful for accurate delimitation between "correct" (e.g., crystallographic) structures and theoretical models that contain severe modeling errors such as wrong packing topology and incorrect conformations of side chains, as well as misaligned regions in homology-built models. Furthermore, identification of functional residues in TM domain and MPs' dimerization sites is possible by detection of regions with abnormally low scores. Finally, "3D-environment profiles" based on membrane score might be applicable for detection of non-obvious homologies in sequence-based applications. The developed technology will certainly be useful for optimization of computer models of MPs, especially GPCRs.

INTERSPECIES RELATIONSHIPS IN THE ACUTE TOXICITY OF XENOBIOTICS

James Devillers[1], Pascal Pandard[2], Eric Thybaud[2], Anne Merle[3]

[1]CTIS, 3 Chemin de la Graviere, 69140 Rillieux La Pape, France; [2]INERIS, Parc Technologique ALATA, BP n° 2, 60550 Verneuil en Halatte, France; [3]ADEME, 2 square La Fayette, BP 90406 Angers Cedex 01, France; E-mail: j.devillers@ctis.fr

LD50 tests on rat and mouse are commonly used to express the relative hazard associated with the acute toxicity of new and existing substances. These tests are expensive, time consuming and actively fought by Animal Rightists. Consequently, there is a need to find alternative methods. If the design of QSAR models can be used as surrogate, the search for interspecies correlations also represents a valuable alternative to the classical mammalian laboratory tests. This study investigates quantitatively the interspecies relationships of the acute toxicity of several hundreds of organic and inorganic chemicals to rat, mouse, *Vibrio fischeri* (Microtox), and *Daphnia magna* accounting for the experimental conditions of the different tests. For the first time, attempts have been made to test the usefulness of nonlinear methods for deriving interspecies toxicity correlations.

COMPUTER PREDICTION OF SUBSTRATES, INHIBITORS AND INDUCERS OF CYTOCHROME P450

<u>Alexander V. Dmitriev</u>, Alexey A. Lagunin, Alexey V. Zakharov, Olga A. Filz, Dmitry A. Filimonov, Vladimir V. Poroikov

Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: a.v.dmitriev@mail.ru

CYP2C19, CYP3A4, CYP2D6, CYP2C9, CYP1A2 are the main human cytochrome P450 isoforms important for a xenobiotic metabolism. These isoforms metabolize a plethora of exogenous substances, which are taken by human beings. Inhibition or induction of this isoforms will allow to control drugs biotransformation, and thus to control drugs therapeutic action and side effects. We have collected the information concerning substrates, inhibitors and inducers of cytochrome P450. The database contains structural formulae of compounds and data about their substrate/inhibition/induction properties. The attention was concentrated on five prevailing human cytochrome P450 isoforms. The current data set contains information about ~1000 structures of chemical compounds and their interaction with the targeted cytochromes P450. For prediction of substrate/inhibitor/inducer activity of cytochrome P450 for drug-like compounds we used the same algorithm that is used by computer program PASS. Accuracy of prediction was estimated on the basis of leave-one-out cross-validation procedure. It was shown that PASS-based statistical algorithm is able to predict the studied activities with average accuracy of prediction about 87%. We also evaluated accuracy of prediction of substrate/inhibitor/inducer activity for drugs from "The Top 100 Drug Interaction 2006 H&H Publications". The test set contain 275 chemical structures of drugs. Interaction with the targeted cytochrome P450 isoforms was predicted. Obtained results of prediction look reasonable to use them for solution of the practical tasks.

USER CUSTOMIZED MOLECULAR DIVERSITY ANALYSIS USING THE MOLDIA SOFTWARE

Ana G. Maldonado, <u>Jean-Pierre Doucet</u>, Michel Petitjean, Bo-Tao Fan[†]

† This poster is dedicated to the memory of our late colleague Prof. Bo Tao Fan

ITODYS, Universite Paris 7 – Denis Diderot, CNRS UMR-7086, 1 rue Guy de la Brosse, 75005, Paris, France; E-mail: doucet@paris7.jussieu.fr

Computational query of virtual molecular databases (DB) by chemists, in order to enrich a collection of potentially interesting compounds, is not atypical any more. Everyday, hundreds of molecules are virtually or combinatorially generated, increasing the current data workflow. Tools dealing with the management, the analysis and the retrieval of chemical data in DBs have been developed since few years as an answer to the new demanding needs. User customization of these tools responding to particular domain criteria is, however, absent or too long to implement because of a lack of flexibility in the current systems.

MolDiA (Molecular Diversity Analysis) is a tool which has been designed to answer to these new user customization needs. This structure-based approach uses customizable weights on molecular descriptors to compute similarity and diversity measures of given DBs. The underlying software principles and architecture have already been published. Here, we present some tests run under different user conditions and DBs, as well as, some MolDiA implementation features including the deployment of XML web technologies as an important key for the integration, exchange and analysis of chemical information.

QUANTUM CHEMICAL SIMULATION OF CYTOCHROME P450 CATALYZED AROMATIC OXIDATION: METABOLISM, TOXICITY, AND BIODEGRADATION OF BENZENE DERIVATIVES

Pavel N. Dyachkov[1], Nina V. Kharchevnikova[2], Alexandr V. Dmitriev[3], Alexei V. Kuznetzov[1], Vladimir V. Poroikov[3]

[1]N.S. Kurnakov Institute of General and Inorganic Chemistry, Russian Academy of Sciences, Moscow, Russia; [2]Institute of Human Ecology and Environmental Health, Russian Academy of Medical Sciences, Moscow, Russia; [3]Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Moscow, Russia

The dependences of biological oxidation and toxicity of the mono- and multisubstituted benzene derivatives on the nature of substituents are studied using an oxenoid model and the quantum chemical calculations. According to this model, the P450 enzyme breaks the dioxygen molecules and generates the active atomic oxygen species (oxens); these species readily react with substrates. Using MO LCAO MNDO approach, we calculated the differences ΔE of the total energies of aromatic compounds and corresponding arene oxides containing tetrahedrally coordinated carbon atoms. We obtained that the ΔE values determine the positions of the enzyme mediated oxidation, rate of substrate biotransformation, and toxicity of the benzene derivatives. In addition to the "dynamic" reactivity index ΔE related to the enzyme-mediated substrate biotransformation, we calculated many standard "static" reactivity indices, corresponding to the substrate molecules in the starting equilibrium geometry (the energies of the occupied and unoccupied MOs, the effective atomic charges, the free valence indices, and the superdelocalizabilities). The arene oxide stability ΔE parameter is shown to be the most adequate characteristic of both the biological oxidation process and toxicity of benzenes. The ΔE parameters were also used successfully to describe the features of di- and trichlorinated biphenyls bacterial metabolism.

MODELING THE TOXICITY OF BENZENE DERIVATIVES TO THE TADPOLE (*Rana japonica*) WITH THE CHARACTERISTIC ROOT INDEX AND SEMIEMPIRICAL MOLECULAR DESCRIPTORS

Doga Erturk, Melek Turker Sacan

Bogazici University, Instıtute of Environmental Sciences, 34342, Bebek/Istanbul/Turkiye

A previous QSPR study with the toxicity of 39 aromatic compounds to *Scenedesmus obliquus* having strong relationship with the Characteristic Root Index (CRI) and the energy of the Lowest Unoccupied Molecular Orbital (E_{LUMO}) [1] encouraged us to look for the similar relationship between the same descriptors and the toxicity of benzene derivatives to *Rana japonica*. The 12 h acute toxicity values of 18 compounds used in this study to develop a QSAR model are taken from literature [2].

The CRI was calculated using a computer program developed by our group and written in TURBO PASCAL for PC and SCIENTIFIC WORKPLACE 3.0 software. Semiempirical molecular descriptors were calculated by the quantum chemical PM3 method. The geometries of 18 benzene derivatives were fully optimized. Aqueous solvation energies of all the conformers were estimated by using the SM54 and added to the gas phase energies. For every compound, molecular orbital energies and dipole moments were calculated for the conformer having the lowest total energy. SPARTAN 04 software was used for quantum chemical computations.

A QSPR model based on the CRI and E_{LUMO} ($-\log EC_{50}$ (mol L^{-1}) = 0.930 ($\pm$0.100) CRI $-$ 0.429 ($\pm$0.047) E_{LUMO} + 1.561 ($\pm$0.207)   (n =18; r^2 = 0.903; *S.E.* = 0.146; $F_{(2,16)}$ = 69.54) to predict acute toxicity of benzene derivatives was obtained without any outliers. Initially, the energy of the Highest Occupied Molecular Orbital (E_{HOMO}) and dipole moment (μ) were also used as descriptors. However, the stepwise multiple linear regression excluded these parameters from the equation. The meaning of the descriptors appeared in the model was discussed. E_{LUMO} reflects electronic properties, whereas the CRI reflects hydrophobicity, molecular size, and branching. The statistical robustness of the developed model was validated by the modified jackknife test. The predictive accuracy of the proposed model was compared with the recently published study in which a toxicity model was developed for the same tadpole. Because of its high statistical significance, the validated model has been used to predict $-\log EC_{50}$ values of compounds for which there are no experimental measurements.

1. M.T. Sacan, M. Ozkul, S.S. Erdem. "QSPR analysis of the toxicity of aromatic compounds to the algae (*Scenedesmus obliquus*). *Chemosphere*, **2007**, 68: 695-702.
2. H. Huang, X. Wang, W. Ou, J. .Zhao, Y. Shao, L. Wang. "Acute Toxicity of benzene derivatives to the tadpoles (*Rana japonica*) and QSAR analyses. *Chemosphere*, 53: 963-970.

COMPUTATIONAL STUDY OF ARTEMISININ INTERACTION WITH HEME AND ITS POSTERIOR DECOMPOSITION

Marcia M. C. Ferreira, Mirian C. S. Pereira, Rudolf Kiralj

Instituto de Quimica, Universidade Estadual de Campinas, Campinas 13083-862, SP, Brazil; E-mail: marcia@iqm.unicamp.br

Malaria is still one of the diseases with serious morbidity and mortality statistics worldwide, especially in countries in development. Another problem related to malaria is appearance of multidrug resistance strains of the *Plasmodium* species, especially of the most mortal *P. falciparum*. A class of novel substances based on artemisinin is a promising tool in combating malaria. There are sufficient experimental and theoretical evidences to consider that artemisinin interacts with iron from hemoglobin, forming free radicals that undergo various decomposition routes in interactions with parasite molecular architecture, thus killing the parasite and disabling its usual defense mechanisms.

The purpose of the first part of this work is to study theoretically stereoelectronic aspects of the interaction between heme and artemisinin in the transitional heme-artemisinin complex. The stability of this complex is important for the artemisinin activation. Through semi-empirical calculations using the PM3 method, the potential energy barrier of artemisinin rotation relative to heme in the heme-artemisinin complex was studied in vacuum and in the partially solvated state. The purpose of the second part of this work is to study computationally artemisinin decomposition routes, which, after being activated by Fe from heme, can undergo different decomposition mechanisms *via* free C- or O-centered radicals. The artemisinin activation is crucial for its biological activity [1]. *Ab initio* method HF with 6-31G** basis set was used in these calculations, whilst heme was simulated by an electron. Electronic and Gibbs energies were the criteria to identify the most probably decomposition pathway.

The minimum heat of formation for the complex with and without water molecules is -702 and -101 kcal mol^{-1}, respectively, which corresponds to the dihedral angle C-Fe-O1-O2 of 44ϵ and 52ϵ around the iron-oxygen O1 bond, respectively. The water molecules bind to heme *via* several hydrogen bonds and O-H...O and C-H...O interactions, which accounts for -67 kcal mol^{-1}. It is observed that the inclusion of water molecules does not affect significantly the stability of the heme-artemisinin complex. Comparing O1 and O2 radical routes of the artemisinin decomposition scheme, the O2 route is more preferable when C3-O13 and C12-C12a bond scissions occur in the seven-membered ring. This route has ΔG more negative than all other routes by 11-21 kcal mol^{-1}, what is not known in the literature up to our knowledge.

1. Krisnha, S. et al., *Drug Resist Updates*. **2004**, *7*, 233 and references cited therein.

COMPUTER-AIDED PREDICTION OF HERG CHANNELS REGULATION

Olga Filz, Alexey Lagunin, Dmitriy Filimonov, Vladimir Poroikov

Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 10, Pogodinskaya Str, Moscow, 119121, Russia; Russian State Medical University, 1, Ostrovityanova Str, Moscow, 117997, Russia; E-mail: olga.filz@gmail.com

Acquired Long QT syndrome is defined as a prolongation of QT interval on electrocardiogram leading to the severe abnormality (high risk of fibrillation). One of molecular mechanisms of acquired QT-prolongation is inhibition of potassium (IKr) HERG-channel by pharmaceuticals or their metabolites. IKr channels interact with wide spectra of pharmaceuticals. Recently it has been reported that there are different mechanisms of HERG channels regulation. Using computer approach, we try to identify which indirect mechanismsmight lead to inhibition of HERG channels. Computer programs PASS and PharmaExpert provide the means for identification of compounds caused the QT prolongation. PASS predicts more than 2800 kinds of biological activity. PharmaExpert program contains the knowledgebase about effect-mechanism(s) and mechanism-effect(s) relationships. For PASS training set we collected 105 substances from literature. For this training set mean accuracy obtained in leave one out cross-validation was 82.5% for activity «QT-prolongation» and 86.5% for «HERG-channel antagonist». As a result of this method application to the evaluation set of 8457 compounds from CMC database we found 583 ones that caused QT interval prolongation and inhibited IKr channel with Pa > 0.5. Using Pharma Expert we determined that 62,8% of 583 chemicals had «1-Phosphatidylinositol-4-phosphate 5-kinase inhibitor» activity in its spectra. Then, we suppose that 1-Phosphatidylinositol-4-phosphate 5-kinase can be involved into the mechanism of QT-prolongation. The thorough studying of HERG channel regulation allows to explain our finding. It was shown experimentally that phosphatidylinositol phosphate (PIP2) participates in HERG channel regulation (Sauviat M. P, Pages N, 2002). Mechanisms of activation are not clear enough, but different assumptions exist. Perhaps, thyrotropine-releasing hormone (TRH) modifies the current kinetics of IKr channels whose activity is regulated via the protein kinase C pathway linked to a G protein-coupled receptor and is regulated by changes in the PIP2 concentration in the membrane. Thus, inhibition of phosphatidylinositol-4-phosphate 5-kinase leads to decrease of PIP2 concentration in plasma, resulting to reduction of HERG currents.

MOLKERN AS NEW EFFECTIVE ENGINE FOR DRUG DISCOVERY SOFTWARE

Eduard S. Fomin[1], Nikolay A. Alemasov[2], Artem S. Chirtsov[2], Arseniy E. Fomin[3]

[1]Institute of Cytology and Genetics SB RAS; [2]Novosibirsk State University; [3]Moscow Institute of Physics and Technology; E-mail: fomin@bionet.nsc.ru

MOLKERN is a template library for all software components required for the analysis, modeling and optimization of spatial macromolecular structures for proteins, cofactors, ligands and their complexes within the force field approximation. Today's computer architecture - hierarchical structure of the memory subsystem, pipelining - is respected by the code to achieve efficient program execution. All algorithms implemented have a linear scaling O(N) both in memory and time requirements. The following operations can be performed:

- modeling and edition of molecular structures: restoration of missing atoms, protonation of proteins at a given pH value, addition of disulfide bonds, and residue replacement;

- optimization of molecular complexes with and without constraints;

- search for hydrogen bonds, salt and water bridges, and surface atoms;

- calculation of the molecule surface area, and the contact interface area;

- energy calculation in vacuum and in water within general Born approximation;

- search for active protein sites, rigid docking, flexible docking with regard to the mobility of protein atoms at the contact site.

The library codes, written in C++, use the STL and BOOST library. The library runs under WINDOWS and LINUX. Some modules can run in parallel using LAM MPI.

Table. Run-time for energy minimization of proteins in water environment. All atoms of proteins are free. We used a Pentium IV computer (2.1 GHz) with the gcc 4.1.1 compiler.

PDB ID	# freedom	E [kCal/mol]	# iteration	time [s]	time / (#iteration * #freedom) [s]
1aie	1566	326.8	72	5.8	5.1e-5
2bin	8649	16279.4	6	5.0	9.6e-5
1f58	19821	14201.9	25	52.7	1.0e-4

QSAR ANALYSIS OF IMBALANCED TOXICITY DATASETS USING K-NEAREST NEIGHBORS APPROACH

<u>Alexander Golbraikh</u>, Hao Zhu, Kun Wang, Mei Wang Bell, Alexander Tropsha

School of Pharmacy, Beard Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC, 27599, USA; E-mail: golbraik@email.unc.edu

One of the challenges of modern QSAR analysis especially as applied to toxicity data is building models for imbalanced datasets, i.e., those with uneven distribution of compounds between different classes or categories. Typically, standard QSAR methodologies do not work well when applied to imbalanced datasets. We have developed novel QSAR methodology and software for studies of imbalanced datasets based on k-nearest neighbor methodology. Our approach includes: (i) automatic detection and removal of outliers; (ii) filtering, i.e., excluding compounds of each class or category dissimilar from all compounds of other classes or categories; (iii) division of a dataset into training, test, and external validation sets ensuring sufficient representation of each class in each set (division into multiple diverse training and test sets is carried out using sphere-exclusion algorithm); (iv) four different target functions; (v) different weights for different classes or categories; (vi) penalizing discrepancies of accuracies between classes or categories; (vii) moving threshold for assigning a compound to a class; (viii) model applicability domains; (ix) consensus prediction by multiple models. We will illustrate this approach using several datasets including EPA Fathead Minnow Acute Toxicity Dataset (EPAFHM) with compounds divided into four categories based on LC50 data, Carcinogenic Potency DataBase (CPDB), aquatic toxicity dataset of compounds tested against *Tetrahymena Pyriformis*, and a few other sets. We will demonstrate that our procedure increases the predictive power of QSAR models for validation sets of compounds in all considered cases.

CLASSIFICATION SAR MODELS OF CHEMICALS TOXICITY TO GUPPY, FATHEAD MINNOW AND RAINBOW TROUT ON THE BASIS OF STRUCTURE SIMILARITY

Veniamin Yu. Grigor'ev[1], Eric E. Weber[2], Oleg A. Raevsky[1]

[1]Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds, Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia; [2]National Exposure Research Laboratory, U.S. Environmental Protection Agency, 30605, Athens, GA, USA; E-mail: beng@ipac.ac.ru

The concept of structure similarity to construct stable predictable models of mode of action (MOA) was applied for training sets of pollutants toxicity to Guppy, Fathead Minnow and Rainbow Trout.

First nearest neighbor (estimated by means of Cosine indexes) was used in one approach to classify MOA of chemicals-of-interest. A probability approach was considered to take a possibility of few MOA for any compound into account. Means maximal Cosine index was applied also to estimate a difference in structure similarity between different MOA for any subset.

The training set of chemicals toxicity to Guppy contained the data for 293 toxicants (90 compounds had nonpolar narcosis mechanism, 121 chemicals had polar narcosis MOA, 51 reactivity compounds and 31 specify chemicals). Right recognition by means of first nearest neighbor was equal to 92.2% for nonpolar narcosis, 91.7% for polar narcosis, 72.6% for reactivity chemicals and 58.1% for specify compounds. The probability approach gave else better results: 96.7%, 96.7%, 86.3 and 54.8% accordingly.

The training set of chemicals toxicity to Fathead Minnow contained the data for 249 toxicants (139 compounds had narcosis I mechanism, 16 chemicals had narcosis II MOA, 42 reactivity compounds, 6 neurodepressants, 16 inhibitors of acetylcholinesterase (AChE) and compounds with other MOA). Right recognition by means of first nearest neighbor was equal to 81.3% for narcosis I, 31.3% for narcosis II, 59.5% for reactivity chemicals, 83.3% for neurodepressants and 75.0% for AChE inhibitors. Again the probability approach gave better results: 87.1%, 43.8%, 66.7, 83.3% and 75.0% accordingly.

The training set of chemicals toxicity to Rainbow Trout contained the data for 337 pollutants (58 fungicides, 116 herbicides, 117 insecticides, 46 microbiocides). Total right recognition was equal to 76.9% by means of first nearest neighbor.

PBTs SCREENING BY MULTIVARIATE ANALYSIS AND QSAR MODELING

Paola Gramatica, Ester Papa

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Structural and Functional Biology, University of Insubria, via Dunant 3, 21100 Varese, Italy; E-mail: paola.gramatica@uninsubria.it

The limited availability of experimental data, necessary for risk assessment of chemicals, and the derived general lack of knowledge of the properties and activities of existing substances has led the European Commission to adopt the REACH (Registration, Evaluation and Authorization of CHemicals) legislation. The high expected cost of REACH has increased the interest on development and validation of alternative methods, such as the Quantitative Structure-Activity Relationships (QSARs), able to predict data and to provide priority lists.

Persistent Bioaccumulative and Toxic (PBT) chemicals are among those of higher concern requiring authorization, thus methods for early identification of these behaviour are needed.

A structurally-based multivariate approach is proposed here for the screening of potential Persistent Bioaccumulative Toxic (PBT) chemicals into the environment.

Overall half lives data in air, soil, water and sediment, as well as BCF and fish acute toxicity data, were collected for a structurally heterogeneous set of more than 200 compounds, and combined by Principal Component Analysis in order to rank the chemicals according to their potential PBT behaviour. The PC1 score, able to identify PBT-like compounds, is modelled by QSAR approach by Multiple Linear Regression (MLR) and molecular descriptors. The here proposed QSAR model of this PBT index, also externally validated for its predictivity, is applicable to chemicals even without *a priori* knowledge of the toxicological and environmental behaviour, being it based on easily calculable theoretical molecular descriptors. This model could be successfully applied as screening tool for the identification of new potential PBTs, also before their synthesis.

The here proposed model was applied to predict the PBT-potential of different classes of compounds also included in the ORATs and in The Water Framework Directive priority lists.

Results from this approach were compared with the US EPA PBT-profiler screening tool.

PREDICTION OF POLYCYCLIC AROMATIC COMPOUNDS' MUTAGENICITY AND GENOTOXICITY BY QSAR CLASSIFICATION MODELS

Paola Gramatica, Pamela Pilutti, Ester Papa

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Structural and Functional Biology, University of Insubria, via Dunant 3, 21100 Varese, Italy; E-mail: paola.gramatica@uninsubria.it

Polycyclic aromatic hydrocarbons (PAHs), absorbed mainly on urban aerosol, are ubiquitous pollutants of high environmental concern. The experimental data of a mutagenicity test on human B-lymphoblastoid cells (alternative to the Ames bacterial test) for a set of 70 oxo-, nitro- and unsubstituted PAHs, detected in particulate matter (PM), were modeled by Quantitative Structure-Activity Relationships (QSAR) classification methods (k-NN, k-Nearest Neighbour, and CART, Classification and Regression Tree) based on different theoretical molecular descriptors selected by Genetic Algorithms. The best models were validated for predictivity both externally and internally. For external validation, Self Organizing Maps (SOM) were applied to split the original data set. The best models, developed only on reduced training sets, show good predictive performance also on the prediction set chemicals (sensitivity: 76.9-92.3%, specificity: 55.6-87.5%). The classification of PAHs according to their mutagenicity, using only few theoretical molecular descriptors, allows a preliminary assessment of the human health risk, and the prioritization, of these compounds.

Analogously, the genotoxicity of 276 heterogeneous Polycyclic Aromatic Compounds (PACs) was modeled by QSAR classification methods. The studied response consisted of *a priori* defined classes of genotoxicity obtained on the basis of SOS Chromotest values. In order to propose only externally validated models, three different splitting approaches, D-optimal Experimental Design, SOM and Random Selection, were applied to the original data set for methodology comparison and selection of the best predictive model, independently of the splitting. QSAR models were developed applying CART and k-NN methods on a training set of 174 compounds. The best models, based on two 1D- and 2D-descriptors, have high sensitivity and specificity also for chemicals of an external prediction set of 102 chemicals, not participating to model development.

QUANTITATIVE STRUCTURE-ACTIVITY RELATIONSHIP STUDY OF SELECTIVE LIGANDS FOR THE THYROID HORMONE RECEPTOR β

Paola Gramatica, Huanxiang Liu

QSAR Research Unit in Environmental Chemistry and Ecotoxicology, Department of Structural and Functional Biology, University of Insubria, via Dunant 3, 21100 Varese, Italy; E-mail: paola.gramatica@uninsubria.it

An accurate and reliable QSAR model of 87 selective ligands for the thyroid hormone receptor beta 1 (TRβ1) were developed, based on theoretical molecular descriptors to predict the binding affinity of compounds with receptor. The structural characteristics of compounds were described wholly by a large amount of molecular structural descriptors calculated by DRAGON. The built model was fully accessed by various validation methods, including internal and external validation, Y-randomization test, chemical applicability domain and all the validations indicates that the QSAR model we proposed is robust and satisfactory. According to these validations, the best QSAR model based on the training set by SOM was selected with six most relevant structural descriptors as the inputs among those with a smaller number of response outliers and structurally influential chemicals. For the training set, this model has the correlation coefficient R^2, Q^2_{LOO} and RMSE are 0.836, 0.793 and 0.550, respectively. By interpreting the molecular descriptors in the regression model, we can conclude that the activity of the studied compounds mainly depends on molecular polarity, size, shape and nucleophilicity. When the best model was applied to the prediction set, the correlation coefficient R^2_{pred}, Q^2_{EXT} and RMSE are 0.730, 0.711 and 0.702, respectively. Thus, the built QSAR model can be used to fast and accurately predict the binding affinity of compounds (in the defined applicability domain) to TRβ1. At the same time, the model proposed could also identify and provide some insight into what structural features are related to the biological activity of these compounds and provide some instruction for further designing the new selective ligands for TRβ1 with high activity.

3DPL – AN ULTRA-FAST, INTERNET ENABLED PROTEIN-BASED SEARCHING/DOCKING SYSTEM FOR BIOLOGISTS AND CHEMISTS

Tad Hurst

ChemNavigator, Inc., 6126 Nancy Ridge Dr., San Diego, California 92121, USA; E-mail: thurst@chemnavigator.com

3DPL is an ultra fast protein-based searching/docking system that is designed to find an enhanced number of active molecules from large set of potential ligands. 3DPL uses a protein 3D structure as the query, and can search as many as 30 structures/second on a single CPU, and can thus search about 1,000,000 structures/day/computer. The hits that result from such a search are reproducibly 15-30 times more likely to be active than those randomly selected. Because of the speed of 3DPL's searching technique, called "Tweak-Docking", searching is NOT limited to the known active site, but can include the entire protein. Thus, 3DPL is appropriate for studies of orphan receptors and studies in which allosteric binding sites are involved. When the known binding site is not specified in a 3DPL run, the program consistently finds the known binding site. 3DPL is internet-enabled, and this allows the researcher to specify the query as a PDB code only. 3DPL will download the 3D structure of the receptor from the PDB. The ligands to be searched can be specified by the researcher in a variety of formats, or can be downloaded at run time from ChemNavigator's online iResearch system. The latter contains more than 25 million samples that can be ordered for subsequent efficacy testing. 3DPL has been used to find lead compounds in as few as 25 ordered samples.

CHEMOINFORMATICS TOOLS FOR PREDICTION OF VARIOUS PROPERTIES OF RELEVANCE TO DRUG DISCOVERY

<u>Svava O. Jonsdottir</u>[1], Irene Kouskoumvekaki[1], Olivier Taboureau[1,2], Niclas T. Hansen[1], Jens E. P. Larsen[1], Preeti Singh[1], Flemming S. Jorgensen[2], Soren Brunak[1]

[1]Center for Biological Sequence Analysis (CBS), BioCentrum, Technical University of Denmark, DK-2800 Kongens Lyngby, Denmark; [2]Department of Medicinal Chemistry, Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK-2100 Copenhagen O, Denmark; E-mail: svava@cbs.dtu.dk

The Chemoinformatics Group at CBS, DTU works with the development of new and innovative computational tools for use in drug discovery and lead optimization [1]. The possibility of using computational methods for screening out compounds at an earlier stage can significantly improve the success rate among drug candidates, as many late drug failures due to toxicity and bioavailability factors can thus be avoided.

By organizing information for chemical receptors and their ligands using self organizing maps (SOM) we have developed a useful virtual screening platform for pre-selecting compounds to be tested in drug discovery pipelines. We have also developed prediction models for pH-dependent aqueous solubility of drugs [2] and models for drug toxicity. Prediction tools for three different toxicity end points have been made, including molecular structure based models for hERG ion channel blocking (cardiotoxicity) and for mutagenicity (genotoxicity), and biomarkers based model for acute organ damage using metabonomics NMR data from rat urine.

1. S.O. Jonsdottir, F.S. Jorgensen, Brunak, S., *Bioinformatics*, **2005**, 21, 2145.
2. N.T. Hansen, I. Kouskoumvekaki, F.S. Jorgensen, S. Brunak, S.O. Jonsdottir, *J. Chem. Inf. Model.*, **2006**, 46, 2601.

FINDING BINDING SITES FOR LOW-MOLECULAR LIGANDS ON THE SURFACE OF C1Q WITH MOLECULAR DOCKING

David M. Karlinsky[1,2], Michael E. Popov[1], Alexander P. Kaplun[2]

[1]M.V. Lomonosov Moscow state academy of fine chemical technology (MITHT), 119571 Moscow, prospect Vernadskogo 86; [2]Shemyakin and Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences (IBCH RAS), 117997 Moscow, ul. Miklukho-Maklaya 16/10; E-mail: karlinskyd@rambler.ru

The classical pathway of the complement system activation is triggered by binding of the globular head of first component of complement (C1q) protein to Fc-domains of antibodies IgG or IgM. A range of negatively-charged low-molecular organic compounds can inhibit this interaction and thus be used as drugs to treat such diseases as heart stroke, asthma, Alzheimer's disease, immune conflicts in transplantation, etc. However, the mechanism of these inhibitors' interaction with C1q or antibodies was unknown.

Using the blind molecular docking method, we found possible binding for these compounds sites at the globular head of C1q. One of these sites at the "top" of globular head of C1q has the highest correlation of the estimated binding free energy of low-molecular ligands with their inhibiting ability.

Finding the site, which was presumably responsible for inhibition, allowed us to estimate the inhibitor properties of new low-molecular compounds *in silico* to select the most promising ones for further chemical synthesis. To this moment nearly 90 of low-molecular compounds were docked to the possible binding site. These docking results are now being refined using molecular dynamics methods, which allow to take into account the mobility of the C1q residues.

MOLECULAR DESIGN OF ANTIOXIDANTS BY SUPERPOSITION OF THEIR EFFICIENCY AND LOW LEVEL OF A TOXICITY

Veronika Khairullina[1], Svetlana Kirlan[2], Anatolyi Gerchikov[1], Lidiya Tyurina[3], Garifa Garifullina[1], Alexander Kolbin[3]

[1]Bashkir State University, 32, Frunze Str., Ufa, 450074, Russia; [2]Ufa State Petroleum Technological University, 1, Kosmonavtov Str., Ufa, 450074, Russia; [3]Institute of herbicides and regulators growth plants, 65, Uljanovich Str., Ufa, 450029, Russia; E-mail: Veronika1979@yandex.ru

A mathematical model for prediction of antioxidative activity (AOA) with a recognition level of ~90% was developed using the SARD-21 computer system. Based on this model, structural modification of ionol and 6-methyluracil was carried out. A set including 32 potential antioxidants was generated. The interval levels of toxicity were theoretically predicted and the effect of structural fragments on the toxic properties of the most efficient potential antioxidants was analyzed. Structural attributes characteristic of highly efficient, low toxicity antioxidants were revealed for the first time. Based on complex analyses of the AOA and toxicity, thirteen structures of potentially efficient low toxicity antioxidants were proposed. The results obtained in this work show that mathematical models for prediction of the AOA and the interval levels of toxicity can be used for rapid evaluation of the antioxidant and toxic properties of organic compounds. By applying these models to the structures of the known AOs one can predict the structures of new potentially nontoxic inhibitors of oxidation processes.

THEORY OF PATTERN RECOGNITION IN THE DEVELOPMENT OF THE LOW-TOXIC ANTIINFLAMMATORY DRUGS

Veronika Khairullina[1], Azat Mukhametov[1], Anatoliy Gerchikov[1], Garifa Garifullina[1], Lidija Turina[2], Felix Zarudiy[1]

[1]Bashkir State University, 32, Frunze Str., Ufa, 450074, Russia; [2]Ufa State Petroleum Technological University, 1, Kosmonavtov Str., Ufa, 450074, Russia; E-mail: Veronika1979@yandex.ru

The ulcerogenic effect from the application of nonsteroid anti-inflammatory drugs (NSAIDs) is caused by the inhibition of the fermentative activity of cyclooxygenase 1 (COX-1), while anti-inflammatory action of NSAIDs is caused by the inhibition of the structural isomer COX-2, whose expression rises in the zones of inflammation. An increase in the selectivity of NSAIDs against to COX-2 will make it possible to decrease their gastrotoxic action and to reach the purpose of the present investigations.

By the computer system Sard-21 we have formed the forecast model of the selectivity of COX inhibition by the NSAIDs and the effectiveness of their anti-inflammatory action with the level of recognition ~ 80 %. There are revealed structural signs, characteristic for effective NSAIDs, which contribute the greatest contribution to the selectivity of COX-2 inhibition. On the base of the established regularities ones define the directions of structural modification in the molecules of paracetamol, resveratrol and diclofenac for the purpose of an increase in their anti-inflammatory activity with the retention of the low level of gastrotoxic action.

As the result of the optimization of the structures of the substances in the selected directions being investigated are designed 37 structures with the required combination of therapeutic properties.

QSRR STUDY OF NITROGEN-CONTAINING BIOLOGICAL ACTIVE HETEROCYCLES

Olga V. Kharitonova, Svetlana V. Kurbatova

Samara State University, Acad.Pavlov str., 1, Samara, 443011, Russia; E-mail: curbatsv@ssu.samara.ru

Quantitative Structure Retention Relationship (QSRR) is the important method for the modeling biochromatographic behaviour of biological active substances and its partition between liophilic and hydrophilic phases. So high performance liquid chromatography (HPLC) is the most convenient and perspective method for physicochemical properties determination and simulation of chromatographic process which are result of drug leading into the living organism.

Nitrogen-containing heterocycles and its derivatives attract researchers attention because of its biological activities and as model of theoretical organic chemistry.

We examined chromatographic behaviour of imidazole, indole, thriazole, benzimidazole, behzthriazole and its derivates under condition of reversed phase HPLC. The analysis was carried out by liquid chromatography Biotronik Chromatographiesystem with column packed with Separon SGX– C_{18} and a UV detector operating at 254 nm.

The dipole moment, polarizability of the molecule, as well as the energy of hydration, molecular refraction, and the distribution coefficient for the n-octanol--water system, were calculated by the AM1 semiempirical method with full molecular geometry optimization realized in the HyperChem 7.0.

By comparison of experimental retention data and calculated van-der-vaals volume, polarizability, lipophilicity and topological Kovats and Winer indices it was established regularity of azoles distribution between lipophilic and hydrophilic phases. Correlation between physico-chemical and thermodynamic properties of heterocycles was investigated also.

The influence of chromatography system nature (mobile and stationary phases, the structures of molecules sorbats) on the thermodynamic of partition nitrogen-containing aromatic heterocyclic was investigated. For estimating all types intramolecular interactions influencing retention and partition such parameters as solvent strength $\varepsilon°$, solvent polarity P by Snyder and the parameter of the solubility δ by Hildebrand and Scott were used.

We study substitute's nature, position and its number influence on the physico-chemical and chromatographic properties of azoles. It was shown, that molecule topology influence sufficiently on the distribution between stationary phase and eluent. By comparison of potentiometry determined pK_a and pK_b values of azoles with its retention parameters it was conclude that pK_a and pK_b influence on selectivity and chromatographic behaviour just as pH value.

COMPUTER-AIDED PREDICTION OF PROMISING ANTI-TUMOR TARGETS TAKING INTO ACCOUNT INFORMATION ABOUT PROBABLE SIDE EFFECTS

Olga N. Koborova[1,2], Alexey V. Zakharov[2], Alexey A. Lagunin[2], Dmitrii A. Filimonov[2], Alexander Kel[3], Fedor Kolpakov[4], Ruslan Sharipov[4], Vladimir V. Poroikov[2]

[1]Faculty of Bioengineering and Bioinformatics, Lomonosov Moscow State University, GSP-2, building 73, Leninskiye Gory, Moscow, 119992, Russia; [2]Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., Moscow, Russia; [3]BIOBASE GmbH, Germany; [4] Institute of Systems Biology, Novosibirsk, Russia; E-mail: okoborova@gmail.com

One of the challenging problems of oncology is finding of promising anti-tumor targets. The purpose of this study is to identify such targets that give an opportunity to stop the proliferation activity of cancer cells and switch them to apoptosis. On the basis of system biology approaches and mathematical graphs theory we developed the algorithm for estimation of potential targets and created the appropriate computer program. Cell cycle networks and proliferation cascades as well as information on differentially expressed transcription factors for the networks and microarray data for different cancer types were taken from Cyclonet (http://cyclonet.biouml.org) and Transpath (http://www.biobase.de) databases. We included information taken from literature about the side effects, if potential target is affected. The method was tested on the breast cancer, since this type of tumor is widely spread. We used pathways, which experimentally confirmed as directly connected with tumor development. As a result, many targets known from literature were identified. Also, some new targets are predicted, which have to be validated experimentally. The results of estimation provide the basis for selection of the most promising targets in different types of cancer.

ARYLSULFONYLATION OF α-AMINOACIDS: QSAR AND COMPUTER MODELING OF REACTION'S MECHANISM

Natalya Scheglova, <u>Ludmila Kochetova</u>, Tatyana Kustova, Natalya Kalinina

SEI HPE «Ivanovo State University», 39, Ermak Str., Ivanovo, 153025, Russia; E-mail: kochetova_lb@mail.ru

Products of arylsulfonylation of amines are widely known in pharmacology as medicines and in enzymology as inhibitors of enzymes. For a few years kinetic characteristics of the reactions of acyl transfer with participation of aminoacids have been investigated by us, their *N*-acylation by chloroanhidrides of aromatic carbonic and sulfoacids, in particular. The data of the precise kinetic experiment allow choosing the descriptors of the molecular structure of reactants for the prediction of their reactivity in the process under investigation. We carried out calculations of electronic, structural and energetic characteristics of the molecules of arylsulfonylchlorides, and also of the neutral and anionic forms of α-aminoacids taking part in the acylation (by methods *RHF/6-31 G*** and *UHF/6-31++G*** respectively). The analysis of the results of quantum chemical calculations of geometrical characteristics of α-aminoacids molecules has shown, that significant changes in the structure of the reactive centre – aminogroup – in the row of native aminoacids has not occurred, hence the application of the descriptors of the structural formula for the prediction of the kinetics of the acylation of aminoacids by chloroanhidrides of aromatic sulfoacids was hardly reasonable. Besides, the electronic characteristics of aminoacids molecules (q_N, E_{HOMO}, Cp_z) in the row considered undergo some changes and can be used as descriptors. For establishing the most probable root of the reaction calculation of three-dimensional potential energy surface (PES) was carried out for the model reaction of glycine with benzenesulfonylchloride (pic.). The results of the calculations testify in favour of the frontal attack by nucleophile on the sulfonyl centre.
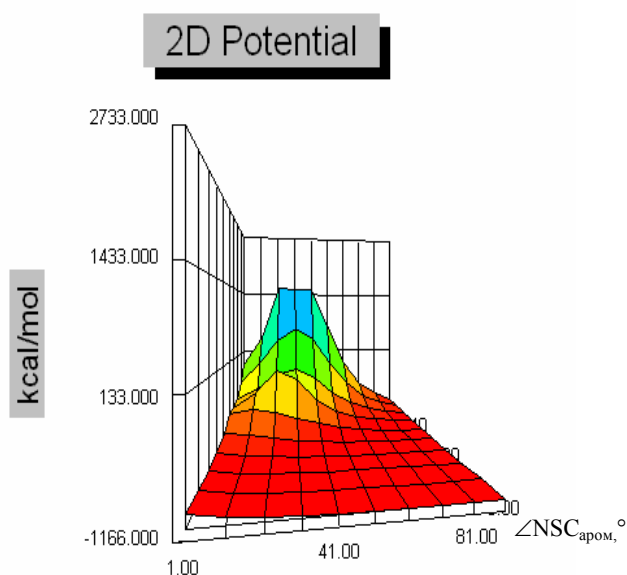
CYCLONET – AN INTEGRATED DATABASE AND COMPREHENSIVE TOOL FOR COMPUTATIONAL PHARMACOLOGY

Fedor Kolpakov[1,2], Vladimir Poroikov[3], Ruslan Sharipov[4,1,2], Luciano Milanesi[5], Alexander Kel[6]

[1]Institute of Systems Biology, 15, Detskiy proezd, Novosibirsk, 630090, Russia; [2]Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia; [3]Institute of Biomedical Chemistry RAMS, Moscow, Russia; [4]Institute of Cytology and Genetics SB RAS, Novosibirsk, Russia; [5]Institute of Biomedical Technologies, CNR, Segrate (MI), Italy; [6]BIOBASE GmbH, Wolfenbuettel, Germany; E-mail: fedor@biouml.org

The aim of the Cyclonet database (http://cyclonet.biouml.org) is to develop an integrative approach to help researchers to understand the key processes of cell life - targets of modern pharmacological agents: cell cycle regulation and carcinogenesis - through modeling and simulation of gene regulatory networks. Cyclonet integrates data of genomics, proteomics, chemoinformatics, and systems biology for their use in drug design. Genomics data are represented by 354 categorized links to available microarray experiments. This section also contains the complete list of human genes linked with 5 microarray experiments and the lists of up- and down-regulated genes for breast cancer revealed by meta-analysis of these microarray experiments. Proteomics section contains information about 2944 proteins, their complexes and interactions. Chemoinformatics – data about 55 key targets for anticancer treatment, 62 anticancer pharmacological activities and 422 related activities. This information is used by the PASS system (http://www.ibmc.msk.ru/PASS) to predict new ligands with anticancer activities. Systems biology section contains more than 200 diagrams and 32 mathematical models of cell cycle and related systems annotated from literature. Novel software technologies were used for development of Cyclonet: 1) BioUML technology (http://www.biouml.org) was used for formal description, visual modeling of eukaryotic cell cycle and for query and editing of the database content. BioUML workbench allows to simulate the described systems behavior using Java or MATLAB simulation engines; 2) BeanExplorer EE (http://www.beanexplorer.com) was used for development of web interface for access to Cyclonet trough the Internet; 3) BMOND database (http://bmond.biouml.org) – important integrated part of Cyclonet - stores diagrams, description diagram elements (genes, proteins, substances, concepts, reactions and semantic relationships), mathematical models and results of simulation. Cyclonet is continually enriching database and will be useful for all scientists working in the field of computational pharmacology and drug design.

SOLUTION STRUCTURES AND DYNAMICS OF THE DIHYDROFOLATE REDUCTASE COMPLEXES WITH ANTIBACTERIAL DRUG TRIMETHOPRIM

Nadezhda Kovalevskaya, Yegor Smurnyy, Berry Birdsall, James Feeney, Vladimir Polshakov

Center for Drug Chemistry, 7, Zubovskaya Str., Moscow, 119815, Russia; E-mail: nadia.k.w@gmail.com

Dihydrofolate reductase (DHFR) catalyses the NADPH-dependent reduction of folate and dihydrofolate to tetrahydrofolate. Since the latter is an important cofactor in the biosynthesis of purines and amino acids, DHFR has proved to be an excellent target for antifolate drugs that act by inhibiting the enzyme in parasitic or malignant cells. The effectiveness of antibacterial drug trimethoprim (TMP) is due to its significantly increased ability of the binding to the bacterial enzyme compared with the vertebrate form.

In order to explore the origins of the selectivity of TMP binding to DHFR, information on the structure and dynamics of the complexes of human and *Lactobacillus casei* forms of enzyme with TMP and/or cofactor NADPH is obtained using the NMR spectroscopy methods and molecular dynamics calculations.

INTEGRATED APPROACH TO ASSESS THE DOMAIN OF APPLICABILITY OF SOME COMMERCIAL (Q)SAR MODELS

Sunil A. Kulkarni, Joel Paterson, Jiping Zhu

Safe Environments Programme, Health Canada, Ottawa, ON, K1A 0L2, Canada; E-mail: sunil_kulkarni@hc-sc.gc.ca

An integrated approach based on structural analysis to assess the domain of applicability (DA) of some well-known commercial (Quantitative) Structure Activity Relationship ((Q)SAR)-based predictive toxicity models is proposed. The first stage consists in identifying key structural features in a query chemical and determining their coverage in model training sets based on a predefined scale of structural feature representation. In the second stage, using a predefined similarity criterion, cluster analysis is applied to generate information on the extent of structural similarity that exists between the query chemical and the training set. Five cancer and mutagenicity models in both CASETOX (Multicase Inc.) and TOPKAT (Accelrys Software Inc.) programs were assessed for their DA in reference to a set of 11 query chemicals that were taken from the Canadian Domestic Substances List. These chemicals do not belong to any of the ten model training sets. Determination of whether a given query chemical lies within or outside the model's DA was carried out for each of the 11 query chemicals in a total of 10 models. Based on key structure feature, in 93 out of the total 110 situations the query chemical was found to lie within the model's DA. However, according to the cluster-based analysis this was true only in 46 cases. In all it was found that there were 36 situations where both key structural feature-based and cluster-based DAs agreed that a query chemical lies within the model's DA whereas in 6 cases they agreed that the query chemical lies outside of it. Therefore, more confidence may be placed in these 42 situations where the DA results from both analyses agree with each other. From a regulatory perspective, the model generated information on chemical safety needs to be as reliable as possible. For this it is important to know if a given query chemical falls within the predictive boundary i.e. DA, of the concerned (Q)SAR model. The methodology presented here could serve as a useful tool for regulators to make preliminary assessment of (Q)SAR-based systems and in turn help the process of hazard-based regulatory assessments of chemicals.

UNIFIED REPOSITORY OF QSAR MODELS AND MOLECULES – THE SYSTEM OF PROPERTY PREDICTION OF THE CHEMICAL COMPOUND

Mikhail I. Kumskov[1], Nikolai S. Zefirov[2], Igor V. Svitanko[1], Alexei M. Zaharov[3], Ludmila A. Ponomareva[3]

[1]Zelinsky Institute of Organic Chemistry, Leninsky prospect, 47, 119991 Moscow, Russia; [2]Chemistry Department, Moscow State University, 119992 Moscow Russia; [3]Mechanics and Mathematics Department, Moscow State University, 119899 Moscow Russia

The System enables chemists to uniformly presented QSPR/QSAR models and consistently deliver scientific views of QSPR/QSAR prediction data that are dynamically tailored to the new created models. First system version will used MATLAB 7.1 with Statistics, Database and Fuzzy Logic Toolboxs to design, investigate and store the QSPR/QSAR models. Molecular structure cannot be represented by a unique formal model; several molecular representations can represent the same molecule, depending on the level of the underlying theoretical approach and these representations are often not derivable from each other. To represent chemical structures for QSPR/QSAR modeling a new approach was developed, based on the analogies between the image and flexible molecular graph representation.

The molecular descriptors are now playing a key role in QSPR/QSAR research. They are derived from several different theories, such as quantum chemistry, information theory, organic chemistry, graph theory, etc. and are applied in modeling several different properties in fields such as toxicology, physical chemistry, medicinal and pharmaceutical chemistry, environmental and toxicological studies and regulatory tools.

Our features space, in which description of the molecules is conducted, - the structural symbolic spectra of marked graphs, - are formed automatically with consequent complicating a detail description and the different presentation level of the molecules [1]. The features space types are defined by the different level of the structure presentation. There are explored the prognostic ability of two types of classifying functions - linear and statistical weighted voting, - applying to the "molecule-feature" tables, formed for topological and spatial presentation of molecules, which are stored in the training molecules database.

We construct a three-level 'descriptor alphabet'. Our idea is to construct a label for a group of critical points by joining labels of individual points and labels denoting distances between the points. The first level of our descriptor alphabet consists of single critical points and their labels. For every molecule, we calculate value of such descriptor by calculating number of occurrences of a critical point of a given type on a molecular surface [2]. An advantage of our features space is that molecule description is invariant to space rotations of molecule, which eliminates the need to align molecules in 3D space (as in CoMFA method). We plan to develop the method further by incorporating fuzzy logic for treatment of non-rigid molecules.

1. D.F. Mityushev, M.I. Kumskov, *Pattern Recognition and Image Analysis*, **1996**, 6(4), 809-822.
2. I.V. Svitanko et al. QSAR modeling on the base of electrostatic molecular surface (amber fragrances). Proc. The 16th European Symposium on Quantitative Structure-Activity Relationships & Molecular Modelling, Sept. **2006**.

ANALYSIS OF PROCESSING MECHANISM OF HUMAN 8-OXOGUANINE-DNA GLYCOSYLASE ENZYME OF DNA REPAIR SYSTEM

Nikita Kuznetsov, Vladimir Koval, Olga Fedorova, Yury Vorobjev

Institute of Chemical Biology and Fundamental Medicine of SB RAN, 8 Lavrentiev Ave. Novosibirsk, 630090, Russia; E-mail: Nikita.Kuznetsov@niboch.nsc.ru

Living cells continuously experience a great number of insults from reactive oxygen species that are produced during respiration and generated by UV or ionization radiation.

The effect of oxidative damages to DNA include miscoding and disregulation of gene expression and may lead to cancer and aging. Base excision repair of damaged bases is initiated by DNA glycosylases, enzymes that recognize lesions and excise the damaged base.

The human 8-oxoguanine-DNA glycosylase hOgg1 plays a prominent role and excise the 8-oxoguanine (8-oxoG), a major damaged purine, from DNA. Defects in hOgg1 have been associated with human cancer and enhanced mutagenesis. Recently it was show that recognition and excision of damaged bases by DNA glycosylases is accompanied by several conformational rearrangements that bring the excised damaged base into enzyme catalytic site, i) initial encounter with DNA and forming of enzyme/DNA complex with eversion of the damaged base, ii) insertion of several enzyme residues into DNA, iii) enzyme isomerization to the catalytically competent form, and finally, iv) reaction of β-elimination to break the damaged DNA strand. At the last stage, the 8-oxoG base bound to the active site of hOgg1, acts as a cofactor of the β-elimination reaction . It was shown that the rate of β-elimination reaction increased at least 10-folds by 8-BrG, an 8-oxoG analog [1,2]. In this work we report a result of computer simulations of 8-oxoG and 8-BrG in the active site of enzyme to understand the nature of the difference between the two compounds. Docking, analysis of interactions and binding rate calculations are performed via our BISON molecular modeling program.

1. N.A. Kuznetsov, V.V. Koval, G.A. Nevinsky, K.T. Douglas, D.O. Zharkov, O.S. Fedorova, *J. Biol. Chem,.* **2007**, 282, 1029-1038.
2. J.C. Fromme, S.D. Bruner, W. Yang, M. Karplus, G.L. Verdine, *Nat.Struct.Biol.,* **2003**, 10, 204-211.

A MULTI-MUTAGENICITY ENDPOINT MODEL BUILT ON TOXICOPHORES

Jens E. P. Larsen[1], Olivier Taboureau[1,2], Flemming S. Jorgensen[2], Soren Brunak[1], Svava O. Jonsdottir[1]

[1]Center for Biological Sequence Analysis (CBS), building 208, Technical University of Denmark, DK – 2800 Kgs. Lyngby, Denmark; [2]Faculty of Pharmaceutical Sciences, University of Copenhagen, Universitetsparken 2, DK – 2100 Kobenhavn O, Denmark; E-mail: jepl@cbs.dtu.dk

We have developed a multi mutagenicity model that covers a wide range of mutagenicity endpoints found in the TOXNET database (http://toxnet.nlm.nih.gov). It is built upon a core of toxicophores developed by Kazius *et al.*, 2005 [1] for modelling the Ames test. We built the model upon these so-called Ames toxicophores, because the Ames test is a well established assay used in the drug development process, and the mechanisms of action are known to a satisfactory degree. Chemicals or their metabolites may either react directly with DNA by forming covalent bonds or chemicals may insert themselves non-covalently into the structure of DNA via the process of intercalation. We wanted to model as many mutagenicity endpoints as possible based on the above mentioned mechanisms of action and at the same time to develop additional toxicophores for the individual endpoints. First, we mined the genotoxicity part of the TOXNET database for the most frequently occurring assays. Then we applied the Ames toxicophores to the endpoints and evaluated the prediction accuracy. Endpoints such as Rec-assay DNA effects *E. coli*, Sex-linked recessive lethal gene mutation *D. melanogaster*, and Mitotic recombination or gene conversion *S. cerevisiae*, had prediction accuracies of 78.2%, 76.9%, and 73% respectively. All the Sister Chromatid Exchange (SCE) endpoints showed good accuracies for the predicted positives but not as good accuracies for the predicted negatives. We therefore developed additional toxicophores using a molecular fragmentation miner, MoFA, and thereby improved the prediction accuracies for SCE *in vitro* human lymphocytes, SCE *in vitro* nonhuman and SCE *in vivo* nonhuman from 51% to 73.3%, 55.3% to 63.7%, and from 64.0% to 77.2%, respectively.

1. J. Kazius, R. McGuire, R. Bursi, *J. Med. Chem.,* **2005**, 48, 312-320.

MOLECULAR DYNAMICS OF ZERVAMICIN II AND ITS MUTANTS IN DIFFERENT SOLVENTS

Olga V. Levtsova, Konstantin V. Shaitan

Lomonosov Moscow state university, Biology faculty, Bioengineering department, 119992, Russia, Moscow, Leninskie Gory, 1-12; E-mail: sunely@yandex.ru

In present time investigation of antimicrobal peptides, which have activity against gram-positive bacteria, some endo- and exoparasite and cancer cells, is very relevant. Research of these peptides represents a priority questions in biopharmacology. One of these peptides is zervamicin II, a member of the antibiotic peptaibol-family, that is produced by fungi *Emericellopsis, Trichoderma*. It consists of 16 amino acids and contains a high proportion of helix-promoting α, α-dialkylated amino acids (Aib, α-aminoisobutyric acid; Iva, D-isovaline). The mechanism of action is still unclear, but it's suggested that the peptides interact with the cell membrane and form ion channel. Comparative study of α, α-dialkylated amino acids (Aib, Iva) with alanine and valine was made using Molecular dynamic method. It was shown, that α, α-dialkylated amino acids have steric restrictions because of added alkyl group and make helical structure of pentaibols more stable. Zervamicin dynamic was investigated in water and methanol to find out conformational changes in different solvents. However no conformational changes were found out, so it means that zervamicin II retains its structure during embedding into the membrane. Zervamicin II molecule has very stable helical structure and in contrast to longer peptaibols hasn't any swivel motion, altering molecule length. To determine amino acids residues important for molecular stability the number of zervamicin II mutants was investigated (Aib replaced by Ala in 7 and 9 positions). Computer research showed that replacement in 7 position didn't influence on the molecular stability, but zervamicin II with replacement in 9 position altered its length in different solvent. So that the molecular length in water was 2,4 nm and in methanol – 1,6 nm. When alanine amino acid was added in 8 position, swivel motion of zervamicin II molecule appeared in methanol solvent, and molecule in water was as stable as native one. It is significant, that directed amino acid replacement is important not only for understanding action mechanism and role of each amino acid but also in pharmacology. So results of this work are directed to create new generation of antimicrobial peptides.

e-EXPERIMENTAL ENDOCRINE DISRUPTOR BINDING ASSAY (e³dba) DATABASE

Nathalie Marchand-Geneste[1], Alain Carpy[1], James Devillers[2], Jean-Marc Porcher[3]

[1]Universite Bordeaux 1 ; CNRS UMR 5255 ISM-LPTC, 351 cours de la Liberation, 33405 Talence, France; [2]CTIS, 3 Chemin de la Graviere, 69140 Rillieux La Pape, France; [3]INERIS, BP2, 60550 Verneuil-en-Halatte, France; E-mail: n.geneste@ism.u-bordeaux1.fr

Due to the abundance of data related to endocrine disruptors present in the environment, the free database "e–experimental endocrine disruptor binding assays" (e³dba) was designed to improve and to centralize the flow of information on experimental endocrine disruptor *in vitro* binding assays on nuclear receptors. It was designed with the aim to evaluate the validity of the existing QSARs and to establish new models. The e³dba database offers numerous interesting functionalities which are presented through selected graphical displays.

e-ENDOCRINE DISRUPTING CHEMICAL DATABASES FOR DERIVING QSAR MODELS

Nathalie Marchand-Geneste[1], James Devillers[2], Jean-Christophe Dore[3], Jean-Marc Porcher[4]

[1]Universite Bordeaux 1 ; CNRS UMR 5255 ISM-LPTC, 351 cours de la Liberation, 33405 Talence, France; [2]CTIS, 3 Chemin de la Graviere, 69140 Rillieux La Pape, France; [3]Departement Regulations, Developpement et Diversite Moleculaire, USM 0502, UMR 8041 CNRS, Museum National d'Histoire Naturelle, 63 rue de Buffon, 75005 Paris, France;.[4]INERIS, BP2, 60550 Verneuil-en-Halatte, France; E-mail: n.geneste@ism.u-bordeaux1.fr

There is increasing evidence that numerous chemicals released into the environment by human activities have the potential to alter the normal functions of the endocrine system in wildlife. These xenobiotics, mimicking natural hormones, are called endocrine disrupting chemicals (EDCs). The aim of this study was to catalogue the different EDC database (biological data, chemical descriptor data, etc.) resources available on the Internet for deriving structure-activity models. *In vitro* and *in vivo* experimental data on nuclear receptor binding assays, *in vitro* test methods reliability and variability assessment for detecting EDCs and comparison between biological assays will be presented and critically analysed.

QSAR STUDY OF FLAVONOIDS' IMMUNOMODULATORY ACTIVITY

Marica Medic-Saric[1], Ivona Jasprica[1], Zeljko Debeljak[2], Ana Mornar[1]

[1]Department of Medicinal Chemistry, Faculty of Pharmacy and Biochemistry, University of Zagreb, Croatia, A. Kovacica 1, 10000 Zagreb, Croatia; [2]Osijek Clinical Hospital, J. Huttlera 4, 31000 Osijek, Croatia; E-mail: bebamms@pharma.hr

Flavonoids are a group of phytochemicals ubiquitous in photosynthesizing organisms that display a remarkable array of biochemical and pharmacological activities. As human immune system can be modified by diet, pharmacologic agents, environmental pollutants, it is assumed that certain members of this group of compounds significantly affect the function of the immune system. Using appropriate cell culture model (human macrophage-like cells - differentiated THP-1 cells), the effects of 31 flavonoid representative on the production of selected cytokines that play crucial roles in innate and adaptive immune responses was determined (TNFα, IL-1β) [1]. TNFα and IL-1β production was measured using ELISA 24 hrs after the addition of chosen flavonoids in 30 μM concentration. *E. coli* lipopolysacharide was used as a positive control (TNFα and IL-1β production was considered as maximal, 100%). Un-stimulated cells did not produce cytokines and were used as a negative control. Some of tested flavonoids exhibited immunomodulatory effects and quantitative structure-activity relationships could be established. 26 Mulliken charges and sums there of corresponding to selected substituents were calculated based on B3LYP DFT single point approach. PM3 optimized geometries were used for these calculations. Random forests (RF) and genetic algorithms coupled with support vector machines (GA/SVM) were used for the development of quantitative relationships between percentages of activation of TNFα and IL-1β production and selected partial atomic charges. RF parameters were kept at their default values while GA/SVM parameters were fine-tuned by internal leave-one-out validation loop. Models obtained this way were characterized by average R^2 greater than 0.65 for calculated and experimental values while corresponding CV was less than 5%. These values were calculated based on external leave-one-out validation approach repeated ten times for different random seeds. Descriptors highly ranked by both, RF and GA/SVM approaches were extracted and their influence, as well as the influence of different substituents on the TNFα and IL-1β production was evaluated.

1. I. Jasprica, J. Dumic, A. Mornar, M. Medic-Saric, *Planta Med.,* **2006***, 72, 1036.

QSAR STUDY OF ANTHOCYANINS, ANTHOCYANIDINS AND CATECHIN AS INHIBITORS OF LIPID PEROXIDATION

Marica Medic-Saric[1], Vesna Rastija[2]

[1]Department of Medicinal Chemistry, Faculty of Pharmacy and Biochemistry, University of Zagreb, Croatia, A. Kovacica 1, Zagreb, 10 000, Croatia; [2]Faculty of Agriculture, Josip Juraj Strossmayer University of Osijek, Trg Sv. Trojstva 3, Osijek 31000, Croatia; E-mail: bebamms@pharma.hr

Catechins, anthocyanidins and their glycoside – anthocyanins, compounds that belong to the natural antioxidant - polyphenols, present in foods of plant origin, are potentially beneficial to human health. Antioxidative effects of those phytochemicals are based on the ability to scavenge different free radicals or the protection of biological molecules against oxidation. In the present study lipid peroxidation inhibitory effect [1] of mentioned polyphenols is correlated with molecular descriptors and physicochemical parameters calculated from three-dimensional structure, since studied compounds contain stereoisomers with different activities. Six groups of 3D descriptors have been used to generate a quantitative structure-activity relationship (QSAR) models: geometrical; GETAWAY (Geometry, Topology, and Atom Weights AssemblY); 3D-MoRSE, RDF (Radial Distribution Function) descriptors; Randic molecular profiles and WHIM (WeigHted Covariance Matrices) descriptors. Geometries of molecules were optimized using the molecular mechanic MM+ force field method applying the HyperChem 8.0 Evaluation software package. The 3D molecular descriptors used in this study have been calculated applying the on-line software Parameter Client (PCLIENT), an extension of E-Dragon. Physicochemical parameters (connolly accessible area, connolly molecular area, connolly solvent excluded volume, ovality, volume and surface area of molecule and hydration energy) were calculated using HyperChem 8.0 and Chem3D Ultra 10 Trial Version. Only 3D molecular descriptors and physicochemical parameters with ability to discriminate stereoisomers were chosen for multiple regression analysis. The selection of predictor variables for multiple regression was performed by best-subset method. Present work should provide the better understanding of structural factors that relate the antioxidant activity of polyphenols.

1. N. P. Seeram, M.G. Nair., *J. Agric. Food Chem.*, **2002**, 50, 5308.

SYNTHESIS AND POTENTIAL BIOACTIVITY OF NEW PYRAZOLYL-SYM-TRIAZINE DERIVATIVES

Svetlana N. Mikhaylichenko[1], Vladimir V. Dounin[1], Vladimir N. Zaplishny[1], Alexey A. Chesnyuk[2]

[1]University of Toronto Scarborough, 1265 Military Trial, Toronto, M1C 1A4 Canada; [2]Kuban State Agricultural University, 13 ul.Kalinina, 350044 Krasnodar, Russian Federation; E-mail: mikhay@utsc.utoronto.ca

Pyrazole and *sym*-triazine derivatives possess a wide spectrum of biological activities. It was of our interest to synthesize new compounds containing both of these heterocycles. The synthesis was successfully done using quaternary salts of *sym*-triazine by following scheme:
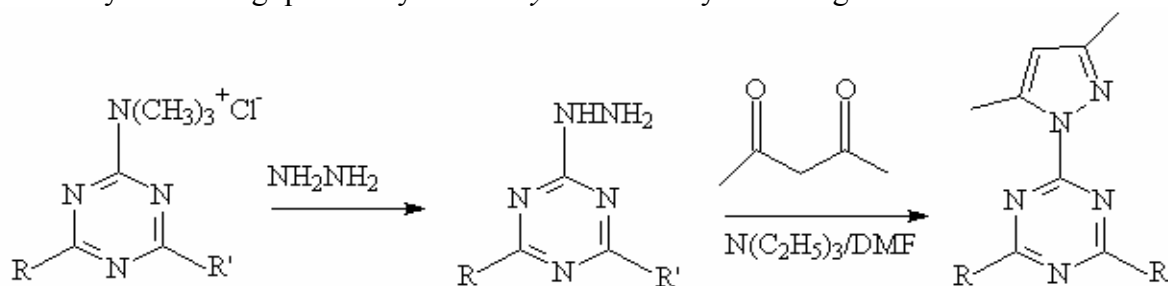
Where R and R'= morpholyl, piperidyl, OCH_3, NEt_2, NPh_2, the same or different.

These new compounds can be described as small white crystals, with melting points ranging from 107 to 188ºC. They are soluble in regular organic solvents and insoluble in water. The structures of all obtained compounds were proven by IR, H^1NMR, and mass spectroscopy.

We have checked for potential bioactivity of the synthesized compounds using the PASS program. The prediction of biological activity is indicated by the probability for the compound to be active or inactive, which are expressed as Pa and Pi respectively [1]. When Pa is greater than 0.7, the compound is most likely to reveal the predicted biological activity in experiments. Apparently, the cyclic AMP phosphodiesterase inhibitor, GABA receptor antagonist, dopamine D4 agonist, neuroprotector, and hypothermic activities for all of these products have been predicted. The two compounds with diethylamino and pyperidyl radicals have yielded predictions of vasodilator and thromboxane B2 antagonist activities, as well as antiischemic activity for the compounds with a methoxy radical. The PASS program predicts myocardial activity if the compound has diethylamino and pyperidyl radicals in its structure. For all these listed biological activities the Pa walue is greater than 0.7

Thereby, all synthesized 2-[3,5-dimethylpyrazol-1-yl]-4,6-disubstituted-1,3,5-triazines are potentially bioactive compounds. Currently we are planning on checking the real biological activity of these compounds.

1. "Interpreting the Results of Prediction." PASS Inet. V.N Orechovich Institute of Biomedical Chemistry. 9 Apr. **2007**. <http://195.178.207.233/PASS/IR.html>.

EFFECTIVITY OF THE PASS PROGRAM IN PREDICTING OF BIOLOGICAL ACTIVITY FOR SYM-TRIAZINE DERIVATIVES

Svetlana N. Mikhaylichenko, Vladimir V. Dounin, Vladimir N. Zaplishny

University of Toronto Scarborough, 1265 Military Trial, Toronto, M1C 1A4 Canada; E-mail: mikhay@utsc.utoronto.ca

Sym-triazine derivatives have a huge spectrum of different biological activity. One of such compounds is the bis-iodomethylate aminoester of *sym*-triazine acid previously synthesized by one of us [1]:
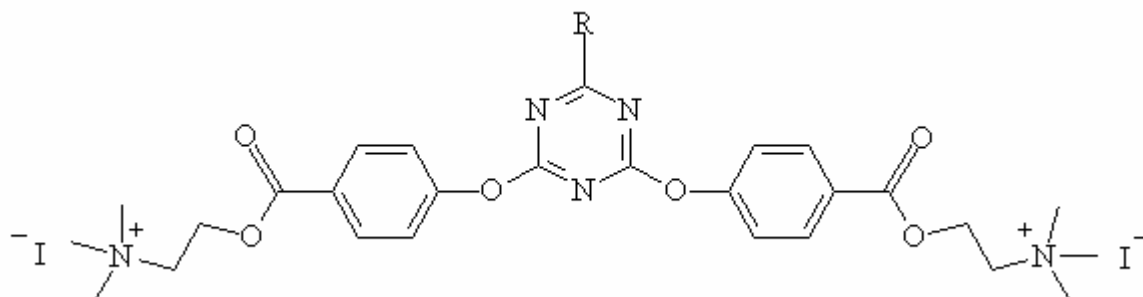
Where R = $NAlk_2$, OAlk, OPh.

In vitro experiments for all these compounds the high inhibitor activity toward acetylcholine esterase and phosphodiesterase has been discovered. The most highly active compound with R = OPh was patented [2].

Currently, with the help of the PASS program the cyclic AMP phosphodiesterase inhibitor activity for all listed above compounds was confirmed [3]. This illustrates to the high reliability of the PASS program in studying the dependence of biological activity on molecular structure.

The program also predicts high cardinitinamidase, cholinephosphate cytidylyltransferase inhibitor, and hematotoxic activities for those compounds, which is still a subject to test by experiments.

1. S.G. Antonyan, V.N. Zaplishny et al., *Khim. Pharm. Journal*., **1986**, Vol. 2, 172-175.
2. USSR Patent, # 1231834, **1986.**
3. "Interpreting the Results of Prediction." PASS Inet. V.N Orechovich Institute of Biomedical Chemistry. 9 Apr. **2007**. <http://195.178.207.233/PASS/IR.html>.

A GENERAL PHYSICOCHEMICAL QSPR MODELS OF CHEMICALS PARTITIONING IN SYSTEM SOLVENT–WATER–AIR

Oleg A. Raevsky, <u>Dmitry N. Moiseev</u>, Olga E. Raevskaya, Elena V. Bovina, Klaus-Jorgen Schaper

Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds, Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia; Borstel Research Center, 245845, Borstel, Germany; E-mail: Kschaper@borstel-fz.de, raevsky@ipac.ac.ru

Chemicals partitioning in system octanol-water is usually considered as very important parameter for quantitative description of transport processes in biochemical and pharmacological systems. However chemical partitioning in different solvent-water systems where a solvent can have essential other physicochemical parameters compare with octanol may be also useful to construct models of chemicals compounds distribution in systems water-diverse membranes.

The poster contains the results of QSPR modeling of chemicals partitioning in 13 systems solvent-water (hexane, cyclohexane, heptane, octane, hexadecane, diethyl ether, dibutyl ether, butyl acetate, chloroform, 1,2-dichloroethane, octanol, triol, PGDP were used as solvents), 13 systems solvent-air (the same solvents as in the case of systems solvent-water) and system water-air on the basis of application of HYBOT descriptors (describing volume related term, hydrogen bond factors and electrostatics), structural similarity conception and physicochemical relationships between partitioning coefficients in systems solvent-water, solvent-air and water-air.

It is estimated that the used HYBOT descriptors completely and quantitatively describe a chemicals partitioning in studied systems. A coefficient values at molecular polarizability in equations for all systems are in rather narrow limits (0.270±0.050). That means that a contribution of volume-related term in chemicals partitioning for all studied systems solvent-water and solvent-air are approximately the same. The most important system for life on the earth water-air is exclusive one where a contribution of volume-related term is practically absent. On the basis of analysis of all regression equations obtained it is possible to consider the chemicals partitioning is aftereffect of competition of solute H-bond acceptor with solvent H-bond donor interactions and solute H-bond donor with solvent H-bond acceptor interactions.

GENERATION OF NOVEL EQUATION FOR THE LOGALITHMIC RATIO OF THE OCTANOL-WATER PARTITION FUNCTION: LOG P ESTIMATION USING PERCEPTRON TYPE NEURAL NETWORK - DEVELOPMENT OF COMPUTER AIDED MOLECULAR DESIGN SYSTEM: MolWorks

Sumie Tajima[1], Makoto Haraguchi[1], <u>Umpei Nagashima</u>[2]

[1]Bestsystems Inc., Tsukuba-Kenkyu-Shien Center C-B-2, Sengen 2-1-6, Tsukuba, Ibaraki 305-0047, Japan; [2]Research Institute for Computational Science, National Institute of Advanced Industrial Science and Technology, Umezono 1-1-1, Tsukuba, Ibaraki 305-8568, Japan; E-mail: u.nagashima@aist.go.jp.jp

We have developed a novel property estimation equation with the group contribution scheme for molecular properties: LogP, in the standard condition using a three layers perceptron type neural network and equipped it into MolWorks™ [1].

127 descriptor (groups) are newly defined as a set to realize more accurate prediction than usual methods. 426 data of LogP are selected for education of the neural network. 202 data were applied to evaluate the efficiency of the equation.

The correlation of observed and predicted LogP by this work is better than the values by several previous reports [2-4]. A novel equation for highly accurate LogP prediction has been obtained using NN. The novel NN equation requires only information 2D chemical formula not MO calculation. It is easy to apply new molecules. Prediction by the equation obtained by NN is better than that by multi regression analysis.

1. MolWorks: http://www.molworks.com
2. H. Chuman, A. Mori, and H. Tanaka, *Analytical Sciences*, **2002**, 18, 1015-1020.
3. A. Breindl, B. Beck, T. Clark, and R. C. Glen, *J. Mol. Model.*, **1997**, 3, 142-155.
4. Albert Leo , D. H. Hoekman, Corwin Hansch, "Exploring Qsar: Hydrophobic, Electronic, and Steric Constants", Acs Professional Reference Book, Amer Chemical Society, **1995**.

THE «MICROCOSM» INFORMATION TECHNOLOGY SYSTEM FOR BIOLOGICAL ACTIVITY PREDICTION OF ORGANIC COMPOUNDS: THE PREDICTION OF HEMORHEOLOGICAL ACTIVITY

Pavel M. Vassiliev, <u>Ludmila V. Naumenko</u>, Alexander A. Spasov

Volgograd State Medical University, 1, pl. Pavshikh Bortsov, Volgograd, 400131, Russia; E-mail: farm@vlpost.ru

The «Microcosm» information technology system for biological activity prediction of organic compounds is considerate. Its basic theoretical conception, principles of fragmental encoding of compound structure, methods and strategies of activity prediction is put; the packet of special computer programs is described. The examples of effective using of technology for searching of new substances with high hemorheological activity are given. IT «Microcosm» programs system, version 4.2 allows to carry out complex analysis of «structure - activity» relationships and to decide following practical tasks: to fulfill the prediction of existence of given biological activity type and its level for compounds of any chemical structure; to predict of activity of mixtures, consisting of several individual compounds, with consideration of its quantitative composition and components synergism; to predict of activity of complexes, that formed as a result of no-covalent intermolecular interactions of several compounds; to generate 2D-pharmacophoric hypothesis, to value its significance and to discover pharmacophores; to form pharmacophore-patterns of given biological activity type; to make conclusions about probable mechanisms of action of different biological activity compounds; to formulate suppositions about structure of receptors interacting with these compounds. Experimental testing of prediction results of hemorheological activity of 37 new derivatives of condensed nitrogen-contained heterocycles given following results to three series of investigations. The prediction precision of existence of hemorheological activity totaled 84 % and the prediction precision of high level of hemorheological activity totaled 65 %. 20 compounds with more activity than the pentoxiphilline standard drug was founded.

NEW MATHEMATICAL MODELS AND SOFTWARE TOOLS FOR COMPLEXITY AND SIMILARITY ANALYSIS OF MOLECULAR GRAPHS

Alexey A. Neznanov, Victor A. Kokhov, Sergey V. Tkachenko

Moscow Power Engineering Institute (Technical University), 14, Krasnokazarmennaya Str., Moscow, 111250, Russia; E-mail: neznanovaa@mpei.ru

The stratified system of mathematical models of structural complexity has been suggested. These models (universal g-model of graph and its submodels) cover wide range of graph properties. Implementation of building and researching of g-models based on the last achievements in the structural spectral analysis of graphs, initially developed by V.A. Kokhov. G-models underlie generalized substructure approach to graphs similarity analysis. G-model is a powerful structural invariant (bipartite graph with structural weights on vertices and edges) that characterized the mutual placement of fragments (selected set of structural descriptors) in graph topology. Current implementation supports unrestricted types of structural descriptors (chemist selects any descriptors which he is interested in) and extensive parameterization of models. Suggested models are used in the software package "Graph Model Workshop" (www.graphmodel.com) for implementing new methods of graphs similarity analysis, ordering and clustering, QSAR-analysis, multistage structural search in large graphs databases. The new subsystem "G-Studio" is made up. It has been used to obtain new results on large set of samples (more than 1000000 graphs have been processed). Practical using of the new methods and software in prediction properties of novel substances, computer drug design and toxicology is proposed and discussed.

NEW STRUCTURAL DESCRIPTORS OF MOLECULES ON THE BASIS OF SYMBIOSIS OF THE INFORMATIONAL FIELD MODEL AND SIMPLEX REPRESENTATION OF MOLECULAR STRUCTURE

Lyudmila N. Ognichenko, Victor E. Kuz'min, Anatoly G. Artemenko

A.V.Bogatsky Physico-Chemical Institute of the National Academy of Sciences of Ukraine, 86 Lustdorfskaya doroga, Odessa 65080, Ukraine; E-mail: ognichenko@mail.ru

In the given work the scheme of calculation of new structural parameters is offered on the basis of symbiosis of the informational field model and simplex presentation of molecular structure. Such structural characteristics can be used for the solution of the different QSAR tasks. In the framework of the information field model every atom depends on influencing of all other atoms of given molecule. Character of such influencing depends on atom properties, fixed in basis of construction of informational field and mutual position of atoms in molecule. Some peculiarities determined by atom nature and geometry of molecule on the whole for mutual influencing of atoms were found out for the different model systems. In the simplex representation of molecular structure a molecule is represented as the system of different simplex descriptors (tetratomic fragments with fixed composition, structure, chirality and symmetry). In the given approach it is offered to differentiate atoms in simplex on the basis of the informational field characteristics, which are calculated in points of the atom location. It is noteworthy that it is possible to use potentials of the informational fields, weighed by different atomic properties (charge, lipophilicity, refraction etc). The efficiency of method was demonstrated on the example of analysis of data sets of inhibitors for angiotenzin converting enzyme and acetylcholinesterase. The researched compounds are structurally homogeneous. It relieves "molecular alignment" procedure in the framework of lattice methods. QSAR tasks have been solved using PLS-method. Resulting PLS-models, received with the use of simplex-informational descriptors and also descriptors, generated in the followings QSAR approaches: COMFA; COMSIA; EVA; HQSAR; Cerius 2, were compared. The advantage of the developed by us method has been revealed by the comparison of statistical characteristics of QSAR models. For example, for our methods determination coefficient calculated in the cross-validation Q^2=0.72-0.88 and for the other methods: $Q^2 = 0.31$-0.72.

SECONDARY STRUCTURE OF SIGNAL PEPTIDES OF NS2 PROTEIN OF HCV AND ITS STABILITY

Igor A. Orshanskiy, Ksenia B. Tereshkina, Yegor V. Tourleigh, Konstantin V. Shaitan

Lomonosov Moscow state university, Biology faculty, Bioengineering department, 119992, Russia, Moscow, Leninskie Gory, 1-12; E-mail: ingarr@yandex.ru

NS2 protein of hepatitis C virus is a transmembrane signal peptide. It plays an important role in cleavage of NS2-3 protein to NS2 and NS3 proteins. NS3 protein is a serine protease, which cleaves all non-structural proteins form polyprotein-precusor. It was shown, that transmembrane localization of NS2 protein is a necessary for NS2-3 protein autocleavage activity. NS2 protein contains four transmembrane domains. Two of them are signal peptides. Research of signal peptides represents a priority questions in biomedicine. In our work we studied their secondary structure and its stability. As transmembrane domains signal peptides must contain mostly alpha-helical structure. To determine amount of such type of secondary structure we conducted molecular dynamics experiments at 300K. Results showed that both peptides are mostly alpha-helical, but first peptide consists from one helix, and second one - from two helixes, separated with one aminoacid residue. To determine stability of these two peptides we conducted molecular dynamics experiments at 700K. At this temperature peptides lost part of their secondary structure, and first peptide was divided into three helixes. Second peptide at 700K consisted of two helixes, but they were separated by several aminoacid. To compare thermostability of two peptides we counted the percentage of secondary structure at 700K. At the result first signal peptide seems to be more stable than second one. Comparing their thermostability and secondary structure of both peptides, amino acid replacement was proposed. It is significant, that directed amino acid replacement is important not only for understanding membrane interaction mechanism of the peptide and role of each amino acid but also in pharmacology. So results of this work are directed to create new generation of antiviral drugs.

THE GEOMETRICAL APPROACH TO FORECASTING EFFECTIVE NARCOTIC DEPENDENCE REMEDIES

Evgenija Osipova, Vladimir Kuz'min, Natalija Kuz'mina

The State Scientific Research Institute of Organic Chemistry and Technology, 23, Sh. Enthusiastov, Moscow, Russia; E-mail: arc.sioct@mail.magelan.ru

The general 3D-model opiate pharmacophore (OPh) is created on the basis of generalization of experimental data about a parity the agonistic and the antagonistic properties of stereoisomers of ligand's molecules from various structural classes for opiate receptors (OR). It represents the set of molecular areas in which three types of interactions "ligand-receptor" are realized: the agonistic interactions transforming ligand to pure agonist; the antagonistic interactions transforming ligand to the pure antagonist; the affine interactions influencing exclusively on ligand's affinity to OR.. The nature of the interactions realized in various areas OPh is established, and the set of geometrical parameters is determined unequivocally describing a relative positioning of molecular pharmacophore areas in space. It is shown, that the nature of antagonistic interactions defines selectivity of antagonistic ligand's interactions with various types OR. It is established, that typical "antagonistic" substitutes at N atom do not interact with OR. Their role is to reduce the "shielding" agonistic interactions in occurrence of antagonistic ligands properties. Developed 3D-model of OPh has been put in a basis of the computation method for a ligands antagonistic properties estimation. This method consists in an establishment of exact conformity between a spatial structure of a substance molecule and 3D-model opiate pharmacophore.

DEVELOPMENT OF THE INFORMATION SUBSYSTEM ON MATHEMATICAL MODELS IN TOXICOLOGY AND PHARMACOLOGY

Ruslan Ostapchuk, Vladimir Zatsepin, Vladimir Ivanchenko

All-Russian Institute of Scientific and Technical Information, 20, Usievich St., Moscow, 125190 Russia; SYSTECH R&D Center, 17/20 Kolokolnikov Per., office 16, Moscow, 107045, Russia; E-mail: systech@aha.ru

The Integral Chemical Informational Analytical Systems (IAS) developed in SYSTECH R&D Center includes Toxicology Subsystem, which contains multi-aspect documentary-factual information in chemical toxicology region. The development of new subsystem of IAS for aggregating and using of information on published and new created mathematical models for data treatment in toxicology, pharmacology and chemical ecology regions is here considered.

Models creating is big part of scientific investigations, so development of the data bases (DB) on mathematical models is very actual problem as additional resource for intensive using of knowledge and data contained in scientific and technical publications. The developed software supports all technological operations of computer-added processing of primary and secondary (DB, monographs, reviews etc.) information resources in printable and electronic forms: primary formalization of documents in text facts form (document fragments in aspects, including model description and abstracts/bibliography data and other aspects); indexation, extraction and formalization of factual data (tuned models in computer executed modules, statistical characteristics of models, learning and test data, substructure and physicochemical descriptors and/or programs for its calculation); loading factual and text data in DB; input, editing and import/export data in traditional data exchange formats and/or as report/publication. As applications of new IAS subsystem are presented popular models: bayesian, regression and comparative models (special cases of last models are interspecific and interapplicational models widely used in toxicological investigations).

DECISION TREES FOR CLASSIFICATION OF ENDOCRINE DISRUPTORS

Annick Panaye[1], James Devillers[2], Jean-Pierre Doucet[1], Nathalie Marchand-Geneste[3], Jean-Marc Porcher[4]

[1]ITODYS, Universite Paris 7 Denis Diderot, UMR7086 CNRS, 1 rue Guy de la Brosse, 75005 Paris, France; [2]CTIS, 3 chemin de la Graviere, 69140 Rilleux la Pape, France; [3]Universite de Bordeaux I, LPTC, ISM-UMR5255 CNRS, 351 Cours de la Liberation, 33405 Talence CEDEX, France; [4]INERIS, Parc Technologique ALATA, BP n°2, 60550 Verneuil en Halatte, France; E-mail: panaye@univ-paris-diderot.fr

With the current concern of limiting experimental assays, increased interest now focuses on *in silico* models able to predict toxicity of chemicals. Endocrine disruptors cover a large number of environmental and industrial chemicals which may affect the functions of natural hormones in humans and wildlife.

In this study a large set of about 200 chemicals covering a broad range of structural classes was considered in order to categorize their relative binding affinity (RBA) to the androgen receptor.

Classification of chemicals into three activity groups, with respect to their RBA value, was carried out in a cascade of recursive partitioning trees, from descriptors calculated from CODESSA software and encoding topological, geometrical and quantum chemical properties. The hydrophobicity parameter (log P), Balaban index, and indices relying on charge distribution (max. partial charge, nucleophilic index on oxygen atoms, charged surface area, etc.) appeared to play a major role in the chemical partitioning. Separation of strongly active compounds was rather straightforward. Similarly, about 90% of the inactive compounds were identified. More intricate was the separation of active compounds into subsets of moderate and weak binders, the task requiring a more complex tree.

SULFONAMIDES AS A SUBJECT TO STUDY MOLECULAR INTERACTIONS IN CRYSTALS AND SOLUTIONS: SUBLIMATION, SOLUBILITY, SOLVATION, DISTRIBUTION

G.L. Perlovich[1], N.N. Strakhova[1], V.P. Kazachenko[1], K.-J. Schaper[2], O.A. Raevsky[1]

[1]Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds, Russian Academy of Sciences, 142432, Chernogolovka, Russia; [2]Research Centre Borstel, Centre for Medicine and Biosciences, D-23845 Borstel, Germany; E-mail: germanper@yandex.ru

Sulfonamides are drugs extensively used for the treatment of certain infections caused by Gram-positive and Gram-negative microorganisms, some fungi, and certain protozoa. Although using of antibiotics has diminished the usefulness of sulfonamides, they still occupy a relatively small but important place in the therapeutic resources of physicians. It should be mentioned, that there are some attempts to correlate different physic-chemical characteristics of these compounds with chemotherapeutic activity: pK_a, protein binding, and electronic charge distribution. Unfortunately, the action of sulfonamides is complicated and can not be described in simple way. There is not enough information to propose suitable mechanisms for the transfer process of sulfonamides between immiscible liquid phases, and between aqueous media and biological membrane models, in order to explain the differences in the pharmacological power as a function of the molecular structure.

The correlations between the sublimation Gibbs energies and the melting points and between the sublimation enthalpies and the fusion enthalpies at 298 K have been derived. These dependencies give opportunity to predict the sublimation thermodynamic parameters on basis of fusion experiments only. The thermodynamic functions of solubility and solvation processes have been analyzed using temperature dependencies of solubility in water and n-octanol and sublimation characteristics of the compounds. The enthalpic term contributes a dominant part to the solvation Gibbs energy. Studying the transfer processes has been carried out by diagram method with analysis of the enthalpic and entropic terms. Distinguishing between enthalpy and entropy, as is possible through the present approach, leads to the insight that the mechanism is different for the different molecules (entropy- or enthalpy determined). Thus, in contrast to interpretation of Gibbs energy of transfer, being excessively used for pharmaceuticals in the form of the partition coefficient and logP, analysis of thermodynamic functions of the transfer process, as outlined in the present work, provides additional mechanistic information [1].

1. G.L. Perlovich, N.N. Strakhova, V.P. Kazachenko, T.V. Volkova, V.V. Tkachev, K.J. Schaper, O.A. Raevsky, *Int. J. Pharm.,* **2007**, 334, 115–124.

NEW APPROACH OF QUANTITATIVE INTERPRETATION OF DECISION TREE MODELS AND ITS APPLICATION FOR SOLUTION OF QSAR TASKS

Pavel G. Polishchuk, Victor E. Kuz'min, Anatoly G. Artemenko, Svetlana Yu. Makan, Sergey A. Andronati

O.V. Bogatsky Physico-Chemical Institute of the National Academy of Sciences of Ukraine, 86, Lustdorfskaya doroga, Odessa, 65080, Ukraine; E-mail: pavel_polishchuk@ukr.net

The decision trees method analysis is one of the most convenient and effective methods of analysis when an investigated property (activity) represented in the rank scale. However it has limited application for solution of QSAR tasks due to difficulties of quantitative interpretation of its models. A new approach to solve this problem was proposed. It is based on the procedure of the trend-vector method. Applying of this approach to the each node of the resulting decision trees model we can estimate the relative influence of each descriptor used in the model on an investigated property. Moreover, each value of descriptor relative influence has a corresponding range of descriptor value, inside which this influence is implemented. Proposed approach was tested in the solution of QSAR task for set of 359 ligands of serotonin 5-HT$_{1A}$ receptors. Simplex representation of molecular structure was used which can simplify of model interpretation. Program SPSS Answer Tree 3.0 (trial version) and C&RT (Classification and Regression Trees) algorithm were used for construction of the decision trees model. The proposed approach made it possible to calculate values of relative influences of all descriptor used in the model. Using of simplex representation of molecular structure allowed selecting molecular fragments which had positive and negative influence on affinity of investigated compounds to 5-HT$_{1A}$ receptors. Obtained information can be useful for design of new high-affinity ligands of 5-HT$_{1A}$ receptors.

TEXT MINING TOOLS IN ANALYSIS OF HIGH-THROUGHPUT DATA

Elena A. Ponomarenko, Andrey V. Lisitsa, Alexander I.Archakov

Institute of Biomedical Chemistry RAMS, Moscow, Russia; E-mail: pon@ibmh.msk.su

The analysis of high-throughput data is the main bottleneck of present-day proteomics. Scientists have to deal with a list of plenty protein names as a result of the experimental chemical compounds in vivo testing. It is rather embarrassing to assign these proteins to the molecular biological process. It is anticipate to allocate the groups of related proteins, which have similar intercellular localization, same metabolic pathway or related function. Such information can be extracted from the different sources, however comparison and studying of characteristics of thousands various proteins is impossible to perform manually.

For this reason, we developed the software, which can distinguish a number of groups of related proteins from high-throughput experiments. The system consists of the three modules: there is a part for text analysis, a part for comparison of the protein descriptors and visualization part. Input data is submitted as a list of protein names. Texts of relevant articles are downloaded for each protein name. Frequencies of the words' occurrence in the texts of relevant articles are used as descriptions of each protein. We used the assumption that the articles for the related proteins will be characterized by a set of terms, specific to every group with similar frequencies of words' occurrence. Processing of the data is realized with PCA method (Principal Component Analysis is using for multivariate data). It is supposed that the proteins which have got inside of one cluster are interconnected. The type of the interrelation between proteins inside of one group (cluster) is defined from the texts of attached articles. The system was tested on experimental data of cytochromes P450 induction after administration of phenobarbital or 3-methylcholantrene to mice.

ANTIINFLAMMATORY AND ANALGESIC ACTIVITY OF NEW DERIVATIVES OF 2-AMINOBENZOTHIAZOLE

Violetta L. Kovaleva[1], Athina Geronikaki[2], Oxana V. Proskurina[1], Elena V. Shilova[1], Vladimir V. Poroikov[3], Dmitry S. Blinov[4]

[1]National Research Center on Biologically Active Compounds, Kirov st. 23, Staraya Kupavna, 142450, Moscow region, Russia; [2]School of Pharmacy, Aristotelian University of Thessaloniki, University Campus, 54006 Thessaloniki, Greece; [3]Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, Pogodinskaya st. 10, Moscow, 119121, Russia; [4]Mordovian Ogarev State University, Bol'shevistskaya st. 68, Saransk, 43000, Russia; E-mail: vnc@pc-club.ru

Creating of new drugs with anti-inflammatory (AI) activity including COX-2 inhibitors is still actually because of numerous side effects of NSAID. We studied AI and analgesic properties of 10 new derivatives of 2-aminobenzothiazoles (ABT) on the models of acute exudative inflammation and analgesia using diclofenac, ketoprofen, meloxicam as references drugs. Spectrum of biological activity for tested substances was estimated with computer program PASS (http://www.ibmh.msk.su/PASS). Probability of presence of such effects and mechanisms as analgesic, AI, antiviral and COX inhibitor was very high especially for substances A1 and A3 for which Pa (probability to be revealed) were 0.607, 0.606, 0.512 and 0.581, respectively. AI activity studied in rat paw edema assay. Substances were dosed i.p. (5 mg/kg) or i.g. (10 mg/kg) in saline solution or in Twin 80. 1 h later, animals were injected with 1% carrageenan or Freund's adjuvant. Hind paw volumes were measured 4 h after carrageenan and 24 h after adjuvant injection using plethysmograph. Ulcerogenic activity of new substances was also estimated. Analgesic activity was determined in model of body twisting. Body twisting numbers were recorded within 15 min after 0.85% acetic acid i.p. injection (0.2 ml/mouse). Tested substances A1, A3, A4, A5 (10 mg/kg p.o.) inhibited carrageenan-induced paw edema by 38, 34, 30.5 and 50%, respectively. A5 showed the most considerable effect comparable with diclofenac and ketoprofen. Other substances (A6, A7, A9) at the same dose and route of administration exhibited moderate AI activity: inhibition of swelling was 18.8%, 18.5%, 21%, respectively. A8 was inactive. At a dose of 5 mg/kg i.p., novel substances showed marked effect. They reduced also adjuvant-induced paw edema. All tested substances showed 26.2%-42.7% analgesic activity at a dose of 50 mg/kg p.o. A5 again showed excellent potency comparable with to diclofenac (42.7% versus 48.4%). Tested substances can be arranged in AI and analgesic activity in the following: meloxicam > diclofenac = ketoprofen = A5 > A1 = A3 > A4 > A9 = A2 = A6 = A7 > A8. The most active substances are characterized by benzyliden group and we suppose this group to play important role in realization of AI and analgesic activity. These results indicate that novel derivatives of ABT are AI agents with analgesic properties.

PREDICTION OF PROTEIN-PROTEIN INTERACTIONS: PHYLOGENETIC PROFILES AND CLUSTER ANALYSIS

Mikhail A. Pyatnitskiy, Andrey V. Lisitsa, Alexander I. Archakov

Institute of Biomedical Chemistry of RAMS, Pogodinskaya str, 10, Moscow, Russia; E-mail: mpyat@ibmh.msk.su

The advent of whole-genome sequencing has led to computational methods that infer protein function and linkages. One of the most promising approach for prediction of protein-protein structural and functional interactions is studying of phylogenetic profiles. A phylogenetic profile of a protein is a binary vector, representing the presence or absence of homologs to that protein across a set of organisms. It was shown that proteins with similar patterns of co-occurrence across many organisms tend to participate in the same protein complex, biochemical pathway or have similar sub-cellular location. In the present work we explored the application of cluster analysis to phylogenetic profiling in order to improve performance of the method. We applied several standard techniques of cluster analysis including hierarchical clustering, kNN, PAM. We have also proposed to use ART1 clustering, which is based on neural networks and was intentionally designed to handle binary vectors. This algorithm also has an advantage of automatically determining required number of clusters. By comparing resulting predictions to expert annotations in KEGG database we obtained that the most accurate methods were Ward's hierarchical clustering and ART1. This results lead to improved version of phylogenetic profiling method and hence, more precise prediction of protein-protein interactions.

LIGAND-SPECIFIC SCORING FUNCTIONS: IMPROVED RANKING OF DOCKING SOLUTIONS

Timothy V. Pyrkov, Yuri A. Kosinsky, Roman G. Efremov

M.M. Shemyakin & Yu.A. Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Ul. Miklukho-Maklaya, 16/10, 117997 GSP, Moscow V-437, Russia; E-mail: pyrkov@nmr.ru

Molecular docking method has become an integral part of many biomolecular studies aimed at understanding the mechanism of enzyme functioning or drug discovery programmes. Nevertheless, significant challenges related to the problem of proper ranking of docking solutions still remain. While the search algorithms are almost always able to find correct conformation of a ligand in the binding site, the scoring methods often fail to discriminate such a conformation among many false variants. One way to treat this problem is to apply more precise scoring filters to re-rank docking solutions. Unlike general-purpose scoring functions we propose several new ligand-specific ranking criteria. The distinctive feature of the approach is that the weighting coefficients corresponding to different terms in the scoring function can vary depending on the properties and topology of the ligand molecule. In that way peculiarities of interaction between protein and different types of compounds are implicitly taken into account. New scoring functions were constructed including hydrogen bonds, hydrophobic and hydrophilic complementarity terms and embedding of a ligand. These scoring functions also discriminate ligands by the size of the molecule, the total hydrophobicity, and the number of peptide bonds for peptidic ligands. Using the training set of 60 protein – ligand complexes, the weighting coefficients were adjusted. Then the proposed method was tested on the results of docking obtained for additional 70 complexes. In both cases the success rate was 10% - 5% better as compared to the standard functions implemented in popular docking software.

KNOWLEDGE-BASED POTENTIALS ALLOW PLACEMENT OF WATER AND IONS IN MACROMOLECULAR INTERFACES

Sergei V. Rahmanov, Vsevolod J. Makeev

GosNIIGenetika, 1 Dorozhny proezd, 1, Moscow 117545, Russia; E-mail: vsevolod_makeev@mail.ru

A new method for calculation of knowledge-based atom contact potentials is proposed. It uses a stochastic scheme for simulating the non-interacting (reference) state for derivation of potentials, thus called Monte Carlo Reference State (MCRS). Due to the use of the MCRS, the resulting potentials are very detailed, continuous, and extend over a broader range of contact distances, including very close contacts. We obtained empirical atomic contact potentials for interaction of all types of protein and nucleic acid atoms with structure-bound water molecules and various ions, based on atom contact statistics in a training structure set comprising 2105 high resolution 3D structures. We tested the atomic hydration potentials (AHP) for prediction of hydration sites in protein structures, and have shown that using AHP it is possible to reproduce experimentally determined locations of individual bound water molecules in proteins and protein interfaces, with higher precision and several times lower over-prediction level, than by using other contemporary methods. This feature is important for modeling macromolecular ligand binding, since specific protein-protein and protein-DNA interaction is often mediated by ordered water molecules [2]. Predictions of ion binding in proteins using empirical potentials for calcium, zinc, and other ions also allow successful recognition of precise binding sites and ion specificities. Keeping in mind that the interaction with aqueous solvent commonly provides about 90% of the total structure stability energy [1], we applied protein atomic hydration potentials in fold recognition tests. For 27 out of 41 proteins of the Rosetta decoy set, the native fold was selected as the one having the lowest estimated solvation free energy among 1855, on the average, misfolded decoy structures. This performance is achieved without any consideration of the protein internal interaction, only by estimate of the structure solvation energy, and it is on par with modern protein folding potentials used for protein structure prediction.

1. X. Li, J. Liang, Knowledge-based energy functions for computational studies of proteins. In: Computational Methods for Protein Structure Prediction and Modeling. Edited by Y. Xu, D. Xu, J. Liang, Springer-Verlag, **2006**.
2. B. Jayaram, T. Jain, The role of water in protein-DNA recognition. *Annu Rev Biophys Biomol Struct,* **2004**, 33,343-361.

COMPUTATIONAL DRUG DESIGN BASED ON A NOVEL APPROACH TO LOCAL SIMILARITY OF PROTEIN BINDING SITES

Vasily E. Ramensky[1,2], Natalia I. Zaitseva[2], Anatoly R. Rubinov[2,3], Victor V. Zosimov[2]

[1]Engelhardt Institute of Molecular Biology of Russian Academy of Sciences, Vavilova 32, Moscow, 1199911, Russia; [2]Applied Acoustics Research Institute, 9 May St. 7a, Dubna-1, 141981, Russia; [3]The Institute for Information Transmission Problems of Russian Academy of Sciences, Bolshoi Karetny per. 19, Moscow, 127994, Russia; E-mail: ramensky@imb.ac.ru

We present a novel notion of binding site local similarity based on the analysis of complete protein environments of ligand fragments. Comparison of a query protein binding site (target) against the 3D structure of another protein (analog) in complex with a ligand enables ligand fragments from the analog complex to be transferred to positions in the target site, so that the complete protein environments of the fragment and its image are similar. The revealed environments are similarity regions and the fragments transferred to the target site are considered as binding patterns. The set of such binding patterns derived from a database of analog complexes forms a cloud-like structure (fragment cloud), which is a powerful tool for computational drug design. It has been shown on independent test sets that application of fragment clouds to self-docking and screening dramatically improves the results and enables reliable reproduction of experimental ligand optimization results.

CLASSIFICATION OF ORGANIC COMPOUNDS BY THEIR MODE OF TOXIC ACTION ON *DAPHNIA MAGNA* BY THREE DIFFERENT METHODS

A.N. Rasdolsky[1], V.A. Gerasimenko[1,4], D.A. Filimonov[2], V.V. Poroikov[2], E.J. Weber[3], O.A. Raevsky[1]

[1]Department of Computer-Aided Molecular Design, Institute of Physiologically Active Compounds, Russian Academy of Sciences, 142432, Chernogolovka, Moscow region, Russia; [2]Institute of Biomedical Chemistry of Russian Academy of Medical Sciences, Pogodinskaya Street 10, 119121, Moscow, Russia; [3]National Exposure Research Laboratory, U.S. Environmental Protection Agency, 30605, Athens, GA, USA; [4]InterBioScreen Ltd., Institutsky Prospect 7a, 142432, Chernogolovka, Moscow region, Russia, http://www.ibscreen.com; E-mail: rasd@ipac.ac.ru

The *Daphnia magna* is the conventional test aqueous organism for assessment of acute toxicity of chemical compounds owing to its high sensitivity to pollutants and to omnipresent prevalence in lakes and rivers. The toxic action of chemical compounds on live aqueous organisms is classified in accordance with a behavioral response of biological organisms to seven types of action: 1) base-line narcosis or narcosis I, 2) polar narcosis or narcosis II, 3) ester narcosis or narcosis III, 4) oxidative phosphorylation uncoupling, 5) respiratory inhibition, 6) electrophile / proelectrophile reactivity, 6) AChE inhibition, 7) CNS seizure responses. Computer approaches of *a priori* reference of new organic compounds to one of the above formulated classes have been subjects of this research. The training set was taken from publication of P. C. Ohe, *at al*., Chem*. Res. Toxicol.* 2005, 18, 536-555. Additionally, 24 chemicals of anticholinesterase activity were used. In total the training set was comprised of 323 compounds: 148 narcotics I, 50 narcotics II, 13 narcotics III, 8 compounds of oxidative phosphorylation uncoupling, 56 with respiratory inhibition MOA, 43 chemicals with anticholinesterase MOA and 5 of CNS. The following computer methods of MOA discrimination of organic compounds were used: 1) Method of structure similarity with the use of atom-centered spherical structural fragments (MOLDIVS software). 2) Bayesian probability method, with the use of MNA structure descriptors of 1st and 2nd levels and B-statistics (PASS software). 3) Nearest neighbor analysis (NNA) with the use of 2048-bit Daylight fingerprints (Daylight toolkit program). The MOA classification accuracy of compounds was: 83% for MOLDIVS, 90% for PASS, and 79 % for Daylight toolkit program.

COMFA/COMSIA AND 2D QSAR STUDIES OF DIARYLPYRIMIDINE HIV-1 REVERSE TRANSCRIPTASE INHIBITORS

Joseph Rebehmed, Florent Barbault, Catia Teixeira, Francois Maurel, BoTao Fan

ITODYS, Universite Paris Diderot – Paris 7 – CNRS UMR 7086 ; 1 rue Guy de la Brosse 75005 Paris, France; E-mail: joseph.rebehmed@paris7.jussieu.fr

Nowadays, reverse transcriptase (RT) inhibitors come out to be the first drug class with potent activity against HIV, inhibiting one the first stages in the viral life cycle [1]. NNRTIs (non-nucleoside reverse transcriptase inhibitors) inhibit the enzyme by binding to an allosteric site which is a lipophilic cavity situated at 10Å from the catalytic site [2]. This inhibitor family presents the advantage of high potency, low toxicity and excellent selectivity for HIV-1 RT. Among them, Diarylpyrimidine (DAPY) compounds represent the most evolved family of NNRTI. Unfortunately, these compounds induce viral resistance [3]. The aim of this study is to realize chemical substitution on the DAPY scaffold in order to obtain new inhibitors of which their activities against mutant strains will be tested afterwards.

In this contribution, 2D and 3D quantitative structure-activity relationship (QSAR) studies were applied on a set of 28 DAPY analogs to model and understand their HIV-1 RT inhibitory activities. Special cares were taken to build our set of molecules according to their bioactive conformations. This criterion is crucial to elaborate good QSAR models and available in our system through the 3D experimental structure of DAPY/RT complex. 2D-QSAR was performed using the heuristic method in CODESSA which had led to a linear model between the inhibitory activity and five descriptors. CoMFA and CoMSIA models were established using SYBYL. A systematic search was performed to elucidate the best predictive model. The better predictive ability of the CoMSIA model over the CoMFA model was assigned to the large contribution of hydrogen-bonding interactions to the inhibitory activity. The CoMSIA PLS contour surfaces were mapped to the NNIBP of the RT to study the protein-inhibitor interaction mechanism. These results will guide further structural modification and prediction of new HIV-1 RT inhibitors.

1. Y. Van Herrewege, *Antimicrob Agents Chemother*, **2004**, 48, (10), 3684-9.
2. P.A. Janssen, *J Med Chem*, **2005**, 48, (6), 1901-9.
3. M.A. Wainberg, *J Acquir Immune Defic Synd*, **2003**, 34 Suppl 1, S2-7.

AIDA: SOFTWARE FOR MODELING OF GLUCOSE-INSULIN INTERACTIONS IN PATIENTS WITH DIABETES MELLITUS TYPE 1

Irine Sarvilina, Anna Krishtopa, Yuri Maklyakov, Daria Gordienko

South Scientific Centre of Russian Academy of Sciences, 41, Chehov str., Rostov-on-Don, 344006, Russia; Rostov State Medical University, 29, Nahichevansky str., Rostov-on-Don, 344022; Russia; E-mail: isarvilina@mail.ru

Diabetes mellitus (DM) is a disease that has been known to exist for thousands of years and that afflicts 6% of the population. DM is a heterogeneous group of diseases, which is characterized high level of the glucose in blood in consequence of defects of insulin secretion, insulin signalizings or their combined defects. Multiple roles of broken regulations in the immune system and inflammatory reactions may play important role in development of DM. Major form of DM is type 1 DM (T1 DM). These patients frequently have very low or undetectable insulin levels, are dependent on exogenous insulin for survival, and are ketosis-prone. The software AIDA contains a simple model of glucose-insulin interactions in the human body. It is intended for simulating the effects on the 24 - hour blood glucose profile of changes in differents insulins for patient with diabetes mellitus type 1 (T1 DM). A clinical model of glucose-insulin interaction in T1 DM has been developed for patients. The model attempts to reflect the underlying pathophysiology of insulin action and absorption of glucose in quantitative terms such as insulin sensitivity, volume of glucose and insulin distribution and maximal rate of gastric emptying. The model's predictions allow a 24 hour simulation of blood glucose profiles for 39 patients with T1 DM to be generated. We have got predictive curves of 24-hours of glucose concentrations in patients with T1 DM, which took gene-engineering insulin – insulin soluble, insulin isophan,-and analogue of the insulin of the long action - insulin glargine. Thereby, computer program AIDA, which can provide advice as to what the next step in improving glycaemic control might be for a given case scenario. The modeling of the glucose – insulin interaction has allowed to voice hypothesis about mechanism of low efficiency of insulin in patients with TI DM.

COMPUTER-AIDED DESIGN OF POLYKETIDES WITH THE REQUIRED PROPERTIES

Anastasia Sergeyko[1], Alla Stepanchikova[1], Boris Sobolev[1], Sergey Zotchev[2], Dmitry Filimonov[1], Alexey Lagunin[1], Vladimir Poroikov[1]

[1]Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 10, Pogodinskaya str., Moscow, 119121, Russia; [2]Norwegian University of Science and Technology, Trondheim, Norway; E-mail: sergeikonp@mail.ru

Polyketides are secondary metabolites with diverse structures and biological activities, thus representing a rich source of potentially valuable pharmacological agents. Polyketides are assembled biosynthetically by the modular polyketide synthase enzymes. The modular polyketide synthase architecture and their mode of action theoretically allow for an enormous number of macrolides to be produced upon combinatorial manipulation with these enzymes. Engineering of all possible variants of a certain polyketide synthase system, isolation and testing of the resulting substances in the laboratory are therefore technically unfeasible. We propose an approach to rational design of new polyketides with the required biological activity spectra. It is based on *in silico* generation of polyketides' structures in correspondence with the appropriate biosynthetic pathways. These virtual libraries can be further analyzed using computer-aided prediction of biological activities, physicochemical properties, drug-likeness, etc. Such analysis helps to select the most relevant compounds with respect to pharmacodynamic and pharmacokinetic properties. We developed BioGenPharm software for generation of polyketides' combinatorial libraries, prediction of activity spectra for generated structures and selection of molecules with the required properties on the basis of user's defined input parameters and selection criteria. For prediction of polyketides' activity spectra we used PASS algorithm (http://www.ibmc.msk.ru/PASS). Validation of PASS prediction ability for polyketides was performed vs. the evaluation set containing 242 natural macrolides from Dictionary of Natural Products. The mean prediction accuracy was 75,6%. To illustrate use of the described method we generated virtual library of 3072 erythromycin analogues and selected 17 substances for which the probability of hepatotoxic action is low.

MODELING CATALYTIC MECHANISMS OF HYDROLYSIS OF NUCLEOSIDE TRIPHOSPHATES (NTP)

Maria Shadrina, Bella Grigorenko, Alexander Nemukhin

Laboratory of Chemical Cybernetics, Department of Chemistry, M.V. Lomonosov Moscow State University, 1/3, Leninskie Gory, Moscow, 119992, Russia; E-mail: My-boxic@yandex.ru

Our work is devoted to the studies of nucleoside triphosphates (NTP) hydrolysis reactions in different biochemical systems. Hydrolysis reactions of NTP are very important for the cell life because small changes in the course of these reactions result in a critical damage of the cell such as a cancerous growth. Detailed mechanisms of these reactions are still not understood. We use an approach to the hybrid quantum mechanical and molecular mechanical (QM/MM) theory based on the effective fragment potential (EFP) technique for modeling properties and reactivity of large molecular systems of biochemical significance. Our computer program is based on the modified version of the PC GAMESS quantum chemistry package (A. A. Granovsky) and the TINKER molecular modeling package (J. Ponder). This method allows us to model mechanisms in biochemical systems taking into account protein environment. The entire system under consideration is devided into QM and MM parts. QM part includes atoms participating in the reaction. Energies and forces are computed by the quantum equations. MM part includes surrounding protein chains. Energies and forces are computed by classical equations. In the present work, we study the mechanisms of nucleoside triphosphates hydrolysis, particularly, guanosine triphosphate (GTP) hydrolysis in G-proteins, by QM/MM method. X-ray data for the enzyme – substrate complexes are used as initial coordinates for the QM/MM calculations. As a result, we construct and analyze energy profiles for the reactions $E \bullet (NTP + H_2O) \rightarrow E \bullet (NDP + P_i)$. Currently, we have found mechanisms of guanosine triphosphate (GTP) hydrolysis in Ras, Ras-GAP, elongation factor EF-Tu, adenosine triphosphate hydrolysis in myosin, methyl triphosphate hydrolysis in water. We have shown these reactions have dissociative mechanisms.

EXPRESSION ANALYSIS OF NF-KB-REGULATED GENES IN BREAST CANCER. META-ANALYSIS OF MICROARRAY DATA

Ruslan Sharipov[1,2,3], Yuriy Kondrakhin[2,3], Fedor Kolpakov[2,3]

[1]Institute of Cytology and Genetics SB RAS, 10, Lavrentyev aven., Novosibirsk, 630090, Russia; [2]Institute of Systems Biology, Novosibirsk, Russia; [3]Design Technological Institute of Digital Techniques SB RAS, Novosibirsk, Russia; E-mail: shrus79@gmail.com

Increased activity of transcription factor NF-κB was shown to play important part in human cancer development, besides other pathologies like autoimmune diseases, inflammation, and various viral infections. Blocking apoptosis NF-κB induces cell survival and proliferation. Elevated levels of NF-κB were revealed in many types of cancer (including breast cancer) and were often associated with phenomenon of chemotherapy resistance. In such way, NF-κB pathway is the subject of inquiry and targeting with various types of inhibitors. Using IDURO method, which was developed by our group for meta-analysis of five sets of microarray data, we performed investigation of genes up- or down-regulated in breast cancer focusing attention on NF-κB downstream targets. Totally we found in microarrays 311 such genes: expression of 25 was greatly changed (11 – up-regulated, 14 – down-regulated), of 107 was significantly changed (64 – up-regulated, 43 – down-regulated), and of 179 was not changed. Expression of NF-κB subunits was not changed. Analysis of revealed genes using EXPLAIN system and Proteome BKL Disease™ database (BIOBASE GmbH) and Gene Ontology demonstrated that they comprise 12,53% of all diagnostic markers, 19,94% - therapeutic targets, 16,30% - associated with molecular mechanisms of breast cancer, and 22,73% - diagnostic markers of ductal subtype of breast cancer. Investigated genes are associated with development of 80 different types of tumors and 170 other human diseases. The strongest association was observed for breast cancer. These data emphasize importance of NF-κB pathway for drug targeting. All obtained data are available in the Cyclonet (http://cyclonet.biouml.org) and the BMOND (http://bmond.biouml.org) databases. The first one contains gene tables and the second - diagrams created using BioUML technology (http://www.biouml.org) describing formally networks of genes with greatly changed expression (e.g., FN1, GATA3, GADD45B, ERBB2, FOS, EGFR, MYC, PTGS2, SOD2, etc). These new data significantly enrich NF-κB section of our databases formed in previous projects. All collected data about NF-κB and related pathways will be used in computational pharmacology for prediction of new targets to design more effective and safe anti-cancer drugs.

COMPUTER MODELING OF INTERACTIONS BETWEEN HIV-1 INTEGRASE AND STYRYLQUINOLINES

Dmitry Shcherbinin, Alexander V. Veselovsky

Institute of Biomedical Chemistry of RAMS, 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: amid_dima@list.ru

HIV-1 integrase is a very attractive target for new drugs against AIDS because this enzyme controls one of the most important processes in virus lifecycle and has not equivalents in human cells. Nowadays only two compounds are in clinical trial. So, searching for new nontoxic inhibitors of HIV-1 integrase is an issue of the day. In our research behavior of core domain in complexes with its inhibitors and pure domain of integrase has been studied using molecular modeling techniques. Analyzing results of molecular dynamics it was found out that in complexes with ligands active site was fixed in "open" conformation. This conformation was used for further analysis of interaction of styrylquinoline derivatives with integrase. Thirty-two compounds were used as a training set. These compounds were docked to the active site of integrase and obtained positions of ligands were used in the further 3D-QSAR with CoMFA and CoMSIA analysis. 3D-QSAR models were designed using steric, electrostatic and H-bonds fields for CoMFA analysis and steric, electrostatic, hydrophobic and donor-acceptor for CoMSIA. Although designed models had significant statistical characteristics, the prediction of inhibitory activity of compounds from test sets was inappropriate. It was suggested, that the main reason of such unsuccessful predictions can be the difference between experimental conditions that had been used in different laboratories.

CISOC-CHIMS: CHEMICAL INFORMATION MANAGEMENT SYSTEM

Tianxiang Shen, Feng Li, Jianhua Yao

Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, China 354, Fenglin Road, Shanghai, 200032, China; E-mail: yaojh@mail.sioc.ac.cn

Up to now, there is a large of chemical information because of development of chemistry and corresponding specialties. In order to make good use of the information efficiently, it is necessary to manage it by a chemical information management system. ISIS_BASE is a good and well known product of Elsevier MDL [1]. However, it has a limitation in treatment of Chinese and expression of multi-step reactions in a scheme. Here, we present CISOC-ChIMS, Chemical Information Management System developed by our group, authorized by National Copyright Administration of China in 2005.

CISOC-ChIMS can be used to browser, edit and search information about chemical structure and text in the database. The framework of the system is showed in Figure 1. Applications' examples of CISOC-ChIMS are showed in Figure 2.
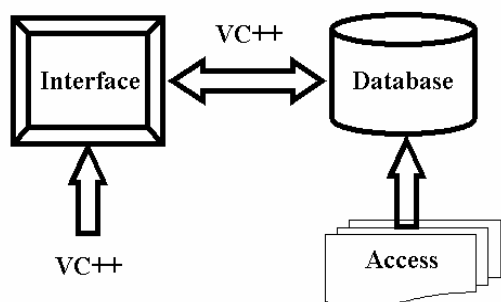
Figure 1. Framework of CISOC-ChIMS
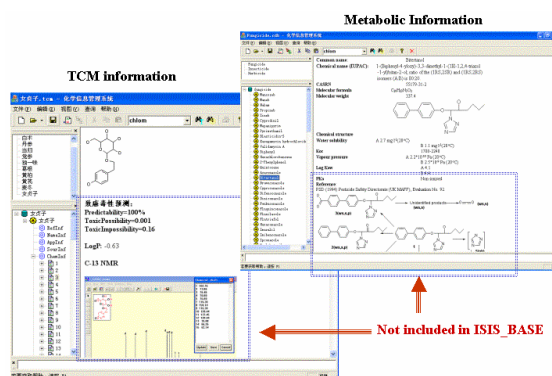
Figure 2. Applications' examples

1. http://www.mdli.com
2. Registration No. 040003, **2005**, National Copyright Administration of the People's Republic of China

FAST TOOLS FOR CALCULATION OF ATOMIC CHARGES WELL SUITED FOR DRUG DESIGN

<u>Dmitry Shulga</u>[1], Alexander Oliferenko[2], Sergey Pisarev[1], Vladimir Palyulin[1,2], Nikolay Zefirov[1,2]

[1]Institute of Physiologically Active Compounds of Russian Academy of Sciences, 1, Severny Proezd, Chernogolovka, 142432, Russia; [2]Department of Chemistry, Moscow State University, Leninskie Gory, Moscow, 119992, Russia; E-mail: shulga@qsar.chem.msu.ru

Partial atomic charges have been extensively used in both directly characterizing electrostatic interactions within molecular mechanics (MM) force fields (FF) and estimating various electrostatics related properties by means of scoring functions and descriptors for QSAR/QSPR studies. In many respects the charges that best reproduce *ab initio* molecular electrostatic potential (MEP) are most advantageous, except the high computational burden they impose due to quantum chemical (QC) calculations. On the basis of our previous efforts we 1) propose new fast empirical topological charging schemes and 2) optimize their parameters to simultaneously reproduce RHF/6-31G*//6-31G* MEP around the structures from a diverse training set of 235 organic molecules containing the functional groups found in stable drug-like structures. The achieved quality in MEP reproduction by our models exceeds that by MMFF94 bond charge increment (BCI) charges although significantly less parameters are used, and approaches the MEP quality based on AM1-BCC charges while not requiring *semi empirical* AM1 optimization, which is needed to produce the latter. The choice of the level of QC theory and the MEP reproduction as a criterion makes the charges computed by our schemes directly suitable for applied molecular dynamics (MD), Monte Carlo, and MM energy calculations using classical force fields, since the analogous premises charges had been employed to parameterize those force fields (e.g. RESP-charges in AMBER). In the realm of QSAR/QSPR research the proposed charge schemes should also perform well as the MEP-derived descriptors have been successfully used in different studies reported in literature. The theoretical background behind the charging schemes, optimization protocol as well as the results obtained will be presented.

COMPUTER DESIGN OF LOW-MOLECULAR COMPOUNDS OF A PEPTIDE KIND – ACTIVATORS OF DIRECTIONAL DIFFERENTIATION OF STEM CELLS

Irina V. Shutova, Olga V. Mel'nik, Vladimir P. Golubovich

The Institute of Bioorganic Chemistry of the National Academy of Sciences of Belarus, Kuprevich str., 5/2, Minsk, 220141, Republic of Belarus; E-mail: shutova2001@mail.ru

The purpose of research is the computer design of low-molecular compounds of a peptide kind – functional centers of human granulocyte colony-stimulating factor (G-CSF) which participate in binding receptors on the surface of hemopoietic stem cells. In order to realize the above aim, a program complex designed by the author was used to provide computer simulation of a protein/receptor interaction and design of low-molecular compounds responsible for biological function of protein. Input data for the program complex were presented by 3D-structures of proteins and protein/receptor complexes imported from the Protein Data Bank (http://www.rcsb.org). In the Protein Data Bank database a search was performed for 3D-structures of granulocyte colony-stimulating factor and complexes, G-CSF/receptor, as a result of which search 10 structures were retrieved. Of interest for computer simulation are granulocyte colony-stimulating factor structure with PDB-bank identifiers 1RHG and the complex G-CSF/receptor with PDB-bank identifier 2D9Q. The above computer simulation allowed to identify G-CSF protein residues Glu19, Lys23 and Asp109, which participate in G-CSF-receptor interaction and, presumably, form the functional center of G-CSF protein. Based on the obtained results, a prognosis was given for olygopeptides which should stimulate mitotic division of stem cells, as well as shaping by hemopoiesis precursor cells of hemopoietic sprout colonies; synthesis and biological tests of the offered frames was carried out.

DATABASE OF ENVIRONMENT POLLUTANTS FOR HEALTH RISK ASSESSMENT AND PROGNOSIS OF THE NEW CHEMICAL SUBSTANCE UNKNOWN CHARACTERISTICS

Sergey M. Novikov, Natalya S. Skvortsova

A.N. Sysin Research Institute of Human Ecology and Environmental Health RAMS, Pogodinskaja str., 10/15, Moscow, RF 119992; E-mail: skvnata@mail.ru

Development of the modern databases that support performance of informational, modeling and computational tasks within health risk assessment as well as working out the new methods of prediction the new chemical substance unknown characteristics are important tasks in improving the methodology of health risk assessment. The developed database contains information about 12578 chemical substances and compounds. The information include the following characteristics of a chemical substance: chemical structure and its code; identification characteristics, ecologically significant physicochemical properties, indices of specific impact, values of acute, chronic and sub-chronic toxicity for different impact paths, national and international hygienic standards, indices of biological toxicity for different biosystems, parameters of risk, toxicity and carcinogenicity estimates using classifications of IARC, U.S.EPA, NTP, ACGIH and the Russian classification of chemical carcinogens, generalized estimates of mutagenicity, embryotoxicity, sensitizing effect, teratogenic and gonadotropic effects. The list of affected organs and systems for each chemical substance is included in the database as well. It contains practically all published carcinogen slope factors and reference doses and concentrations for acute, chronic and sub-chronic effects. The substance identification in the database can be done using CAS, RTECS or the state registration numbers, chemical compound name, its synonym or commercial name, emission code, gross formula, chemical class, outward appearance, chemical structure. The substance chemical structure is stored as a SMILES linear record or as a MOL file in a chemical library.

The prediction of various ecologically significant physicochemical properties is performed using specialized forecasting computer modules or EpiSuite package recommended by US EPA. The system provides unit conversion. In addition, the database contains information that enables prediction using chemical structure and substructural codes, physicochemical properties (partition coefficient air/water, octanol or oil/water and others).

INTEGRATED SYSTEM FOR PREDICTION OF THE BIOLOGICAL ACTIVITY, TOXICITY AND PHYSICOCHEMICAL PROPERTIES

Vladlen S. Skvortsov[1,2], Alexej A. Lagunin[2], Dmitrij A. Filimonov[2], Oleg A. Raevsky[1]

[1]Institute of Physiologically Active Compounds RAS, Chernogolovka, Russia; [2]Institute of Biomedical Chemistry RAMS, Moscow, Russia

In this communication we present the new computer system that is uniting a set of our early programs. It includes the features of such programs as HYBOT+/HYBOT 3D, CRATOX, PRETOX, PASS, MedChemDesk, etc. The main purpose of our system is the prediction for wide variety of organic compounds the biological activity spectrum, the side effects, including mechanisms of toxicity, acute toxicity, and a set of physicochemical properties, which are responsible for distribution in environment and living organisms. The physicochemical properties include lipophilicity, solubility, volatility, interstitial absorption etc. The predictions are providing high accuracy in most cases. The access to the program server is possible through TCP/IP network with using standard set of service queries. This approach allows easily adapting an access to server of different applications. In present time, there is a set of already created JAVA client programs having the possibilities to be executed as completely independently and within WWW sites. The integration of system access functions library in Konstanz Information Miner is an example of using capabilities of our program. The results of calculations can be use for solving diverse problems: molecular modeling, QSAR, ecologies, risk assessment, effective experiment planning.

ONLINE RESOURCE FOR THEORETICAL STUDY OF HYDRATION OF BIOPOLYMERS

Egor V. Sobolev[1], Oleg V. Sobolev[2], Dmitry A. Tikhonov[1]

[1]Institute of Mathematical Problems of Biology RAS, 4 Institutskaya str., Pushchino, Moscow Reg. 142290, Russia; [2]Izhevsk State Technical University, 7 Studencheskaya str., Izhevsk, 426069, Russia; E-mail: egor@impb.psn.ru

An online resource has been developed for theoretical study of hydration of biopolymers by the method of the integral equations theory of liquids in the RISM (Reference Interaction Site Model) approximation. The RISM method is used to take into account the influence of the solvent and requires less computational effort than the Monte-Carlo and molecular dynamics explicit solvation methods do. At the same time it allows a more thorough and rigorous consideration as compared to implicit solvation models such as the Poisson-Boltzmann reaction field method or the generalized Born theory. The software developed includes all steps in studying a biopolymer with a given spatial structure and force field. It prepares the input data and carries out the RISM calculation yielding the atom-atom correlation functions of the biopolymer with water molecula. From these functions the algorithm finds atomic partial contributions to the hydration free energy using various free energy expressions of the integral equation theory. The calculation results are automatically recorded in a database and become available on the website as tables of partial thermodynamical quantities. In addition, the site displays an interactive 3D model of a molecule the atoms of which can be painted in different colors in accordance with their partial contributions to the thermodynamical quantity chosen by the user. The user can choose atoms on this molecule and their correlation functions will be displayed. The molecular model on the site is visualized by a Java-applet "jmol". Java-script was used to implement the interaction with the user and to draw the correlation functions. This approach has enabled us to create a cross-platform resource and reduce the amount of data to be transmitted. The resource enables researchers to assess the potentialities of the RISM method since they can carry out the calculations by themselves and then exchange the results. Besides, it will assist in accumulation of data on the behavior of biologically important molecules in water environment and stimulate the search for additional empirical regularities which may enhance the accuracy of estimates of the free energy in the framework of the method.

APPLICATION OF SET-ASSOCIATION APPROACH TO SEARCHING FOR SINGLE NUCLEOTIDE POLYMORPHISMS OF ″ENVIRONMENTAL GENES″ RESPONSIBLE FOR COMPLEX DISEASE SUSCEPTIBILITY

Mariya A. Solodilova, Alexey V. Polonikov, Vladimir P. Ivanov

Department of Medical Biology, Genetics and Ecology, Kursk State Medical University, 3, Karl Marx Str., Kursk, 305041, Russia; E-mail: polonikov@rambler.ru

Polymorphisms in genes of xenobiotic-metabolizing enzymes (XME) also known as ″environmental genes″ are largely responsible for interindividual differences in ability to activate and detoxify chemical agents and therefore may influence individual disease susceptibility. In complex diseases, multiple disease loci presumably interact to produce the pathologic phenotype, and it has been difficult to map susceptibility loci by conventional locus-by-locus methods. Hoh et al. (2001) developed a set association approach (SAA) that captures the simultaneous effects of multiple gene loci and thereby achieve a more global view of gene action and interaction than is possible by traditional gene-by-gene analysis. In our pilot study a SAA was applied to searching for key single nucleotide polymorphisms (SNPs) of ″environmental genes″ responsible for susceptibility to complex diseases.. A total of 1018 unrelated Russian subjects were recruited in this study, including 295 healthy controls, 216 asthmatics, 203 hypertensives and 304 patients with peptic ulcer disease. The blood samples were genotyped for 28 SNPs of 18 XME genes and 29 SNPs of 19 known candidate genes of the diseases by polymerase chain reaction followed by restriction fragment length polymorphism analysis in all study subjects. We used the SUMSTAT computer program (http://linkage.rockefeller.edu/ott/sumstat.html) to calculate a sum of single-marker statistics, which results in a single genome-wide test statistic with high power (Hoh J. and Ott J., 2004). Genome-wide significance of this statistic is calculated by permutation tests (10 000 permutations). We have found high-order gender-related patterns of interactions between various sets of SNPs of ″environmental genes″ and other genes that may demonstrate either a genetic heterogeneity of complex human diseases studied and an involvement of specific toxicogenetic mechanisms into their etiology.

INHIBITION OF THE HILL REACTION BY ORGANOPHOSPHORUS INSECTICIDES AND SIMPLE ORGANIC MOLECULES: A QSAR STUDY

Milan Soskic

Faculty of Agriculture, University of Zagreb, Zagreb, Croatia; E-mail: msoskic@agr.hr

In an endeavor to accomplish optimal weed control, with a minimum of ecological damage, numerous QSAR studies have been performed for herbicides inhibiting photosynthetic electron transport in chloroplasts (the Hill reaction). There is no such study dealing with plant-protecting agents (e.g. insecticides) which are not intended to target the plant itself. However, in case of inadequate application, even these compounds can exert phytotoxicity. The main objective of this work was to explore in quantitative terms, the interactions of some nonspecific inhibitors of the Hill reaction with isolated chloroplasts. The study encompassed organophosphorus insecticides (which are widely used in agriculture), chlorinated benzenes (industrial pollutants), n-aliphatic acids (which accumulate in the plants exposed to stress) and n-aliphatic alcohols. Highly significant correlations (as a rule, r > 0.95) between pI_{20} (the negative logarithms of the concentrations causing 20% inhibition of the Hill reaction) and log P values (the logarithms of the partition coefficients in the n-octanol/water system) were established for each of the above mentioned groups of compounds. From the obtained regression models, it may be concluded that all four analyzed groups of compounds inhibit the Hill reaction by a similar, nonspecific mechanism. On the other hand, their inhibitory potencies were somewhat different and increased in the following order: aliphatic alcohols < aliphatic acids < chlorinated benzenes < organophosphorus insecticides. A model comprising all analyzed compounds could be constructed by introducing indicator variables (I=1 for a particular group of compounds and zero for the others). We also generated, for the more active compounds, a combined model based on pI_{50} values (negative logarithms of the concentrations which inhibit the Hill reaction by 50%) which facilitated comparisons with published QSAR models for specific inhibitors of the Hill reaction.

ELECTRON AND SPATIAL STRUCTURE PECULIARITIES OF MACROHETEROCYCLIC COMPOUDS

Mikhail Stryapan, Mikhail Islyaikin

Ivanovo State University of Chemistry and Technology, 7, Engels St., Ivanovo, 153000, Russia; E-mail: islyaikin@isuct.ru

Macroheterocyclic compounds (Mc) – structural analogs of porphyrin are the objects of great interest as substances with potential biological activity. Geometric characteristics and charges distribution can be used for studying the activity – structure relationships. The DFT method with B3LYP electronic correlation at 6-31G(d,p) basis set (Gaussian03) was applied to study the electron and spatial structure peculiarities of Mc, namely triazoleporphyrazine **1**, hemiporphyrazine **2**, dicarbohemiporphyrazine **3** and trithiadiazoletriizoindolemacrocycle **4** (table 1).

Table 1

1

2 X = N
3 X = CH

4
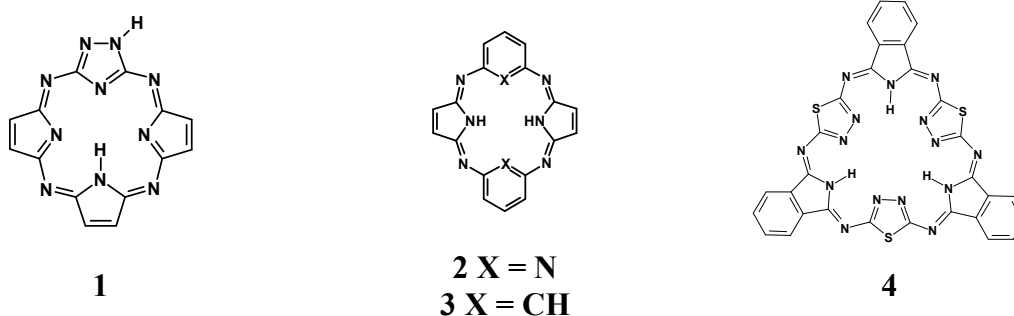
The aromaticity of various contours of conjugation that can be highlight in the molecules of **1-4** was evaluated by HOMA (Harmonic Oscillator Model of Aromaticity) and NICS (Nucleus Independent Chemical Shifts) criteria. Tautomerism and probable protonated forms of triazoleporphyrazine were studied. It was shown that the protonation of **1** gives an efficient influence on aromaticity. The positive values of NICS criteria that were found at the center of **2** and **3** correspond to the nonaromatic molecules.

"PROPELLER" CHIRALITY AND SPONTANEOUS RESOLUTION OF TRISTYRYLPHOSPHINE CHALCOGENIDES INTO ENANTIOMERS

Boris G. Sukhov[1], Aidar T. Gubaidullin[2], Igor A. Litvinov[2], Svetlana F. Malysheva[1], Alexander V. Vashchenko[1], Vladimir I. Smirnov[1], Yulia A. Grigor'eva[1], Nina K. Gusarova[1], Boris A. Trofimov[1]

[1]A.E. Favorsky Institute of Chemistry SB RAS, Favorsky St., 1, Irkutsk, 664033, Russia; [2]Institute of Organic and Physical Chemistry of RAS, Kazan, Russian Federation. 420008, Arbuzov str. ,8; E-mail: aidar@iopc.knc.ru

Chiral compounds are of great fundamental and practical importance. The crystallization of C_3-symmetric compounds from solution often affords enatiomerically pure single crystals.

The present work deals with the detail investigations of spatial structure of tris(*E*-styryl)phosphine oxide (**a**), tris(*E*-styryl)phosphine sulfide (**b**) and tris(*Z*-styryl)phosphine oxide (**c**).

Quantum chemical calculations [B3LYP, D95++(d,p)] have shown that phosphine oxides **a**, **c** possess "propeller" C_3-chirality, that breaks down for phosphine sulfide **b**. In the latter one of the styryl radicals turns out to be spatially extended and, hence the molecule looses all elements of the symmetry (Fig.).
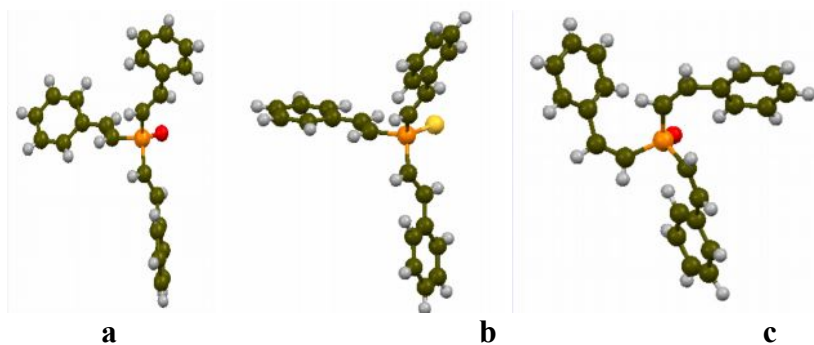
Fig. Geometry of the phosphine chalcogenides **a-c** in crystals.

Such symmetry of tris(*E*-styryl)phosphine sulfide (**b**) is caused by the presence of bulky lone pair at the sulfur atom, which effects on abnormal position of one of the radicals.

The X-ray diffraction analysis has completely proved the calculations data. It has been found that tris(*E*-styryl)phosphine oxide (**a**) crystallize as conglomerate in chiral spatial group R3, it means that crystal is built up of only molecules of one enantiomer, and hence we have spontaneous resolution of the racemate.

Thus, C_3-chiral tris(styryl)phosphine chalcogenides **a-c**, prepared from elemental phosphorus, [1] are new promising materials for special asymmetric syntheses and applications.

1. B.A. Trofimov, T. N. Rakhmatulina, N.K. Gusarova, S. F. Malysheva, *Usp. Khim.* **1991**, *60*, 2619.

WEB-ORIENTED SYSTEM KEENBASE FOR NEW DRUGS DESIGN

<u>Aleksey V. Sulimov</u>[1], Vladimir B. Sulimov[1], Alexey N. Romanov[1], Fedor V. Grigoriev[1], Olga A. Kondakova[1], Peter A. Bryzgalov[1], Denis A. Ostapenko[2]

[1]Reseach Computer Center at M.V. Lomonosov Moscow State University, 119992, Russia, Moscow, Leninskie Gori 1-4; [2]Faculty of Computational Mathematics and Cybernetics at M.V. Lomonosov Moscow State University, 119992, Russia, Moscow, Leninskie Gory, MSU, 2-nd educational building; E-mail: sulimovv@mail.ru

New inhibitors search for a given target-protein is the initial stage of a new drug development. Fast and effective decision of this problem defines minimization of expenses at all following stages and the overall drug development duration. Computer molecular modeling helps to increase new inhibitors elaboration effectiveness. System Keenbase has been developed for this purpose on the base of original programs for docking, scoring, atom typification, filtering of undesirable molecular structures, molecular visualization, and auxiliary programs conducting interactive user access to the system resources and computing nodes. The docking program SOL and metacomputing system X-Com play the central role in Keenbase performance. SOL docks ligands into the target-protein active site using genetic algorithm. The docking procedure takes into account the electrostatic, Van-der-Waals interaction of ligand with proteins, along with protein, ligand and protein-ligand complex desolvation. The docking process takes into account possible flexibility of ligand and calculates the corresponding rotamers energies. Program SOL can perform the correct docking of ligands, having up to 12 inner rotational degrees of freedom with the accuracy near 1.5 E. X-Com governs access to computing resources through Internet. Keenbase contains some target-proteins prepared for the ligand docking. User can upload his own target-proteins and ligands into the system, dock latter into the active site of a given target-protein, and rank them in respect with their binding energy. Validation (www.keenbase.ru) shows that Keenbase docking and scoring performance is comparable and better than ones of AutoDock, and Keenbase virtual screening facilities are much better due to Internet access to the system and to the computing nodes. The Keenbase programs have been used for new thrombin inhibitors design. In collaboration with Hematology Scientific Center of Russian Academy of Medicinal Sciences (Laboratory of F.I.Ataullakhanov) new class of nanomolar synthetic low molecular thrombin inhibitors having high anticoagulation activity has been discovered.

SYSTEMATIC CONFORMATIONAL STUDIES OF NEW POTENTIAL HIV-1 ENTRY INHIBITORS TARGETING CD4-GP120

Catia Teixeira[1], Florent Barbault[1], Joseph Rebhemed[1], Karen Urgin[2], Nawal Serradji[2], Francoise Heymans[2], Francois Maurel[1]

[1]ITODYS, Universite Paris Diderot – Paris 7, CNRS UMR 7086; 1 rue Guy de la Brosse 75005 Paris, France; [2]Pharmacochimie moleculaire et systemes membranaires, Universite Paris Diderot – Paris 7, EA 2381 ; 2 place Jussieu 75005 Paris, France; E-mail: catia.teixeira@paris7.jussieu.fr

Attachment of the human immunodeficiency virus (HIV-1) to the cell surface is the first step of the virus cycle. It is accepted that the fusion of the virus membrane with that of the target cell involve the binding of the viral glycoprotein gp120 to the cell receptor CD4. This recognition process induces several conformational changes that lead to the viral penetration in the cell [1,2]. Therefore, development of entry inhibitors represents an important avenue of drug therapy.

Several years ago some of us discovered a new series of entry inhibitors. Among the number of compounds synthesized in this series, PMS601 (see Figure 1) was found to exhibit micromolar invivo HIV inhibition activities. They demonstrated that the carbamate group is a key point in activity. It was supposed that its role relies on the virus entry process and this hypothesis was strengthened with the discovery of a high potency antiviral compound, BMS_378806, who interacts specifically to gp120 [3]. The overall goal of this project is to design new original compounds that take benefit from both previous compounds.
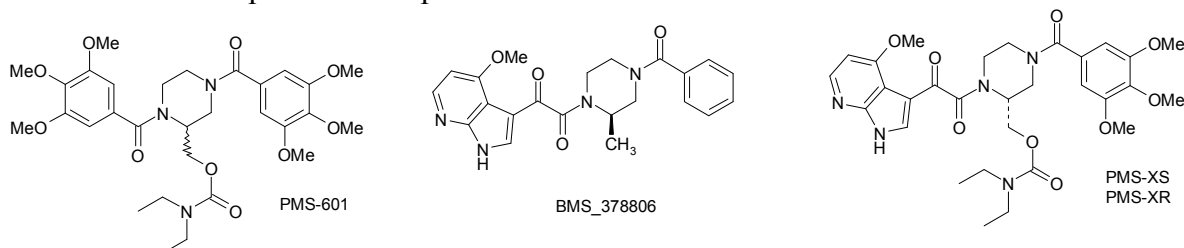
Figure 1: Molecular structures of PMS-601, BMS_378806, PMS-XS and PM-XR

Accurate prediction of a ligand binding position in a protein active site is a difficult task and especially when the target is highly flexible, like it is for gp120. It is well known that binding process, more generally, involves significant conformational changes in the ligand as well as of the protein active site. Therefore the study of the conformational properties of the ligand is an important step in the design of new compound. Moreover, the energies barriers between conformations are necessary for optimal overview of ligands plasticities.

For this reason a protocol of systematic conformational search was designed. First, 10ns molecular dynamics simulations were produced to generate several ligands conformations of compounds PMS-XR and PMS-XS (Figure 1). Typical ligands ring conformations were then extracted from the trajectories and were minimized with AM1 semi-empirical method. Finally, systematic searches for torsion angles were performed and gave the access to the ligands potential energies surfaces.

1. G. Borkow, et al , *Current drug targets,* **2005**, 5, 3-15.
2. S. C. Harrison, *Advances in virus research*, **2005**, 64, 231-261
3. R. Kong, et al, *Biochimica et biophysica acta*, **2006**, 1764, 766-772.

2D-QSAR STUDIES OF SOME BENZOXAZOLES AGAINST *S. AUREUS*

Betul Tekiner-Gulbas, Ozlem Temiz-Arpaci, Ilkay Yildiz, Esin Aki-Sener, Ismail Yalcin

Pharmaceutical Chemistry Dept., Faculty of Pharmacy,Ankara University, Tandogan 06100, Ankara, Turkey

Benzoxazoles are the structural isosters of natural nucleotides and interact easily with the biopolymers. So that benzoxazoles possess potential antimicrobial and antibiotic activities [1-6].

In this study, a congeneric set of some benzoxazole compounds 1-19 (Formula 1) were tested for their antibacterial activity against Staphylococcus aureus ATCC 25923 and the QSAR analysis by using Hansch analysis method were studied.
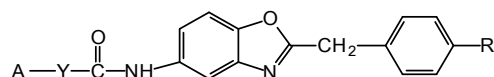
A= 4-chlorophenyl, 4-bromophenyl, 4-fluorophenyl, 4-methylphenyl, 4-ethylphenyl, 4-terbutylphenyl, 4-nitrophenyl, phenyl
Y= —, CH_2
R= H, Cl

Formula 1

The best equation was found in below;

$$\text{Log } 1/C = +0.357(\pm0.19) \text{ Ix} -0.374(\pm0.25) \text{ лR}_1 +0.293(\pm0.16) L_{R1} -1.150 (\pm0.69) R_{R1} +2.757(\pm0.52)$$

1. M. Prudhomme, J. Guyot, G.J. Geminet, *Antibiotics* **1986**, 39, 934.
2. H.K. Sarma, B.K Sharma, S.C. Tiwari, *Current Science*, **2003**, 85, 1401.
3. J.P. Haansu, K.D. Klika, V.V. Ochaveranko, K. Pihlaja, K.K. Haatela, P.M. Vourela, J. of Indst. *Mic. & Biotech.*, **2001**, 27, 62.
4. O.T. Arpaci, E.A. Sener, I. Yalcin, N. Altanlar, *Arch. Pharm. Pharm. Med. Chem.* **2002**, 6, 283.
5. I. Yıldız-Oren, B. Tekiner, I. Yalçın, O. Temiz-Arpacı, E. Akı-Şener, N. Altanlar, *Archiv der Pharmazie* , **2004**, 337(7), 402.
6. O. Temiz-Arpacı, A. Ozdemir, I. Yalçın, I. Yıldız, E. Akı-Sener, N. Altanlar, *Archiv der Pharmazie*, **2005**, 338 (2-3),105.

PROGRAM-APPARATUS SOLUTIONS AT ION CHANNEL COMPUTER MODELLING BY THE EXAMPLE OF GABA-A CHANNEL

Ksenia B. Tereshkina, Konstantin V. Shaitan, Eduard V. Tereshkin

Moscow M.V. Lomonosov State University, Faculty of Biology, Department of Bioengineering, Leninskie gory, 1-12, Moscow, 119992 Russia; E-mail: ksenia@moldyn.org

Currently problems of investigation of molecular design and functioning of ion channels are very relevant. Among various membrane receptor types special attention should be given on ligand-gated ion channels. Mutations in membrane receptors cause various hereditary diseases as epilepsy, Parkinson disease, hyperplexia, cardiac diseases etc. Furthermore, effects of anesthetics are bound to influence on ion channels. Studying of ion migration through receptor channel it is possible to detect of conformational transition by ligand binding, including anesthetic, that can help in development of new kinds of biosensors. In this work program-apparatus complex (PAC) for biological macromoleculas study was developed and tested. PAC can be used in areas of bioengineering, biomedicine, new materials. Using PAC GABA-A receptors were investigated [1]. In this work theoretical model of channel part of homomeric GABA-A receptor of $\alpha 1$-subunits was received. Using standard MD protocol [2] diffusion coefficients of migration particles were found. It was found that delicate balance of electrostatic interaction influences on ion and particles migrational dynamics. Point mutations was found switch the channel from anion to cation-selective. Key role of solvation was verified. Fluctuation parameters of inner channel interior were received. All results are in good agreement with experimental data. Obtained results can be applied at drug design, biosensor and nanobiodevices development.

1. E. E. Benarroch. GABA-A receptor heterogeneity, function, and implications for epilepsy, *Neurology*, **2007**, 68(8), 612-614.
2. K.V. Shaitan, E.V. Tourleigh, D.N. Golik, K.B. Tereshkina et al. – Dynamic molecular design of bio- and nanostructures, *Rossiyskiy Himicheskiy Zhurnal* [In Russian], **2006**, 50(2), 53–65.

3D-QSAR MODELS OF IRREVERSIBLE AND REVERSIBLE INHIBITORS OF ACETYLCHOLINESTERASE

Olga V. Tikhonova[1,2], Vladlen S. Skvortsov[1,2], Veniamin Yu. Grigor'ev[2], O.A.Raevsky[2]

[1]Institute of Biomedical Chemistry of Rus.Acad.Med.Sci., 10, Pogodinskaya Str., Moscow, 119121, Russia; [2]Institute of Physiologically Active Compounds of Rus.Acad.Sci., 1, Severnyi proezd, Chernogolovka, Moscow region, 142432, Russia; E-mail: olga.tikhonova@ibmc.msk.ru

It is well known that one of the modes of toxic action is associated with acetylcholinesterase (AChE) inhibition. Irreversible inhibition of AChE leads mainly to acute toxicity of different chemicals, while reversible inhibitors of this enzyme do not exhibit such toxicity and some of them can be even used for the treatment of neurological disorder. The 3D-QSAR models of irreversible and reversible inhibitors that include organophosphorus and carbamate compounds were established using the combination of receptor-based alignments and CoMFA, CoMSIA methods. Automated docking to the crystal structure of AChE and partial manual alignment were performed for a dataset of reversible inhibitors. Since no docking programs are capable of reliable representing the covalent bond formation the manual alignment (fitting to a reference molecule) based on previous investigation was performed for irreversible inhibitors. Predictivity of the models was validated by cross-validation ($q^2 > 0.7$). Comparison of different binding mode based on obtained models was performed. A good correlation between experimental inhibition activity and toxicity data was shown for *Daphnia magna* and *Fathead minnow*. Therefore obtained 3D-QSAR models can be used for aquatic toxicity prediction of some pollutants.

BANKS of MODELS AND KNOWLEDGE BASE OF RESULTS SAR - BASIS OF MOLECULAR DESIGN AND PREDICTION

Lidija Tyurina[1], Alexander Kolbin[1], Liana Sementeeva[2], Rail Valitov[1]

[1]Institute of herbicides and regulators growth plants, 65, Uljanov Str., Ufa, 450029, Russia; [2]State Petroleum Technological University, 1, Kosmonavtov Str., Ufa, 450074, Russia; E-mail: tjurina@anrb.ru

The researches fulfilled on the system SARD and SARD-TOX are carried out as for the conditional structurally related compounds enough and for bioisosteres. Pharmacological and pesticidal types activities (more than 30 types) were investigated. Each researched set included from 40 up to 1000 molecules. Broncholiticals, immunosuppresal, immunostimulatoral, antioxidantal, embryotrophical, antiviral, ulseric, uretonical antihelminthical, herbicidal, fungicidal and etc. types activities are investigated. Besides the toxicological characteristics are researched: an acute toxicity - DL_{50}, metric of a zone of acute action -Zac and approximately safe levels influence. The results of researches on each direction are brought in bank of models and knowledge base. The bank of models includes mathematical predict models and initial and designed structural formulas of the chemical compounds with their experimental and prognosticated estimations of properties. The knowledge base includes: a) of an estimation of influence of structural fragments and other descriptors on the researched properties; b) the data (concrete the formulas and their coordinates in many-dimensional space of tags describing model of the forecast) for four types of the calculated structural standards S1- S4 (S1-one, S2-other properties, S3-combination of properties, S4-with maximal indeterminacy of properties); c) measure of similarity and difference of molecules with the calculated standards; d) estimation of the contributions of all fragments for each individual molecule on their property and sequences of replacement of fragments at the design; e) probability of replacement of fragments to all set of the researched molecules with the given activity. Moreover others dates there are, for example, fragments estimation of the toxicological characteristics. Application these data provide extensive opportunities for combined researches on search substances with the given properties.

CHARGE-TRANSFER INDICES FOR MULTIFUNCTIONAL AMINO ACIDS

Francisco Torrens[1], Gloria Castellano[2]

[1]Institut Universitari de Ciencia Molecular, Universitat de Valencia, Edifici d'Instituts de Paterna, P. O. Box 22085, 46071 Valencia, Spain; [2]Instituto Universitario de Medio Ambiente y Ciencias Marinas, Universidad Catolica de Valencia *San Vicente Martir*, Guillem de Castro-94, 46003 Valencia, Spain; E-mail: francisco.torrens@uv.es

Valence topological charge-transfer (CT) indices are applied to the calculation of pH at the pI isoelectric point. The combination of CT indices allows the estimation of pI. The model is generalized for molecules with heteroatoms. The ability of the indices for the description of molecular charge distribution is established by comparing them with the pI of 21 amino acids. Linear correlation models are obtained. The CT indices improve multivariable regression equations for pI. The variance decreases by 95%. No superposition of the corresponding G_k–J_k and G_k^V–J_k^V pairs is observed in most fits, which diminishes the risk of collinearity. The inclusion of heteroatoms in π-electron system is beneficial for the description of pI, owing to either the role of the additional p orbitals provided by heteroatom or role of *steric* factors in π-electron conjugation. The use of only CT and valence CT indices $\{G_k, J_k, G_k^V, J_k^V\}$ gives limited results for modelling pI of amino acids. Furthermore, the inclusion of the numbers of acidic and basic groups improves all models. The effect is specially noticeable for amino acids with more than two functional groups. The fitting line obtained for the 21 amino acids can be used to estimate the isoelectric point of lysozyme and its fragments, by only replacing $(1+\Delta n/n_T)$ with $(M+\Delta n)/n_T$. For lysozyme, the results of smaller fragments can estimate that of the whole protein with 1–13% errors. Provisional conclusions follow. (1) The inclusion of heteroatoms in the π-electron system was beneficial for the description of the isoelecric point, owing to either the role of the additional p orbitals provided by the heteroatom or the role of steric factors in the π-electron conjugation. (2) The use of only charge-transfer and valence charge-transfer indices $\{G_k, J_k, G_k^V, J_k^V\}$ gave limited results for modelling the isoelectric point of amino acids. Furthermore, the inclusion of $(1+\Delta n/n_T)$ improved all the models. The effect is especially noticeable for those amino acids with more than two functional groups, *viz*. Arg, Asp, Glu, and, specially, His, and Lys. Moreover, the fractional index casts some light on the importance of the side-chain functional groups in the pI simulations of functional-rich molecules.

COMPUTATIONAL DESIGN OF ACETYLCHOLINESTERASE INHIBITORS BY FRAGMENT LINKING

Daria A. Tsareva, Alexander A. Oliferenko, Vladimir A. Palyulin, Nikolai S. Zefirov

Department of Chemistry, Moscow State University, 1/3 Lenin Hills, Moscow, 119992 Russia

Acetylcholinesterase (AChE) has been an important target for decades. Currently the most interesting use of AChE inhibitors is in the treatment of Alzheimer's disease (tacrine, donepezil). The resolved structure of the enzyme exhibits two binding sites: the active one and the peripheral one. It was found experimentally that simultaneous binding to both sites enhances the inhibitory activity very much. The effect is even stronger when two intervening ligands are tied together, or connected to each other with a spacer. Several successful studies are known in which such tethered ligands were identified following the methodology of fragment linking. However, the discovery potential of this powerful approach is not yet used in full, because organic spacers for linking the two parts of tethered ligands are normally picked up in a nonsystematic, empirical way.

Here, we suggest a 3D structure generation method for the identification of prospective AChE inhibitors based on using software HostDesigner, which is in use in coordination chemistry. This software assembles candidate structures by linking user-defined fragments (recognition elements) with organic spacers taken from a large fragment library. This combinatorial procedure is well developed and documented; it can quickly generate many thousands of candidates and effectively screen those using rather simple scoring functions. As the primary (for the active site) and auxiliary (for the peripheral site) recognition elements we selected the tetrahydroacridine moiety and several N- and O-bases, respectively. The final pool of the processed structural data consisted of a few dozens of compounds. At a first instance, the inhibitory activity of the generated structures was evaluated using QSAR methodologies of neural networks and multiple linear regression analysis.

FRAGMENTAL DESCRIPTORS IN QSAR MODELS: SELECTION AND INTERPRETATION

Irina G. Tsygankova

Institute of Experimental and Theoretical Biophysics RAS Pouschino, Moscow reg., 142290; E-mail: tsygan@iteb.ru

At present thousands descriptors of molecular structure derived in QSAR and QSPR approaches may be used for model building in drug design. The fragmental molecular descriptors depict a molecule as a system of distinct blocks, corresponding to some atoms, chemical bonds or functional groups. Being simple and representing molecular structure clearly, these descriptors are widely used. In the developed approach based on the fragmental descriptors, QSAR equation is built as a sum of contributions by single fragments of different types, by the pairs of different fragments, or by the more large groups of fragments:

$$A \approx \sum_i n_1^i A_i + \sum_{(i,j)} n_2^{i,j} A_{(i,j)} + \sum_{(i,j,k)} n_3^{i,j,k} A_{(i,j,k)} + \cdots$$

Here A – is some biological activity, A_i , $A_{i,j}$, $A_{i,j,k}$ – one-, two-, three-fragment contributions (i.e. parameters of QSAR equation, which are determined by linear regression using training set of compounds). Molecular descriptors n_1^i, $n_2^{i,j}$, $n_3^{i,j,k}$ are equal to the number of definite fragments and fragment groups in a molecule. The types of fragments, designated as $i, j, k, \ldots,$ are user defined. Such approach was applied to the QSAR modeling of the inhibitory activity on influenza neuraminidase by some cyclopentane and cyclohexene derivatives (N=30). After removing the multicollineated descriptors (r>0.98) and descriptors with low variance (v < 0.1) from the pool of molecular descriptors, there were still many variables (p=78). So, for getting an effective model variable subset selection (VSS) was performed. The resulting QSAR equation comprised only six variables and gave activity estimations with good accuracy (standard deviation was 0.75 in the units of ln IC_{50}, that was no more than 10% of the whole activity range). The special algorithm for VSS with iterative use of stepwise regression was designed for quick and exhaustive search of descriptor space. The structural meaning of the descriptors kept in the model was investigated.

GRID-SERVICES FOR EVIDENCE-BASED PHARMACOLOGY

Alexey Zhuchkov[1], Nikolay Tverdokhlebov[2], Boris Alperovich[2], Alexander Kravchenko[1]

[1]Telecommunication Centre "UMOS", Russia, Moscow, Profsoyuznaya 144; [2]Institute of Chemical Physics RAS, Russia, Moscow, Kosygina, 4; [3]Sechenov Moscow Medical Academy, Russia, Moscow, Trubeckaya 8, bld. 2; E-mail: nickhard@chph.ras.ru

Accurate risk assessment and patient safety are among the most valuable goals of information technology in Health Care. A prospective way to reach these goals is to provide medical professionals with complete and actual data about drug usage in clinical practice and this is the subject of evidence-based pharmacology. The problem is that these data are scattered over numerous diverse databases and web-pages. The Grid offers three important technological advantages to overcome this problem by providing integrated information space, controlled access to data and enabling full-scale usage of multi-agent technology. In the Russian ChemBioGrid meta-models and grid-services have been developed to provide biomedical researchers with an ability to build personal and corporative collections of data of clinical usage of drugs and known adverse events. The meta-models are XML-structures with basic tags Drug Name, Adverse Event, Risk Factor and Risk Probability aimed to store orderly the data which have been found in accessible information resources. These information structures are being filled with facts and formalized semantic links by means of web- and grid-services (agents) of data search and lexical analysis then are being stored in Repository of Meta-Descriptions which is a distributed data warehouse and is maintained by means of "intermediate" grid-services. Owing to the use of grid-technology the intellectual agents are able to continue data search and analysis (delegating initial access rights) even after they fulfill the initial request and thus data collections in the Repository are always up-to-date. Data and semantic links are being presented to users in the form of ontologies via a GUI. With these services and meta-models members of biomedical virtual organization of ChemBioGrid construct personal and collaborative collections of drug clinical usage data from any information resource to which the virtual organization has access rights.

COMPUTER BASED DRUG DESIGN OF LIGANDS FOR CRD DOMAIN OF FRIZZLED RECEPTOR AS PROSPECTIVE ANTICANCER COMPOUNDS

Andrew. E. Voronkov, Igor I. Baskin, Vladimir A. Palyulin, Nikolai S. Zefirov

Department of chemistry, Moscow State University; E-mail: andreyvoronkov@yandex.ru

The Wnt-Frizzled signaling pathway is known to be active in many types of cancer. The inhibition of Wnt-Frizzled signaling pathway in many cases was shown to lead to tumor growth arrest. The interaction of Wnt signaling proteins with Frizzled receptors is conducted through cystein rich domain (CRD) of Frizzled receptor. We have constructed the first models of dimeric CRD-domains of Frizzled receptors and analyzed the binding sites for Wnt protein. The homology modeling method with the use of X-ray structure of CRD domain of mouse Frizzled 8 receptor as template was used. We used the data of binding site analysis and our models for the prediction of potential inhibitors of Wnt-Frizzled interactions which can interact with the binding site on the surface of the dimeric CRD domain. The virtual screening procedure was employed for the selection of the best compounds from the dataset extracted from ZINC database of focused ligands. These compounds were modified with *de novo* drug design software and docked to the binding site on the surface of the CRD dimer of Frizzled receptors. These compounds can be used as potential effective inhibitors of Frizzled-Wnt signaling pathway.

THE 3D-MODELING, COMPUTER PREDICTION AND EXPERIMENTAL TESTING OF 5-HT$_3$-ANTISEROTONIN ACTIVITY OF NEW CHEMICAL COMPOUNDS

Pavel M. Vassiliev[2], Dmitry S. Yakovlev[1], Vladimir V. Poroikov[3], Dmitry A. Filimonov[3], Alexander A. Spasov[2], Maxim V. Chernikov[2], Alexandra N. Zheltukhina[2]

[1]Volgograd Science Centre of Rus. Acad. Med. Sci. and of Volgograd Region Administration, 1, pl. Pavshikh Bortsov, Volgograd, 400131, Russia; [2]Volgograd State Medical University, 1, pl. Pavshikh Bortsov, Volgograd, 400131, Russia; [3]Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: farm@vlpost.ru

The computer prediction of high level of 5-HT$_3$-antiserotonin activity of 175 new derivatives of condensed nitrogen-contained heterocycles was carried out, by joint using of the IT «Microcosm», of the PASS system and of the method of 3D-geometric similarity to standards. Three compounds with high 5-HT$_3$-antagonistic activity were found as the result of following experimental testing. At the first step, with applying of IT «Microcosm», on the training set of 98 new experimental investigated substances, the calculation of decision rules for prediction of high level of 5-HT$_3$-antiserotonin activity was made. On the basis of four methods and three strategies of prediction, the computation of predictive evaluations spectrum of high level of 5-HT$_3$-antagonistic activity was fulfilled for 175 new untested compounds. After checking on uncontradictoriness of the predictive evaluations spectrum, 25 compounds with reliable calculated evaluations was selected. At the second step, with using of PASS system, the calculation of existence of high level of 5-HT$_3$-antagonistic activity for these structures was produced. Ten structures with highly reliable calculated evaluations of existence of 5-HT$_3$-antiserotonin activity were selected. At the third step, 3D-models for these ten compounds was build by molecular mechanic methods, and the evaluation of 3D-geometric similarity degree with conformations of four most active compounds-leaders (from number of earlier tested) was produced. Tree compounds with most high indices of similarity to standards of their 3D-models were selected. The substances was tested, all of them were highly active. Thus, the joint consistent using of two prediction systems of biological activity on 2D-structure formula and of prediction method on 3D-geometric similarity to standards allows to optimize the searching of new chemical compounds with high 5-HT$_3$-antiserotonin activity practically with 100 % efficiency.

QSAR MODELLING OF RAT'S CARCINOGENIC TOXICITY

Alexey Zakharov, Alexey Lagunin, Dmitry Filimonov and Vladimir Poroikov

Institute of Biomedical Chemistry of Rus. Acad. Med. Sci., 10, Pogodinskaya Str., Moscow, 119121, Russia; E-mail: alexey.zakharov@ibmc.msk.ru

There are more than 70 000 chemicals in use today and many more being synthesized. It is vital that there are efficient methods to assess the effect of these compounds on the environment and on human health. Experimental testing is both time-consuming and expensive, and accordingly, there is a pressing requirement for accurate *in silico* methods to provide an initial screen that generates alerts for carcinogenic toxicity. The QSAR studies of the environmental fate of chemicals have become a necessary tool for carcinogenic risk assessments. The Distributed Structure-Searchable Toxicity (DSSTox) database from the U.S. Environmental Protection Agency (www.epa.gov) provides various databases Carcinogenic Potency Database (CPDB), which currently contains structures of 1481 chemicals (CPDBAS_v3b) of which 527 structures have designated carcinogenic toxicity for rats (male and female together). The toxicity end-points are based on the TD_{50} (mmol/L) value for the rats. All complex, non-organic structures and structures with value of carcinogenic toxicity more then 3 of standard deviation were deleted. 512 structures remained after this preparation was used for QSAR modeling. We used QNA (Quantitative Neighbourhoods of Atoms) descriptors and Self-Consistent Regression for QSAR modeling of rat's carcinogenic toxicity. The statistical parameters of the correlation are the follows: Number = 512, R^2 = 0.584, F = 11.814, SD = 0.965, Q^2 = 0.512. It was excluded 43 structures after out layers analysis and QSAR modelling was repeated. Results of modelling: Number = 469, R^2 = 0.703, F = 16.945, SD = 0.776, Q^2 = 0.631. This set was divided randomly 20 times on two sets: test and training. Test set contain 10 structures and training set 459. R^2 average of test set prediction is 0.51. These results show satisfactory predictive ability of our method, hence, it may be used for assessment carcinogenic toxicity of compounds in virtual screening.

QSAR STUDY OF BILITRANSLOCASE TRANSPORT ACTIVITY

Spela Zuperl[1], Sabina Passamonti[2], Marjana Novic[1]

[1]Laboratory of Chemometrics, National Institute of Chemistry, Hajdrihova 19, POB 3430, SI-1115 Ljubljana, Slovenia; [2]Department of Biochemistry, Biophysics and Macromolecular Chemistry, University of Trieste, 34127 Trieste, Italy; E-mail: spela.zuperl@ki.si

Bilitranslocase is plasma membrane carrier involved in uptake of diverse organic anions, some phthaleins and dietary anthocyanins. The latest study showed that, contrary to anthocyanins, most flavonols did not interact with bilitranslocase. [1] In this study a set of endo- and xenobiotics was tested as bilitranslocase inhibitors. The aim was to discover whether bilitranslocase played role in the transport of these, so far unknown ligands, through the cell membrane. We have tested the interaction of the carrier with purines, pyrimidines and their derivatives (nucleosides, nucleotides and congeners). The interaction was assessed by evaluating the kinetics of inhibition of bilitranslocase transport activity. The experiments for the determination of bilitranslocase inhibition constants were performed with a series of substrate (sulphobromophtalein) concentrations, while the investigated molecules were added in stoichiometric concentrations. The experimental results were the basis for a data-driven modelling study using artificial neural networks. 3D chemical structures (minimal energy conformation from Mopac) were represented by molecular descriptors calculated with the Codessa software. The molecular descriptors together with the corresponding bilitranslocase inhibition constants were used to train the counter-propagation neural network, designed for classification and prediction purposes. Optimised models were validated by previously determined validation set. Genetic algorithm for variable selection was introduced and detailed investigation of selected descriptors was performed.

1. A. Karawajczyk, V. Drgan, N. Medic, G. Oboh, S. Passamonti, M. Novic, Properties of flavonoids influencing the binding to bilitranslocase investigated by neural network modelling, *Biochem. Pharm.*, **2007**, 73, 308-320.

IN SILICO PREDICTION OF CYP450-MEDIATED METABOLISM PROFILE

Sehan Lee

Yonsei University, Seoul 120-749, Korea; E-mail: seni0206@hotmail.com

CYP450 is considered to be the most important single enzyme family in drug metabolism. This superfamily of enzymes is thought to metabolize approximately 90% of all marketed drugs. In this study, we tried three approaches using statistical methods to classify substrate, empirical model to predict the activation energy of CYP450 reaction and combination of docking method and semi-empirical molecular orbital calculations to determine the binding mode of CYP450 enzyme-substrate complex. Each model gives a lot of insights to describe CYP450-mediated system in vivo. Statistical method is useful to separate potential substrates and non-substrate with only two dimensional descriptors. Empirical model can explain aliphatic hydroxylation and aromatic hydroxylation which altogether constitute most reactions mediated by CYP450 using AM1 quantum mechanical calculations. Third model gives major metabolic positions of substrates in CYP450. Further, collection knowledge of appropriate ligand and its binding site of CYP450 will help to follow up pharmacokinetics of novel compounds.

A QSAR STUDY OF ACUTE TOXICITY TO FATHEAD MINNOW (*PIMEPHALES PROMELAS*)

Youngyong In[1], Sung Kwang Lee[1], Pil Je Kim[2], Kyoung Tai No[1,3]

[1]Bioinformatics and Molecular Design Research Center, Seoul 120-749, Korea; [2]National institute of Environmental Research, Incheon, 404-170, Korea; [3]Department of Biotechnology, Yonsei University, Seoul 120-749, Korea

The evaluation of toxicity is important to researchers who have many chemicals to test and authorize. Toxicity prediction models provide useful information to these people. We applied several statistical methods of QSARs to fathead minnow acute toxicity (EPAFHM). Multiple linear regression (MLR), artificial neural network (ANN), and preclassification-MLR were adapted to the QSARs for 555 chemicals. The preclassification method was used a recursive partitioning (RP) for classification of simplified mode of action as narcosis and others. In this study, we used 445 chemicals for the training and 110 chemicals for the test set for MLR study. The developed model has correlation coefficients (r^2) of 0.71, 0.8, 0.74/0.76, and 0.76/0.70 for MLR, ANN, class-reactive, and class-narcosis of the training set, respectively. And the applicability method was introduced to the developed MLR models for confirmation of predictability. The developed model is useful in evaluating whether chemical toxicity was present or absent.

THE INTEGRATED COMPUTER SYSTEM «SARD-TOX–OVOLMEP» FOR DESIGN AND PREDICTION OF CHEMICAL COMPOUNDS WITH AN ESTIMATION of VARIOUS ASPECTS Of THEIR ACTION

Lidija Tyurina[1], Liana Sementeeva[2], Alexander Kolbin[1], Olga Tyurina[3], Tatjana Solominova[1], Vladislav Kirlan[2]

[1]Institute of herbicides and regulators growth plants, 65, Uljanov Str., Ufa, 450029, Russia; [2]Ufa State Petroleum Technological University, 1, Kosmonavtov Str., Ufa, 450064, Russia; [3] Ufa State medical University, 3, Lenin Str., Ufa, 450000, Russia; E-mail: tjurina@anrb.ru

The joined system «SARD-TOX–OVOLMEP» to the prediction and design of biologically active compounds) includes three subsystems: 1) computer system of analysis «structure-property» relationship, molecular design and prediction - «SARD-21»; 2) system for prediction of the toxicological characteristics - «SARD-TOX»; 3) system of definition of similarity and difference of molecules on the basis of the quantum-chemical characteristics – «OVOLMEP». 1. «SARD-21» allows: to estimate the influence of structural descriptors on the action, make mathematical models, to determine the perspective directions of designing and carry out design structures with the given type of activity. 2. A basis «SARD-TOX» are the complexes of hierarchical models realizing sequential contraction of prognosticated intervals. The boundaries of intervals for each model are installed automatically at a stage of creation of models. The selection of compounds is made for creation of models also automatically according to initial boundaries of intervals. The optimal models oriented to one or two algorithms (the geometrical approach or a method of voting) are units of the prognosis complex (UPC). They are shaped on the basis of bank of all generated models, which answer the given criteria. The constitution of the prognosticating complex is specified by the logic scheme reflecting intervals. A way of the prognosticated object under the scheme depending on recognition on everyone UPC. On a course of the scheme the intervals are narrowed down. On escaping of the prognosticating system the maximum narrow values of intervals of the toxicological characteristics are designated. 3. The system «OVOLMEP» carries out comparison van-der-vaals sizes and cards of molecular electrostatic potentials of structural samples with different molecules calculated by using methods of quantum chemistry.

COMPUTER-ASSISTED PREDICTION OF CYTOTOXICITY FOR HIV-1 INTEGRASE INHIBITORS

<u>Dmitry S. Druzhilovskiy</u>, Dmitry F. Filimonov, Alexey A. Lagunin, Tatyana A. Gloriozova, Vladimir V. Poroikov

Institute of Biomedical Chemistry, Russian Academy of Medical Sciences, Pogodinskaya Street 10, Moscow, 119121 Russia; E-mail: dmitry.druzhilovsky@ibmc.msk.ru

Because of the insufficient safety of known remedies for therapy of HIV infection, finding of new anti-HIV agents with low probability of side/toxic effects is necessary. The purpose of this work is development of computer approach for prediction of cytotoxicity and optimization new integrase inhibitors on this basis. We analyzed the peculiarities of action of known integrase inhibitors and applied the identified structure-activity relationships for search of cytotoxic effects. Computer program PASS has been used for the analysis of structure-activity relationships [1, 2]. For this purpose the set of 122 compounds representing various chemical classes of integrase inhibitors with experimentally determined value CC_{50} has been created. The estimation of prediction quality for cytotoxicity action was made on the basis of leave one out cross-validation for the total training set. The prediction was carried out on the basis of the structural formula of compound both for 16 new substances, found by computer-aided prediction of anti-integrase activity, and for known integrase inhibitor. Also, the structural elements that have a positive and negative impact on the cytotoxicity activity were identified. For a complex estimation of biological activity spectra computer program PharmaExpert was used. By means of PharmaExpert compounds with low probability of the cytotoxic activity have been selected.

1. V. Poroikov, D. Filimonov (**2005**). In: Predictive Toxicology. Ed. by Christoph Helma. N.Y.: Taylor & Frensis, p. 459-478.
2. http://ibmc.msk.ru/PASS.
3. V. Poroikov, A. Lagunin, D. Filimonov. Pharmaexpert: Diseases, Targets and Ligands – Three in One. Proceedings of the 15th European Symposium on Structure-Activity Relationships (QSAR) and Molecular Modeling, Ed. by Esin Aki (SENER), Ismail Yalcin, Istanbul, September 05-10, **2004**, p.514-515.

Godjaev, N.M., 73
Golbraikh, A., 98
Golovin, A., 32
Golubeva, I.S., 76
Golubovich, V.P., 150
Gorb, L.G., 78
Gordienko, D., 143
Gorunova, O.V., 76
Gosh, P., 63
Gramatica, P., 40, 53, 100, 101, 102
Grigoriev, F.V., 158
Grigor'ev, V., 99, 162
Grigor'eva, Yu.A., 157
Grigorenko, B., 145
Gubaidullin, A.T., 157
Gusarova, N.K., 157

H

Hansen, N.T., 51, 104
Haraguchi, M., 125
Hege, A.I., 26
Henrick, K., 32
Heymans, F., 159
Hu R., 82
Hudson, B.D., 44
Hurst, T., 103

I

In, Y.Y., 34, 173
Ihlenfeldt, W.-D., 17
Islyaikin, M., 156
Ivanchenko, V., 131
Ivanenkov, Y.A., 66, 80
Ivanov, V.P., 154
Ivashchenko, A.A., 66, 80

J

James, T., 79
Jasprica, I., 120
Javahery, B., 64
Jiang, H., 28
Jonsdottir, S., 51, 104, 116
Jorgensen, F.S., 51, 104, 116
Judson R., 15

K

Kalinina, N., 110
Kalliokoski, T., 62
Kaplun, A.P., 105
Karamzin, S.S., 81
Karlinsky, D., 105
Kassam, K., 22
Kazachenko, V.P., 133
Kel, A., 25, 109, 111
Kel-Margoulis, O., 25
Khairullina, V., 106, 107

Kharchevnikova, N.V., 93
Kharitonova, O.V., 108
Kim, P.J., 34, 173
Kiralj, R., 67, 95
Kirlan, S., 106
Kirlan, V., 174
Koborova, O., 109
Kochetova, L., 110
Kokhov, V.A., 127
Kolbin, A., 106, 163, 174
Kolpakov, F., 109, 111, 146
Kondrakhin Yu., 146
Kondratieva, E., 88
Kondakova, O.A., 158
Konovalova, T., 25
Kos, A., 19
Kosinsky, Yu.A., 138
Koval, V., 115
Kovaleva, V., 83, 136
Kovalevskaya, N., 112
Kouskoumvekaki, I., 51, 104
Kravchenko, A., 167
Krishtopa, A., 143
Kulkarni, S., 113
Kumskov, M., 114
Kurbatova, S.V., 108
Kustova, T., 110
Kuz'min, V.E., 78, 128, 134
Kuz'min, V., 130
Kuz'mina, N., 130
Kuznetsov, A.V., 93
Kuznetsov, N., 115

L

Lagunin, A., 23, 91, 96, 109, 144, 152, 170, 175
Larsen, J., 104, 116
Lee, S.K., 34, 173
Lee, S., 172
Leopoldino, A.M., 33
Leszczynski, J., 78
Levtsova, O., 117
Li, A., 45
Li, F., 148
Liao, Q., 50
Lisitsa, A., 18, 135, 137
Litvinov, I.A., 157
Liu, H., 40, 102
Livingstone, D., 44
Lloyd, D.G., 79
Lucic, B., 69

M

Makan, S.Yu., 134
Makeev, V.J., 139
Maklyakov, Yu., 143
Malan, S.F., 54
Maldonado, A.G., 92
Malysheva, S.F., 157

Shaitan, K.V., 117, 129, 161
Schaper, K.-J., 124, 133
Sharipov, R., 109, 111, 146
Scheglova, N., 110
Shcherbukhin, D., 147
Shen, T., 50, 148
Shilova, E., 83, 136
Shorina-Zhabereva, A., 88
Shulga, D., 149
Shutova, I., 150
Silva C.H.T.P., 33
Silva, V.B., 33
Singh, P., 104
Sitzmann M., 16
Skvortsov, V., 84, 152, 162
Skvortsova, N., 151
Smirnov, V.I., 157
Smurnyy, Y., 112
Sobolev, B.,144
Sobolev, E., 153
Sobolev, O.V., 153
Solodilova, M., 154
Solominova, T., 174
Soskic, M., 155
Spasov, A.A., 126, 169
Stegmaier, P., 25
Stein, M., 27
Stepanchikova, A., 144
Strakhova, N.N., 133
Stryapan, M., 156
Sukhov, B., 157
Sulimov, A., 158
Sulimov, V.B., 158
Svitanko, I.V., 114

T

Taboureau, O., 104, 116
Tajima, S., 125
Teixeira, C., 47, 142, 159
Tekiner-Gulbas, B., 55, 160
Temiz-Arpaci, O., 52, 55, 74, 160
Tereshkin, E.V., 161
Tereshkina, K.B., 129, 161
Tetko, I., 53
Threadgill, D.W., 26
Thybaud, E., 90
Tikhonov, D.A., 153
Tikhonova, O., 162
Tkachenko, S.V., 127
Torrens, F., 164
Tourleigh, Y.V., 129
Trinajstic, N., 69
Trofimov, B.A., 148
Tropsha, A., 53, 98
Tsareva, D., 165
Tsygankova, I., 166
Tverdokhlebov, N., 167
Tyurina, L., 106, 107, 163, 174
Tyurina, O., 174

U

Urgin, K., 159

V

Valitov, R., 163
Varnek, A., 53
Vashchenko, A.V., 157
Vassiliev, P.M., 126, 169
Veselovsky, A.V., 84, 147
Vinnik, A.A., 46
Vorobjev, Yu., 58, 115
Voronkov, A., 168
Voss, N., 25
Vracko, M., 42, 63

W

Wade, R.C., 27
Waleev, T., 25
Wang, B., 45
Wang, K., 98
Watkins, P.B., 26
Weber, E.E., 99, 141
Wingender, E., 25
Williams, R.V., 24
Williams-Devane, C.-L., 15
Wolf, M., 15

Y

Yakovlev, D., 169
Yalcin, I., 52, 55, 74, 160
Yamakawa, M., 65
Yao, J., 50, 148
Yavorskaya, N.P., 76
Yildiz, I., 52, 55, 74, 85, 160
Yuan, S., 50

Z

Zafer, B., 55
Zaitseva, N.I., 140
Zakharov, A., 21, 91, 109, 170
Zaharov, A.M., 114
Zaplishny, V.N., 122, 123
Zarudiy, F., 107
Zatsepin, V., 131
Zefirov, N., 59, 70, 114, 149, 165, 168
Zhang R., 82
Zheltukhina A.N., 169
Zhu, H., 53, 98
Zhu, J., 113
Zhuchkov, A., 167
Zosimov, V.V., 140
Zotchev, S., 144
Zou, H., 28
Zuaboni, D., 36
Zuperl, S., 171