Abstracts of publications related to QSAR

Editor: Ferenc Darvas, Budapest
Technical Editor: Zsuzsanna Darvas, Budapest

| Title                                         | Abstract No. |
|-----------------------------------------------|--------------|
| Review                                        | 220 - 239    |
| Correlation Analysis, Application: Theoretical Papers | 240 - 247    |
| Correlation Analysis, Application: Pharmacology | 248 - 264    |
| Correlation Analysis, Application: Agricultural Chemistry | 265 - 273    |
| Correlation Analysis, Application: Physical Organic Chemistry | 274 - 279    |
| Correlation Analysis, Application: Chromatography | 280 - 285    |
| Correlation Analysis, Application: Environmental Sciences | 286 - 287    |
| Correlation Analysis, Application: Others     | 288 - 295    |
| Multivariate Analysis                         | 296 - 303    |
| Quantum Chemistry                             | 304 - 307    |
| Molecular Graphics, Theoretical Papers        | 308 - 311    |
| Molecular Graphics, Applications              | 312 - 328    |
| Proteins                                      | 329 - 341    |
| Software                                      | 342 - 354    |
**Review**

**1990/220**

**Title:** Progress in the design of bioactive molecules. (Review)

**Authors:** Block, J.H.

College of Pharmacy, Oregon State University
Corvallis OR 97331 – 3507, USA.

**Source:** ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.2-25. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

**Results:** An overview is given on the approaches for the discovery and design concepts of bioactive molecules:

a) Natural products derived from plant extracts and their chemically modified derivatives (cardiac glycosides, atropine, cocaine, penicillins, cephalosporins, tetracyclines and actinomycins, pyrethrins and cycloporsin);
b) Biochemically active molecules and their synthetic derivatives: acetylcholine, histamine, cortisone/hydrocortisone, indole-3-acetic acid (phenoxyacetic acid herbicides);
c) Principles of selective toxicity is discussed exemplified by trimethoprim/methotrexate, tetracyclines, acyllovir, azidothymidines, antifungal agents;
d) Metabolism of xenobiotics;
e) Exploitation of secondary effects (serendipity);
f) Receptor mapping;
g) Quantitative structure-activity relationship studies;
h) Empirical screening (shotgun approach).

**1990/221**

**Title:** QSAR. Theory and application. (Review)

**Authors:** Trinajstic*, N.; Nikolic, S.; Carter, S.
The Ruder Boskovic Institute
HPOB 1016, YU-4101 Zagreb, Croatia, Yugoslavia.

**Source:** Kern. Ind. 1989, 38(10), 469–484.

**Results:** An overview is given on three computer aided design (CAD) systems developed by IBM Tokyo Scientific Center:

a) Molecular Design Support System providing a strategic combination of simulation programs for industrial research and development optimizing computational time involved and the depth of the resulting information;
b) MolWorld on IBM Personal Systems intended to create an intelligent visual environment for rapidly building energetically stable 3D molecular geometries for further simulation study;
c) Molecular Orbital Graphics System designed to run on IBM main-frame computers offering highly interactive visualization environment for molecular electronic structures;
d) The systems allow interactive data communication among the simulation programs for their strategically combined use;
e) The structure and functions of MolWorld is illustrated on modeling the alanine molecule: (i) data model of molecular structures; (ii) chemical formula input; (iii) generation of 3D molecular structure; (iv) formulation of bonding model; (v) interactive molecular orbital graphics; (vi) methods of visualizing electronic structures; (vii) use of molecular orbital graphics for chemical reactions.

**1990/222**

**Title:** How does a key fit a flexible lock? Structure and dynamics in receptor function. (Review)

**Authors:** Neubig*, R.R.; Thomsen, W.J.

Department of Pharmacology, University of Michigan Medical School
M6322 Medical Science Bldg. 1, Ann Arbor MI 48109-0626, USA.

**Source:** BioEssays, 1989, 11(5), 136–141.

**Results:** A review is given on recent observations about receptor structure and the dynamic nature of drug receptors and the significance of receptor dynamics for drug design:

a) Receptors are classified according to structure and function (i) ion channels (nicotinic acetylcholine, GABA, Glycine); (ii) G protein linked (adrenergic (α,β), muscarinic acetylcholine, angiotensin, substance K, rhodopsin); (iii) tyrosine kinase (insulin, IGF, EGF, PDGF); (iv) guanylate cyclase (atrial natriuretic peptide, speractin);
b) Protein conformational changes can be best studied on allosteric proteins whose crystal structure is available (e.g. hemoglobin, aspartate transcarbamylase, tryptophan repressor (no high resolution structure of a receptor structure is known));
c) Receptor conformational changes can be studied by several indirect approaches (i) spectral properties of covalent or reversibly bound fluorescent reporter groups; (ii) the sensitivity of the receptor to various enzymes; (iii) the sedimentation of chromatographic properties of the receptor; the affinity of binding of radioligands; (iv) the functional state of the receptor;
d) There are many unanswered questions: e.g. (i) are there relatively few conformational states for receptors with fluctuations around them or many stable conformational states; (ii) How can static structural information be used in drug design when multiple receptor conformations exist.

**1990/224**

**Title:** Interfacing statistics, quantum chemistry, and molecular modeling. (Review)

**Author:** Magee, P.S.

BIOSAR Research Project
Vallejo CA 94591, USA.

**Source:** ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.37–56. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

**Results:** A review is given on the application and overlap of quantum chemical, classical modeling and statistical approaches for the
understanding of binding events at the molecular level. A new complementary method called statistical docking experiment is also presented:

a) Insights obtained using energy-minimized structures;
b) Activation in the bound state, types and energies of interactions at the receptor site and in crystal;
c) Four successful examples (significant regression equations) are given for the modeling of binding events using physico-chemical descriptors and correlation analysis: (i) binding of a diverse set of pyridines to silica gel during thin-layer chromatography; (ii) binding of meta-substituted N-methyl-arylcarbamates to bovine erythrocyte AChE; (iii) binding of meta-substituted N-methyl-arylcarbamates to AChE obtained from susceptible and resistant green rice leafhoppers; (iv) activity of phenols inhibiting oxidative phosphorylation of ADP to ATP in yeast;
d) A new statistical method for mapping of binding sites has been developed based on the hypermolecule approach, identifying key positions of binding and nature of the energy exchange between the hypermolecule atoms and the receptor site;
e) Two examples are given on the successful application of statistical modeling (statistical docking experiment) based on the hypermolecule approach: (i) inhibition of housefly head AChE by meta-substituted N-methyl-arylcarbamates (n = 36, r = 0.841, s = 0.390, F = 25.82); (ii) inhibition of housefly head AChE by ortho-substituted N-methyl-arylcarbamates (n = 46, r = 0.829, s = 0.485, F = 14.24).

1990/225

Title: A rapid method for the computation, comparison and display of molecular volumes. (Review)
Authors: Bohacek*, R.S.; Guida, W.C.
Pharmaceuticals Division, CIBA-GEIGY Corporation
Summit NJ 07901, USA.
Source: J. Mol. Graphics 1989, 7(June), 113 – 117.
Results: A rapid method is described for the calculation and visualization of approximate van der Waals volumes of molecules:
a) The individual volume elements of a molecule are mapped into an array storing the total molecular volume;
b) The relevant template for each atom in the molecule is mapped into a bit array and the appropriate atomic position is marked;
c) Volume comparisons (e.g. common volume or excluded volume) are made by bit-wise Boolean operations;
d) The algorithm for the visualization of the molecular surface comprising the calculated van der Waals volume is given;
e) Comparisons of CPU times required for the calculation of the van der Waals molecular volumes of various compounds using the methods of Stouch and Jurs, Pearlman, Gavazotti and the new method showed that similar or better results can be achieved using the new algorithm with VAX-class computers on molecules containing up to several hundred atoms.

1990/226

Title: QSAR studies on drugs acting at the central nervous system. (Review)
Author: Gupta, S.P.
Department of Chemistry, Birla Institute of Technology and Science Pilani-333031, India.
Source: Chem. Rev. 1989, 89(8), 1765 – 1800.
Results: An overview is given of QSAR studies on drugs exerting their primary effects upon the central nervous system (CNS):
a) QSAR of CNS drugs has been systematically discussed according to the following classes: (i) general (nonspecific) CNS depressants: general anesthetics; hypnotics and sedatives; (ii) general (nonspecific) CNS stimulants; (iii) selective modifiers of CNS functions: anticonvulsants, antiparkinsonism drugs, analytes and psychopharmacological agents; (iv) miscellaneous: drugs interacting with central α-adrenoreceptors, drugs interacting with histamine receptors, cholinergic and anticholinergic drugs;
b) The review indicates that the fundamental property of the molecules which mostly influence the activity of CNS drugs is hydrophobicity (they have to pass the cell membrane and the blood-brain barrier);
c) Electronic parameters, indicative of dipole-dipole or charge-dipole interactions, charge-transfer phenomena, hydrogen-bond formation, are another important factor governing the activity of most CNS agents;
d) Topographical, lipophylic and electronic structures of CNS pharmacophores are reviewed;
e) 191 QSAR equations, 24 tables and 3 figures from 294 references are shown and discussed.

1990/227

Title: Contribution of quantitative agrochemical design strategies to mechanism-of-action studies. (Review)
Authors: Plummer*, E.L.; Dixon, J.A.; Kral, R.M.
Agricultural Chemical Group, FMC Corporation
P.O. Box 8, Princeton NJ 08543, USA.
Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.157 – 168. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.
Results: Quantitative structure-activity relationship approaches offer many benefits besides quickly identifying the most effective compound in a congeneric set:
a) QSAR studies may provide insight into the active site requirements of the compounds;
b) Active site interactions can be separated from other factors such as uptake, transport and metabolism;
c) Activity mechanism can be better understood and mechanism based assays can be designed;
d) Outliers may give clue to the discovery of new lead compounds;
e) A detailed example is presented, where QSAR investigations lead to the conclusion that the site of insecticidal action of the benzyl esters and α-cyanobenzyl esters of cis-trans-dichlorovinyl-2,2-dimethylcyclopropane carboxylic acid is not identical;
f) The second example described the elucidation of structural requirements of chitin synthesis inhibitory benzoyl urea larvicides by QSAR techniques and their similarity to the natural substrate, UDP-N-acetylglucosamine.

1990/228

Title: Modelling in peptide design. (Review)
Author: Jameson, B.A.
Fels Institute for Cancer Research and Molecular Biology, Temple University School of Medicine
3420 North Broad Street, Philadelphia PA 19140, USA.
Source: Nature 1989, 341(6241), 465-466.
Results: The availability of an increasing number of crystalized protein structures and the development of powerful modeling software for the visualization of the tertiary structures of biologically important proteins have important implications for protein engineering and rational drug design:
a) One of the important goals of protein engineering is the design of iso-steric analogues of proteins;
b) Molecular software packages are available for molecular modeling and are among others developed by (i) BioDesign, Inc., Pasadena, California; (ii) Biosym Technologies, San Diego, California; (iii) Tripos, St. Louis, Missouri; (iv) Polygen, Waltham, Massachusetts; (v) Chemical Design Ltd. Oxford;
c) The molecular modeling packages use three basic parameters: (i) descriptive energy field; (ii) algorithm for performing molecular mechanics calculations; (iii) algorithm for performing molecular dynamics calculations;
d) Modelling study of the binding events occurring between the envelop protein (gp120) of the AIDS (HIV) virus and its cellular receptor (CD4) protein supported the hypothesis that this domain was directly involved in binding the gp120 envelop protein leading to the design of conformationally restricted synthetic peptides binding to CD4.

1990/229

Title: Finding lead structures from amino acid sequence similarities of target proteins. (Review)
Authors: Nishioka*, T.; Sumi, K.; Oda, J.
Institute for Chemical Research, Kyoto University
Uji, Kyoto 611, Japan.
Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.105 – 122. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.
Results: A new technique called "homology graphing" has been developed for the analysis of sequence-function relationships in proteins which can be used for sequence based drug design and the search for lead structures:
a) As target protein is inhibited by the ligands of other proteins having sequence similarity, computer programs have been developed for the search of the similarity of proteins;
b) Proteins are organized into hierarchical groups of families and superfamilies based on their global sequence similarities;
c) Global sequence similarities were used to find inhibitors of acetolactate synthase (ALS) and resulted in a quinone derivative as a lead structure of new ALS inhibitors;
d) Local sequence similarities of bacterial and mammal glutathione synthase (GSH) were used to find inhibitors of GSH;
e) It was shown that the sequence segment of GSH was similar to dihydrofolate reductase (DHFR) is part of the ATP-binding site;
f) Biological bases of local similarity between sequences of different proteins were indicated: molecular evolution of proteins and functionally important local regions;
g) Homology graph, as a measure of sequence similarity was defined;
h) Sequence-chemical structure relationship based on homology graph and the procedure to find lead structures was illustrated by an example resulting in a list of 33 potential inhibitors selected by the procedure based on the sequence segment from residue 150 to 210 of the sequence of tobacco ALS.

1990/230

Title: Molecular design and target site analysis in fungicide development. (Review)
Authors: Sisler*, H.D.; Ragsdale, N.N.
Department of Botany, University of Maryland
College Park MD 20705, USA.
Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.198 – 214. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.
Results: A review is given on the molecular design of the following major types of antifungal compound in relation to biochemistry, molecular modeling and target site fit:
a) Squalene epoxidase inhibitors (allilamines and thiocarbanilates) blocking conversion of squalene 2,3-oxidosqualene;
b) Inhibitors of sterol C-14 demethylation by cytochrome P-450 (piperazines pyridines, pyrimidines, imidazoles and triazoles);
c) Inhibitors of sterol Δ4→Δ2 isomerization and/or sterol Δ4→Δ2 reductase inhibitors (morpholines);
d) Benzimidazoles specifically interfering with the formation of microtubules and the activity phenylcarbamates on benzimidazole resistant strains;
e) Carboxamides specifically blocking the membrane bound succinate ubiquinone oxidoreductase activity in the mitochondrial electron transport chain in Basidiomycetes;
f) Melanin biosynthesis inhibitors selectively interfering with the polyketide pathway to melanin in Pyricularia oryzae by blocking NADPH dependent reductase reactions of the pathway (thialide, PCBA, chlobentiazone, tricyclosal, pyroquilon, PP389).

1990/231

Title: Quantitative modeling of soil sorption for xenobiotic chemicals. (Review)
Author: Sabljic, A.
Theoretical Chemistry Group, Department of Physical Chemistry, Institute Rudjer Boskovic
HPOB 1016, YU-4101 Zagreb, Croatia, Yugoslavia.
Source: Environ. Health Perspect. 1989, 83(2), 179 – 190.
Results: The environmental fate of organic pollutants depends strongly on their distribution between different environmental compartments. A review is given on modeling the soil sorption behavior of xenobiotic chemicals:
a) Distribution of xenobiotic chemicals in the environment and principles of its statistical modeling;
b) Quantitative structure-activity relationship (Q SAR) models relating chemical, biological or environmental activity of the pollutants to their structural descriptors or physico-chemical properties such as logP values and water solubilities;
c) Analysis of the QSAR existing models showed (i) low precision of water solubility and logP data; (ii) violations of some basic statistical laws;
d) Molecular connectivity model has proved to be the most successful structural parameter modeling soil sorption;
e) Highly significant linear regression equations are cited between KOM values and the first order molecular connectivity index (ξ) of a wide range of organic pollutants such as polycyclic aromatic hydrocarbons (PAHs) and pesticides (organic phosphates, triazines, acetanilides, uracils, carbamates, etc.) with r values ranging from 0.976 to 0.986 and s values ranging from 0.202 to 0.300;
f) The molecular connectivity model was extended by the addition of a single semiempirical variable (polarity correction factor) resulting in a highly significant linear regression equations between the calculated and measured KOM values of the total set of compounds (n = 215, r = 0.969, s = 0.279, F = 3291);
g) Molecular surface areas and the polarity of the compounds were found to be responsible for the majority of the variance in the soil sorption data of a set of structurally diverse compounds.

1990/232

Title: Strategies for the use of computational SAR methods in assessing genotoxicity. (Review)
Authors: Richard*, A.M.; Rabinowitz, J.R.; Waters, M.D.  
Genetic Toxicology Division, MD-68, U.S. Environmental Protection Agency  
Research Triangle Park NC 27711, USA.  
Source: Mutat. Res. 1989, 221 (3), 181 – 196.  
Results: A review is given on the overall strategy and computational SAR methods for the evaluation of the potential health effects of chemicals. The main features of this strategy are discussed as follows:  
a) Generalized SAR model outlining the strategy of developing information for the structure-activity assessment of the potential biological effects of a chemical or a class of chemicals;  
b) Models for predicting health effects taking into account a multitude of possible mechanisms;  
c) Theoretical models for the mechanism of the key steps of differential activity at the molecular level;  
d) SAR strategies using linear-free energy methods such as the Hansch approach;  
e) Correlative SAR methods using multivariate techniques for descriptor generation and an empirical analysis of data sets with large number of variables (SIMCA, ADAPT, TOPKAT, CASE, etc.);  
f) Data base considerations describing three major peer-reviewed genetic toxicology data bases (i) National Toxicology Program (NTP) containing short term in vitro and in vivo genetic tests; (ii) data base developed by the EPA Gene-Tox Program containing 73 different short term bioassays for more than 4000 compounds, used in conjunction with ADAPT, CASE and TOPKAT; (iii) Genetic Activity Profile (GAP) in form of bar graphs displaying information on various tests using a given chemical.

1990/233  
Title: Quantitative structure-activity relationships. Principles, and applications to mutagenicity and carcinogenicity. (Review)  
Authors: Benigni*, R.; Andreoli, C.; Giuliani, A.  
Laboratory of Toxicology and Ecotoxicology, Istituto Superiore di Sanita  
Rome, Italy.  
Source: Mutat. Res. 1989, 221(3), 197 – 216.  
Results: Methods developed for the investigation for the relationships between structure and toxic effects of compounds are summarized:  
a) The extra-thermodynamic approach: the Hansch paradigm, physical chemical properties that influence biological activity and their parametrization, originality of the Hansch approach, receptors and pharmacophores: the natural content of the Hansch approach, predictive value of QSARs, a statistical tool: multiple linear regression analysis, the problem of correlations among molecular descriptors, other mathematical utilizations of extrathermodynamic parameters;  
b) The substructural approach: when topological (substructural) descriptors are needed, how to use topological descriptors;  
c) QSAR in mutagenicity and carcinogenicity: general problems, specific versions of the substructural approach used for mutagenicity and carcinogenicity, applications to mutagenicity and carcinogenicity.

1990/234  
Title: Linking structure and data. (Review)  
Author: Bawden, D.  
Address not given.  
Source: Chem. Britain 1989, 25(Nov), 1107 – 1108.  
Results: The integration of information from different sources, particularly linking structural with non-structural information is an important consideration in chemical information technology. A review is given on integrated systems:  
a) SOCRATES chemical/biological data system for chemical structure and substructure searching combined with the retrieval of biological and physicochemical data, compound availability, testing history, etc.;  
b) Psidom suite of PC based structure handling routines combining chemical structure with the retrieval of text and data;  
c) Cambridge Crystal Structure Databank on X-ray data of organic compounds integrating information on chemical structure, crystal conformation, numerical information on structure determination, bibliographic reference and keywording;  
d) computer aided organic synthesis for structure and substructure search, reaction retrieval, synthetic analysis and planning, stereochemical analysis, product prediction and thermal hazard analysis.
binding, unfolding kinetics, conformational equilibria between different conformational states, fast and slow internal dynamics and other phenomena.

Title: Predicting mechanism and activity. The trendy computational soothsayer. (Review)

Authors: Henry, D.R.
Molecular Design Limited
2132 Farallon Drive, San Leandro CA 94577, USA.

Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.26 – 36. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

Results: An overview is given on the current trends and possible future impact of Computer-Aided Molecular Design (CAMD). The potential benefit of applying CAMD techniques is to increase the success rate of finding new drugs compared to the success rate of empirical research:

a) Hardware trends (performance and prices of low-end, mid-range and high-end 3D graphics workstations and optical disk and fiber optic technology);
b) Software trends: (i) growing use of 3D databases (Cambridge Crystallographic Database, Brookhaven Protein Databank, THOR, SMILES, ALADDIN, MACCS-I1 and associated programs such as WIZARD, CONCORD) (ii) merging of statistical and molecular modeling techniques (COMFA) for finding solution to CAMD problems.

c) Results: Supercomputers have been much more accessible by scientists and engineers in the past few years in part as a result of the establishment of National Science Foundation (NSF) supercomputer centers. The most powerful class of supercomputers have program execution rates of 100 million to 1 billion floating-point operations per second, memory storage capacities of some ten million to 100 million computer words and a standard digital word size of 64 bits, the equivalent of about 15 decimal digits. The following examples are given for the use of supercomputer resources for chemical calculations and modeling:

a) Modeling of key chromophores in the photosynthetic reaction center of Rhodopseudomonas viridis showing the heme group, the iron atom and the chlorophyll which absorbs light and causes rapid transfer of electron to pheophitin and then to the quinone; modeling includes a significant part of the protein having about 2000 atoms out of a total of some 12,000;

b) ALADDIN (Daylight Chemical Information Systems) is also searching databases of 3-D structures to find compounds that meet biological, substructural and geometric criteria such as ranges of distances, angles defined by three points (dihedral angles) and plane angles that the geometric object must match. Aladdin is one of a number of menus working within the framework provided by daylight's chemical information system.

Title: Software adds new dimension to structure searching. (Review)

Author: Borman, S.
C&EN
1155 Sixteenth St., N.W., Washington DC 20036, USA.

Source: C&EN 1989, 67(38), 28 – 32.

Results: Lately a number of chemical information systems based on three-dimensional (3-D) molecular structures have been developed and used in many laboratories:

a) CONCORD uses empirical rules and simplified energy minimization to rapidly generate approximate but usually highly accurate 3-D molecular structures from chemical notation or molecular connection table input;
b) Chemical Abstracts Service (CAS) has added 3-D coordinates for some 4 million organic substances to the CAS registry file;
c) Cambridge Structural Database System contains X-ray and neutron diffraction crystal structures for tens of thousands of compounds;
d) MACCS-3D developed by Molecular Design Ltd., contains the standard MACCS-I structures to which additional 3-D data, such as Cartesian coordinates, partial atomic charges and molecular mechanics energy are added; MACCS-3D allows exact match, geometric, submodel and substructure searching of 3-D models with geometric constraints specified to certain degree of tolerance; two 3-D databases are also available from Molecular Design that can be searched using MACCS-3D [Drug Data Report (10,000 models) and Fine Chemicals Directory (90,000 models)];
e) ALADDIN (Daylight Chemical Information Systems) is also searching databases of 3-D structures to find compounds that meet biological, substructural and geometric criteria such as ranges of distances, angles defined by three points (dihedral angles) and plane angles that the geometric object must match. Aladdin is one of a number of menus working within the framework provided by daylight's chemical information system.

Title: Improved access to supercomputers boosts chemical applications. (Review)

Author: Borman, S.
C&EN
1155 Sixteenth St., N.W., Washington DC 20036, USA.

Source: C&EN 1989, 67(29), 29 – 37.

Results: Supercomputers have been much more accessible by scientists and engineers in the past few years in part as a result of the establishment of National Science Foundation (NSF) supercomputer centers. The most powerful class of supercomputers have program execution rates of 100 million to 1 billion floating-point operations per second, memory storage capacities of some ten million to 100 million computer words and a standard digital word size of 64 bits, the equivalent of about 15 decimal digits. The following examples are given for the use of supercomputer resources for chemical calculations and modeling:

a) Modeling of key chromophores in the photosynthetic reaction center of Rhodopseudomonas viridis showing the heme group, the iron atom and the chlorophyll which absorbs light and causes rapid transfer of electron to pheophitin and then to the quinone; modeling includes a significant part of the protein having about 2000 atoms out of a total of some 12,000;
b) Modeling of transition state of reaction between chloride and methyl chloride including electron clouds and water molecules surrounding the reaction site;
c) Analysis of nucleic acid and protein sequences to evaluate the secondary structure of these biopolymers;
d) Construction of a graphical image of hexafluoropropylene oxide dimer, a model for DuPont Krytox high performance lubricant;
e) Calculation of the heats of formation of diaminobenzene isomers indicated that the target para isomer was 3 kcal/mol less stable than the meta isomer byproduct therefore the development for its large scale catalytic synthesis was not undertaken (saving was estimated to be $1 to $2 million).

**Correlation Analysis, Application: Theoretical Papers**

**1990/240**

**Title:** The analysis of the ortho effect.

**Author:** Sotomatsu, T.

**Department of Agricultural Chemistry, Kyoto University**

Kyoto 606, Japan.

**Source:** Doctoral thesis, Department of Agricultural Chemistry, Kyoto University, 1989.

**Results:** The thesis devoted to the quantitative analysis of the effect of mono- and di-ortho substituents of aromatic molecules:
a) A new steric parameter $E_o^o$, has been defined for ortho-substituents by measuring the acidic hydrolysis of substituted benzamides in water and the N-methylation of substituted N,N-dimethylanilines by methyl iodide in methanol;
b) Comparison of the newly defined $E_o^o$ parameter with the Taft-Kutcher-Hansch $E_o$ (TKH $E_o$) parameter showed characteristic steric effects of ortho-alkoxy and $\pi$ bonded planar type substituents (e.g. NO$_2$, Ph);
c) In various correlation analyses using retrospective data $E_o^o$, satisfactorily represented the steric effects of ortho-substituents on reactivity and biological activity of various organic compounds;
d) Semi-empirical AM1 calculations using a hydrocarbon model to study the steric effects of a number of ortho-substituents resulted in the calculation of the ES value (difference in the heat of formation between ortho-substituted toluene and t-butylbenzene) which linearly correlated with the $E_o^o$ and the TKH $E_o$ parameters;
e) Effects of di-ortho substitution on lipophilicity could be mostly expressed by the summed effect of the 2- and 6-position substituents;
f) Highly significant regression equations were calculated for the $pK_a$ values of di-ortho-substituted benzoic acids using various substituent parameters;
g) Quantitative analysis of the effect of ortho-substitution is difficult because it is a result of overlapping steric and electronic effects.

**1990/241**

**Title:** Calculation of partition coefficient of N-bridgehead compounds. Comparison of Rekker and Hansch-Leo methods.

**Authors:** Novák*, K.; Kollár, Gy.; Szász, Gy.

**Institute of Pharmaceutical Chemistry, Semmelweis Medical University**

Hőgyes Endre u. 9, Budapest, H-1092, Hungary.

**Source:** Magy. Kém. Foly. 1989, 95(2), 73-78.

**Compounds:** 56 Bicyclic and tricyclic pirido[1,2-a]pyrimidine derivatives.

**Data determined:**

$\log P_{\text{max}}$, $\log P_{\text{2.4}}$ (maximum value for the lipophilicity for a zwitterion and the net lipophilicity at physiological pH).

---

*Additional content not included in the image.*
Chemical descriptors:

\( pK_a \) (negative logarithm of the acidic dissociation constant).

Results: Sulphate conjugates are expected as having highly increased hydrophilic properties. Zwitterionic sulphate conjugates have, however, particular properties: (i) the lipophilicity of zwitterionic conjugates have a constant value between their \( pK_a \) values as the opposite charges of the anionic and cationic groups neutralize each other; (ii) compound (II) is more lipophilic than (I) at pH values below 7 despite the presence of the sulphate group; (iii) intramolecular interaction of the oppositely charged groups increase lipophilicity [propanol-0-sulphate (II) is more lipophilic than propanol-4-sulphate (IV)].

Fig. 1 shows the relationship between lipophilicity and pH for the compounds (circle represents (I), triangle (II), rhomboid (III) and square (IV)).

\[ \Sigma \sigma \] (sum of the Hamnett's constants of the substituents characterizing the electron-withdrawing power of the substituents).

Results: Highly significant linear equations are presented for the relationships between \( S_{298}^o(g) \) values of di-, tri-substituted and the monosubstituted benzene derivatives, \( r \) values ranging from 0.991 to 1.000. Based on these equations \( S_{298}^o(g) \) values of ortho-, meta-, para-substituted derivatives can be estimated. \( \sigma_{cal} \) values of the disubstituted derivatives calculated by the \( S_{298}^o(g) \) values agreed well with the observed ones (Eq. 1):

\[
\sigma_{obs} = 1.019 \pm 0.042 \sigma_{cal}
\]

Highly significant linear equations describe the relationships between \( \log \) and the three other descriptors for ortho, meta, para substituents, \( r \) values ranging from 0.994 to 0.998. The best equation was calculated for the meta derivatives (Eq. 2):

\[
\log \sigma_{meta} = 4.160 \pm 0.328 \sigma_{o} - 0.460 \pm 0.175 \left( \Sigma \sigma + 1.356 \pm 0.152 \Sigma \sigma_{cal} \right)
\]

The presented equations are suggested for estimating QSAR descriptors.

1990/243

Title: Determination of the novel quantitative structure-activity relationships descriptor \( \sigma_{cal} \) for di- and trisubstituted benzene derivatives.

Authors: Kawaki*, H.; Sasaki, Y.; Takagi, T.; Fujii, S.; Masuda, F.

Faculty of Pharmacy, Kinki University
Kowakae 3-4-1, Higashi-Osaka 577, Japan.

Source: Chem. Pharm. Bull. 1989, 37(12), 3268 – 3271.

Compounds: 53 Disubstituted benzene derivatives, with NO\(_2\), CN, CO\(_2\)Me, CO\(_2\)Et, CO\(_2\)Et, F, Cl, Br, I, Me, Et, OMe, OEt, OH, NH\(_2\), NM\(_2\) substituents in ortho, meta and para positions and 53 trisubstituted benzene derivatives with the above listed substituents in 1,2,3-; 1,2,4-; 1,3,5-positions and the corresponding monosubstituted benzenes.

Data determined:

\( S_{298}^o(g) \) [observed absolute entropy (e.u.) of the compounds].

Chemical descriptors:

\( \sigma_{obs} \); \( \sigma_{cal} \) [substituent entropy constant, representing both the dispersion and repulsion interaction energies, observed or calculated by the \( \sigma_{cal} = \log[S_{298}^o(g)B/S_{298}^o(g)A] \) equation, where A and B denote the reference benzene and the disubstituted benzene derivatives, respectively];

\( \mu^2/\epsilon \) (descriptor, characterizing the repulsion interaction energy);

\( \log \) (gas-liquid chromatographic relative retention of the compound);

\( \Sigma \sigma \) (sum of the Hamnett's constants of the substituents characterizing the electron-withdrawing power of the substituents).

Results: Highly significant linear equations are presented for the relationships between \( S_{298}^o(g) \) values of di-, tri-substituted and the monosubstituted benzene derivatives, \( r \) values ranging from 0.991 to 1.000. Based on these equations \( S_{298}^o(g) \) values of ortho-, meta-, para-substituted derivatives can be estimated. \( \sigma_{cal} \) values of the disubstituted derivatives calculated by the \( S_{298}^o(g) \) values agreed well with the observed ones (Eq. 1):

\[
\sigma_{obs} = 1.019 \pm 0.042 \sigma_{cal}
\]

Highly significant linear equations describe the relationships between \( \log \) and the three other descriptors for ortho, meta, para substituents, \( r \) values ranging from 0.994 to 0.998. The best equation was calculated for the meta derivatives (Eq. 2):

\[
\log \sigma_{meta} = 4.160 \pm 0.328 \sigma_{o} - 0.460 \pm 0.175 \left( \Sigma \sigma + 1.356 \pm 0.152 \Sigma \sigma_{cal} \right)
\]

The presented equations are suggested for estimating QSAR descriptors.

1990/244

Title: Some physicochemical parameters of 11H-indolo[3,2-c]quinoline.

Authors: Lin, G.M.; Lan*, N.T.

Department of Pharmacy, National University of Singapore
Lower Kent Ridge Road, 0511, Republic of Singapore.

Source: Heterocycles 1989, 29(12), 2353 – 2359.

Compounds:
a) 11H-indolo[3,2-c]quinoline of type I; b) 2 Quinoline derivatives of type II and III.

Chemical descriptors:

\( P, P_{app} \) [true and apparent partition coefficients, respectively, where \( P = P_{app}(1 + [H^+]/K_a]) \);

\( \log P \) (logarithm of the partition coefficient in 1-octanol/water in a buffer of pH 3.0);

\( pK_a \) (negative logarithm of the acidic dissociation constant);

\( S_0 \) [limiting solubility (mol/L) calculated from a plot of solubility of the substance at different pH and at constant ionic strength);
The hydrophobicity of I was found to be significantly lower than that calculated logP values of f (Rekker's constant, characterizing hydrophobicity). Since interesting pharmacological activities have been reported for several derivatives of this type of compounds, the hydrophobicity of the unsubstituted 11H-indolo[3,2-c]quinoline has been calculated to be 2.22:

\[
\log P = \text{obs. log}P \text{ of I} = f(\text{Cl}) - f(\text{O}) - f(\text{Me}) - 2f(b) + 2f(H) = 2.78 - 0.94 - (-0.57) - 0.89 - 2(-0.12) + 2(0.23) = 2.22
\]

1990/245

Title: Voronoi binding site model of a polycyclic aromatic hydrocarbon binding protein

Authors: Boulu, L.G.; Crippen*, G.M.; Barton, H.A.; Kwon, H.; Marletta, M.A.

College of Pharmacy, University of Michigan Ann Arbor MI 48109, USA.

Source: J. Med. Chem. 1990, 33(2), 771 - 775.

Compounds: 16 Polycyclic aromatic hydrocarbons (PAHs): benzo[al]pyrene, dibenzo[a,c]anthracene, chrysene, pyrene, cyclopenta[c,d]pyrene, fluoranthene, 3-methylcholanthrene, 1-aminonaphthalene, 2-aminofluorene, 1-aminanthracene, 9-hydroxybenzo[a]pyrene, 7a,8@-dihydroxy-7,8-dihydrobenzo[a]pyrene, 9b,10a-dihydroxy-9,10-dihydrobenzo[a]pyrene, 4a,5@-dihydroxy-4,5-dihydrobenzo[a]pyrene, 2a,3@-dihydroxy-2,3-dihydrofluoranthene.

Biological material: Polycyclic aromatic hydrocarbon binding protein (PBP) from mouse liver which binds polycyclic aromatic hydrocarbons with high affinity.

Data determined:

\[
\text{IC}_{50} \quad \text{[concentration of the inhibitor (} \mu \text{mol/L) required for displacing 50 % of the } [\text{H}]\text{benzo[al]pyrene ([H]}\text{B[a]P)} \text{ from PBP];}
\]

\[
K_i \quad \text{[Michaelis-Menten affinity constant of the inhibitor calculated by } K_i = \text{IC}_{50}(1 + [L]/K_D)\text{, where } K_D \text{ and } [L] \text{ are the enzyme ligand dissociation constant and the concentration of } [\text{H}]}\text{B[a]P];}
\]

Molecular modeling (atomic coordinates of the molecules were taken from the Cambridge Structural Database or obtained using the QUANTA molecular modeling package and minimized using a molecular mechanics potential function such as CHARMM);

\[
\Delta G_{m,\text{calc}} \quad \text{(calculated binding energy for the for the ligand molecule m of the set falling within its respective experimental range: } \Delta G_m \leq \Delta G_{m,\text{calc}} \leq \Delta G_{m*});
\]

Interaction energy [interaction energy between a molecule and the binding site model was assumed to be the sum of its atomic contributions according to the expression \(\Delta G(b) = \Sigma_{\text{regions}} \Sigma_{\text{atoms}} \times \varepsilon_{r,\text{type}}(a)\), where \(\varepsilon_{r,\text{type}}(a)\) was the interaction energy parameter between the site region r and the atom-type of atom a and \(\Delta G(b)\) was the total interaction energy for the binding mode b (binding mode was regarded as feasible when the molecule was in its energetically most favorable conformation)].

Results: For development of the binding site model, first a simple geometry was proposed and \(\Delta G_m \leq \Delta G_{m,\text{calc}} \leq \Delta G_{m*}\) was calculated for the whole set of compounds. If the calculated binding energy of any of the compounds was outside of the above boundary, the proposed site geometry was rejected and a more complex one was considered. This procedure had been repeated until all molecules in the set were fitted within the experimental data range. As a result a 3D, five-region Voronoi binding site model has been developed for the PAHs containing a trigonal pyramid (r1) in the center and portions r2-r5 having infinite volumes and indicated by boundary planes. Region r2 represented access to the solvent and regions r3-r5 were blocked for binding (Fig. 1):

![Fig. 1](image)

In Fig. 2 benzo[al]pyrene is shown in its optimal binding mode with its atom barely touching the boundary surfaces and edges:

![Fig. 2](image)

Calculations showed that benzene and other monoaromatic ring compounds should be very weak competitors for the B[a]P site. The model correctly predicted the binding energy of nine competitors outside of the training set.

1990/246

Title: A novel approach to molecular similarity. (Review)

Authors: Cooper*, D.L.; Allan, N.L.

Department of Chemistry, University of Liverpool

P.O. Box 147, Liverpool L69 3BX, England.

Source: J. Comput.-Aided Mol. Design 1989, 3(3), 253 - 259.

Compounds: CH3CH2CH3, CH3OCH3, CH3SCH3, H2O, H2S.

Data determined:

\(\rho(p)\) [momentum density (dimension not given) of a molecule in space p];
S_{AB}(n) [similarity index (%) in terms of the momentum densities of two molecules, where n gives the position of the nuclei in plane X of the position space].

Results: The novel approach examines the similarity of electron densities as functions of their moments rather than the positions of the electrons. Fourier transform of the wave function (i.e. the momentum function) is compared for two molecules, by using essentially Carbo's similarity index. Fig. 1 shows the total momentum density calculated for H\textsubscript{2}O in the plane px = 0:

![Fig. 1](image)

The S_{AB}(0) values were calculated for each pair of molecules ranging from 47.7% to 99.8%. High similarity indices were calculated between all three molecules of general formulae (CH\textsubscript{3})\textsubscript{n}X and the highest values of S_{AB}(n) for all n occurred in the comparison of CH\textsubscript{3}CH\textsubscript{2}CH\textsubscript{3} and CH\textsubscript{3}SCH\textsubscript{3} and were virtually independent of n. The least similar molecules were as expected CH\textsubscript{3}SCH\textsubscript{3} and H\textsubscript{2}O [S_{AB}(0) values between 43.9 and 54.5%]. It was noted that no single index can be expected to be a universal solution and it was suggested that the simultaneous examination of a variety of measures of similarity, each emphasizing different aspects of molecular shape, bonding topology and electronic structure effects is necessary.

Correlation Analysis, Application: Pharmacology

1990/247

Title: On the three-dimensional Wiener number.
Authors: Bogdanov*, B.; Nikolic, S.; Trinajstic, N.
Department of Chemistry, The University of Skopje, YU-9100 Skopje, Macedonia, Yugoslavia.
Source: J. Math. Chem. 1989, No.3, 299 - 309.
Compounds: The first 21 alkanes excluding methane and including n-pentane.
Chemical descriptors:
- EF [enthalpy function (cal/K·mol)];
- 2D\textsubscript{W} (Wiener index calculated as the sum of all unique shortest distances between atoms in the hydrogen suppressed graph of the compound);
- 3D\textsubscript{W} (Wiener index calculated as the sum of all geometric distances between atoms in the hydrogen suppressed molecule of the compound).

Results: The traditional 2D Wiener number is defined as the sum of the lengths of all possible routes in the molecular graph. Here the length is proposed to be calculated as the real three-dimensional length between atoms: this is the 3D Wiener number. This number has many of the advantageous features of the related and very much studied 2D Wiener number. Additionally, it is highly discriminative and its use in quantitative structure-property relation studies (QSPR) appears to be encouraging, according to the preliminary calculations. Of these the most convincing is the set of statistical parameters for the linear correlation between the experimental and calculated enthalpy functions of the lower alkanes not shown here. Three different models have been tried and in all cases the 3D Wiener number seemed to be superior to the 2D one as it is reflected in (Eqs. 1 - 6).

\begin{align*}
\text{EF} &= 0.252(\pm0.017)(2\text{D}\textsubscript{W}) + 12.384(\pm0.599) \quad (1) \\
\text{n} &= 21 \quad \text{r} = 0.960 \quad s = 1.278 \quad F = 226.027 \\
\text{EF} &= 0.205(\pm0.013)(3\text{D}\textsubscript{W}) + 12.157(\pm0.578) \quad (2) \\
\text{n} &= 21 \quad \text{r} = 0.965 \quad s = 1.211 \quad F = 253.578 \\
\text{EF} &= -0.0034(\pm0.001)(2\text{D}\textsubscript{W})^3 + 0.450(\pm0.054)(2\text{D}\textsubscript{W}) + 10.508(\pm0.677) \quad (3) \\
\text{n} &= 21 \quad \text{r} = 0.978 \quad s = 0.981 \quad F = 198.877 \\
\text{EF} &= -0.002(\pm0.001)(2\text{D}\textsubscript{W})^3 + 0.358(\pm0.038)(3\text{D}\textsubscript{W}) + 10.246(\pm0.617) \quad (4) \\
\text{n} &= 21 \quad \text{r} = 0.982 \quad s = 0.881 \quad F = 248.639 \\
\ln \text{EF} &= 2.171(\pm0.033) + 0.257(\pm0.010) \ln (2\text{D}\textsubscript{W}) \quad (5) \\
\text{n} &= 21 \quad \text{r} = 0.986 \quad s = 0.044 \quad F = 676.456 \\
\ln \text{EF} &= 2.054(\pm0.035) + 0.272(\pm0.010) \ln (3\text{D}\textsubscript{W}) \quad (6) \\
\text{n} &= 21 \quad \text{r} = 0.988 \quad s = 0.041 \quad F = 775.024
\end{align*}

Biological material: Serotonin receptor 5-HT2.
Data taken from the literature:
- K\textsubscript{1} [affinity of the compound (nM) for the [\textsuperscript{3}H]ketanserin-labeled 5-HT\textsubscript{2} receptors sites].

Chemical descriptors:
- \(\pi\) (Hansch-Fujita's substituent constant characterizing hydrophobicity);
- \(\sigma_\text{m}, \sigma_\text{p}\) (Hammett's constants, characterizing the electron-withdrawing power of the substituent in meta- and para-position, respectively);
- R, F (Swain-Lupton's electronic parameters, characterizing the resonance and field effect, respectively);
- L, B, B\textsubscript{s} (STERIMOL steric parameters (Å)).
Equations were, on the other hand, calculated for the log(1/Ki) values of the compounds (Eq. 1, Eq. 2, Eq. 3):

\[
\log(1/K_i) = 1.06(\pm 0.12) \pi - 0.28(\pm 0.07) L + 6.96(\pm 0.31) F
\]

\[
\log(1/K_i) = 0.50(\pm 0.13) \pi + 1.56(\pm 0.42) B_i + 3.81(\pm 0.48) F
\]

\[
\log(1/K_i) = 2.40(\pm 0.34) B_i - 1.28(\pm 0.41) F + 3.10(\pm 0.52)
\]

It was shown that the central 5-HT2 receptor affinity of 2.5-DMAs is related mainly to lipophilicity and length of R. The importance of the electronic contribution of R was not so evident. Significant regression equations were, on the other hand, calculated for \(\sigma_p\) and \(\pi\) (\(r = 0.894\) and \(r = 0.842\)).

1990/249

Title: Analysis of binding at 4-aminobutyric acid receptor sites by structure-activity relationships.

Authors: Magee*, P.S.; King, J.W.

BIOSAR Research Project
Vallejo CA 94591, USA.

Source: ACS Symposium Series 1989, No.413. In: Probing Bioactivity Mechanisms p.281-290. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

The binding site is a sterically restricted lipophilic cleft exhibiting chiral selection [e.g. Eq. 3 (uptake in mouse brain)].

\[
\text{IC}_{50} = 1.10(\pm 4.18) \text{ IAME} + 0.649(\pm 2.18) \text{ IR} - 0.585(\pm 2.22) \text{ IRNG} + 4.16
\]

As in the equations nearly all indicator variables had negative regression coefficients it was concluded that instead of searching for better analogs, the research should be directed toward degradable pro-GABA or pro-muscimol derivatives that are efficiently taken up into the central nervous system (CNS).

1990/250

Title: Synthesis and QSAR of 1-aryl-4-(β-2-quinolyl/1-isooquinolyl)ethyl)piperazines and some related compounds as hypotensive agents.

Authors: Murti, A.; Bhandari, K.; Ram, S.; Prabhakar, Y.S.; Saxena, A.K.; (Late) Jain, P.C.; Gulati, A.K.; Srimal, R.C.; Dhawan, B.N.; Nityanand, S.; Anand*, N.

Central Drug Research Institute
Lucknow 226 001, India.

Source: Indian J. Chem. 1989, 28B(Nov), 934-942.

Compounds: 57 Various 2-piperazine-quinoline derivatives.

Biological material:

a) Cats;

b) Rats;

c) Mice.

Data determined:

\[
\text{LD}_{50} = \text{dose (mg/kg) of the compound required to kill 50% of the test organisms)}
\]

\[
\text{BPH} = \text{hypotensive effect on blood pressure (mm Hg) administered 2.85 \, \mu \text{mol/kg i.v. of the compounds).}
\]

Chemical descriptors:

\[
\log P \quad \text{(logarithm of the partition coefficient in 1-octanol/water)}
\]

\[
\sigma \quad \text{(Hammett's constant, characterizing the electron-withdrawing power of the substituent)}
\]

\[
R \quad \text{(Swain-Lupton's electronic parameter, characterizing the resonance effect)}
\]

Results: A highly significant parabolic relationship between log(BPH) and logP is presented for 23 2-substituted quinoline analogues (Eq. 1):

\[
\text{log(BPH)} = 4.789(\pm 0.391) \log P - 0.569(\pm 0.046) (\log P)^2 - 8.269
\]

Based on Eq. 1, an optimal logP is predicted (logPe = 4.23). The highest activity was produced by the 1-(3-methylphenyl)-4-(β-2-quin-
nonylphenyl) piperazine, its logP value being near to the optimal value (4.52). Log(BPH) values calculated by Eq. 1 agree well with the observed ones.

1990/251

Title: Analgesic and antipyretic activities of 1,2,6-thiadiazin-3-one 1,1-dioxides. SAR design of a new analgesic agent.
Authors: Giraldez, A.; Nieves, R.; Ochoa*, C.; Vara de Rey, C.; Cenarruzabelitia, E.; Lasheras, B.
Instituto de Quimica Medica, CSIC
Juan de la Cierva, 3, E-28006 Madrid, Spain.
Source: Eur. J. Med. Chem. 1989, 24(5), 497-502.
Compounds: 9 Thiadiazinone derivatives of type I and 6 pyrazolone derivatives of type II, where R' = H, Me or Ph; R2 = Me or Ph; R3 = H, Me, Cl, Br, I, Ph, N(Me)2 and acetylsalicylic acid (ASA) and phenylbutazone as standards.

\[
\text{log PRAp} (c_{ak}) = -0.65 \times \log k' - 0.21 \times \log k' - 1.54 \times C - 0.08 \times E + 1.87
\]

\[ n = 8 \quad r = 0.933 \quad s = 0.28 \quad F = 5.01 \] (2)

Significant linear relationships between the observed and calculated log(PRAp) (Fig. 1), log(PRAp) (Fig. 2) values are presented:

It was established that the increasing FRAn potency is associated with bulky 4-substituents and an optimum k' value (k' = 0.43) is supposed for the highest analgesic potency. For the highest antipyretic potency ratio the optimal k' is 0.69 and a 4-substituent with shielding effect on C-4 is supposed.

1990/252

Title: Correlation between hydrophobicity of N-alkylxanthine derivatives and their biological activities on guinea-pig isolated tracheal smooth muscle.
Authors: Miyamoto, K.; Takagi, K.; Sakai, R.; Wakusawa, S.; Koshiura, R.; Nadai, M.; Apichartpichean, R.; Hasegawa*, T.
Department of Hospital Pharmacy, Nagoya University School of Medicine
65 Tsumura-cho, Showa-ku, Nagoya 466, Japan.
Source: J. Pharm. Pharmacol. 1989, 41(12), 844 - 847.
Compounds: 8 Xanthine derivatives of type I, where R' = H or Me, R2 = Me, Et, Pr or Bu.

\[
\text{log PRAp} (c_{ak}) = -0.45 \times \log k' - 0.33 \times \log k' - 0.31 \times E + 0.43
\]

\[ n = 8 \quad r = 0.970 \quad s = 0.07 \quad F = 21.18 \] (1)

Significant linear relationships between log(PRAp) (Fig. 1), log(PRAp) (Fig. 2) values are presented:

Biological material: Guinea-pig.
Data determined:
Ki [cyclic (c)AMP phosphodiesterase (PDE) inhibition constant (µM)];
EC50 [relaxant effect, concentration of the compound (µM) producing 50 % tracheal smooth muscle relaxation in-vitro].
Chemical descriptor:
logP (logarithm of the apparent partition coefficient in 1-octanol/water).

Results: Highly significant linear relationships are presented for the relationships between - logKi and -logEC50 for the carbachol-induced contraction (Fig. 1) or on the resting tone (r = 0.902 and 0.892, respectively):
A significant linear relationship (r = 0.872) between - logKi and logP was calculated (Fig. 2):
Based on the relationships it was established that the potency of the relaxant effect of a xanthine derivative depends on the cell membrane permeability and its affinity for cAMP PDE.
Title: Correlation between the lipophilicity of substituted phenols and their inhibition of the Na⁺/K⁺-ATPase of Chinese hamster ovary cells.

Authors: Cascorbi*, I.; Ahlers, J.
Freie Universität Berlin, Institut für Biochemie und Molekularbiologie
Ehrenbergstr. 26–28, D-1000 Berlin 33, Federal Republic of Germany.

Source: Toxicology 1989, 58(2), 197–210.
Compounds: 3,5-Dimethoxyphenol, 4-chlorophenol, 2,6-dichlorophenol, 4-methyl-2-nitropheno1, 2,4dichlorophenol, 2,4,6-trichlorophenol, 2,3,4,5-tetrachlorophenol, 2,4,6-triiodophenol, pentachlorophenol.

Biological material: Chinese hamster ovary (CHO) cells.

Data taken from the literature:
E_{20C}; EC_{50C}; EC_{20aA}; EC_{50aA} [concentration (mmol/L) of the compound leading to a 20 or 50 % inhibition of the cell growth or adenosine uptake, respectively].

Data determined:
EC_{20C}; EC_{50C} [concentration (mmol/L) of the compound leading to a 20 or 50 % inhibition of the Na⁺/K⁺-ATPase activity, respectively].

Chemical descriptors:
logP (logarithm of the partition coefficient in 1-octanol/water);
s (Hammel’s constant, characterizing the electron-withdrawing power of the substituent);
Eₐ (Taft’s constant, characterizing steric effects of the substituent);
x (molecular connectivity index, calculated by Koch’s method).

Results: Highly significant linear relationships were calculated between log(EC_{20a}) and logP (r = 0.963), the relationship between log(EC_{50a}) and s being less good (r = 0.767). Combining the two parameters the relationship has improved (Eq. 1):

\[
\log(EC_{20a}) = 0.725 \sigma - 1.182 \log P - 0.220, \quad n = 9, \quad r = 0.990, \quad s = 0.143, \quad F \text{ not given}
\]

Introducing Eₐ into Eq. 1, the relationship did not improve (r = 0.991). The measured and calculated log(EC_{20a}) values agreed well (Fig. 1):

A similarly good equation was calculated between log(EC_{20aA}) or log(EC_{50a}) and logP, s (r = 0.984, 0.993, respectively). A highly significant linear relationship is presented for log(EC_{20C}) and log(EC_{50a}) (Eq. 2):

\[
\log(EC_{20}) = 0.692 \log(EC_{20A}) - 0.124, \quad n = 8, \quad r = 0.991, \quad s = 0.127, \quad F \text{ not given}
\]

The equations are suggested for predicting toxicity of the phenol derivatives.

Title: Distance geometry of α-substituted 2,2-diphenylpropionate antimuscarinics.

Authors: Gordon*, R.K.; Breuer, E.; Padilla, F.N.; Smejkal, R.M.; Chiang, P.K.
Department of Applied Biochemistry, Walter Reed Army Institute of Research
Washington DC 20307-5100, USA.

Source: Mol. Pharmacol. 1989, No.36, 766–772.
Compounds: a) 15 Compounds of type I, where R = N-ethyl-3-piperidyl-, N-methyl-2-piperidylmethyl-, N-methyl-3-pyrrolidinyI-, 1-azabicyclo[2.2.2]-octane-3a-ol-, 2-methyl-2-azabicyclo[3.2.1]-octa-7a-yl-, N-methyl-2-pyrolidinethyl-, N-methyl-4-piperidylmethyl-, N-ethyl-4-piperidyl, 6-methyl-6-azabicyclo[3.2.1]-octane-3a-ol- (Azaprophen), N-methyl-4-piperidinethylmethyl-, N-methyl-4-piperidinethyl-

b) △[H]NMS (muscarinic inhibitor).

Biological material:
a) Muscarinic receptors of N4TG1 neuroblastoma cells;
b) α-Amylase of rat pancreas acini;
c) Guinea pig ileum.
Data determined:

IC_{50}([H]NMS binding) [concentration of the compound (nmol/L) required for 50 % inhibition of [H]NMS to the N4TG1 neuroblastoma cells;]

K_{i} (inhibitor constants derived from the IC_{50} values for the neuroblastoma cells;)

pA_{2} [measure of the inhibition of acetylcholine-induced contraction of guinea pig ileum by the compounds (dimension not given);]

IC_{50}(α-amylase release) [concentration of the compound (dimension not given) required for 50 % inhibition of carbochol-induced α-amylase release from pancreas acini;]

Molecular modeling

(D = distance (Å) between the carbonyl oxygen (constant electronegative locus) and the protonated nitrogen (center of cationic charge) of the compounds ranging between 4.4 Å and 5.9 Å).

Results: Three significant parabolic regression equations were calculated for IC_{50}([H]NMS binding), IC_{50}(α-amylase release) and pA_{2} (Eq. 1, Eq. 2 and Eq. 3, respectively):

\[ IC_{50} = 32.8 D - 3.15 D^{2} - 76.7 \] \hspace{1cm} (1)

\[ IC_{50} = 35.9 D - 3.42 D^{2} - 85.8 \] \hspace{1cm} (2)

\[ pA_{2} = 22.9 D - 2.19 D^{2} - 51.6 \] \hspace{1cm} (3)

Maximum antimuscarinic potency was achieved with a D value of about 5.2 Å in all three assays. A hypothetical binding site model of the muscarinic receptor was suggested. Fig. 1 shows the compound with R = N-methyl-4-piperidyl group, interacting with the muscarinic receptor’s anionic site that may contain an aspartic acid residue and the carbonyl oxygen, where the ester site of the antagonist can interact in hydrogen bonding with the receptor:

Data determined:

logk' [logarithm of the capacity factor, measured by reversed-phase liquid chromatography (RPLC) in methanol].

Chemical descriptors:

logP (logarithm of the partition coefficient in 1-octanol/water);

\( \pi \) (Hansch-Fujita’s substituent constant characterizing hydrophobicity);

\( \sigma \) (Hammett’s constant, characterizing the electron-withdrawing power of the substituent);

B_{3,3} (STERIMOL steric parameter, characterizing the steric effect of the meta substituents);

\( r_{w} \) (RPLC derived hydrophobic substituent constant, defined by Chen and Horváth, and extrapolated to 0 % methanol);

I_{1} (indicator variable 1 for the present 0 for the absence of hydrogen bonding substituents).

Results: Logk' values were determined for the benzenesulfonamides and correlated with chemical descriptors. A highly significant linear relationship between logk' and logP was calculated (Eq. 1):

\[ \text{logk'} = 0.72(±0.06) \text{logP} + 0.15(±0.07) \] \hspace{1cm} (1)

A highly significant linear equation is presented for the relationship between \( r_{w} \) and \( \pi \) (Eq. 2):

\[ r_{w} = 0.83(±0.10) \pi + 0.32(±0.11) \] \hspace{1cm} (2)

Introducing an indicator variable (I_{1}) which takes into account the possibility of the substituent to make hydrogen bondings, the relationship significantly improved (Eq. 3):

\[ r_{w} = 0.91(±0.08) \pi + 0.41(±0.18) I_{1} + 0.03(±0.16) \] \hspace{1cm} (3)

Introducing \( \sigma \) and B_{3,3} descriptors, the equation does not improve significantly. Eq. 3 is suggested for predictive calculations.

1990/256

Title: Lipophilicity measurements of benzenesulfonamide inhibitors of carbonyl anhydrase by reversed-phase HPLC.

Authors: Altomare, C.; Carotti*, A.; Cellaire, S.; Ferappi, M.

Department Farmaco-chimico, University of Bari

Via Amendola 173, I-70126 Bari, Italy.

Source: J. Pharm. Pharmacol. 1989, 41(12), 856–858.

Compounds: Chlorpromazine (CPZ) and seven of its metabolites: Nor1CPZ, Nor2CPZ, 7-OHCPZ, CPZNO, CPZSO, Nor1CPZSO, Nor2CPZSO.

Biological material: Horse serum cholinesterase.

Data determined:

K_{i} [horse serum cholinesterase inhibiting ability (M)].

Chemical descriptors:

pK_{a} (negative logarithm of the acidic dissociation constant);

logP (logarithm of the partition coefficient in 1-octanol/water).

Results: Relationships between K_{i} values and the chemical descriptors were investigated for CPZ and its listed metabolites. Relationship between log(1/K_{i}) and logP was calculated (Eq. 1) no numerical intercept (c) is given:

\[ \log(1/K_{i}) = 0.25 \text{logP} + c \] \hspace{1cm} (1)

The equation shows that 78 % of the variation in log(1/K_{i}) is accounted for by the change in logP. This correlation for CPZ metabolites may indicate that their ability to inhibit cholinesterase is due in part to hydrophobic bonding to the enzyme. The K_{i} values are relatively high which suggests that the bonding is not very specific. The influence of pH on the determination of logP is discussed.
1990/257

Title: Design and synthesis of 4H-3,1-benzoxazin-4-ones as potent alternate substrate inhibitors of human leukocyte elastase.
Authors: Krantz*, A.; Spencer, R.W.; Tam, T.F.; Liak, T.J.; Copp, L.J.; Thomas, E.M.; Rafferty, S.P.
Syntex Research (Canada)
2100 Syntex Court, Mississauga, Ontario L5N 3X4, Canada.
Source: J. Med. Chem. 1990, 33(2), 464 - 479.
Compounds: 175 Benzoxazin-4-one derivatives of type I, where R¹ - R³ are various substituents.

\[
\begin{align*}
&\text{(1)} \\
&\text{Biological material: Human leukocyte elastase (HLE).} \\
&\text{Data taken from the literature:} \\
&\text{Crystal structure (acyl-enzyme crystal structure was determined by X-ray diffraction measurements).} \\
&\text{Data determined:} \\
&Kᵢ \text{ (binding constant (mol/L) of the test compound with HLE);} \\
&\log(K_{OH-}) \text{ (logarithm of the alkaline hydrolysis rate (mol}^{-1}\text{L}^{-1}\text{s}^{-1});} \\
&k_{ac}, k_{de} \text{ (acylation rate (mol}^{-1}\text{L}^{-1}\text{s}^{-1}) and deacylation rate (s}^{-1}) \text{ of HLE by the compounds);} \\
&p\logP \text{ (logarithm of the partition coefficient in 1-octanol/water approximated by HPLC measurement).} \\
&\text{Chemical descriptors:} \\
&\sigma \text{ (Hammett's constant, characterizing the electron-withdrawing power of the substituent);} \\
&MR \text{ (molar refractivity);} \\
&l(C₂) \text{ (indicator variable 1 for carbon atom attaching to the 2-position of the benzoxazin ring);} \\
&I(2) \text{ (indicator variable 1 for R¹ containing aryl group).} \\
&\text{Results: Significant linear regression equations were calculated for } \\
&\log(K_{OH-}) \text{ and } p\logP \text{ (Eq. 1, Eq. 2):} \\
&\log(K_{OH-}) = 2.64(±0.07) \sigma + 0.15(±0.03) MR_S + 1.64 (1) \\
n = 123 \quad r = 0.965 \quad s = 0.27 \quad F = 881 \\
&\log(p\logP) = 2.98(±0.42) (R²) + 1.30(±0.34) \sigma_{ring} - 0.82(±0.25) l(C₂) - 0.85(±0.22) l(2) + 0.26(±0.10) MR_S + 4.36 (2) \\
n = 135 \quad r = 0.625 \quad s = 0.84 \quad F = 16.6 \\
&\text{In spite of the complexity of the full mechanism of inhibition involving at least six transition states and five distinct intermediates, a significant linear regression equation was calculated for } Kᵢ \text{ (Eq. 3):} \\
pKᵢ = 3.46(±0.30) \sigma(R²) - 1.66(±0.15) l(C₂) - 0.72(±0.15) l(2) + 0.67(±0.07) MR_S - 0.77(±0.12) MR_R + 7.21 (3) \\
n = 162 \quad r = 0.814 \quad s = 0.71 \quad F = 75.6 \\
&\text{Since the crystal structure of the acyl-enzyme complex, the acylation and deacylation rate were available, it was concluded that the inhibition begins with the histidine 57 catalyzed attack of serine 195 O₆ at the benzoxazinone C₄, while the carbonyl oxygen occupies the oxygen hole formed by glycine 194 and serine 195.}
\]

1990/258

Title: Quantitative structure-activity relationships of antibacterial compounds based on the nalidixic acid structure.
Authors: Block*, J.H.; Yu, Y.; King, J.W.; Veerloop, A.
College of Pharmacy, Oregon State University
Corvallis OR 97331 - 3507, USA.
Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.301 - 353. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.
Compounds:
\begin{enumerate}
\item a) 31 Compounds of type I, where R = H, CONH₂, SO₂Me, CONHC₆H₅CH₂CH₃, COOMe, NO₂, SO₂Ph, NH₂, (CH₂)₅CH₃, O, NMe₂, N(CH₂)₅CH₃CH₃, OH, OMe, OEt, O-i-Pr, O-n-Pr, OPh, O-ν-Bu, OHex, OBz, F, Cl, Br, I, SEt, 5-n-Pr, CH₂CH₂Ph, SPh;
\item b) Trimetroprim, methotrexate.
\end{enumerate}

\[
\text{Biological material: Dihydrofolate reductase (DHFR) from Escherichia coli, Streptococcus aureus, bovine liver.}
\]

Data determined:
\begin{enumerate}
\item δ₁V₈, δ₁H₄ [¹H NMR chemical shifts (ppm) of the 2-NH₂ or 4-NH₂ amino groups of the compounds groups (FB = free base, HCl = hydrochloride salt)];
\item MIC [minimum inhibitory concentration (units of 10⁻⁶ mol) of the test compounds required to inhibit bacterial growth];
\item I₅₀ [concentration of the compound (units of 10⁻⁶ mol) required for 50 % inhibition of DHFR].
\end{enumerate}

Chemical descriptors:
\begin{enumerate}
\item p\logP (logarithm of the partition coefficient in 1-octanol/water).
\end{enumerate}

Results: It was shown earlier that binding of diaminooquinazolines to DHFR correlated with the torsional angle of the 4-amino group of the quinazoline nucleus. It was postulated that the interaction between the adjacent 5-substituent and the 4-amino group was very important in determining DHFR binding of the compounds possibly, because of the influence on the hydrogen bond formed between the 4-amino group and a residue at the active site. The existence of such interaction in 5-substituted 2,4-diaminoquinazolines were shown by measuring δ₁V₈ and δ₁H₄ values. The σ₁ and σ₉ₑlectron parameters correlated well with chemical shifts of the 2-NH₂ groups (Eq. 1) but showed poor correlation for the 4-NH₂ group (Eq. 2), respectively:
\begin{align*}
\delta₁V₈ &= 0.55(±0.09) \sigma₁ + 0.46(±0.10) \sigma₉ₑ + 5.97 \quad (1) \\
n &= 8 \quad r = 0.97 \quad s = 0.06 \quad F = 39.0 \\
\delta₁H₄ &= 0.24(±0.40) \sigma₉ₑ + 0.49(±0.43) \sigma₉ₑ + 8.77 \quad (2) \\
n &= 8 \quad r = 0.56 \quad s = 0.26 \quad F = 1.1
\end{align*}

The equations suggest that the through-ring resonance interactions between the 5-substituent and the adjacent 4-amino group are disrupted by some other effects which might have significance for binding.

1990/259

Title: Antifolate and antibacterial activities of 5-substituted 2,4-diaminoquinazolines.
Authors: Harris, N.V.; Smith, C.; Bowden, K.
Rhone-Poulenc Ltd., Dagenham Research Centre
Rainham Road South, Dagenham, Essex, England.
Source: J. Med. Chem. 1990, 33(1), 434 - 444.
Compounds:
\begin{enumerate}
\item a) An extensive set of compounds based on the nalidixic acid structure.
\item b) Subset of (I) (set A) containing fifty two 6,7-disubstituted 1-alkyl-1,4-dihydro-4-oxoquinoline-3-carboxylic acids;
\end{enumerate}
c) Subset of (I) (set B) containing one hundred and sixty two
1,6,7-trisubstituted-1,4-dihydro-4-oxo-1,8-naphthyridine-3-carboxylic acids;

d) Subset of (I) (set C) containing eighty five 1,4-dihydro-4-oxo-
1,8-naphthyridine-3-carboxylic acids with substituted azetidinyl,
pyrrolidinyl and piperidinyl rings at position 7, fluorine at position 6 and ethyl, vinyl or 2-fluoroethyl substituent at position 1.

Biological material: Ps. aeruginosa V-1, E. coli NIH JC-2, S. aureus 209P.

Data taken from the literature:

MIC [minimum inhibitory concentration (µg/mL) of the test compound against Ps. aeruginosa, E. coli and S. aureus].

Chemical descriptors:

logP (logarithm of the partition coefficient in 1-octanol/water estimated using CLOGP program);

r (Hansch-Fujita’s substituent constant characterizing hydrophobicity);

FR (calculated lipophilicity of the substituent);

L, B1, B2 (STERIMOL steric parameters (Å⁻¹));

MR (molar refractivity);

F (Swain-Lupton’s electronic parameter, characterizing the field effect);

Ea (Taft’s constant, characterizing steric effects of the substituent);

I7 (indicator variable 1 for R7 is other than hydrogen);

IF, IBr, IS, ICO, INO, ICI, IM, IE (indicator variable 1 for the presence of Br, SMe, COMe, CN, NO₂, Cl, Me or Et respectively);

l(7NCO) (indicator variable 1 for the presence of carbonyl function on the 7-N heteroatom);

R11, R12 (indicator variable 1 for the presence of -N(CH₂CH₂)₂NH or -N(CH₃)₂ ring, respectively);

API (σ⁻⁻⁻ electronic potential interactions between position 6 and 7);

B7R7 (B7 STERIMOL parameter of the substituent of the R7 ring in set C).

Results: The original aim of developing a unified QSAR model for the antimicrobial activity of the compounds for Ps. aeruginosa, E. coli, S. aureus could not be achieved. However, significant regression equations were calculated for the MIC values of various subsets of compounds A mixed model using the physicochemical parameters and Free-Wilson indicator variables was also examined [Eq. 1 (set A, S. aureus), Eq. 2 (set A, P. aeruginosa), Eq. 3 (set B, E. coli), Eq. 4 (set C, P. aeruginosa)]:

\[
\log(1/\text{MIC}) = 0.699(±0.242) \text{FR}(6) + 1.116(±0.220) \text{MR}(6) + 0.812(±0.119) \text{FR}(7) - 0.314(±0.073) \text{FR}(7)^2 + 0.899(±0.194) \text{MR}(7) - 0.103(±0.031) \text{MR}(7)^2 + 0.549(±0.319) \text{FR}(7) \\
\text{n} = 43 \quad r = 0.878 \quad s = 0.419 \quad F = 20.13
\]

\[
\log(1/\text{MIC}) = -0.524(±0.170) \text{IF}(1) - 0.904(±0.371) \text{B1}(6) + 0.239(±0.112) \text{FR}(7) - 0.259(±0.069) \text{FR}(7)^2 + 0.584(±0.186) \text{MR}(7) - 0.081(±0.028) \text{MR}(7)^2 + 0.479(±0.181) \text{R11}(7) + 2.082(±0.176) \\
\text{n} = 41 \quad r = 0.872 \quad s = 0.375 \quad F = 14.96
\]

\[
\log(1/\text{MIC}) = 1.278(±0.181) \text{IF}(6) - 1.262(±0.191) \text{R12}(7) + 1.992(±0.119) \text{FR}(7) \\
\text{n} = 22 \quad r = 0.928 \quad s = 0.389 \quad F = 58.53
\]

The study showed that the most active compounds have fluorine in position 6, R7 can be a wide variety of nitrogen containing substituent and the best predictor for R7 is its lipophilicity.

Title: Lipophilicity-antifungal activity relationships for some isoflavonoid phytoalexins.

Authors: Arnold*, A.; Melirini, L.

Dipartimento di Scienze Molecolari Agroalimentari, Sezione di Chimica, Universita di Milano
Via Celoria 2, I-20133 Milano, Italy.

Source: J. Agric. Food Chem. 1990, 38(3), 834-838.

Compounds: 17 Phytoalexins: pisatin, 3,6a-dihydroxy-8,9-(methylenedioxypertocarpan, 6a,11a-dehydropisatin, 3-hydroxy-8,9-(methylenedioxypertocarpan, 6a,11a-dehydropisatin, 3-hydroxy-8,9-methyloxypertocarpan, (+)-3-hydroxy-9-methyloxypertocarpan, (-)-3-hydroxy-9-methyloxypertocarpan, vestitol, sativan, formonenetin, coumestrin, 4′-O-methylcoumestro1, phaseollinisoflavan, 2′-methoxyphaseollin-isoallavan, glyceollin, 6a,11a-dehydroglyceollin, 5-tuberis, 6,11a-dehydrotuberisorin.

Biological material:

a) Aphanomyces euteiches;

b) Fusarium solani f.s. tuberisorbin.

Data taken from the literature:

I (inhibition of growth (%)) of the fungi caused by 0.1 mol/L solution of the test compound.

Data determined:

logP (logarithm of the partition coefficient in 1-octanol/water determined by reversed-phase HPLC);

k’ (capacity factor determined RP-HPLC).

Results: A highly significant linear regression equation was calculated for logP of six reference compounds using their k’ values (Eq. 1):

\[
\log P = 1.79 \log k' + 2.18 \quad (1)
\]

The lipophilicity of the phytoalexins were within the range of log P = 1.5 - 4.2. It was found that the antifungal activity of similar compounds positively correlated with antifungal activity but no equation could be calculated for the whole set of compounds. It was suggested, however, that compounds with logP values higher than 3.5 were retained in the membranes, therefore phytoalexins with slightly lower lipophilicity, as well as greater fungitoxicity and systemic activity should be searched. Certain structural features seemed to correlate with antifungal activity such as the presence of phenolic OH and benzyl hydrogen.

Title: Design, synthesis, testing and quantitative structure-activity relationship analysis of substituted salicylaldehyde Schiff bases of 1-amino-3-hydroxyguanidine tosylate as new antiviral agents against coronavirus.

Authors: Wang, P.-H.; Keck, J.G.; Lien*, E.J.; Lai, M.M.C.

Section of Biomedical Chemistry, School of Pharmacy, University of Southern California
Los Angeles CA 90033, USA.

Source: J. Med. Chem. 1990, 33(2), 608 – 614.
Compounds: 15 Compounds of type $\text{RCH=NNH} (=\text{NH}) \text{NOHCH}_2\text{C}_6\text{H}_4\text{SO}_3\text{H}$, where R groups are various aromatic moieties.

Biological material: Mouse hepatitis virus (MHV)(type coronavirus).

Data determined:

$\text{TCID}_{50}$ [concentration of the test compound (mol/L)] required for the inhibition of MHV replication.

Chemical descriptors:

- $\pi$ (Hansch-Fujita's substituent constant characterizing hydrophobicity);
- $\Sigma \upsilon_{3,5}$ [sum of the Hansch-Fujita's substituent constants of the substituents in 3- and 5-position of the R aromatic group];
- $\upsilon_{1}$ [measure of lipophilicity ($\upsilon_{1} = \log (1/R_{T} - 1)$ where $R_{T}$ is the retention factor) obtained by reversed-phase thin-layer chromatography];
- $\mu$ (molar refractivity);
- $\Delta V_w$ [difference of the van der Waals volume of the compound and a reference compound ($R = 2\cdot\text{HO-Ph}$)];
- $\sigma_m$ (Hammett's constants, characterizing the electron-withdrawing power of the substituent in meta-position);
- $E_C^*$ (corrected Taft's constant, characterizing steric effects of the substituent);
- $\mu_R$ [dipole moment (debye) of R];
- $\Sigma \mu_{3,5}$ (vector summation of group dipole moments of the groups in meta-positions).

Results: A highly significant linear regression equation was calculated for a subset of the compounds without ortho hydroxy group (Eq. 1):

$$\log(1/\text{TCID}_{50}) = 0.29(\pm0.19) \pi_R + 3.82(\pm0.52)$$  \hspace{1cm} (1)

Further significant linear regression equations were calculated for a larger subset and for the complete set of guanidine derivatives (Eq. 2 and Eq. 3, respectively):

$$\log(1/\text{TCID}_{50}) = 0.67(\pm0.31) \pi_{3,5} - 0.34(\pm0.50) \Sigma \sigma_m + 4.30(\pm0.32)$$  \hspace{1cm} (2)

$$\log(1/\text{TCID}_{50}) = -0.42(\pm0.38) \mu_R - 1.56(\pm1.59) \upsilon_{1} + 4.32(\pm1.55)$$  \hspace{1cm} (3)

It was suggested that the ability of the ortho OH group to form fairly stable intramolecular hydrogen bond may contribute to the greater stability of the Shiff base functional group and the higher biological activity of the substances (various subsets required different equations). Results showed that compounds with increasing lipophilicity and electron donating substituents at the 3- and 5-positions have high inhibitory activity.

**Biological material:**

- Rabbits;
- Rats;
- Guinea pig.

Data taken from the literature:

- $\text{RPS}$ (relative platelet stimulation, $R_{50} = EC_{50} C_{18} \text{PAF/EC}_{50}$ analogue).
- $\text{PRP}$ (platelet aggregation (M) in platelet-rich plasma (PRP));
- $\text{EC}_{50}$ (dose (M/kg) required to lower the arterial diastolic blood pressure);
- $\text{EC}_{50}$ (dose (M/kg) inducing bronchoconstriction of 3 cm H$_2$O);
- $\text{EC}_{50}$ (dose (M/kg) inducing 50 % thrombocytopenia).

Chemical descriptor:

- $\Sigma f$ (sum of the Rekker's constants of the etheroxid chain, characterizing hydrophobicity).

Results: $\text{RPS}$, $\text{EC}_{50}$, $\text{EC}_{50}$, $\text{EC}_{50}$ values were measured and presented for the C$_{12}$ PAF analogue and compared with that of other analogues. C$_{12}$ PAF analogue was less potent than the C$_{16}$ or C$_{18}$ PAF analogues and equivalent to C$_{12}$ PAF analogue, showing that the activity decreased with lipophilicity. A highly significant parabolic relationship was calculated between log($R_{50}$) and $\Sigma f$ (Eq. 1):

$$\log(\text{RPS}) = 1.22(\pm0.24) \Sigma f - 0.09(\pm0.02) (\Sigma f)^2 - 4.07(\pm0.54)$$  \hspace{1cm} (1)

The maximum activity was calculated $\Sigma f = 6.78$, this corresponds to the C$_{16}$ PAF.

1990/262

Title: QSAR and Molecular shape analysis of aryl-substituted alanine analogs as antigelling agents.

Authors: Yuan, C.J.; Hopfinger, A.J.; Johnson*, M.E.

Department of Medicinal Chemistry and Pharmacognosy, University of Illinois at Chicago

Box 6998 (MC 781), Chicago IL 60680, USA.

Source: J. Theor. Biol. 1989, 141(1), 41 – 52.

Compounds:

- 16 Substituted phenylalanine analogs, with the following substituents: H, 4-NH$_2$, o-F, m-F, 4-F, m-OCONH$_2$, p-NO$_2$, N-Me, 3-1, o-Cl, p-Cl, 3-NO$_2$, p-1, p-Br, $\alpha$-Me, $\beta$-Me;
- 11 Substituted tryptophane analogs, with the following substituents: H, 1-Me, 4-F, 5-F, 5-Me, 6-Me, 6-F, 7-Me, 5-OH, 5-MeO, 5-Br.

Biological material: Deoxyxygenated human sickle cell hemoglobin (HbS).

Data determined:

- RA [relative antigelling activity of the compound inhibiting polymerization of HbS defined as the slope of the $C_{50}$ versus inhibitor concentration plot, where $C_{50}$ is the equilibrium solubility of HbS];
- Molecular modeling (molecule structures were built using the molecular modeling program CHEMELAB II);
- Conformational analysis (energy minimization of the compounds were calculated using the free valence geometry energy minimization method);
- MSA (Molecular shape analysis according to Hopfinger was used to quantitatively compare the shape similarity of analogs in their minimum energy conformer states (within 8 kcal/mol of their global minimum energy) (Fig. 1 shows the superposition of the reference conformations of the phenylalanine and tryptophane analogues)).
Chemical descriptors:

- \( \log P \) (logarithm of the partition coefficient in 1-octanol/water);
- \( \pi, \pi' \) (Hansch-Fujita’s substituent constant characterizing hydrophobicity of a substituent on the aromatic ring and the hydrophobicity of the aromatic ring itself, respectively);
- \( V_0 \) (common overlap steric volumes (Å³) between pairs of superimposed molecules in a common low energy conformation);
- \( \mu, \mu_0 \) (dipole moment (Debye)) of the whole molecule and of the aromatic ring, respectively, calculated using the CNDO/2 method);

Quantum chemical indices (partial atomic charges calculated by the CNDOI2 method);

\( \Theta_1 - \Theta_4 \) (torsion angles (deg) (Fig. 1) rotated during the conformational analysis of the compounds).

Results: Significant parabolic regression equations were calculated for the antigelling activity of the phenylalanine and tryptophan analogues (Eq. 1 and Eq. 2, respectively):

\[
RA = 45(\pm10) \sum (\pi + \pi') + 1.1(\pm0.3) \mu^2 + 1.3(\pm0.5) V_0^{25} - Br - Phr = 213(\pm72)
\]

(1)

\[
RA = 4.7(\pm1.0) \mu^2 + 134(\pm15)
\]

(2)

The different QSAR for the phenylalanine and tryptophan analogues indicated that they interact with hemoglobin in different ways or at different sites. For the phenylalanine analogues the hydrophobicity of the side chain, the aromatic dipole moment and the steric overlap volume explained approximately 50%, 20% and 10% of the variance in antigelling activity, respectively. For the tryptophan analogues the square of the dipole moment or the steric overlap volume explained 70% or 60% of the variance in RA, respectively, being the two descriptors highly correlated. The results show that the tryptophan analogs have a relatively tight fit with the receptor site.

Title: S-Aryl (tetramethyl) iso thiouronium salts as possible antimicrobial agents, IV.

Authors: Tait, A.; Gamberini, G.; Giovannini, M.G.; Di Bella, M.

Dipartimento di Scienze Farmaceutiche, University of Modena, Italy.

Source: Aminonucleotides, 1989, 44(12), 1129–1140.

Compounds: 55 Compounds of type I, where \( R = H, Me, Et, OMe, Ob, COOH, COMe, COOMe, F, Cl, Br, I, NO_2; \) \( R^1 = H, Me, Et, OMe, OEt, COOH, COMe, COOMe, F, Cl, Br, I, NO_2, CN; \) \( R^2 = H, Me, Et, OMe, OEt, OH, COOH, COMe, COOMe, F, Cl, Br, I, CN, \)

NHCOMe, SO,NHMe, SO,NMe, \( R^3 \) = H, Cl; \( R^4 = H, Cl; X \) was not specified.

Biological material: Staphylococcus aureus and Streptococcus pyogenes.

Data taken from the literature:

- MIC (minimum inhibitory concentration (mol/L) of the compound against S. aureus and S. pyogenes);
- log(1/C) (antimicrobial activity of the compound (mol/L) against S. aureus and S. pyogenes (details not given)).

Data determined:

- logKw (capacity factor of the compound determined by HPLC).
- Chemical descriptors:
  - logP (logarithm of the partition coefficient in 1-octanol/water);
  - \( \pi \) (Hansch-Fujita’s substituent constant characterizing hydrophobicity);
  - \( \Sigma_{\chi_{obsd}} \) (hydrophobicity of the compound calculated determined experimentally in 1-octanol/water);
  - \( \sigma^+ \) (Hammett polar substituent constant, characterizing the electron-withdrawing power of a substituent which has resonance interaction with the reaction center).

Results: Highly significant linear regression equations were calculated for logKw and \( \Sigma_{\chi_{obsd}} \) (Eq. 1, Eq. 2):

\[
\log K_w = 0.885(\pm0.08) \Sigma_{\chi_{obsd}} + 0.874(\pm0.06)
\]

(1)

\[
\Sigma_{\chi_{obsd}} = 0.710(\pm0.079) \Sigma \pi - 0.158(\pm0.128) \Sigma \sigma^+ + 0.136(\pm0.071)
\]

(2)

Significant linear regression equations were calculated for the log(1/C) values of S. aureus and S. pyogenes the compounds (Eq. 3, Eq. 4, respectively):

\[
\log(1/C) = 0.465(\pm0.262) \Sigma \sigma^+ + 0.303(\pm0.186) \Sigma_{\chi_{obsd}} + 3.355(\pm0.180)
\]

(3)

\[
\log(1/C) = 0.520(\pm0.392) \Sigma \sigma^+ + 0.314(\pm0.269) \Sigma_{\chi_{obsd}} + 3.330(\pm0.258)
\]

(4)

In both Eq. 3 and Eq. 4, log(1/C) depended primarily on electronic factors (\( \Sigma \sigma^+ \)) and only secondarily on hydrophobicity (\( \Sigma_{\chi_{obsd}} \)). A threshold logP value for the active iso thiuronium salts was indicated, as the compounds with logP values between -0.70 and -1.58 were found to be totally inactive with the exception of the nitro-derivatives.

Correlation Analysis, Application: Agricultural Chemistry

Title: Comparative QSAR study of the chitin synthesis inhibitory activity of benzoyl-ureas versus benzoyl-biurets.
Authors: Bordás*, B.; DeMilo, A.B.; Lopata, A.; Haught, S.B.

Department of Chemistry, Plant Protection Institute, Budapest, Hungarian Academy of Sciences

POB 102, H-1525 Budapest, Hungary.

Source: Tagungsbericht 1989, No.274. In: Insecticides - mechanisms of action and resistance p.157 - 165. Akad. d. Landwirtschaftswiss. d. DDR, 1989.

Compounds:
a) 18 Compounds of type I, where

\[ R' = \text{H, F, CI; } R^1 = \text{H, F, CI; } R^2 = \text{H, Me, CF} \]

b) 18 Compounds of type II, where

\[ R^1 = \text{H, F, CI; } R^2 = \text{H, F, CI; } R^3 = \text{H, Cl, CF} \]

Biological material: 8 Insect species: Aedes aegypti, Musca domestica, Chilo suppressalis, Hylemya platura, Oncopeltus suppressalis, Oncopeltus fasciatus, Pieris brassicae, Leptinotarsa decemlineata.

Data taken from the literature:

\[ \text{LC}_{50} \text{ [concentration of the biuret analogue (ppm) required to kill 50% of insect larvae (A. aegypti, M. domestica, C. suppressalis, H. platura, } \]

\[ O. suppressalis, O. fasciatus, P. brassicae, L. decemlineata]. \]

Data determined:

\[ \text{LC}_{50} \text{ [concentration of the biuret analogue (ppm) required to kill 50% of insect larvae (A. aegypti or M. domestica).} \]

Molecular modeling (models of the compounds were built using MOLIDEA).

Conformational analysis (minimum energy conformations of the compounds were calculated using molecular mechanics method).

Chemical descriptors:

\[ \pi \text{ (Hansch-Fujita's substituent constant characterizing hydrophobicity); } \]

\[ \sigma_p \text{ (Hammett's constant, characterizing the electron-withdrawing power of the substituent in para-position); } \]

\[ F \text{ (Swain-Lupton's electronic parameter, characterizing the field effect); } \]

\[ MR \text{ (molar refractivity); } \]

\[ B_6 \text{ (Verloop's STERIMOL parameter representing the maximum width of the substituent perpendicular to its } \alpha \text{-atom). } \]

Results: Highly significant linear regression equations were calculated for \( \text{LC}_{50} \) values of the biuret analogues (I) against A. aegypti, M. domestica (\( n = 7, r = 0.993, F = 37.3 \)) and \( n = 7, r = 0.997, F = 88.05 \)). Further 9 significant linear or parabolic regression equations were calculated for the retrospective \( \text{LC}_{50} \) values of the urea analogues (II) against 8 insect species (\( r \) values ranging from 0.973 to 0.999). The equations had very similar form in terms of descriptors and the sign of the regression coefficients. Principal component analysis of the activity data matrix of the biological activity data of the biuret and urea derivatives showed high intercorrelation with principal components explaining 69.61 %, 19.02 % and 9.30 % of the variance. Fig. 1 shows the minimum energy conformation of a highly active representative of the urea analogs (dimilin) with 5.6 Å distance between the 1 and 15 carbon atoms. Fig. 2 shows the low energy conformation of the corresponding biuret analog with the two benzene rings in approximately the same plane and with the same C1-C18 distance (5.6 Å) allowing to fit a hypothetical benzoylurea pharmacophore.

The similarity of the regression equations and the modelling study supported the hypothesis that the benzoylbiurets act by the same mechanism as the benzoylureas.

1990/266

Title: Quantitative structure-activity analyses and mode of action studies of phenolic uncouplers.

Author: Miyoshi, H.

Department of Agricultural Chemistry, Kyoto University Kyoto 606, Japan.

Source: Doctoral thesis, Department of Agricultural Chemistry, Kyoto University, 1989.

Results: The thesis is devoted to the quantitative analysis of the uncoupling activity of substituted phenols using chemical descriptors in order to obtain further information on the mode of action of phenol uncouplers:

a) The study of the partition coefficient of substituted phenols in liposome/water system \([\text{PL/W}]\) showed that (i) \([\text{PL/W}]\) depended primarily on the logP value; (ii) influence of steric and electronic parameters depended on the type of the lipid involved;

b) QSAR analysis of uncoupling phenols in rat-liver mitochondria identified the relevant physicochemical parameters required for phenols being protonophore in inner mitochondrial membrane and quantitatively separated the potency as the protonophore in the inner mitochondrial membrane and the incorporation factor (logP);

c) Protonophoric potency of substituted phenols was linearly related to uncoupling activity when certain critical physicochemical parameters of the experiment were taken into account;

d) Linear relationship was calculated between uncoupling activities of substituted phenols and related uncouplers in the mitochondria from the flight muscles of house flies and in spinach chloroplasts;

e) The results indicated a shuttle type mechanism for the uncoupling action of substituted phenols.

1990/267

Title: Uncoupling properties of a chlorophenol series on Acer cell suspensions. A QSAR study.

Authors: Ravanel*, P.; Taillandier, G.; Tissut, M.

Laboratoire de Physiologie cellulaire végétale, Université Joseph Fourier

BP 53X, F-38041 Grenoble cédex, France.

Source: Ecotoxicol. Environ. Safety 1989, 18(3), 337 - 345.
Compounds: 22 Chlorinated phenols substituted with 2-Cl, 3-Cl, 4-Cl, 2,3-Cl, 2,4-Cl, 2,5-Cl, 3,4-Cl, 3,5-Cl, 2,3,6-Cl, 2,4,5-Cl, 2,4,6-Cl, pentachlorophenol, 4-Cl-2-Me, 4-Cl-3-Me, 4-Cl-2,3-Me, 4-Cl-3,5-Me, 4-Cl-2-allyl, 4-Cl-2-Pr-5-Me, 2-Cl-6-NO₂, 2,4-Cl-6-NO₂, 2-Cl-4,6-NO₂.

Biological material: Acer pseudoplatanus L. cell suspensions.

Data determined:
\[ D_{50} = \frac{\text{concentration of the compound (µmol/L) required for 50% uncoupling effect registered by measuring the oxygen consumption rate by polarography}}{2} \]

\[ D_{100} = \frac{\text{minimal concentration of the compound (µmol/L) required for giving a full uncoupling effect}}{2} \]

Chemical descriptors:
\[ \log \left( \frac{l}{D_{50}} \right) = 1.813(±2.145) \log P - 0.160(±1.349) \log \Sigma D - 0.792(±1.820) A + 0.744(±3.069) a_1 + 1.2606 \]

\[ n = 22 \quad r = 0.832 \quad s = 0.388 \quad F = 9.53 \]

\[ \log (1/D_{100}) = 0.466(±0.1014) \Sigma D - 1.843(±5.318) A + 0.202(±2.049) a_2 - 4.250 \]

\[ n = 18 \quad r = 0.939 \quad s = 0.222 \quad F = 34.74 \]

The equations for the uncoupling effects in the whole cells and those calculated previously for isolated mitochondria or chloroplasts possess similar structures.

1990/268

Title: Effects of 3' substituents on diphenyl ether compounds.

Authors: Yoshimoto*, T.; Funakoshi, Y.; Fujita, Takashi; Enomoto, Y.; Okano, K.

Life Science Business Development Department, Mitsui Toatsu Chemicals, Inc.

Chiyoda-ku, Tokyo 100, Japan.

Source: J. Pestic. Sci. 1989, 14(4), 465 - 474.

Compounds: 50 Various diphenyl ether derivatives.

Biological material: Echinocloa crus-galli.

Data determined:
\[ \text{BA} \quad \text{[biological activity, the herbicidal activity of the compound (M) required for 90% inhibition in test tube]} \]

\[ \text{POT, FIE} \quad \text{[herbicidal activity of the compound (M) in greenhouse and outdoor potted plant tests, respectively]} \]

Chemical descriptors:
\[ \sigma_n, \sigma_p \quad \text{(Hammett's constants, characterizing the electron-withdrawing power of the substituent in meta- and para-position, respectively)} \]

\[ \text{C}^3 \quad \text{[water solubility (mol/m³) of the compound]} \]

\[ \text{GRT} \quad \text{[retention time (min) measured by gas chromatography]} \]

\[ \text{Rm} \quad \text{[measure of lipophilicity (Rm = log (I/RF - 1) where RF is the retention factor) obtained by reversed-phase thin-layer chromatography]} \]

Results: The influence of the 3' substituents (ortho position in the phenyl ring) on the chemical structure and biological activity was studied. Relationships are presented between the biological activities and the physicochemical properties, leading on the other hand, no significant linear relationships. The highest correlation was found between POT and FIE, r = 0.927. A poor equation was calculated (r = 0.487) for the relationship between FIE and BA for compounds having N in the substituent. The relationship could not be improved by introducing Rm or GRT. The multiple correlation coefficient markedly improved introducing C^3 into the equation (Eq. 1):

\[ \text{FIE} = 0.298 \text{BA} - 0.218 \text{C}^3 + 0.028 \quad (1) \]

Based on Eq. 1, outdoor activity can be predicted by the biological activity and water solubility for diphenyl ethers containing N atom in the substituents.

1990/269

Title: Substituent effect on photosynthesis of N²-α-substituted benzyl-N⁴-alkyl-2,4-diamino-6-chloro-s-triazines and their phytotoxic property.

Authors: Omokawa, H.; Kobayashi*, I.; Konnai, M.

Central Research Laboratories, Idemitsu Kosan Co., Ltd.

Sodegaura-cho, Chiba 299-02, Japan.

Source: Agric. Biol. Chem. 1989, 53(10), 2723 - 2729.

Compounds:

a) 29 Compounds of type I, where R'¹ = H, Me, Et, Pr, Bu, Pent, Hex, 1-Me-Et, 1-Me-Pr, 2-Me-Pr, 1-Et-Pr, 1-Me-Bu, 1-Me-Pent, 1-Me-Hex, 1,1-Me₂-Et, 1,2-Me₂-Pr, 2,2-Me₂-Pr, 1,3-Me₂-Bu, c-Pr, c-Pen, c-Hex; R² = Me, Et, Pr, 1-Pr; N in the substituent.

b) Atrazine.

Biological material:

a) [Barnyard millet (Echinochloa crusgalli Beauv. var. frumentacea Trin.)];

b) (Plant preparations of E. crusgalli: abaxial epidermis peeled leaf disks, mesophile cells, intact and broken chloroplasts).

Data determined:
\[ \text{pI} = \text{[negative logarithm of the concentration of the compound (mol/L) required for 50% inhibition of the above-ground part of the plant]} \]

\[ \text{pI} = \text{[negative logarithm of the concentration (mol/L) of the compound required for 50% inhibition of the rate of oxygen evolution by the plant preparation under illumination (inhibition of photosynthesis)]} \]

\[ \text{ΔpI} = \text{[difference of the pI values of the compound pairs with R¹ = Me and R² = H]} \]

Chemical descriptors:

\[ n_c \quad \text{[Number of the carbon atoms in the R² alkyl chain (R² = H)} \]

Results: SAR suggested that the space for the N' and NZ substituents of the binding site is relatively large. The variation of the number of the carbon atoms of R² alkyl chain (R² = H) is studied. The change in phytotoxic activity of the compound towards the whole plant correlated with their permeability through the plant membranes.
Title: Inhibition of susceptible and resistant green rice leafhopper acetylcholinesterase by N-methylcarbamate and oxadiazolone insecticides.

Authors: Ohta*, H.; Kyomura, N.; Takahashi, Y.; Magee, P.S. Mitsubishi Kasei Corporation

Source: ACS Symposium Series 1989, No. 413, in Probing Bioactive Mechanisms p.136-146. edited by Magee, P.S., Henry, D.R., Yokohama, Japan.

Compounds:
- a) 20 N-methylcarbamates of type I, where R = Me, Et, i-Pr, s-Bu, CH2=CH=CH2, OMe, OEt, O-i-Pr, O-S-Bu, OCH2CH=CH2, OCH2C-CH, Cl, Br, NH-i-Pr, NMe2, N(CH2CH=CH2), CN, Ph, CH=NOMe;
- b) 81 Oxadiazolones of type II, where R1 = Me, Et; R2 = H, Cl, Br, Me, OMe, OEt, O-i-Pr, O-S-Bu; R3 = H, i-Pr, s-Bu, t-Bu, Cl, OMe, NO2, diCl; R4 = H, Me, F, Cl, Et, NO2, CF3; R5 = = H, Me, Et, i-Pr, s-Bu, t-Bu, F, Cl, diCl, OMe, NO2, CF3, SMe.

Chemical descriptors:
- \( \pi \) (Hansch-Fujita’s substituent constant characterizing hydrophobicity);
- \( \sigma_p \) (Hammett’s constant, characterizing the electron-withdrawing power of the substituent in para-position);
- \( \sigma_i \) (Charton’s electronic constant, characterizing the electron-withdrawing power of the substituent);
- \( \pi R \) (sigma resonance value calculated by the formula \( \sigma_R = \sigma_P - \sigma_I \));
- \( \nu_0 \) (steric parameter introduced by Charton);
- IOR [indicator variable 1 for R1 = OEt and 0 for R1 = Me in (II)];
- HB [indicator variable 1 for R2 = RO in (II)];
- IX0 [indicator variable 1 for R2 halogen substituent].

Results: Similar linear regression equations were calculated for the \( p_{190} \) values of the N-methylcarbamates against (R)-AChE and (S)-AChE. The difference between the correlation coefficients and the electronic parameters indicate, however, a shift in mechanism (Eq. 1, Eq. 2):

\[
p_{190}(S) = 0.34(\pm 0.35) \pi + 1.65(\pm 0.87) \nu_0 - 0.74(\pm 0.51) \sigma_R + 3.74(\pm 0.45) \\

n = 19 \quad r = 0.900 \quad s = 0.320 \quad F = 21.62
\]

\[
p_{190}(R) = 0.56(\pm 0.23) \pi + 0.99(\pm 0.59) \nu_0 + 1.63(\pm 0.68) \sigma_I + 2.57(\pm 0.36) \\

n = 20 \quad r = 0.916 \quad s = 0.233 \quad F = 27.66
\]

Two very similar linear regression equations were calculated for the \( p_{190}(S) \) and \( p_{190}(R) \) values of S- and R-AChE (Eq. 3, Eq. 4):

\[
p_{190}(S) = -1.50(\pm 0.69) \nu_0 - 0.55(\pm 0.49) \Sigma + 1.66(\pm 0.65) \Sigma 1X0 - 0.32(\pm 0.29) IOR + 0.41(\pm 0.43) HB + 5.62 \\

n = 81 \quad r = 0.608 \quad s = 0.627 \quad F = 8.80
\]

\[
p_{190}(R) = -0.83(\pm 0.64) \nu_0 - 0.82(\pm 0.46) \Sigma + 1.99(\pm 0.61) \Sigma 1X0 + 0.56(\pm 0.28) IOR + 0.77(\pm 0.41) HB + 7.77 \\

n = 81 \quad r = 0.694 \quad s = 0.494 \quad F = 13.94
\]

A comparison of the two sets of equations (Eq. 1, Eq. 2 and Eq. 3, Eq. 4) shows that the inhibitors of type (I) and (II) bind differently at the target sites of AChE. The equations for the (S)- and (R)-AChE indicate that the resistance mechanism of green rice leafhopper is the result of a modified AChE receptor site.

Title: Structure-activity relationships in diphenyl ethers for insect growth regulating activity against mosquitoes.

Authors: George, N.; Vasuki, V.; Kalyanasundaram, M.

Source: Indian J. Med. Res. 1989, 89(Sept), 344-349.

Compounds: 30 Substituted diphenyl esters (DPEs) and A-23, methoprene (standard).

Biological material: Three vector species of mosquitoes: Culex quinquefasciatus, Aedes aegyti, Anopheles stephensi.

Data determined:
- IGR [insect growth regulating activity (% of inhibition)];
- LC90, LCw [concentration (mg/L) of the compound required to kill 50 and 90% of the test organisms, respectively];
- EL50, ELw [concentration of the compound (mg/L) required for 50 and 90% inhibition in adult emergence].

Chemical descriptor:
- logP (logarithm of the partition coefficient in 1-octanol/water).

Results: The biological activity of three out of the 30 (DPE-16, 19 and 28) substituted diphenyl esters were measured and listed. IGR values were measured for the three compounds and compared with that of A-23 and Methoprene. It was found that the position of acetamido group in the phenol moiety when it is in the ortho position (DPE-16)
increases the lipophilicity of the compound with a logP value of 2.54. If the same group is in meta or para position, the logP values are 2.16 and 1.99, respectively and they are comparatively ineffective. When both the ortho positions are substituted with tertiary butyl groups (DPE-28) the logP value is relatively higher (3.30) which increases the lipophilicity of the compound and explains the pronounced IDR activity at relatively low concentrations.

1990/272

Title: Quantitative structure-activity relationship study of aromatic trifluoromethyl ketones. In vitro inhibitors of insect juvenile hormone esterase.

Authors: Székács*, A.; Bordás, B.; Matolcsy, G.; Hammock, B.D. Department of Entomology and Department of Environmental Toxicology, University of California Davis CA 95616, USA.

Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.169 – 183. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

Compounds:

41 Substituted arylthio-trifluoropropanones of type I, where R

Chemical descriptors:

\[ r \] (Hansch-Fujita’s substituent constant characterizing hydrophobicity);

\[ \sigma_m, \sigma_p \] (Hammett’s constant, characterizing the electron-withdrawing power of the substituent in meta or in para position, respectively);

MR (molar refractivity);

H-Do, H-Ac (indicator variable 1 for hydrogen donor or hydrogen acceptor substituent, respectively).

Results: A highly significant linear regression equation was calculated for \( \text{pI}_50 \) using stepwise regression analysis (Eq. 1):

\[
\text{pI}_50 = 0.960(\pm 0.296) \text{H-Do} + 0.165(\pm 0.034) \text{MR} - 1.633(\pm 0.529) \sigma_m - 0.608(\pm 0.300) \pi_n + 0.117(\pm 0.026) \text{MR}_n + 0.115(\pm 0.064) \text{MR}_v + 0.984(\pm 0.141) \Sigma \pi + 4.803
\]

The contribution of the individual variables to the regression model according to their partial r value was \( \Sigma \pi (0.754) > \text{MR}_n (0.627) > \text{MR} (0.403) > \pi_n (0.341) > \sigma_m (0.252) > \text{H-Do} (0.189) > \text{MR}_v (0.138) \).

Fig. 1 shows the plot of calculated versus measured \( \text{pI}_50 \) values of the compounds (Fig. 1):

Based on the regression model, a number of para-substituted compounds of type I have been predicted to possess a 2–3 magnitude higher JHE inhibitory potency than those involved in the study [e.g. \( R = 4-\text{PhC-CH}, 4-(\text{Ph}_2\text{N}), 4-\text{PhCHJ} \)].

1990/273

Title: Organosilane insecticides. 2. Chemistry and structure-activity relationships.

Authors: Sieburth*, S.McN.; Langevine, C.N.; Dardaris, D.M. Agricultural Chemical Group, FMC Corporation Princeton NJ 08543, USA.

Source: Pestic. Sci. 1990, 28(3), 309 – 319.

Compounds: 18 Compounds of type I, where \( R' = \text{EtO, MeO, i-Pr, H, Cl, Me, CF,; R}_2 = \text{Me, Et, c-Pr, CH-CH,, CF=CF,, F; Z = CH,, CO, NH, NCHO, Me, 0.} \)

Biological material:

a) Third instar southern armyworm (Spodoptera eridania);

b) Mexican bean beetle (Epilachna varivestis);

c) Pinto bean plants.

Data determined:

\( \text{LC}_{50} \) [Concentration of the compound (10 g/L a.i. in the spray mixture) required to kill 50 % of the insect placed on pinto bean in a 48 h experiment].

Chemical descriptors:

\[ x \] (Hansch-Fujita’s substituent constant characterizing hydrophobicity);

\( \sigma \) (Hammett’s constant, characterizing the electron-withdrawing power of the substituent);

MR (molar refractivity).

Fig. 1
Results: A highly significant linear regression equation was calculated for the descriptors of R' (R' = i-PrO was eliminated as an outlier) (Eq. 1):

$$\log(1/LC_{50}) = 0.155(\pm0.012) MR - 0.421(\pm0.145) \sigma - 3.78 \quad (1)$$

The compound with R' = EtO, R'' = Me and Z = O was found to be an effective, broad spectrum insecticide. The replacement of the quaternary carbon with a silicon atom can simplify the synthesis of test compounds and thus can be advantageously utilized for the preparation of large compound sets for QSAR studies.

**Correlation Analysis, Application:** Physical Organic Chemistry

1990/274

**Title:** Reactivity of biologically important reduced pyridines. 5. Relative importance of electron versus proton loss in ferricyanide-mediated oxidation of dihydroxycinnamamides.

**Authors:** Brewster, M.E.; Kaminski, J.J.; Gabanyi, Z.; Czako, K.; Simay, A., Bodor, N.

Center for Drug Design and Delivery, College of Pharmacy, Box J-497, J. Hillis Miller Health Center, University of Florida, Gainesville FL 32610, USA.

Source: Tetrahedron 1989, 45(14), 4395 - 4402.

**Compounds:** Compounds of type I where R = NH₂, H, F, Cl, Br, I, NO₂, CF₃, CH₃, C₂H₅, OCH₃, H(CH₂)₂.

The data suggest that the initial electron loss from the given compounds is the preeminent factor effecting the reaction rate. A single mechanism is suggested over the entire range of reactivities, where a transition state with a considerable positive charge is involved.

**Chemical descriptors:**
- logP (logarithm of hydrophobicity).
- Results: Calculations for electronoacceptor and electronodonor enthalpic and free energy factors on the base of functional groups were made according to the principle of independence of active centers:
  - \(\Delta H_e = |\Delta H_{i1}| E_i E_j\) (enthalpy of connection (kJ/mol));
  - \(\Delta G_e = |\Delta G_{i1}| C_i C_j\) (free energy of connection (kJ/mol));
  - \(\Delta S_e = |\Delta S_{i1}| C_i C_j\) (entropy of connection (kJ/mol/K));

Linear correlation was found between the calculated and measured characteristics:

$$\Delta H_e = -0.23(\pm0.65) + 0.99(\pm0.03) \Delta H_m \quad (3)$$

and

$$\Delta G_e = 0.06(\pm0.27) + 1.01(\pm0.02) \Delta G_m \quad (4)$$

The good linear correlations between the measured and calculated data show that the functional group approaches might be used for these compound types.

For Candida albicans the following equations were calculated:

$$\log(I/C) = 2.19 + 1.61 C_i$$

When the test material is Trychophyton gypseum var Kaufman-Wolf:

$$\log(I/C) = -0.01 - 4.82 E_i$$

The accuracy of the fitting was the same as the measurement error of \(\Delta H_m\) and \(\Delta G_m\). The entropy might be calculated from enthalpy, Gibbs energy and temperature:

$$\Delta S_e = -4.70(\pm2.40) + 0.88(\pm0.06) \Delta S_m \quad (5)$$

The good linear correlations between the measured and calculated data show that the functional group approaches might be used for these compound types.
For correlations calculated for B.Subtilis PCI219, and logP, the statistical characteristic were practically the same. Equations show the possibilities of connection between thermodynamic parameters and the biological activities.

1990/276

Title: Substituent effects on the electron affinities of perfluorobenzenes C,F,R.

Authors: Dillow, G.W.; Kebarle*, P.

Department of Chemistry, University of Alberta

Edmonton T6G 2G2, Canada.

Source: J. Am. Chem. Soc. 1989, 111(15), 5592 – 5596.

Compounds: C,F,R where R = F, Cl, C,F, C,F, COCH,F, CHO, CN, Br, I, NO,F, COC,F,F.

Chemical descriptors:

$\Delta G$ (the electron attachment free energy (kcal/mol) at 423 K);

$\Delta H$ (the electron attachment free enthalpy (kcal/mol) at 423 K);

EA [electron affinity (eV)].

Results: EA s of the given compounds have been measured with a pulsed electron high-pressure mass spectrometer in order to examine the substituent effects. Thus, the substituent effects for the $\sigma$-acceptor/$\pi$-donor substituents (F, Cl, Br, I) were found to be very much larger for the C,F,F relative to the nitrobenzenes. These results indicate that the extra electron enters a $\pi^*$-orbital (Fig. 1):

- $\log(S) = 1/3 \, \Delta G + (1/3)^2 \, \Delta G_N - 2.5075$
- $\Delta G \, \Delta G_N$

The substituent effects for the $\sigma$-acceptor/$\pi$-donor substituents (F, Cl, Br, I) were found to be very much larger for the C,F,F relative to the nitrobenzenes. These results indicate that the extra electron enters a $\pi^*$-orbital, which is localized on the C-R atoms.

1990/277

Title: Quantitative structure-activity relationships. 7. Influence of water solubility of alcohols of the general formula C,H$_{2n+1}$,OH on the para-substituent, the ortho-di-tert-butyl groups and the peroxide radical.

Authors: Héberger*, K.; Lopata, A.; Müller, J.

Central Research Institute for Chemistry, Hungarian Academy of Sciences

POB 17, H-1525, Budapest, Hungary.

Source: Int. J. Chem. Kinetics 1989, No.21, 1181-1193.

Compounds: 21 Phenols with H, NO,F, OH, Et, t-Bu, Br, Me, OMe, Cl, CN, COMe, O-T-Bu, Ph, COOH, NO,F, NO$_2$, CHO, COO-, t-Bu in para-position and with H, H or t-Bu, t-Bu in 2,6-di-ortho positions.

Results: A new linear regression model (Eq. 1) has been developed for the structure-solubility relationship of aliphatic alcohols. The study indicates that solubility of aliphatic alcohols depends primarily on molecular connectivity (v$_1$), the number of carbon atoms in the aliphatic chain (n), the number of hydrogens on the $\alpha$-carbon atom (normal, iso, tertiary, tertary) and the degree of branching (VG):

- $\log(S) = 1/3 \, \Delta G + (1/3)^2 \, \Delta G_N - 2.5075$

The result support Kier’s, furthermore Kier and Hall’s earlier models on the structural dependence of water solubility of alcohols.

1990/278

Title: Linear free energy relationships for peroxy radical-phenol reactions. Influence of the para-substituent, the ortho-di-tert-butyl group and the peroxide radical.

Authors: Héberger*, K.; Lopata, A.; Müller, J.

Central Research Institute for Chemistry, Hungarian Academy of Sciences

POB 17, H-1525, Budapest, Hungary.

Source: Int. J. Chem. Kinetics 1989, No.21, 1181-1193.

Compounds: 21 Phenols with H, NO,F, OH, Et, t-Bu, Br, Me, OMe, Cl, CN, COMe, O-T-Bu, Ph, COOH, NO,F, NO$_2$, CHO, COO-, t-Bu in para-position and with H, H or t-Bu, t-Bu in 2,6-di-ortho positions.

Results: Highly significant linear regression equations were calculated by stepwise regression analysis for logK in spite of the diverse data set originating from different laboratories using different peroxy radicals (Eq. 1, Eq. 2):
log \( k \) = \(-0.801\sigma_1 - 2.483\sigma_R + 3.766\) (1)
\( n = 32 \quad r = 0.851 \quad s = 0.429 \quad F = 38.0 \)

log \( k \) = \(-0.932\sigma_1 - 2.302\sigma^* + 3.802\) (2)
\( n = 32 \quad r = 0.848 \quad s = 0.432 \quad F = 37.2 \)

1Hα was not selected by stepwise regression indicating that the ortho-di-t-Bu substitution had no significant effect on the rate of hydrogen abstraction from phenols by the radicals. The form of the equations for different subsets of the phenols and radicals indicated that the reaction mechanism was the same for the different peroxy radicals.

**Correlation Analysis, Application:**

**Chromatography**

1990/279

Title: A fractal study of aliphatic compounds. A quantitative structure-property correlation through topological indices and bulk parameters.

Authors: Mishra*, B.K.; Guru, B.K.; Mishra, R.K.

Chemical Physics Group, Department of Chemistry, Sambalpur University

Jyoti Vihar, Burla 768 019, India.

Source: Indian J. Chem. 1989, 28A(Nov), 927-931.

Compounds: Short chain (1 ≤ n ≤ 10) and long chain (n > 10) alkanes and alcohols, acids, nitriles, ketones.

Data taken from the literature:

- \( V_p \) [vapor pressure (atm)];
- \( D_t \) [refractive index of the compound];
- \( p \) [density of the compound (dimension is not given)].

Data determined:

- \( V_w \) [van der Waals volume, calculated from the van der Waals radii of the atoms];
- MW [molecular weight];
- SDV [steric density of the functional group].

Results: Highly significant equations are presented for calculating \( V_w, SDW, MW \) and \( r \) values ranging from 0.92 to 1.00, other statistics and the number of investigations are not given. \( \alpha \) and \( \beta \) values calculated by these equations were introduced to the equation given above and the physicochemical properties were calculated. The observed and calculated \( V_p, D \) and \( p \) values are presented and compared for the alkanes, alcohols, acids and nitriles. The observed and calculated physicochemical parameters agreed well. Fractal nature of the alkyl chain length was discussed and a relationship was presented between the fractal-dimensioned alkyl chain length and a generalized topological index.

Correlation Analysis, Application: Chromatography

1990/280

Title: Application of micellar liquid chromatography to modeling of organic compounds by quantitative structure-activity relationships.

Further experiments using various surfactant types in the mobil phase suggested that logk' values generated on a lamellar phase may be better predictors of hydrophilicity than logP obtained from binary solvent systems.

Correlation Analysis, Application: Chromatography

1990/281

Title: Isoxazolinyldioxepins. 2. The partitioning characteristics and the complexing ability of some oxazolinyldioxepin diastereoisomers.

Authors: Camilleri*, P.; Munro, D.; Weaver, K.; Williams, D.J.; Rzepa, H.S.

Shell Research Limited, Sittingbourne Research Centre

Sittingbourne, Kent, ME9 8AG England.
Source: J. Chem. Soc. Perkin Trans. II 1989, No.11, 1935 – 1937.
Compounds: 10 Oxazolinyldioxepin derivatives of type I and II, where \( X = H, F, Cl, CF_3 \) or \( CH_3 \).

Data determined:
log\( k' \) [logarithm of the capacity factor, measured by reversed-phase liquid chromatography (RPLC)];
MEP (molecular electrostatic computed by Geesner-Prettre and Pullman's VSSPOT procedure).

Chemical descriptor:
log\( P \) (logarithm of the partition coefficient in 1-octanol/water).

Results: The log\( k' \) and log\( P \) values were measured for the two type of diastereomers and a highly significant linear relationship between log\( k' \) and log\( P \) was presented (\( r = 0.995 \)).

The complex forming ability of the diastereoisomers with mono-cations was investigated and explained in terms of the structures and electronic properties of the compounds.

1990/282

Title: Retention prediction of analytes in reversed-phase high-performance liquid chromatography based on molecular structure. III. Monosubstituted aliphatic compounds.

Authors: Smith*, R.M.; Burr, C.M.
Department of Chemistry, Loughborough University of Technology, Loughborough, Leics. LEI 3TU. England.

Source: J. Chromatogr. 1989, 481(Nov), 71-84.
Compounds: 25 various monosubstituted aliphatic compounds.

Data determined:
log\( k' \) [logarithm of the capacity factor, measured by reversed-phase liquid chromatography (RP-HPLC) in methanol and acetonitrile eluents];
\( I = I_{IP} + I_{IS,R} + I_{IS,Al-x} + I_{IS,AI-x} + I_{IS,Al-x} + I_{IS,Al-x} + I_{IS,Al-x} + I_{IS,Al-x} \) are the index contribution of the parent compound, a saturated alkyl chains, substituents on saturated aliphatic carbons, aromatic substituents and interactions between substituents, respectively.

Chemical descriptors:
\( F \) (fragmental constant according to Hansch and Leo characterizing hydrophobicity).

Results: log\( k' \), \( I \) values and the effect of the substituents as a retention index increment were determined by HPLC. Relationship between \( F \) and \( I \) was studied and an approximately linear relationship was found, \( r \) values ranging from 0.869 to 925 in methanol, from 0.860 to 0.924 in acetonitrile. Equations are not given. The developed system is suitable for predicting retentions in RP-HPLC based on the structure of the compound.

1990/283

Title: A chromatographic method for hydrophobicity investigations of sorbents.
Authors: Pasechnik*, V.A.; Solovyova, L.Ya.; Gorbunov, A.A.; Mitrofanov, Ye.V.; Karabanova, Ye.A.
All-Union Research Institute for Highly Pure Biopreparations 7 Pudozhsk, 197110 Leningrad, USSR.

Source: Chromatographia 1989, 28(Sept), 258 – 262.
Compounds: \( C_2-C_8 \) Normal alcohols, alkylamines and alkanediols.

Data determined:
\( \varepsilon_0 \) [free energy (in kT units) of interaction between a CH\(_2\)-group and the sorbent surface, characterizing hydrophobicity];
Kd [distribution coefficients for the sorbents studied: LiChrospher RP-18 (1), Si-100 Polyol RP-8 (2), LiChrosorb RP-8 (3), Spherisorb ODS (4), TSK-TMS-250 (5), Butyl Toyopearl 650-M (6), SOLOZA KG-20 (7), SOLOZA KG-14 (8), SOLOZA KG-7 (9) on Fig. 1];
y [a constant described by the following equation: y = ln[n(Kd - 1)]]

Chemical descriptors:
n (number of the carbon atoms in the investigated compound);
x (the mole percentage of the hydrophobic component in SOLOZA KG sorbents).

Results: Linear relationships are presented plotting y versus n for the 9 hydrophobic sorbents (Fig. 1) and the slopes of these straight lines are suggested for experimental determination of $\varepsilon_0$ values. $\varepsilon_0$ values determined by the suggested method are listed.

Data determined:
lg(K) (logarithm of the capacity factor measured by reversed phase HPLC).
Chemical descriptors:
logP (logarithm of hydrophobicity);
MR (molar refractivity);
$\alpha_X$ (zero order molecular bonding type connectivity index);
lg(cM) (logarithm of molarity);
P (polarity).

Results: Three linear models were fitted with independent variables of log(P), MR and $\alpha_X$. The best fitting parameters (independent of composition) were obtained from the following models (no statistical characteristics is presented):

\begin{equation}
lg(K) = a_0 + a_1 P + a_2 \log P + a_3 P \log P
\end{equation}

\begin{equation}
lg(K) = a_0 + a_1 \lg(cM) + a_2 \log P + a_3 \lg(cM) \log P
\end{equation}

The two types of correlations (with structural and with moving phase parameters) together might be used for the optimization of chromatographic separation of complex mixtures of sulphur-containing substances.

Title: Mathematical description of the chromatographic behaviour of isosorbide esters separated by thin layer chromatography.
4x, AERUD

The experts estimated the biodegradation time that might be required for AERUD. The data sets were collected from 22 biodegradation experts.

Data determined:

Rf, Rtl [retention factors obtained by thin-layer chromatography in benzene/ethylacetate/isopropanol (7:3:1.5) and in dichloromethane/diisopropylether/isopropanol (20:4:2.1) eluent systems, respectively].

Chemical descriptors:

lt (information index, based on the distribution of the elements in the topological distance matrix);

G/W (the geometrical analogue);

λ (Randic connectivity index);

Lmax (maximum geometric distance in the molecule);

v (electroacy index proposed by Yee);

qT (sum of absolute atomic charges in the molecule);

pA (sum of absolute atomic charges in the functional group);

qB (charges of C atoms at β-position to the functional group added to qA);

E (electrocy index);

EHOoMo (energy of the highest occupied molecular orbital, in β units);

ELUMO (energy of the lowest unoccupied molecular orbital, in β units).

Results: A model for predicting Rf values of isosorbide esters is suggested. Relationships between Rf, Rtl and the chemical descriptors were studied. The best equation was calculated between Rf and λ, the correlation coefficient was 0.976. Rf values calculated by equations based on chemical descriptors are also presented and compared with measured ones. The measured and calculated Rf values agreed well, their differences are less than 0.1 Rf unit in all the cases.

Correlation Analysis, Application: Environmental Sciences

1990/286

Title: Screening-level model for aerobic biodegradability based on a survey of expert knowledge.

Authors: Boethling*, R.S.; Sabljic, A.

U.S. Environmental Protection Agency, Office of Toxic Substances Washington DC 20460, USA.

Source: Environ. Sci. Technol. 1989, 23(6), 672 – 679.

Compounds: 46 highly diverse chemicals. Grouped according to the following properties: contains (ester or amide or anhydride) or (heterocyclic N) or (O bound to C) or (unbranched alkyl group with greater than 4 carbons).

Data determined:

AERUD (aerobic ultimate degradation in receiving waters).

Chemical descriptors:

\[ \chi^2 \] (valence second order molecular connectivity index);

\[ \chi^4 \] (fourth order path/cluster connectivity index);

nCl (number of covalently bound chlorine atoms);

M∞ (molecular weight).

Results: The paper has aimed at developing a model for predicting AERUD. The data sets were collected from 22 biodegradation experts. The experts estimated the biodegradation time that might be required for AERUD on the time scales of days, weeks, months and longer. 46 highly diverse chemicals but typical in wastewater treatment systems were examined. Zero to six order molecular and cluster connectivity indexes were calculated using computer programs written in FORTRAN for IBM PC/XT. The best fitted linear regression model is:

\[ \text{AERUD} = 0.60 \ln^2 \chi^2 + 57.25 \text{nCl}/\text{M}_\infty + 17.56 \chi^4/\text{M}_\infty + 1.45 \] (1)

Further on, the time of AERUD was corrected according to a few simple rules, developed according to experts’ opinion. The corrected model correctly classifies 41 of 46 chemicals while the predictions from Eq.1 correctly classifies only 37 of 46 chemicals.

AERUD PREDICTED FROM MODEL 1

| CONTAINS ESTER, AMIDE OR ANHYDRIDE? | YES | CONTAINS HETEROCYCLIC N? | NO - ADD (-0.235) | A |
| CONTAINS 0 BOUND TO C? | NO | CONTAINS HETEROCYCLIC N? | NO - ADD (-0.417) | C |
| CONTAINS UNBRANCHED ALKYL GROUP WITH>4 CARBONS? | YES - ADD (+0.385) | D |
| YES - ADD (+0.156) | E |
| NO - ADD (-0.251) | F |

Fig.1

1990/287

Title: Application of the QWASI (Quantitative Water Air Sediment Interaction) fugacity model to the dynamics of organic and inorganic chemicals in lakes.

Authors: Mackay*, D.; Diamond, M.

Institute for Environmental Studies, University of Toronto Toronto, Ontario, MSS 1A4, Canada.

Source: Chemosphere 1989, 18(7/8), 1343 – 1365.

Chemical descriptors:

C [concentration (mol/m³)];

f [fugacity (Pa)];

Z [fugacity capacity (mol/m³.Pa)];

D [first order rate constant: transport or transformation parameter (mol/Pa.h)].

Results: The QWASI fugacity model describes the fate of a (contaminating) chemical, such as organo-chlorine compounds, pesticides or metals. The lake model consists of water, bottom and suspended sediments, and air. The model includes the following processes: advective flow, volatilization, sediment deposition, resuspension and burial, sediment-water diffusion, wet and dry atmospheric deposition, and degrading reactions. The steady state solution of the model is illustrated by application to PCBs in Lake Ontario using the equilibrium criterion of fugacity as the variable controlling environmental fate of the chemical. The applications are based upon inaccurate data. Use of fugacity is inappropriate for involatile chemicals, such as metals, or ionic species, because fugacities are calculated from a basis of vapor phase concentrations. For these materials the use of the equilibrium concentration activity is more appropriate since activities are calculated from a water phase base. Thus, a new equilibrium criterion, termed the “equivalent” concentration (or equivalent aqueous concentration) is suggested as being preferable. This concentration has the advantage of being applicable in all phases, such as water, air and sediments. The formalism developed in the QWASI approach can also be applied, making possible a ready comparison of the relative rates (and thus, the importance) of diverse environmental fate processes. All these are illustrated by applying the model on a steady state basis to
the PCB example and to the fate of lead in Lake Ontario. The estimated and observed concentrations of PCBs and lead in Lake Ontario agree well: the largest difference in the case of PCBs in rain amounts to a factor of three. In other phases, and especially in the case of lead, the difference is usually less than 30 per cent. Although in order to judge the biological effects of a contaminant it is of fundamental importance to know its transport and transformations, and the present model has been proven to useful to describe this; direct biological implications are not deduced at the present stage.

Correlation Analysis, Application Others

1990/288

Title: Quantitative structure-activity relationships for the cytotoxicity of substituted aniline mustards in tissue culture.

Authors: Denny*, W.A.; Wilson, W.R.; Palmer, B.D.

Cancer Research Laboratory and Section of Oncology, School of Medicine, University of Auckland

Private Bag, Auckland, New Zealand.

Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.391 – 399. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

Compounds: 13 Compounds of type I, where R = H, 3,4-NO2, 4-NO2, 4-SO2Me, 3-SO2Me, 4-CONMe, 3-CONMe, 4-SMe, 4-Me, 3-Me, 4-OME, 3-OME, 4-NH2.

\[ R - \text{N(CH}_{2}\text{CH}_{2}\text{Cl)}_{2} \] (I)

Biological material: Chinese hamster cell line AA8 and a derived mutant line (UV4) which lack the ability to form the incision step of the normal DNA excision repair pathway and is thus hypersensitive to alkylating agents.

Data taken from the literature:

E \[ \text{[reduction potential of the compound (eV)]} \]

Data determined:

CT10 \[ \text{[cytotoxic potency of the compounds (µmol·hr) in the clonogenic assay calculated by multiplying drug concentration with the time of exposure required at that concentration to reduce the surviving fraction of AA8 or UV4 cells to 10%]} \]

IC50 \[ \text{[concentration of the compound (µmol/L) required for 50% inhibition of the proliferation of the cell culture under aerobic or anaerobic conditions]} \]

ratio \[ \text{[hypoxic selectivity of the compound measured in UV4 culture and calculated by formula: ratio = IC50(aerobic)/IC50(hypoxic)]} \]

T1/2 \[ \text{[half-life of the compound (h) incubated in Alpha culture medium]} \]

Chemical descriptors:

\( \sigma \) \[ \text{(Hammett's constant, characterizing the electron-withdrawing power of the substituent)} \]

\( \sigma^- \) \[ \text{(Hammett's polar electronic constant characterizing the electron-withdrawing power of the substituent for anilines)} \]

Results: Significant linear regression equations were calculated for the half-life (T1/2), growth inhibition (IC50) and clonogenicity data (CT10) using Hammett constants (Eq. 1, Eq. 2, Eq. 3):

\[ \log(CT10) = 2.46(±0.40) \sigma + 0.21 \] (2)

n = 12 \( \quad r = 0.97 \quad s = 0.34 \quad F \text{ not given} \)

\[ \log(CT10) = 2.56(±0.27) \sigma^- + 0.07 \] (3)

n = 8 \( \quad r = 0.92 \quad s \text{ not given} \quad F \text{ not given} \)

The similar slopes of the equations show that these compounds exert their cytotoxicity primarily by alkylation. While the majority of the tested compounds showed no hypoxia-selective cytotoxicity (ratio awa 1:0), the 4-NO2 and 3-NO2 substituted compounds were more toxic to UV4 cells under hypoxic conditions (ratio = 3.2 for the compound with R = 4-NO2), indicating cellular selectivity of the nitro-group. The measured hypoxic selectivity of the 3-NO2 and 4-NO2 substituted compounds was a fraction of the calculated ratio (measured 220 fold and calculated 3500 fold by Eq. 2 between the 4-NO2 and 4-NH2 substituted compounds). The main reason for the difference between the calculated and measured hypoxic selectivity is suggested to be the low reduction potential of the 4-NO2 and 3-NO2 groups (E = −500 mV and E = −470 mV, respectively).

1990/289

Title: Correlations and mechanisms of chemical toxicity in animals.

Authors: Magee*, P.S.; King, J.W.

BIOSAR Research Project

Vallejo CA 94591, USA.

Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.390 – 399. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

Compounds: Various sets of phenols, diarylamine rodenticides, aryln-methylcarbamate insecticides, organophosphates, anilines, pyridines with LD50 values from the literature.

Biological material: Rat, dog, mouse, guinea pig, rabbit, cat, pigeon, duck, quail.

Data taken from the literature:

LD50 \[ \text{[dose of the compound (dimension not given) required for killing 50% of the test organisms obtained from toxicity compilations]} \]

Data determined:

\[ \log(MW/LD50) \] (logarithm of LD50 value corrected for molecular weight).

Chemical descriptors:

MW \[ \text{(molecular weight)} \]

\( \pi \) \[ \text{(Hansch-Fujita's substituent constant characterizing hydrophobicity)} \]

\[ \log P \] \[ \text{(logarithm of the partition coefficient in 1-octanol/water)} \]

\( \sigma^- \) \[ \text{(Hammett's constant characterizing the electron withdrawing power of the substituent for anilines)} \]

INO2, IOH \[ \text{(indicator variable 1 for nitrophenols and hydroxyphenols, respectively)} \]

I26 \[ \text{(indicator variable 1 for 2,6-disubstitution)} \]

HB \[ \text{(indicator variable for hydrogen bonding substituent in ortho-position)} \]

IOR, INP \[ \text{(indicator variable 1 for phenolate leaving group in organophosphates and for phosphoramidates, respectively)} \]

v2,6 \[ \text{(indicator variable 1 for steric effect of 2,6-substituents)} \]

Results: Significant linear regression equations were calculated using single laboratory and interlaboratory log(MW/LD50) values of diverse compound sets measured on various test animals. Analyzing the equations, the differences among animals or between different methods of administration could be interpreted in mechanistic terms. Important generalizations could be made by examining the chemical
type, test animals, the mean level of toxicity and the form of the equation. E.g. analysis of the toxicity of phenols showed a transition between simple dependence from logP to exclusive dependence to reactivity factors indicating two separate classes of phenol toxicity (Eq. 1 for mouse i.p. toxicity, and Eq. 2 for rat oral toxicity):

\[
\log(MW/LD_{50}) = 0.429 \Sigma x - 0.546 \\
= 26 \quad r = 0.931 \quad s \text{ not given} \quad F = 156
\]

\[
\log(MW/LD_{50}) = 0.396 \Sigma x^- + 0.520 126 - 0.876 \\
= 71 \quad r = 0.729 \quad s \text{ not given} \quad F = 38.58
\]

Inter-route comparisons were made in terms of mean log(MW/LD50) values for phenol toxicity in rats: iv (0.373) > ip (-0.157) > sc (-0.429) awa dermal (-0.469) > oral (-0.720). Inter-class comparisons could be ordered by the same measure for mouse (oral): diarylamines (1.15) > phosphates (0.487) > carbamates (0.227) > phenols (-0.499) > anilines (-0.723). The study showed that the analysis of inter-laboratory toxicity data may give significant regression equations allowing to classify mechanism. The development of an extensive knowledge base in three dimensions (animals, routes and toxicants) is envisaged as combining mechanistic insight obtained from regression analysis with toxicity orders in terms of log(MW/LD50) values.

1990/290

Title: Structure-activity relationship studies on the toxicity of benzene derivatives. III. Predictions and extension to new substituents.

Authors: Hall**, L.H.; Maynard, E.L.; Kier, L.B.

Department of Chemistry, Eastern Nazarene College

Quincy MA 02170, USA.

Source: Environ. Toxicol. Chem. 1989, 8(5), 431-436.

Compounds: 105 Substituted benzenes.

Biological material: Fathead minnow (Pimephales promelas).

Data determined:

pLC50 [negative logarithm of the concentration of the compound required to kill 50% of the test organisms].

Chemical descriptors:

\( \Delta T \) (toxicity contribution for a substituent);

\( \tau \) (Hansch-Fujita's substituent constant characterizing hydrophobicity);

\( E_s \) (Taft's constant, characterizing steric effects of the substituent);

\( R \) (Swain-Lupton's electronic parameter, characterizing the resonance effect).

Results: An additivity model, pLC50 = \( \sum n_i \Delta T_i + T_0 \), where \( n_i \) is the number of \( i \)th substituents in a benzene derivative, \( \Delta T_i \) is the toxicity contribution of the \( i \)th substituent and \( T_0 \) is the toxicity of the parent compound (benzene), was used for predicting toxicity of 10 compounds not included in the original data set. The predicted toxicity are in reasonable agreement with their experimental values. A 12-variable model including \( \Delta T \) for substituents was developed. The regression statistics are presented (\( n = 105; r = 0.921; s = 0.31, F = 43 \)), equation is not given. No significant relationship was found between \( \Delta T \) a series of and physicochemical parameters, only for the \( E_s \) and \( R \) were found moderate correlations, \( r = 0.75 \) and 0.49, respectively.

1990/291

Title: Structure-activity relationships in mutagenicity and in nucleophilic ring opening of N-(arylalkyl)phenanthrene 9,10-imines.

Authors: Roll, M.; Shetler, S.; Stark, A.-A.; Blum*, J.

Department of Organic Chemistry, The Hebrew University Jerusalem 91904, Israel.

Source: Mutagenesis 1990, 5(1), 25–30.
1990/292

Title: Mutagenicity of dimethyl heteroaromatic triazines in the Ames test. The role of hydrophobicity and electronic effects.

Authors: Shusterman*, A.J.; Debnath, A.K.; Hansch, C.; Horn, G.W.; Fronczek, F.R.; Greene, A.C.; Watkins, S.F.

Department of Chemistry, Pomona College
Claremont CA 91711, USA.

Source: Mol. Pharmacol. 1989, 36(6), 939 – 944.

Compounds: 24 Various phenyl and heterocyclic triazines.

Biological material: Salmonella typhimurium.

Data determined:
C (molar concentration of triazine that causes 30 mutations above background/10^8TA92 bacteria, determined by the Ames test).

Chemical descriptors:
\[ \log P \] (logarithm of the partition coefficient in 1-octanol/water);
\[ q_{\text{HOMO}} \] (energy (eV) of the triazine’s highest occupied molecular orbital (HOMO) calculated by the minimum neglect of differential overlap (MNDO));
\[ q_{\text{HOMO}} \] (electron density on N1 in the HOMO calculated by the MNDO).

Results: Highly significant linear relationships between \[ \log (1/C) \] and \[ \log P, q_{\text{HOMO}} \] (Eq. 1); \[ \log P, q_{\text{HOMO}} \] are presented indicating that the more hydrophobic and more electron-rich triazines are more active according to the Ames test:

\[ \log (1/C) = a + b \log P + c \log q_{\text{HOMO}} \]

where \( a = 21, b = 0.919, c = 0.631 \) for hydrophobic triazines

\[ \log (1/C) = a + b \log P + c q_{\text{HOMO}} \]

where \( a = 21, b = 0.931, c = 0.585 \) for electron-rich triazines

The \[ \log (1/C) \] values calculated by Eq. 1 and 2 were compared with measured ones and agreed well. It was established that \[ q_{\text{HOMO}} \] and \[ q_{\text{HOMO}} \] are useful as electronic indices for phenyl and heteroaromatic triazines simultaneously. Hydrophobicity and electronic characteristics of the compound are suggested as descriptors to predict mutagenic activity.

1990/293

Title: Structural basis of carcinogenicity in rodents of genotoxicants and non-genotoxicants.

Authors: Rosenkranz*, H.S.; Klopman, G.

Departments of Environmental Health Sciences, School of Medicine, Case Western Reserve University
Cleveland OH 44106, USA.

Source: Mutat. Res. 1990, No.228, 105 – 124.

Compounds: A diverse set of 189 organic compounds compiled from the National Toxicology Program (NTP) Cancer Bioassay of which 63% was carcinogenic in rats or mice or both.

Biological material:
a) Mice;
b) Rats.

Chemical descriptors:
Substructures [a total of 32355 fragments were generated from the 189 compounds using the program CASE (Computer-Automated Structure Evaluation) system].

Results: A comparative classification of the compounds were performed using CASE for identifying molecular fragments associated with carcinogenic activity (biophores) as well as deactivating fragments (biophobes). CASE identified 21 biophores and 2 biophobes from the 32355 fragments of the 189 compounds with a less than 12.5% probability of being associated with carcinogenicity as a chance. The sensitivity and specificity of the analysis was unexpectedly high: 1.00 and 0.86, respectively. The predictive power of CASE was tested using the identified biophores and biophobes on a group of chemicals not present in the data base. The ability of CASE to correctly predict carcinogens and presumed non-carcinogens was found to be very good. It was suggested that non-genotoxic carcinogens may act by a broader mechanism rather than being chemical specific.

1990/294

Title: The structure-activity relation analysis. A topological descriptor model for odor.

Authors: Vitjuk, N.V.; Pozigun, V.V.; Kuzmin*, V.E.

Institute of Physical Chemistry of USSR, Institute of Navy’s Engineers in Odessa Odessa, USSR.

Source: Khim. Farm. Zh. 1989, 23(5), 607 – 610 (Russ., with Engl. summary).

Compounds: 9 compounds of type I where \( R = H, \text{CH}_3, \text{C}_2\text{H}_5, \text{CH}_2\text{H}_4, \text{C}_2\text{H}_5, \text{CH}_3, \text{CH}_3, \text{C}_2\text{H}_5, \text{C}_2\text{H}_5, \text{SC}_4\text{H}_9; 3 \) compounds of type II where \( R = H, \text{C}_2\text{H}_5\text{OH}, \text{SC}_4\text{H}_9\text{OH}. \)

\[
\begin{align*}
\text{CH}_3=\text{C}-\text{CH}_2-\text{S}-\text{R} \\
\text{CH}_3
\end{align*}
\]

\[
\begin{align*}
\text{CH}_2-\text{CH}-\text{CH}_2-\text{S}-\text{R} \\
\text{OH} \quad \text{CH}_3
\end{align*}
\]

Data determined:
\( p_i^a \) (a priori probability of appearance of the i-th active compound);
\( p_i^n \) (a priori probability of appearance of the i-th nonactive compound).

Chemical descriptors:
\( x_1 \) (the first order molecular connectivity index);
\( x_2 \) (the second order molecular connectivity index);
\( I_x \) (information-theoretic index on graph distances calculated by the Wiener index according to Gutmann and Platt);
\( W \) (Wiener index);
\( R \) (topological distance);
\( v \) (rank of smell, where the rank is defined to equal with one for the most active compound).

Results: The authors' previously proposed structure-activity relationship approach was applied for structure-odor relationship. 13 different compounds of groups I and II were examined using the topological indices \( W, R, I, x \) as independent variables and \( v \) as the dependent variable. The best correlation was obtained between \( R \) and \( v \) (Fig. 1):
Title: Correlation of partitioning of nitroimidazoles in the n-octanol/saline and liposome systems with pharmacokinetic parameters and quantitative structure-activity relationships (QSAR).

Authors: Betageri, G.V.; Rogers*, J.A.
Faculty of Pharmacy and Pharmaceutical Sciences, University of Alberta
Edmonton, Alberta T6G 2N8, Canada.

Source: Pharm. Res. 1989, 6(5), 399 – 403.

Compounds: Nine nitroimidazole derivatives: RO-07-0741, RO-07-2044, SR-2508, SR-2555, misonidazole, desmethylmisonidazole, azomycin riboside, azomycin, iodoazomycin riboside.

Biological material:
a) Dogs;
b) Mice;
c) Murine EMT-6;
d) Chinese hamsters.

Data taken from the literature:
C [concentration of the compound (dimension not given) required to produce a given rate of response];
PPC [peak plasma concentration (mol/L) of the nitroimidazoles];
PCR [plasma clearance rate (L/kg hr)];
UDU [unchanged drug in urine (%)] after i.v. injection of saline solutions of drug in dogs);
OB [oral bioavailability (%) after administration of 400 mg drug in a size 00 hard gelatin capsule];
ALD [acute LD50 (mmol/kg)] after i.v. injection of saline solutions of drug in dogs).

Chemical descriptors:

\[ \log P, \log P_{mr}, \log P_{tr}, \log P_{mr}, \log P_{tr} \]

Results: LogP and logPmr, values were determined for the nitroimidazole derivatives. Significant linear equations were calculated, the best one related for logP and logPmr (Eq. 1):
Biological material: Rats.

Data determined:
Death(P), Death(A), Liver(A), Liver(C), Lung

Conformational analysis
(low energy conformers of the compounds were calculated
using CHEMLAB-II and Allinger's MMP2);
MB (moment of binding calculated using the fragment binding
constants of Andrews);

Chemical descriptors:
logP (logarithm of the partition coefficient in 1-octanol/
water calculated by the CLOGP program);
x [several molecular connectivity descriptors (details
not given) calculated using the ADEPT program];
\( J_x \) [Kier’s molecular shape descriptors (6 descriptors)];
MSA [Hopfinger’s MSA volume descriptors (3 descriptors)];

Molecular shape
[Jurs shadow descriptors (6 descriptors) and
length/breadth descriptors (2 descriptors)].

Results: For modeling the shape of the compounds, SIMCA was us-
ed: the approach was to generate disjoint principal models of clustered
points in a multidimensional space. The number of clusters for each
structure was determined by using hierarchical cluster analysis. Fig.
1 shows the orthogonal views of a schematic representation of the SIM-
CA models for the atom clusters in senecionine:

Each compound in turn was used as a reference structure. Every other
structure was superimposed on the reference using the ends of the cor-
responding binding moment vector plus the ring nitrogen atom.
Canonical correlation analysis was used for calculating the correlation
between the five biological activity data and shape descriptors of 21
structures. The best correlation was observed for Jurs’ shadow
descriptors. The MSA and SIMCA descriptors were comparable. The
model was able to express both the amount and direction of shape the
differences, and also for encoding relevant information for correlation
with the biological activity.

Source: Arzneim.-Forsch./Drug Res. II, 1989, 39(11), 1406–
1410.

Compounds: 6 N-substituted 3-methyl-4-nitropyrazole-5-carboxa-
mites (II), 6 N-substituted 4-amino-3-methylpyrazole-5-carboxamides
(III), 14 N-substituted 3-methyl-4-diazopyrazole-5-carboxamides and
N-piperidinyl-N-(1,3-dimethyl-4-nitrosopyrazol-5-yl)-urea (VII).

Biological material: 15 Representative Gram-negative and Gram-
positive bacteria and intestinal bacterial species.

Data determined:
MIC [minimum inhibitory concentration of the compound
(μg/ml) required for inhibition of the microorganisms in
vitro];
PC1, PC2 [principal component loadings, determined by principal
component analysis (PCA)].

Chemical descriptors:
f (Rekker’s constant, characterizing hydrophobicity);
MR (molar refractivity);
v [infrared absorption peak (cm\(^{-1}\)) of the C=O group].

Results: Univariate and multivariate methods, like principal com-
ponent analysis were applied for the investigation of the antibacterial
structure-activity relationships. It was established that PC1 is mainly
loaded by the Gram-negative bacteria, while PC2 has high loadings on-
ly with the Gram-positive strains. Regressing the scores of PC1 and
PC2 against the physico-chemical variables, it was established that MR
is the most important variable for the Gram-negative antibacterial ac-
tivity (Eq. 1) and v is the most important one for the Gram-positive
antibacterial activity (Eq. 2):

\[
PC1 = -0.63(±0.23) \text{MR} + 2.03(±0.81) \quad (1) \\
n = 13 \quad r = 0.875 \quad s = 0.506 \quad F = 35.86
\]

\[
PC2 = 0.019(±0.009) \nu - 31.51(±15.50) \quad (2) \\
n = 13 \quad r = 0.803 \quad s = 0.622 \quad F = 20.04
\]

Based on the above observations, some new congeners are designed
which possibly maintain the high Gram-positive activity and improve
the more valuable Gram-negative activity.

1990/299

Title: Structure-activity correlations for psychotomimetics. I.
Phenylalkylamines: electronic, volume, and hydrophobicity para-
eters.

Author: Clare, B.W.
School of Mathematical and Physical Sciences, Murdoch University
Murchon, Western Australia 6150, Australia.

Sources: J. Med. Chem. 1990, 33(2), 687-702.

Compounds: 63 Compounds of type \( \text{R}^1 = H, \text{Me}; \text{R}^2 = H, \text{MeO}; \text{R}^2 = H, \text{MeO}, \text{EtO}, \text{PrO}, \text{EtS}; \text{R}^4 = H, \text{MeO}, \text{EtO}, \text{PrO}, \text{BuO}, \text{Me}, \text{Et}, \text{MeS}, \text{EtS}, \text{PrS}, \text{BuS}, \text{BzO}, \text{i-Bu}, \text{Bu}, \text{Am}, \text{i-PrS}; \text{R}^3 = H, \text{MeO}, \text{EtO}, \text{MeS}, \text{EtS}; \text{R}^6 = H, \text{MeO}; \text{R}^2 \text{and } \text{R}^4 \text{together } -\text{OCH}_2\text{O}; \text{R}^3 \text{and } \text{R}^4 \text{together } -\text{OCH}_3\text{O}.\)
Discriminant analysis resulted in a function containing six variables which misclassified only one compound in the training set. When the data was repeatedly split randomly into a training and a test set, the misclassification rate was 9% (15 out of 161 classifications). Fig. 2 shows the plot of the two canonical variables from discriminant analysis visualizing the separation of hallucinogenic and nonhallucinogenic derivatives (meaning of symbols are the same as in Fig. 1).

Multiple regression analysis (MRA) was found to be the most useful for identifying relevant and discarding redundant variables. Highly significant parabolic regression equations were calculated for the human activity data (A) (n = 50, r ranging from 0.9004 to 0.9563, F not given) and for animal data (ED$_{50}$) (n = 16, r = 0.8679 and r = 0.9825, F not given). Eight descriptors were found to be highly significant parabolic regression equations were calculated for the formation of charge transfer complex by accepting charge. Data did not support the hypothesis that the human activity data and animal data determined: 

ED$_{50}$ [dose of the compound (mg/kg) which causes 50% of the rats which were trained on 1 mg/kg reference compound to respond as they would to the training drug];

Conformational analysis g (geometries of the compounds were calculated using MMF2 from starting geometries determined by the program EUCLID).

Chemical descriptors:

$V$ [van der Waals volume (Å) of the substituent];

$\pi$ (Hansch-Fujita’s substituent constant characterizing hydrophobicity);

$\varepsilon_{HOMO}, \varepsilon_{LUMO}$ (energies of the highest occupied and lowest unoccupied $\pi$ orbitals, respectively);

$\varepsilon_{HOMO}, \varepsilon_{LUMO}$ (energies of the highest occupied and lowest unoccupied molecular orbitals, respectively);

$\varepsilon_{HOMO}, \varepsilon_{LUMO}$ (energies of the next highest occupied and next lowest unoccupied molecular orbitals, respectively);

$Q$ (charge on an atom calculated by CNDO/2);

$D$ [dipole moment (debyes)];

$I_{Me}$ (indicator variable 1 for the presence of a methyl group on the a-carbon atom).

Results: Pattern recognition analysis of 63 phenylalkylamine psychotomimetics [K-nearest neighbor (kNN) method in 24D space] indicated of the separability of the data. The clustering of the compounds was visualized in 2D by non-linear mapping (Fig. 1: open circles represent inactivity, crosses low activity and asterisks high activity):
Principal component analysis of the data set extracted 2 principal components, explaining 64% of the variance of the sweet compounds. The sweet compounds clustered in a relatively confined region of the 15D space whereas the tasteless and bitter compounds were scattered around the sweet compounds. A Coomans plot, however, indicated, when plotting $D'$ versus $D^2$, that sweet and nonsweet compounds could be well separated along the $D'$ axis (Fig. 1, sweet: empty circle, tasteless: full circle, bitter: triangle):

It was also shown that sweet potency was a monotonously decreasing function of $D^1$, i.e. the nonsweet compounds had higher $D^1$ values than those of sweet compounds and a sweet candidate must have as low $D^1$ value as possible. On the basis of the zero-principal component SIMCA method 89% (89/9) of the sweet and 88% (15/17) of the nonsweet compounds were correctly classified on two levels of significance ($P = 0.01$ and 0.05).

1990/302

Title: Conformation of cyclopeptides. Factor analysis. A convenient tool for simplifying conformational studies of condensed poly-ring systems. Prolyl-type cyclopeptides.

Authors: Oldziej, St.; Dokurno, P.; Liberek, B.; Kolodziejczyk, A.S.; Ciarkowski*, J.; Gdaniec, M.

University of Gdansk, Institute of Chemistry

Sobieskiego 18, PL-801952 Gdansk, Poland.

Source: J. Mol. Structure (Theochem.) 1990, No.204, 301 – 324.

Compounds:
- 2,6-dioxopiperazine (DOP) family;
- pyrrolidine (PYR) family;
- 2,5-dioxopiperazine/pyrrolidine (DOP/PYR) family.

Chemical descriptors:
- $\Theta_1$ [Endocyclic torsion angle, degree (°)].

Results: Ring puckering theory (RPT) shows that any of some 70 conformations of the six-membered DOP-ring family may be reproduced by means of a superposition of the canonical twist (T), boat (B) and chair (C) forms. Physically, the coefficients have the meaning of relative contributions (amplitudes) of the T, B and C forms into the total conformation of the ring. Here factor analysis (FA) and principal component analysis was used in conformational studies of 30 various X-ray conformers of DOP/PYR. A correspondence was found between factors identified and RPT, when the rings are considered separately. This fact allows a physical interpretation of the FA results: two or three puckering variables were found for the DOP and PYR rings expressing the absolute amplitudes of the basic pucker modes. Subsequent FA treatment of the condensed system revealed five conformational variables necessary and sufficient to describe the two-ring puckering completely. Each of the basic pucker modes defines a unique pattern of conformational variation of the whole two-ring system. The results demonstrate that FA is a powerful technique in analysing condensed poly-ring systems, not amenable to the RPT treatment.

1990/301

Title: Preprocessing, variable selection, and classification rules in the application of SIMCA pattern recognition to mass spectral data.

Authors: Dunn III*, W.J.; Emery, S.L.; Glen, G.W; Scott, D.R.

College of Pharmacy, The University of Illinois at Chicago

833 South Wood, Chicago IL 60612, USA.

Source: Environ. Sci. Technol. 1989, 23(12), 1499 – 1505.

Compounds: A diverse set of 121 compounds observed in ambient air classified as (1) nonhalogenated benzenes; (2) chlorine containing compounds; (3) bromo- and bromochloro compounds; (4) aliphatic hydrocarbons; (5) miscellaneous oxygen-containing hydrocarbon-like compounds (aliphatic alcohols, aldehydes and ketones).

Data determined:

Pattern recognition [SIMCA (Simple Modeling by Class Analogy)] pattern recognition was applied to autocorrelation-transformed mass spectra of the compounds using providing chemical class assignment for an unknown;

$m/z_1$, $m/z_2$, $m/z_3$ (first three principal components scores of SIMCA).

Results: SIMCA pattern recognition method was applied on a training set of 78 toxic compounds targeted for routine monitoring in ambient air. The analysis resulted in very good classification and identification of the compounds (87% and 84%, respectively). However, the training procedure proved to be inadequate as a number hydrocarbons from field samples (GC/MS analysis) were incorrectly classified as chlorocarbons. A new approaches for the preprocessing (scaling the MS data by taking the square root of the intensities followed by autocorrelation transform), variable selection (only the 16 most intense ions in the MS spectrum were taken), and for the classification rules of SIMCA has been introduced to improve results on real data. Fig.1 and Fig. 2 shows the geometric interpretation of the original and modified SIMCA pattern recognition classification rules, respectively, where $d$ and $d^*$ is the distance of object from class model (when object has a principal component score within or outside of the range of the class model), respectively, and $d_m$ is the new distance of the objects to the class mean.
As a result of the revised rules the classification performance has been greatly improved for field data (97 – 94 %).

1990/303

Title: A QSAR model for the estimation of carcinogenicity. Example application to an azo-dye.

Authors: Enslein*, K.; Borgstedt, H.H.

Health Designs, Inc.
183 East Main Street, Rochester NY 14604, USA.

Source: Toxicol. Let. 1989, 49(2-3), 107 – 121.

Compounds:

a) 185 Structurally heterogeneous organic compounds;
b) CI disperse yellow 3 of type I.

Chemical descriptors:

Selected parameters [features with positive contribution to carcinogenicity: saturated primary or secondary aliphatic ester, longest aliphatic chain in molecule, secondary aryl amine, methyl amine fragment, two electron withdrawing groups bound to NH, aliphatic aldehyde, aryl sulfonic acid, two electron-releasing groups on single benzene ring (para), one withdrawing group and 2 releasing groups (1,2,5) on single benzene, primary aliphatic alcohol, aryl alcohol, aryl methoxy, aryl aldehyde, difference path molecular connectivity index order 0 (DIFPAT0), ethane or ethylene between 1 releasing group and 1 withdrawing group, 3-carbon chain between 1 electron-withdrawing group and 1 electron releasing group.

Results: A highly significant discriminant equation was calculated using the quantified features of the compounds (n = 368, F = 43) with an overall accuracy of 98.9% correct classification. The dye of type I was used as a model compound for prediction purposes. Relevant features were present in the database with sufficient frequency to cover the structure of (I) so it was regarded as estimable by the model (aryl amide, azo fragment, longest aliphatic chain in molecule, benzene, aryl alcohol, SUMPATI, DIFPAT0). Validation of the estimate was made by the search for compounds in the database having some of these features. The analysis gave a positive carcinogenicity estimate for compound (I) with a relatively high confidence. Addition of parameters representing biotransformation and pharmacokinetic factors to the database was suggested.

Quantum Chemistry

1990/304

Title: MNDO calculations of highly mutagenic (chloromethyl)benzo[a]pyrenes.

Authors: Bail*, J.C.; Salmeen, I.T.

Scientific Research Laboratory, Ford Motor Company
P.O.Box 2053, Dearborn MI 4821, USA.

Source: Chem. Res. Toxicol. 1989, 2(6), 375 – 378.

Compounds: 12 (Chloromethyl)benzo[a]pyrenes (ClMeB[a]P) of type I, with chloromethyl groups at positions 1, 4 – 6, and 10 – 12.

Biological material: Salmonella typhimurium strains TA100 and TA98.

Data taken from the literature:

logK [log rate constant of the solvolysis (dimension not given) of ClMeB[a]P in 50 % aqueous acetone].
Data determined:
Mutagenicity [number of the revertants (revertants/nmol) caused by the test compound in the Ames test];
LUMO [LUMO (energy of the lowest unoccupied molecular orbital) coefficient of the P₆ orbital on the exocyclic carbon of the CIMeB[a]P carbocation];
$E_{\text{LUMO}}$ [energy of LUMO (eV) of the CIMeB[a]P carbocation];
$B_o$ [bond order of the exocyclic carbon-to-ring bond in CIMeB[a]P carbocation];
$\Delta H_{f, \text{CIMeB[a]P}}$ [heats of formation (kcal/mol) of CIMeB[a]P calculated by MNDO];
$\Delta H_{f, \text{MeB[a]P}}$ [heats of formation (kcal/mol) of MeB[a]P carbocation calculated by MNDO];
$\Delta E$ [energy difference (kcal/mol) between the heats of formation between CIMeB[a]P and the MeB[a]P cation];
IP [ionization potential (eV) of the parent CIMeB[a]P ranging from 7.98 to 8.5 eV];
logK [log rate constant of the solvolysis (dimension not given) of CIMeB[a]P in 50% aqueous acetone];

Geometry optimization (complete geometry optimization was carried out using MNDO).

Results: The (chloromethyl)benzo[a]pyrenes were shown to be highly potent direct-acting mutagens in the Ames test requiring no metabolic activation.
None of the calculated parameters correlated significantly with mutagenicity. LogK showed very poor linear correlation with mutagenicity ($r = 0.49$ for strain TA98 and 0.57 for strain AT100).
LogK values, however, correlated well with $\Delta E$ ($r = 0.95$) supporting the hypothesis that CIMeB[a]P react via carbocations (Fig. 1):

It was suggested that the mechanism of mutagenicity of the CIMeB[a]Ps measured in the Ames test is probably more complex than the simple reactivity of carbocation intermediates.

Compounds: Vasopressin and 3 vasopressin analogues: (1-β-mercaptocaptoproic acid)-arginine-vasopressin [(Mpa')-AVP], (1-β-mercapto-β,β-cyclopentamethylenepropionic acid) arginine-vasopressin [(Cpp')-AVP], (1-thiosalicylic acid)-arginine-vasopressin [(Ths')-AVP].

Biological material: $V_1$ and $V_2$ vasopressor receptor (details not given).

Data taken from the literature:
Conformational analysis [lowest energy conformations of the vasopressin analogues (details not given)].

Data determined:
$[E(R)]$ [dipole interaction potential (dimension not given)];
MEP map [molecular electrostatic potential (kcal/mol) in a plane];
MEF map [molecular electrostatic field map (kcal/mol), mapping the $[E(R)]$ values of a molecule surface in a plane, predicting the directions and energies of the interactions with small polar molecules at distances greater than the van der Waals sphere];
MEP and MEF surfaces (construction of surfaces corresponding to a given value of potential);
3D MEP and MEF maps (3D maps were generated by superimposing the equipotential curves corresponding to a value of 20 kcal/mol in the case of MEP, and 1.5 kcal/mol in the case of MEF, computed in several planes perpendicular to the mean plane of the analogues in low energy conformations, stacking over each other in 1 Å distance).

Results: The three vasopressin analogues differ significantly in their biological activities. Both MEP and MEF maps of the of the biologically active (Mpa')-AVP and (Cpp')-AVP are similar, but they are different from that of the inactive (Ths')-AVP. Fig. 1, Fig. 2 and Fig 3 shows the maps of MEP of (Mpa')-AVP, (Cpp')-AVP and (Ths')-AVP, respectively, in their low energy conformations:

Title: Theoretical studies of the mechanism of the action of the neurohypophyseal hormones. I. Molecular electrostatic potential (MEP) and molecular electrostatic field (MEF) maps of some vasopressin analogues.
Authors: Liwo, A.; Tempczyk, A.; Grzonka*, Z.
Institute of Chemistry, University of Gdansk ul. Sobieskiego 18, PL-80952 Gdansk, Poland.
Source: J. Comput.-Aided Mol. Design 1989, 3(3), 261 – 284.
A new method for calculating the points of the equipotential curves was also presented.

**1990/306**

**Title:** Molecular determinants for drug-receptor interactions. 12. Molecular orbital studies on opioid analgesics N-allyl-normetazocine and nalorphine. Conformations and electrostatic molecular potentials.

**Authors:** Grassi, A.; Pappalardo, G.C.; Marietta, A.

**Source:** J. Theor. Biol. 1989, 140(4), 551 - 570.

**Compounds:** 2 Mixed agonist-antagonist opioid analgesics: N-allyl-N-normetazocine (NAM) and nalorphine (NLP), both as free bases and N-protonated form (NAMH+ and NLPH+, respectively).

**Data taken from the literature:**
Crystal structure (crystal coordinates of the molecules were determined by X-ray diffraction methods).

**Data determined:**
- Electrostatic molecular potential (eV) were calculated using AM-1 type semiempirical MO calculations.
- Minimum energy conformations were calculated using X-ray structures as input geometries followed by AM1-method (Fletcher-Powell algorithm).

**Chemical descriptors:**
- \( \omega_1, \omega_2 \) [rotational angles of the N-allyl group (deg)].

**Results:** Four similar energy minima were located by AM1 calculations for both NAMH + and NLPH +. Fig. 1 shows conformational energy map of NAMH + (energies are in kcal/mol):

The energy minima for the protonated NAM + and NLPH + were the most populated ones with conformational enantiomers relative to the involved N-allyl-piperidine moiety (37 % and 44 %, respectively). It was shown that the equipotential curve localization of EMP contour maps were very similar for the corresponding conformations of both NAMH + and NLPH +, indicating that both molecules should interact at the same anionic sites of the opioid receptor (\( \mu \) morphine receptor).

**Fig. 2** and **Fig. 3** shows the EMP contour maps of NAMH + and NLPH +, respectively, in their preferred conformations:

**1990/307**

**Title:** Conformational analysis of fenvalerate and an ether-type pyrethroid.

**Authors:** Kurita*, Y.; Tsushima, K.; Takayama, C.

**Source:** ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.183 – 197. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

**Compounds:** Esfenvalerate (SS and SR isomers) of type 1, 3-phenoxybenzyl 2-(4-ethoxyphenyl)-3,3,3-trifluoropropyl ether (R
isomer) (II), α-cyano-3-phenoxybenzyl 2-(4-chlorophenyl)-2-methylpropionate (S isomer) (III) and deltamethrin (IV).

Data determined:
Conformational analysis (minimum energy conformations of the compounds in vacuum were calculated using AM1 molecular orbital method and Broyden-Fletcher-Goldfarb-Shanno method integrated into MOPAC);
RMS (root mean square, indicating the goodness of fit between two conformers in 3D);
logP (logarithm of the partition coefficient in 1-octanol/water estimated using CLOGP program);
E [heat of formation of the most stable conformer (kcal/mol)].

Chemical descriptors:
θ₁ = θ₁₄ [dihedral angles in the compound (deg)].

Results: It was assumed that the 3D positions of the benzene rings of the pyrethroids are decisive for good insecticidal activity. The lower energy conformers of (I) (SS and SR isomers), (II) (R isomer), (III) (S isomer) and deltamethrin (IV) were compared by superimposition. In spite of their opposite configuration, esfenvalerate (I) (SS isomer) and the new type pyrethroid II (R isomer) were reasonably superimposed, indicating that the positions of the benzene rings in space are important and the bonds between them are not directly determinant (Fig. 1) [II(R) dotted line and I(SS) solid line]:

Molecular Graphics,
Theoretical Papers

Title: Second-generation computer-assisted inhibitor design method.
Authors: Desjarlais, R.L.; Seibel, G.L.; Kuntz, I.D., Jr.
Department of Pharmaceutical Chemistry, University of California San Francisco CA 94143, USA.
Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.60 – 69. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.
Compound:
a) 2-(2-Quinolyl)cyclohexane phenylhydrazone of type I;
b) Compound of type II.

Biological material: Penicillopepsin, an aspartyl protease, from Penicillium jaunthellum.
Data taken from the literature:
Crystal structure (X-ray crystal coordinates of penicillopepsin was obtained from the Protein Data Bank).

Data determined:
Electrostatic potential [electrostatic potential of the protein atoms (kcal/mol) is calculated using the partial charges in the AMBER united atom force field);
Docking (The DOCK program was used to find molecules that have a good geometric fit to the receptor).

Results: A second generation computer-assisted drug design method has been developed utilizing a rapid and automatic algorithm of locating steroically reasonable orientations of small molecules in a receptor site of known 3D structure. It includes also a scoring scheme ranking the orientations by how well the compounds fit the receptor site. In the first step a large database (Cambridge Crystallographic Database) is searched for small molecules with shapes complementary to the receptor structure. The second step is a docking procedure investigating the electrostatic and hydrogen bonding properties of the receptor displayed by the MIDAS graphics package. The steps of the design procedure is given. The algorithm includes a simple scoring function approximating a soft van der Waals potential summing up the interaction between the receptor and ligand atoms. Directional hydrogen bonding is localized using electrostatic potential of the receptor at contact points with the substrate. The shape search of (I) was described in detail. Fig. 1 shows the electrostatic potential of

Fig.1
penicillopepsin around the substrate (I) ranging from high \([-100\) kcal/mol (bold line)], intermediate \([\text{about} \sim 50\) (densely dotted)] and low \(0\) kcal/mol (sparsely dotted)] values.

The potential field indicated that there were two positions where positive charge would be desirable for tight binding: position 16 and position 23. Taking into account the receptor-ligand fit and the sites and directions of hydrogen bonding a new molecule has been proposed (II) to test the hypothesis. The synthesis of (II) is under way.

1990/309

Title: New tool for the study of structure-activity relationships in three dimensions. The hypothetical active-site lattice.
Authors: Doweyko, A.M.
Uniroyal Chemical Company, Inc., World Headquarters Middlebury CT 06749, USA.
Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.82 – 104. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.
Compounds:
a) [72 Different inhibitors of DHFR (substituted pyrimidine and triazine derivatives)];
b) Methotrexate (MTX).
Biological material: Escherichia coli dihydrofolate reductase (DHFR).
Data taken from the literature:
Ki (Michaelis inhibition constant representing the affinity of the substrate to the enzyme).
Data determined:
Conformational analysis (structures of the DHFR inhibitors were energy minimized using MM2 molecular mechanics program);
H (integer, representing hydrophobicity of the atom in the molecule and taking a value equal to \(-1, 0, +1\), corresponding to low, medium or high electron density and categorized as atom types usual in MM2 method);
HASL (hypothetical 4D active-site lattice comprising the 3D structure (cartesian coordinates) of molecules and hydrophobicity (H values) of their atoms);
FIT (measure of molecular overlap using HASL for quantitation calculated by the formula FIT = L(common)/L(ref) + L(common)/L(molecule), where L(common) is the number of lattice points in common between two molecules, L(ref) is the number of lattice point belonging to the stationary molecule and L(molecule) is the number of lattice points belonging to the molecule to be fitted).

Results: A new method has been developed for the construction of a hypothetical active site (HASL), and the estimation of the binding of potential inhibitors to this site. The molecules were quantitatively compared to one another through the use of their HASL representations. After repeated fitting one molecule lattice to another, they were merged to form a composite lattice reflecting spatial and atomic requirements of all the molecules simultaneously. The total pKi value of an inhibitor was divided to additive values among its lattice points presumed to account for the binding of every part of the molecule. Using an iterative method, a self consistent mathematical model was produced distributing the partial pKi values of the training set in a predicting manner in the lattice. The HASL model could be used quantitatively and predictively model enzyme-inhibitor interaction. A lattice resolution of 2 – 3 Å was found to be optimal. A learning set of 37 E. coli DHFR inhibitors were chosen to test the predictive power of the HASL model at various resolutions. Binding predictions (pKi values) were calculated for the entire inhibitor set at each resolution and plotted separately for the learning and test set members at 2.8 Å resolution (Fig. 1):

The 2.8 Å E. coli DHFR HASL model was tested for predicting the pKi of MTX in its bound conformation. Considering that no MTX analogues were used in developing the HASL model, the predicted pKi value of 10.11 was in quite good agreement with the experimental value (measured pKi = 10.89).

1990/310

Title: GEO. A new tool for molecular modelling based on distance geometry calculations with NMR data.
Authors: Sanner, M.; Widmer, A.; Senn*, H.; Braun, W.
Sandoz AG, Preclinical Research CH-4002 Basel, Switzerland.
Source: J. Comput.-Aided Mol. Design 1989, 3(3), 195 – 210.
Compounds:
a) Cyclosporin A (cyclic undecapeptide containing nonstandard amino acids);
b) Cyclic hexapeptide cyclo(-L-Pro-MeTyr-L-Ala-MeTyr-MeTyr-D-Ala-).
Data determined:
Molecular models (3D structure of the molecules have been constructed and displayed using the program GEO, communicating with Cambridge X-ray data bank, Brookhaven protein data bank, Sandoz X-ray data bank SYBYL and DISMAN;
Distance geometry [nuclear Overhauser enhancements (NOE) and spin-spin coupling constants were measured by 2D NMR methods, semiempirically calibrated as proton-proton distance (Å) and dihedral angle (deg) constraints and used in distance geometry calculations (DISMAN) and/or in restrained molecular dynamics calculations to determine 3D structure of molecules in solution];

Crystal structure (atomic coordinates of the compounds were determined by X-ray crystallography);

RMS [root mean square deviation (Å) of the corresponding atoms of two superimposed molecular structures].

Chemical descriptors: \( \phi, \psi \) [dihedral angles (deg)].

Results: Distance geometry calculations were carried out using GEOM and DISMAN, to identify all conformations of the compounds in solution which were consistent with experimental data obtained by NOE measurements. The application of GEOM was demonstrated by modelling cyclosporin A with and without a limited set of H-bond constraints and with a full NMR data set. In case of cyclosporin A, 100 randomly generated linear analogues of the cyclic structure were formed from the monomers. Geometric cyclization was achieved using DISMAN, resulting in many different but stereochemically correct conformations of cyclosporin A. Superposition of the backbones of the 10 best cyclic conformers showed RMS deviations between 1.8 Å and 3.1 Å. Fig. 1 shows the superposition of a DISMAN generated ring conformation (thick line) with its X-ray structure of cyclosporin A (thin line) with H-bond constraints (RMS = 1.25 Å):

Although the described method is not exhaustive, it explores a much greater variety of initial structures than had been previously possible.

1990/311

Title: A new MODEL parameter set for \( \beta \)-lactams.

Authors: Durkin, K.A.; Sherrod, M.J.; Liotta*, D.
Department of Chemistry, Emory University
Atlanta GA 30322, USA.

Source: J. Org. Chem. 1989, 54(25), 5839 – 5841.

Compounds: 22 \( \beta \)-Lactam antibiotics of diverse structure.

Data taken from the literature:
Crystal structures (crystal coordinates of the \( \beta \)-lactams were determined using X-ray diffraction method).

Chemical descriptors: \( a, b, c, d, e \) [bond angles (deg) of the lactam moiety (Fig. 1), where \( \chi = O, S, SO, SO_2 \)].

Results: Superposition of the X-ray structures and the calculated geometries of \( \beta \)-lactams using the original parameter set in the MM2 force field in MODEL gave satisfactory rms values. A lack of planarity of the \( \beta \)-lactam ring and significant differences in the calculated bond lengths and angles around the \( \beta \)-lactam nitrogen were detected, however.

In order to improve fit, a new atom type with new parameters has been developed for the \( \beta \)-lactam nitrogen (wild atom type 60 coded with symbol 22 in MODEL). The new parameters were evaluated by comparison of the calculated and X-ray geometries of the 22 \( \beta \)-lactams. Using the new parameter set, the X-ray data were satisfactorily reproduced except for the sulfone \( \beta \)-lactams. It was indicated that the AMPAC data were not suitable for the sulfones as the hypervalent sulfur compounds are not well described in the AM1 Hamiltonian. An additional parameters was, however, derived giving good structural data unrelated to the AMPAC information. It is not known which the new parameter sets is the best for the sulfone \( \beta \)-lactams.
Biological material:
a) cDNA of the hamster lung β2-adrenergic receptor and β2-adrenergic receptor kinase;
b) Bacterio-ovine- and bovine-rhodopsin and rhodopsin kinase.

Data determined:
Protein primary sequence (amino acid sequence of the hamster lung β2-adrenergic receptor has been deduced by cloning the gene and the cDNA of the hamster lung β2-adrenergic receptor); (The COSMIC molecular modeling program was used for modeling α-helices in a hydrophobic environment using φ and ω torsion angles of −59° and −44°, respectively, according to Blundell et al.);
(NHOMO, HOMO, NLUMO, LUMO (the two highest lying occupied and the two lowest lying unoccupied orbitals, respectively, calculated using INDO molecular orbital calculation);
Crystal structure (crystal coordinates of noradrenaline has been determined by X-ray diffractometry);
Conformation analysis (minimum energy conformation of the β2-adrenergic receptor model has been calculated using molecular mechanics method).

Chemical descriptors: φ, ω [torsion angles (deg)].

Results: Strong experimental evidences suggested that rhodopsin and β2-adrenergic receptor had similar secondary structure. Thus, it was assumed, that similarly to bacterio-ovine- and bovine-rhodopsins, β2-adrenergic receptor possesses a structure consisting of seven α-helices traversing the cell membrane. Fig. 1 shows the postulated arrangements of the α-helices of rhodopsin and the β2-receptor.

Using the experimental data, a model of the β2-adrenergic receptor has been generated for the study of its interaction with noradrenaline. A possible binding site was created. Successful docking indicated that HOMO and LUMO orbitals contributed to the binding in a charge-transfer interaction between Trp-109 and noradrenaline. A hydrogen bond was detected between the threonine residue of the model receptor and the noradrenaline side chain hydroxyl explaining why chirality was found to be important for the activity of adrenergic substances.

Title: Three-dimensional steric molecular modeling of the 5-hydroxytryptamines receptor pharmacophore.
Authors: Schmidt, A.W.; Peroutka, S.J.
Department of Neurology, Stanford University School of Medicine Stanford CA 94305, USA.
Source: Mol. Pharmacol. 1989, 36(4), 505-511.
Compounds:
a) 19 Different compounds categorized into seven chemical families and six main steric structures, all of them having less than 10 nm affinity for the 5-HT3 receptor;
b) ICS 205-930 (type I);
c) Atropine (type II).

Biological material: Hydroxytryptamine3 receptor (5-HT3) in rat cortices.

Data taken from the literature:
K<sub>i</sub> [binding affinity (nM) of the compounds to the 5-HT<sub>3</sub> receptor].

Data determined:
Molecular modeling (3D molecular models of each 19 compound were made using CAMSEQ/M Molecular Modeling System);
D [distance (Å) from the center of the aromatic ring to the ring-embedded nitrogen, when the nitrogen is placed in the same plane as the aromatic ring].

Results: In order to derive rules for the 5-HT<sub>3</sub> pharmacophore, a molecular graphics-based analysis was made using six core structures. The structures were aligned so as to overlay the aromatic rings and to place the ring embedded nitrogen atom in the same plane as the aromatic ring. Nine steric rules were derived from the analysis common to all 19 potent 5-HT<sub>3</sub> agents. Fig. 1 shows the 3D representation of the six overlaid 5-HT3 core structures using CAMSEQ/M. The 5-HT<sub>3</sub> inactivity of atropine could be explained because its steric properties differed from those the active ICS 205-930 only by a single atom and failed to meet two of the nine hypothetical criteria.
1990/314

Title: Novel glutamic acid derived cholecystokinin receptor ligands.
Authors: Freidinger*, R.M.; Whitter, W.L.; Gould, N.P.; Holloway, M.K.; Chang, R.S.L.; Lotti, V.J.
Merck Sharp & Dohme Research Laboratories
West Point PA 19486, USA.
Source: J. Med. Chem. 1990, 33(2), 591 -595.
Compounds:

a) 16 Compounds of type I, where 
R' = n-Pr, n-Pen, c-Hex; 
RZ = OH, OCH3Ph, OEt, pyrrolidinyl; 
R' = Ph, 3,4-C12-Ph, 2-indolyl, 
4-Cl-Ph-NH, 3-MeO-Ph-NH;

b) MK-329 and L-365,260 of type II, where R = indolyl and m-
Me-phenylamino, respectively.

Data taken from the literature:

IC50(CCK) [concentration of the compound (µmol/L) required 
for 50% inhibition of [125I]CCK-33 binding to rat 
pancreas and guinea pig cortex];

IC50(gastrin) [concentration of the compound (µmol/L) required 
for 50% inhibition of [125I]gastrin binding to guinea 
pig gastric glands];

Crystal structure (atomic coordinates of the compound of type II (R = 
indolyl) determined by X-ray diffraction study).

Data determined:

Molecular modeling (200 Conformations were calculated using a 
distance geometry algorithm and energy 
mimized by a modified MM2 force field in 
MOLEDIT).

Results: All conformers within 5 kcal/mol of the lowest energy con-
former were superposed on the X-ray structure of MK-329. Fig. 1 
shows the overlap of MK-329 (dashed line) and the compound with R' 
= n-Pen, RZ = OH, R' = 2-indolyl (full line):

A model has been created which represent the common CCK receptor 
binding elements of benzodiazepine (MK-239) and two glutamic acid 
analogues (R' = n-Pr, RZ = OH, R' = Ph, and R' = n-Pen RZ = 
OH, R' = 3,4-C12-Ph). This model was found to be useful in designing 
CCK-A selective receptor ligands with increased potency. Compounds 
designed by applying the model did not possess, however, the expected 
activities. The model was not useful for the design of CCK-B/gastrin 
selective analogues.
Correlation of \( \beta \)-bend conformations of tetrapeptides with their activities in CD4-receptor binding assays.

**Authors:** Shah, D.; Chen, J.M.; Carty, R.P.; Pincus, M.R.; Scheraga, H.A.

Cornell University, Department of Chemistry, Baker Laboratory, Ithaca NY 14853-1301, USA.

Source: Int. J. Peptide Protein Res. 1989, 34(4), 325–332.

**Biological material:**
- a) CD4 receptor of monocytes;
- b) 5 Tetrapeptides: Thr-Thr-Asn-Tyr (from peptide T), Thr-Ile-Asn-Tyr (from polio virus coat protein), Ser-Ser-Asn-Tyr (from ribonuclease A), Ser-Ser-Ala-Tyr (from the gp 120 coat protein of human immunodeficiency virus), Asn-Thr-Lys-Tyr (from an active synthetic pentapeptide);
- c) Ribonuclease A.

**Data taken from the literature:**
Crystal structure (crystal coordinates of the proteins were determined using X-ray diffraction method).

**Data determined:**
Conformational analysis [over 200000 low energy conformations were computed for each tetrapeptide using ECEPP (Empirical Conformational Energy Program for Peptides) and the chain buildup procedure];

RMS [root mean square deviation (Å) of the corresponding atoms of two superimposed molecular structures];

Superimposition (A tetrapeptide conformation was regarded as superimposable if the RMS deviation of its backbone heavy atom coordinates from the corresponding coordinates of the ribonuclease A sequence (residues 22-25) was less than 1 Å);

\( P_i \) (probability that a given tetrapeptide sequence is superimposable on the ribonuclease A structure);

\( P_{i,j} \) (probability that the \( i \)th residue (amino acid) will occur in the \( j \)th conformational state of the tetrapeptide which is superimposable to ribonuclease A).

**Results:** It was suggested that the five tetrapeptides were essential components of larger peptides and might be responsible for their biological activity (binding to the CD4 receptor). Earlier it was hypothesized that the critical tetrapeptide located in a segment of ribonuclease A, would assume low energy conformations (residues 22–25, a \( \beta \)-bend, having a segment homologous to the sequence of peptide T). Low energy conformers of the tetrapeptides could be superimposed to the native structure of segment 22–25 of ribonuclease A. Fig. shows the superimposition of peptide T (full square):

Many low energy conformers could be calculated for the tetrapeptides but for the polio sequence. The \( P_i \) value for most tetrapeptides were 5–10 times higher that the value of the less active polio sequence. The results supported the hypothesis that the active peptide T adopts the native ribonuclease \( \beta \)-bend.

---

Potential cardiotonics. 4. Synthesis, cardiovascular activity, molecule- and crystal structure of 5-phenyl- and 5-(pyrid-4-yl)-substituted 1,2-dihydro[pyrid-2-thions.

**Authors:** Hagen*, V.; Rumler, A.; Reck, G.; Hagen, A.; Labes, D.; Heer, S.

Institut für Wirkstoffforschung, Alfred Kowalke-Str. 4, O-1136 Berlin.

Source: Pharmazie 1989, 44(12), 809–813 (German with English Summary).

**Compounds:** 2 Compounds of type I, where \( X = O \) (milrinon) and \( S \).

**Biological material:**
- a) Guinea pig;
- b) Dogs.

**Data determined:**
ED\(_{30}\) [dose of the compound (mol/kg) required for 30 % increase of the heart beat frequency of guinea pig or dog heart];

ED\(_{10}\) [dose of the compound (mol/kg) required for 10 % decrease of systolic or diastolic blood pressure of dog];

Crystal structure (atomic coordinates of the compounds were determined by X-ray diffraction);

Molecular modeling (molecule models were built using MOLPAC);

MEP [molecular electrostatic potential (MEP) (dimension not given) was calculated using CNDO/2].

**Results:** Milrinon and its oxygen containing bioisoster possess highly similar crystal structure and MEP isopotential maps (Fig. 1 and Fig. 2) matching the 5-point model of the for cardiotonically active selective PDE-III inhibitors.

---

Fig.1
Both compounds show strong positive inotropic and vasodilator activity. It was suggested that the negative potential region around the thio carbonyl group such as the carbonyl group in milrinon imitates the negative potential field around the phosphate group of cAMP.

1990/318

Title: Molecular mechanics calculations of cyclosporin A analogues. Effect of chirality and degree of substitution on the side chain conformations of (2S,3R,4R,6E)-3-hydroxy-4-methyl-2-(methylamino)-6-octenoic acid and related derivatives.

Authors: Miller, K.E.; Rich*, D.H.

School of Pharmacy, University of Wisconsin-Madison
425 North Charter Street, Madison WI 53706, USA.

Source: J. Am. Chem. Soc. 1989, 111(22), 8351-8362.

Compounds:
- a) Cyclosporin A (CsA) of type I;
- b) 4 Cyclosporin A analogues obtained by substituting the amino acid in position one by MeBmt, (4S)-MeBmt, MeBth and MeBmt of types II, III, IV and V, respectively;
- c) Concanavalin A.

Data taken from the literature:
Crystal structure (crystal structure of native CsA obtained from the Cambridge Crystallographic Databank);
IS [immunosuppressive activity (%) measured by the degree of inhibition of concanavalin A stimulated murine thymocytes].

Data determined:
Molecular modeling (molecular model of CsA was built using SYBYL Ver. 5.03);
Solution structure [solution conformation of CsA in CDCl₃ has been elucidated via molecular dynamics simulation incorporating 58 distance constrains obtained from IR spectroscopy and nuclear Overhauser effect (NOE) data];
Conformational analysis (conformational analysis was performed using the SEARCH subroutine within SYBYL);
Energy minimization (low energy conformers were calculated using molecular mechanics within MacroModel Ver. 1.5 applying an all-atom version of the AMBER force field).

Results: A total of 12 conformations of CsA have been identified within 4 kcal/mol of the minimum energy conformer. Population analysis showed that one conformer dominates in solution. Fig. 1 shows the superposition of the peptide backbone of the crystal and solution structures of CsA (crystal structure is drawn with thick line and the solution structure with thin line).

1990/319

Title: Molecular modeling study of class Ia and IIb antiarrhythmics. Modulated receptor hypotheses.

Author: Marrer, S.

Pharmaceutisches Institut der Freien Universitat Berlin
Koenigin-Luise Strasse 2 + 4, D-1000 Berlin 33, Federal Republic of Germany.

Source: Pharm. Acta Helv. 1989, 64(12), 338-344. (German with English summary).

Compounds: 3 Antiarrhythmic drugs of class I: chitin (class Ia), lidocaine (class Ia), EO 122 (class IIa) of structural type I, II and III, respectively.
Data determined:

Molecular modeling (models of (I), (II) and (III) were built using SYBYL based on X-ray coordinates of the compounds);

Conformational analysis (minimum energy conformations of the compounds were calculated using the SEARCH option of SYBYL and MNDO method; [conformational energy (kcal/mol)];

E [interaction energy of the molecules (kcal/mol) with a hypothetical receptor probe (negatively charged oxygen atom) calculated by GRID].

Chemical descriptors:

A<sub>a</sub>, B<sub>a</sub> [dihedral angles (deg) of the N(1)-C(8)-C(9)-C(4') and C(9)-N(11)-C(12)-C(13) bonds of (III), respectively].

Results: The specific receptor area of the sodium channel was modeled with a negatively charged oxygen probe (carboxyl group), interacting with the positively charged (protonated) ligand. Fig. 1 shows areas for energetically favorable interaction (areas I, II and III) using isoenergy contours (−4 kcal/mol):

In Fig. 1, lidocaine K1 and K2 are two favorable conformations for binding. The study indicated that the receptor site of the sodium channel in the open state has two, and in the inactivated state one negatively charged areas. Fig. 1 explains the preferred activity of quinidine for the open state of sodium channel and the equal activity of lidocaine for the open and inactivated sodium channel which can interact with each (I, II, III) negatively charged areas. This result supports the "modulated receptor hypothesis" proposed earlier.

Title: Crystal structures and conformational analysis of ochratoxin A and B. Probing the chemical structure causing toxicity.

Authors: Bredenkamp*, M.W.; Dillen, J.L.M.; van Rooyen, P.H.; Steyn, P.S.

Department of Chemistry, Rand Afrikaans University
PO Box 524, Johannesburg 2000, Republic of South Africa.

Source: J. Chem. Soc. Perkin Trans. I 1989, No. 11, 1835 – 1839.

Compounds:

a) Three Aspergillus toxins of type I: ochratoxin A, where R<sup>1</sup> = H; R<sup>2</sup> = Cl; ochratoxin B, where R<sup>1</sup> = R<sup>2</sup> = H; ochratoxin C, where R<sup>1</sup> = Et; R<sup>2</sup> = Cl;

b) Two hydrolysis products of type II: ochratoxin α, where R = Cl; ochratoxin β, where R = H.

1990/320
Biological material:
a) Aspergillus ochraceus;
b) Carbopeptidase A.

Data determined:
- $k_{obs}$ (first order rate coefficient ($10^{-6}$/sec) of the hydrochloric acid hydrolysis of ochratoxin A and B);
- X-ray crystallography (Coordinates of the crystal structure of ochratoxin A and B was obtained using X-ray diffraction);
- Molecular modeling (models of ochratoxin A and B was built using ALCHEMY);
- $\delta$ ($^{13}$C NMR chemical shifts (ppm) of the amide and ester carbonyls of the ochratoxins).

Chemical descriptors:
- $pK_a$ (negative logarithm of the acidic dissociation constant).

Results: A reversal of the hydrolysis rate between ochratoxin A and B was observed comparing the hydrolysis rates obtained in vitro (carbopeptidase A) and in vivo (hydrochloric acid). The difference in hydrolysis rates cannot be due to conformation since the two toxins have the same conformation in both crystal and in solution. Fig. 1 shows the fit of ochratoxin A and B based on superimposing the phenolic carbon atoms.

It is suggested that the relative large steric bulk of the chloro atom hinders the fit between ochratoxin A and the receptor site of carbopeptidase A. Thus, probably the slower metabolism is the reason, why ochratoxin A is more toxic than ochratoxin B.

1990/321

Title: Inhibitors of cholesterol biosynthesis. 1. Trans-6-(2-Pyrrol-1-ylethyl)-4-hydroxyxpyran-2-ones, a novel series of HMG-CoA reductase inhibitors. 1. Effects of structural modifications at the 2- and 5-positions of the pyrrole nucleus.

Authors: Roth*, B.D.; Ortwine, D.F.; Hoefle, M.L.; Stratton, C.D.; Sliskovic, D.R.; Wilson, M.W.; Newton, R.S.
Parke-Davis Pharmaceutical Research Division, Warner-Lambert Company
280 Plymouth Road, Ann Arbor MI 48105, USA.
Source: J. Med. Chem. 1990, 33(1), 21 - 31.

Compounds:
a) 40 Novel trans-6-(2-Pyrrol-1-ylethyl)-4-hydroxyxpyran-2-ones of type I, where $R^1 = 4-F$-Ph, 4-Ph-Ph, 4-MeO-Ph, 4-Cl-Ph, 4-HO-Ph, 3-CF$_3$-Ph, 3-MeO-Ph, 3-HO-Ph, 2-MeO-Ph, 2-HO-Ph, 2-naphthyl, 1-naphthyl, c-hexyl, c-hexyl, aliphatic bicyclo moiety, Ph$_2$CH, 2,4-F$_2$-Ph, 2,6-(MeO)$_2$-Ph, 2,5-Me$_2$-Ph, 2-i-PrO-Ph, 2-CI-Ph, Et$_3$CH, $\beta$-MeO-acylanyl, $R^2 =$ Me, i-Pr, t-Bu, CHEt, c-propyl, c-butyl, c-hexyl, CF$_3$; $X =$ ortho-, meta- and para-phenylene, -CH$_2$(CH$_3$)$_2$, -CH$_2$CH$_2$(CH$_3$)$_2$, -CH(CH$_3$)$_2$. $\delta$

b) Compactin of type II.

1990/322

Title: Synthesis and biological activity of new HMG-CoA reductase inhibitors. 1. Lactones of pyridine- and pyrimidine-substituted 3,5-dihydroxy-6-heptenoic (-heptanoic) acids.

Charge distribution studies showed that compactin had two distinct regions of relatively large partial charges corresponding to the pyrrole ring and the isobutyric acid side chain. Experiments for more closely mimicking the polar regions associated with the high activity of compactin indicated that potency of the new compounds was relatively insensitive to the polarity of the $R^1$ group. It was also suggested that an electron deficient pyrrole ring was required for high potency.
Authors: Beck, G.; Kesseler, K.; Baader, E.; Bartmann*, W.; Bergmann, A.; Granzer, E.; Jendralla, H.; Kerekjarto, B.v.; Krause, R.; Paulus, E.; Schubert, W.; Wess, G.
Hoechst AG
Postfach 80 03 20, D-62630 Frankfurt/M. 80, Federal Republic of Germany.
Source: J. Med. Chem. 1990, 33(1), 52 – 60.

Compounds:

a) 29 Cholesterol biosynthesis inhibitors of type I, where R' = Me, Et, i-Pr, t-Bu, c-Hex, 4-F-Ph; R2 = 4-F-Ph, 4-MeO-Ph, 4-CF3-Ph, i-Pr, 4-CI-Ph; R3 = Me, Ph, i-Pr, t-Bu, c-Hex, 4-F-Ph, 2,5-Me2-Ph, 3,5-Me2-Ph; Y = CH, N; A-B = (E)-CH=CH, (2)-

b) 2 Compounds of type II, where R = H (compactin) and R = Me (mevinolin);

Crystal structure (atomic coordinates of the compounds were determined by X-ray diffraction);

Biological material: 3-Hydroxy-3-methylglutaryl-CoA reductase (HMG-CoA reductase) from rat liver.

Data determined:

[concentration of the compound (nmol/L) required for 50 % inhibition of HMG-CoA];

Conformational analysis (low energy conformations of the compounds were calculated using SYBYL).

Results: Extensive structure-activity relationships studies led to the synthesis of HMG-CoA reductase inhibitors exceeding the activity of mevinolin. Fig. 1 shows the superposition of mevinolin and one of the more active inhibitors (HR 780, where R' = i-Pr; R2 = 4-F-Ph; R3 = Ph; Y = CH; A-B = (E)-CH=CH) occupying the same regions of space except for the phenyl ring (mevinolin thick line, HR 780 thin line):

The standard deviation of the two superposed structures was calculated to be 0.217 Å.

Title: Conformation-activity relationship of sweet molecules. Comparison of aspartame and naphthimidazolesulfonic acids.

Authors: Castiglione-Morelli, M.A.; Lelj, F.; Naider, F.; Tallon, M.; Tancredi, T.; Temussi*, P.A.

Dipartimento di Chimica, Università di Napoli
via Mezzocannone 4, I-80134 Napoli, Italy.
Source: J. Med. Chem. 1990, 33(2), 514 – 520.

Compounds:

a) Aspartame (α-L-Asp-L-PheOMe)(sweet) and its analogue (α-D-Asp-D-PheOMe)(bitter);

b) 2 Sulfonaphthimidazoles: 3-Anilino-2-styryl-3H-naphth[1,2-d]imidazolesulfonate (SSN) (sweet) of type I and 3-anilino-2-phenyl-3H-naphtho[1,2-d]imidazolesulfonate (TSN) (tasteless) of type II.

Data taken from the literature:

Crystal structure (atomic coordinates of SSN and TSN were determined by X-ray diffraction study).

Data determined:

Conformational analysis [low energy conformations of the molecules were calculated using molecular dynamics (MD) calculation (GROMOS package) taking into account stretching, bending and torsion contributions, both proper and improper, and adding nonbonded, van der Waals and electrostatic interactions, and hydrogen bond contributions];

NMR

(The proton and carbon spectra of TSN and SSN obtained by the use of several 1-D and 2-D techniques leading to complete assignments of all resonances).

Results: The previously determined sweetness receptor has been refined with the use of combination of saccharin and two large and nearly completely rigid, but geometrically optimized (ab initio MD calculations) molecules (I and II). Fig. 1 and Fig. 2 show the correct and bad fitting of SSN and TSN, respectively, showing the contour of the new sweet taste site:
500 MHz $^1$H NMR studies indicated that the minimum energy conformation of aspartame is $F_1D_{II}$ instead of $F_1D_{III}$ which is not consistent with the sweetness receptor model, the required interconversion is, however, only 1 kcal/mol. MD simulation in vacuo also showed the high flexibility of aspartame and the availability of the energetically favorable conformation of $F_1D_{II}$ which fits the refined sweet site (Fig. 3):

**Fig. 3**

Title: Synthesis and cardiac electrophysiological activity of aryl-substituted derivatives of the class III antiarrhythmic agent sematilide. Potential class I/III agents.

Authors: Phillips*, G.B.; Morgan, T.K.,Jr.; Nickisch, K.; Lind, J.M.; Gorenz, R.P.; Wohl, R.A.; Argentielti, T.M.; Sullivan, M.E. Departments of Medicinal Chemistry and Pharmacology, Berlex Laboratories, Inc. 110 East Hanover Avenue, Cedar Knolls NJ 07927, USA.

Source: J. Med. Chem. 1990, 33(2), 627 – 633.

Compounds: 12 Compounds of type I (selective class III antiarrhythmic agents), where $R^1 = H$, Ph, 4-Et-Ph, 4-Cl-Ph, 4-MeO-Ph, 4-MeSO$_3$NH-Ph, 2,6-Me$_2$-Ph, 2,6-(i-Pr)$_2$-Ph, naphthalenyl; $R^2 = H$, Ph, naphthalenyl; $R^3 = H$, Ph, naphthalenyl, showing mostly good electrophysiological activity.

Chemical descriptors:

- $N$-$C$-$C$-$N$ [dihedral angle (deg)].

**Results:** An attempt was made to correlate electrophysiological activity with the effect of the position of the aryl group on the conformation of the side chain using molecular modeling. The study suggested that the compounds with class III activity prefer a gauche (A in Fig. 1) and compounds in which class I activity prefer trans relationship of the nitrogens (B in Fig. 1):

**Fig. 1**

The study indicated that the point of attachment of the aryl moiety had an effect on the side chain conformation which appeared to be a controlling factor of the electrophysiological profile of these compounds.

**1990/324**

Title: A molecular mechanics analysis of molecular recognition by cyclodextrin mimics of $\alpha$-chymotrypsin.

Authors: Venanzi*, C.A.; Canzian, P.M.; Zhang, Z.; Bunce, J.D. Department of Chemical Engineering, Chemistry, and Environmental Sciences, New Jersey Institute of Technology Newark NJ 07102, USA.

Source: J. Computat. Chem. 1989, 10(8), 1038 – 1052.

Compounds: $\beta$-Cycloexdrin (B-CD), hepta-$N$-methylformamide $\beta$-cyclodextrin (capped B-CD) (formulas I and II show the glucose unit and the capped glucose unit of the CDs):
Biological material: Chymotrypsin.

Data taken from the literature:
Crystal structure (crystal coordinates of the macrocycles determined using X-ray diffraction analysis).

Data determined:
Molecular modeling
(models of B-CD and in chains by N-methylformamide and N-dimethylformamide substituted (capped) B-CD were built using the AMBER program and the coordinates for building the N-methylformamide substituent were calculated using MNDO in the MOPAC program);
Conformational analysis (energy minimization of the molecules were calculated in vacuo using molecular mechanics program with the AMBER force field);
Structure superposition (energy minimized structures of B-CD and capped B-CD were separately fit to the X-ray structure of the B-CD complex);

Results: B-CD and capped B-CD were analyzed as biomimetic models of the active site of chymotrypsin. Capped B-CD was shown to be the more effective biomimetic catalyst. Capping also altered certain structural features of molecular recognition. The orientation of the secondary hydroxyls were altered due to twisting of some of the glucose units. Secondary hydroxyl oxygen mimics the Ser-195 of chymotrypsin in initiating the acyl transfer event through nucleophilic attack on the substrate. Fig. 1 shows the energy minimized structures of B-CD (A) and capped B-CD (B) (fragment number is given in parenthesis).
The MEP maps of B-CD and capped B-CD showed that the qualitative features of the electrostatic recognition were practically the same in the two mimics.

1990/326

Title: Conformational and intramolecular hydrogen bonding effects on post-emergence and pre-emergence selectivities of herbicidal pyrrole dicarboxylates.
Authors: Andrea*, T.A.; Stranz, D.D.; Yang, A.; Kleier, D.A.; Patel, K.M.; Powell, J.E.; Price, T.P.; Marynick, D.S.
Agricultural Products Department, E.I. Dupont de Nemours, Stine-Haskell Research Center Newark DE 19711, USA.
Source: Pestic. Sci. 1990, 28(1), 49 - 68.
Compounds: 4 Compounds of type I, where R = 3-Cl, 4-Cl, 5-Cl, 6-Cl.

Biological material: Four monocotyledonous (Johnson grass, yellow foxtail, barnyard grass, yellow millet) and four dicotyledonous weed species (velvetleaf, morning glory, prickly sida, sicklepod).

Data determined:
Herbicidal activity [pre-emergence and post-emergence herbicidal activities of the compounds were measured and rated using a scale ranging from 0 (no activity) to 9 (complete kill)];
TSCF [measure of the compound's ability (dimension not given) to translocate upwards in plants through xylem vessels];
Kd [soil sorption coefficient calculated by the formula Kd = C0/C0, where C0 is the concentration of the compound (µg compound/g soil) and C0 is the concentration of the compound (µg compound/mL) in water solution in equilibrium with the soil];

Molecular modeling [models of the compounds were built using MACCS and PRXBLD programs];
Conformational analysis (minimum energy conformations of the compounds were calculated using MM2 molecular mechanics method);

Computer graphics (molecules were visualized using program MOGLI on an Evans and Sutherland Picture System 33);

Electronic structure [total energies, orbital eigenvalues, atomic charges and dipole moments of simple model analogs of type I were calculated using PRDDO (Partial Retention of Diatomic Overlap) level of approximation].

Chemical descriptors:
- logP (logarithm of the partition coefficient in 1-octanollwater).

Results: Conformational analyses and high level quantum mechanical calculations of the conformational preferences showed that the compounds with \( R = 4-\text{CI} \) and \( 5-\text{CI} \) substituents adopt a coplanar structure stabilized by intramolecular hydrogen bond, whereas the 3-C analogue does not (Fig. 1):

\[ \text{Fig.1} \]

Higher logP values (0.6 – 1.0 logarithmic unit difference), higher \( K_d \) and TSCF values of the 4-Cl and 5-Cl substituted compounds relative to the 3-Cl analog were interpreted as the result of the intramolecular hydrogen bond and were consistent with the observation that the 4-Cl and 5-Cl analogs were active as post-emergence but not pre-emergence herbicides while the 3-Cl derivative was active in both modes.

1990/327

Title: Application of molecular modeling techniques to pheromones of the marine brown algae Cutleria multifida and Ectocarpus siliculosus (Phaeophyceae). Metalloproteins as chemoreceptors?

Authors: Boland, W.; Hoever, F.-P.; Krüger, B.-W.

Institut für Organische Chemie Richard-Willstätter-Allee, D-7500 Karlsruhe, Federal Republic of Germany.

Source: Z. Naturforsch. 1989, 44(9-10), 829 – 857.

Compounds:
- a) Algal pheromones: \((+)-(6S)\) ectocarpene (I) and \((+)-(3S,4S)\) multifidene (II);
- b) 4 Analogues of ectocarpene and multifidene.

\[ \text{Fig.1} \]

Biological material:
- a) Male gametes of Ectocarpus siliculosus;
- b) Cutleria multifida males.

Data taken from the literature:
- Crystal structure (crystal coordinates of compounds were obtained from the Cambridge Structural Database);
- Q [threshold concentration (mol/L) of the response of male gametes of E. siliculosus and C. multifida to the pheromones (details not given).]

Data determined:
- Molecular modeling (geometrical models of the compounds were constructed using information from the Cambridge Structural Data Base (CSD) and calculated using molecular mechanics methods in SYBYL);
- Conformational analysis (minimum energy conformations of the compounds were calculated using molecular mechanics method within SYBYL).

Chemical descriptors:
- \( K_{FC72} \) [partition coefficient in FC72/water (FC72 = fluorocarbon solvent)].

Results: As both ectocarpene (I) and multifidene (II) trigger mutual cross reactions between male gametes of Ectocarpus siliculosus and Cutleria multifida males it was supposed that a common mode of binding should exist for the two structurally different pheromones. The active analogue approach was applied to model the pheromone receptor by superposing the minimum energy conformations of active structural analogues (III, IV, V, VI) on ectocarpene and multifidene. The common active conformation of (I) and (II) was extracted by systematic superimposition of the analogues. To explain the function of the double bonds in the pheromones, the presence of a receptor bound metal cation was assumed. Simultaneous optimization of both structures without and with a receptor bound metal cation resulted in virtually the same conformations. Fig. 1 shows the mapping of multifidene onto ectocarpene in their biologically relevant conformations.

1990/328

Title: Critical differences in the binding of aryl phosphate and carbamate inhibitors of acetylcholinesterases.
Author: Magee, P.S.
BIOSAR Research Project
Vallejo CA 94591, USA.

Source: ACS Symposium Series 1989, No.413. In: Probing Bioactive Mechanisms p.147–156. Edited by Magee, P.S., Henry, D.R., Block, J.H., American Chemical Society, Washington, 1989.

Compounds:
a) 18 Organophosphate insecticides of type I, where \( R^1, R^2 = OR, SR (R = C_1 - C_3 \text{ alkyl}), \text{Et, Ph}; R^1 = H, Cl; R^2 = H, Me; R^3 = Cl, Br, I, SMc, CN, NO_2, S(=O)Me; R^4 = H, Cl; X = O, S; \)
b) 14 Commercial carbamate insecticides of type II, where \( R^1 = H, s-Bu, C_1, O-i-Pr; R^2 = H, Me, s-Pen, t-Bu, Et, i-Pr; R^3 = H, SMe, NMe_2, N(CH_2CH=CH_2); R^4 = H, Me, Et, t-Bu; R^1 \text{ and } R^3 \text{ together are } -O-C(CH_3)O-, -O-C(CH_3)CH_2-, -CH=CH-CH_2-CH-; \)
c) Acetylcholine.

Biological material: Acetylcholinesterase.

Data determined:

D values for the meta- and para-methyl substituents of N-methylcarbamate and dimethylphosphate were meta = 6.20, para = 8.10 and meta = 5.48, para = 7.03 Å, respectively.

---

**Protein**

1990/329

**Title:** A comparison of the CHARMM, AMBER and ECEPP potentials for peptides. 1. Conformational predictions for the tandemly repeated peptide (Asn-Ala-Asn-Pro)_9.

Authors: Roterman, I.K.; Gibson, K.D.; Scheraga*, H.A.

Baker Laboratory of Chemistry, Cornell University
Ithaca NY 14853 – 1301, USA.

Source: J. Biomol. Str. Dyn. 1989, 7(3), 391-419.

Biological material: Tandemly repeated peptide (Asn-Ala-Asn-Pro) which is a major immunogenic epitope in the circumsporozoite (CS) protein of Plasmodium falciparum (malaria parasite).

Data taken from the literature:

Conformational analysis [low energy conformations of (Asn-Ala-Asn-Pro)_9];

\( \phi, \psi, \omega, \chi^1, \chi^2, \chi^3 \) [dihedral angles (deg)].

Data determined:

Conformational analysis [minimum energy conformations of (Asn-Ala-Asn-Pro)_9 was calculated using CHARMM (Chemistry at Harvard Macromolecular Mechanics), AMBER (Assisted Model Building with Energy Refinement) and ECEPP (Empirical Conformational Energy Program for Peptides) potential energy functions];

rms [root mean square deviation (Å) of the position of the corresponding atoms of two superimposed molecular structures].

Results: 24 Low energy conformations of (Asn-Ala-Asn-Pro)_9 has been determined using CHARMM, ECEPP and AMBER in order to determine their final conformations and relative energies. The final conformations were compared calculating the rms values of their C\(^\alpha\) atoms and matching the parameters of the energy minimized (Asn-Ala-Asn-Pro)_9 peptide to that of the ideal helix or coiled coil. The similarity of the final conformations obtained by using any two different potentials starting from the same conformation varied from the satisfactory to highly unacceptable. The extent of difference between any pairs of the final conformations generated by two different potential energy functions were not significantly different. The lowest-energy conformation calculated by each of the energy potentials for any starting conformation was a left handed helix and pair-wise superposition of the C\(^\alpha\) atoms in the final conformations showed small rms values (1.0 – 1.3 Å). It was suggested that the native conformation of (Asn-Ala-Asn-Pro)_9 in the CS protein may be a left-handed helix, since all three potential energy functions generated such conformation.

1990/330

**Title:** A computer model to dynamically simulate protein folding.

Studies with crambin.

Authors: Wilson*, C.; Doniach, S.

Graduate Group in Biophysics, Department of Biochemistry and Biophysics, University of California
San Francisco CA 94143-0448, USA.
Simulated annealing

Phi-psi probability plot (probabilities of the occurrences of phi-psi dihedral angle pairs for each amino acid were determined and plotted using the data of approximately 100 proteins from the Brookhaven Protein Data Bank);

Simulated annealing (optimization technique for the reproduction of the folding process converging to the native minimum energy structure by dynamically sampling many different conformations of the simplified protein backbone).

Chemical descriptors:
Phi-psi values [dihedral angles (deg) defined by the bonds on either side of the α-carbon atom of the amino acid residue in a protein].

Results: A simplified model has been developed for the representation of protein structures. Protein folding was simulated assuming a freely rotating rigid chain where the effect of each side chain approximated by a single atom. Phi-psi probabilities were used to determine the potentials representing the attraction or repulsion between the different amino acid residues. Many characteristics of native proteins have been successfully reproduced by the model: (i) the optimization was started from protein models with random conformations and led to protein models with secondary structural features (α-helices and β-strands) similar by nature and site to that of the native protein; (ii) the formation of secondary structure was found to be sequence specific influenced by long-range interactions; (iii) the association of certain pairs of cysteine residues were preferred compared to other cysteine pairs depending on folding; (iv) the empirical potentials obtained from phi-psi probabilities led to the formation of a hydrophobic core of the model peptide.

Chemical descriptors: [accessible surface area (Å²)].

Results: Four kinds of Monte Carlo simulations of about 20,000 steps of the conformations of crambin were carried out by using the second derivative matrix of energy functions (starting from native and unfolded conformations both in two kinds of systems, in vacuo and in solution). Fig. 1 shows the native (a) and the unfolded (b) conformation of crambin.

Starting from native conformation, the differences between the mean properties of the simulated crambin conformations obtained from in vacuo and solution calculations were not very large. The fluctuations around the mean conformation during simulation were smaller in solution than in vacuo, however. Simulation starting from the unfolded conformation resulted in a more intensive fluctuation of the structure in solution than in vacuo indicating the importance of the hydration energy term in the model. The conformations generated in the simulations starting from the native conformation deviate slightly from the X-ray conformation (rms = 0.70 Å and 1.10 Å for in vacuo and solution simulations, respectively). The results indicate that the simulations of the protein with hydration energy are more realistic that the simulations without hydration energy.
Results: Earlier studies overestimated the catalytic rate decrease of the hypothetical D102A point mutant of thrombin (20 orders of magnitude decrease calculated instead of the 4 order of magnitude measured). The source of error was due to an overestimation of V and neglecting the effects of the surrounding water molecules and induced dipoles. To compensate for these errors, a scale factor of 0.12 was introduced into the calculations. As a result of the rescaling, one magnitude increase of $k_{cat}$ for the D121 mutant and two magnitudes decrease of $k_{cat}$ of the K41 mutant of ribonuclease A was predicted. It was shown that the effect of the mutations on the catalytic rate depended almost entirely on steric factors. It was suggested that in mutants of serine proteases where the buried Asp is replaced by Ala or Asp, the $k_{cat}$ value will decrease between 4 - 6 orders of magnitude.

Title: High-resolution structure of an HIV zinc fingerlike domain via a new NMR-based distance geometry approach.

Authors: Summers*, M.F.; South, T.L.; Kim, B.; Hare, D.R. Department of Chemistry and Biochemistry, University of Maryland Baltimore County Baltimore MD 21228, USA.

Source: Biochemistry 1990, 29(2), 329 – 340.

Biological material:

a) Polypeptide with 18 amino acid sequence comprising a zinc fingerlike domain from the gag protein p55 of human immunodeficiency virus (HIV);

b) Rubredoxin.

Data determined:

$^{1}H$ NMR [proton NMR signal assignments were made by first determining scalar connectivities within amino acid residues by COSY (double quantum filtered correlated spectroscopy) and HOHAHA (homonuclear Hartmann-Hahn spectroscopy) and then correlating the signals of adjacent residues by 2D NOESY (two-dimensional nuclear Overhauser effect spectroscopy) and ROESY (rotating-frame Overhauser effect spectroscopy)];

$^{13}C$ NMR [NMR technique for obtaining chemical shift assignments (ppm) for Zn(p55F) and for obtaining broad-band $^{13}C$ decoupling];

2D NOESY back-calculation [two-dimensional nuclear Overhauser effect spectroscopy used to generate simulated spectra for structures obtained by distance geometry (DG) calculations];

RMS [root mean square deviation (Å) of the corresponding atoms of two superimposed molecular structures].

Results: The atomic resolution structure of an HIV zinc fingerlike domain has been generated by a new NMR-based DG method using 2D NOESY back-calculation. The quality of the structures thus obtained were evaluated on the basis of the consistence with the experimental data (comparison of measured and back-calculated NMR spectra) rather than comparing it to structural informations from other sources (e.g. X-ray data). The method provided a quantitative measure of consistence between experimental and calculated data which allowed for the use of tighter interproton distance constraints. The folding of the C(1)-F(2)-N(3)-C(4)-G(5)-K(6) residues were found to be virtually identical with the folding of the related residues in the X-ray structure of the iron domain of rubredoxin (RMS values 0.46 and 0.35 Å). The backbone folding of the peptide was found to be significantly different from that of the "classical" DNA-binding Zn-finger. Fig. 1 shows the wire frame model of all the backbone atoms and certain side chain atoms of the peptide (DG structure) (dashed lines indicate hydrogen atoms):

![Fig.1](image_url)
active site of the protease dimer in an extended conformation with extensive van der Waals and hydrogen bonding and was more than 80% excluded from contact with the surrounding water (Fig. 1, where the inhibitor is shown in thicker lines and the hydrogen bonds in dashed lines):

![Fig. 1](image1.png)

The binding of the inhibitor induced substantial movement in the enzyme around the residues 77 to 82 in both subunits at places exceeding 2 Å. Fig. 2 shows the superimposed native (thin line) and inhibited (thick line) protease dimers with the inhibitor in the binding pocket.

![Fig. 2](image2.png)

The structure of glutamine synthetase is discussed. It was established that hydrophobic interactions are important for the intersubunit interactions, and the hydrophobic interactions between the two rings of subunits are stronger than between the subunits within a ring. The C-terminal helix contribute strongly to the inter-ring hydrophobic interaction. ASPs are suggested to estimate the contribution of the hydrophobic energy to protein folding and subunit assembly and the binding of small molecules to proteins.

**Title:** Determination of the complete three-dimensional structure of the trypsin inhibitor from squash seeds in aqueous solution by nuclear magnetic resonance and a combination of distance geometry and dynamical simulated annealing.

**Authors:** Holak*, T.A.; Gondol, D.; Otlewski, J.; Wilusz, T.

Max-Planck-Institut für Biochemie
D-8033 Martinsried bei München, Federal Republic of Germany.

**Source:** J. Mol. Biol. 1989, No.210, 635–648.

**Biological material:** Trypsin inhibitor (CMTI-I) from pumpkin seeds Cucurbita maxima.

**Data taken from the literature:**
Crystal structure (atomic coordinates of CMTI-I was determined by X-ray diffraction method).

**Data determined:**
(34 three-dimensional structures of CMTI-I in aqueous solution was determined by molecular mechanics (XPLOR))
constrains, supplemented by 22 backbone angle constrains, all obtained from $^2$D $^1$H NMR measurements); Stereospecific assignments (stereospecificity was taken into account for 17 of the 48 prochiral centers of CMTI-I using floating chirality assignment during the dynamical simulated annealing calculations); RMS [root mean square deviation (Å) of the corresponding atoms of two superimposed structures]; $\phi$ [backbone torsion angle constrains (deg) derived from NMR coupling constants data].

Results: In order to obtain information of the 3D structure of the free CMTI-I in solution, a total of 34 inhibitor structures were calculated by a combination of distance-geometry and dynamical simulated annealing methods, resulting in well defined 3D positions for the backbone and side-chain atoms. Fig. 1 shows the superposition of the backbone (N, C', C, O) atoms of the structures best fitted to residues 2 to 29 (binding loop): The average RMS difference between the individual structures and the minimized mean structure was 0.35(±0.08) Å for the backbone atoms and 0.89(±0.17) Å for all heavy atoms.

The model was consistent with the requirements of steric factors, complementary electrostatic interactions and NMR data of the complex. Comparison of the new model and the NMR data of the previously proposed flavodoxin-cytochrome $c_3$ complex showed that both proteins interacted with the same heme-group of cytochrome $c_3$.
detected in the segments from the residues 16 to 18 and 25 - 25. Fig. 1 shows the best superposition (residues 2 to 29) of the NMR and crystal structure of CMT-I indicating the backbone C, C*, N, O, as well as the disulfide C8 and S atoms:

Fig.1

It was demonstrated that uncertainty in NMR structure determination can be eliminated by including stereospecific assignments and precise distance constraints in the definition of the structure.

1990/339

Title: Coenzyme B12 chemistry. The crystal and molecular structure of Cob(II)alamin.

Authors: Kräutler*, B.; Keller, W.; Kratky, C.
Laboratory of Organic Chemistry, Eidgenössische Technische Hochschule (ETH) - CH-8092 Zürich, Switzerland.

Source: J. Am. Chem. Soc. 1989, 111(24), 8936 - 8938.

Compounds:
a) Coenzyme B12;
b) Cob(II)alamin.

Biological material: Apoenzyme, binding the coenzyme B12 and the homolysis product cob(II)alamin.

Data determined:
Crystal structure (coordinates of the crystal structure of the compounds were determined by X-ray crystallography).

Results: The speed of the homolysis of the organometallic bond is 10^{16} times higher in the apoenzyme bound coenzyme B12 than in a homogenous solution. Structural changes occurring during the Co-C bond homolysis of the coenzyme B12 leading from cobalt(III) corrin to cobalt(II) corrin were investigated. Fig. 1 shows the superposition of structures of the cobalt corrin part of the B12 (dotted line) and of cob(II)alamin (solid line):

Fig.1

The crystal structure of B12 and cob(II)alamin are strikingly similar and offers no explanation for the mechanism of the protein-induced activation of homolysis. It was suggested that the Co-C bond may be labilized by the apoenzyme itself and in addition to a substrate-induced separation of the homolysis fragments (which might be supported by a strong binding of the separated fragments to the protein).

1990/340

Title: Solution conformation of endothelin determined by nuclear magnetic resonance and distance geometry.

Authors: Endo*, S.; Inooka, H.; Ishibashi, Y.; Kitada, C.; Mizuta, E.; Fujino, M.
Tsukuba Research Laboratories, Takeda Chemical Industries Ltd. 7 Wadai, Tsukuba, Ibaraki 300-42, Japan.

Source: FEBS Lett. 1989, 257(Oct), 149 - 154.

Compounds: Human Endothelin I (ET-I) with the following sequence:

Biological material: Endothelin (ET)(21-amino acid peptide isolated from cultured porcine endothelial cells.

Data determined:
\(^1\)H NMR (complete stereospecific assignments were carried out and proton-proton distance constrains were determined by the analyses of DQF-COSY, HOHAHA and NOESY spectra);

NH, \(\alpha H, \beta H\) [chemical shifts (ppm) of proton resonances of human ET];

3D structure [3D structure of ET was calculated using the distance geometry program DADAS based upon the NOESY proton-proton distance constrains determined by NMR spectroscopy];

RMS [root mean square distance (Å) between 5 ET conformers calculated by distance geometry (DADAS)].

Results: The solution conformation of ET has been determined by the combined use of 2D \(^1\)H NMR spectroscopy and distance geometry calculations. Five structures of ET have been calculated from different initial conformations. The superposition of the backbone atoms the calculated structures is shown in Fig. 1. The average RMS value in the core region for the main-chain atoms was 0.46 Å.
The lack of specific interactions between the core and tail portions of ET and a characteristic helix-like conformation in the region from Lys$^9$ to Cys$^{15}$ was shown. Literature data indicated that neither the ET$_{15}$ nor the ET$_{14-21}$ truncated derivatives of ET showed constricting or receptor binding activity suggesting that the ET receptor recognizes an active conformation consisting of both the tail and core portion. The present study, however, suggested that the receptor bound conformation of ET is probably different from that in solution because the lack of interaction between tail and core. The hydrophobic nature of the tail suggested the importance of a hydrophobic interaction with the receptor.

1990/341

Title: The selectivity filter of a ligand-gated channel. The helix-M2 model of the ion channel of the nicotinic acetylcholine receptor.
Authors: Hucho, F.; Hilgenfeld*, R.
Institut für Biochemie, Freie Universität Berlin Thielallee 63, D-1000 Berlin 33, Federal Republic of Germany.
Source: FEBS Lett. 1989, 257(0kt), 17–23.
Compounds: Triphenyl-methyl-phosphit cation (TPMP$^+$$)$. Biological material: Nicotinic acetylcholine receptor (AChR), a prototype of the type I of membrane receptor protein from the electric tissue of Torpedo and Electrophorus.
Results: A computer model of the AChR ion channel has been proposed. Fig. 1 shows the side view of the ion channel model with five pore-forming M2-helices and the channel blocking photoaffinity label (TPMP$^+$$) represented by the dotted sphere:

![Fig.1](image1)

It was supported by electronmicroscopy, electrophysiological- and biochemical experiments that the M2-helices were formed by homologous amino acid sequences containing negatively charged amino acid side chains which act as the selectivity filter. The amino acid side chains may undergo conformational changes during the permeation of the cation. The predicted transmembrane folding of four transmembrane α-helices of type I receptors is shown in Fig. 2:

![Fig.2](image2)

Energy profile calculations indicate that other transmembrane sequences of the receptor protein besides M2 may affect the ion channel.

1990/343

Title: New QSAR techniques eyed for environmental assessments.
Author: Borman, S.
C&EN
1155 Sixteenth St., N.W., Washington DC 20036, USA.
Source: C&EN 1990, 68(8), 20–23.
Results: A new expert system SPARC is being developed at EPA and at the University of Georgia to develop quantitative structure-activity relationships for broad compound classes:

a) Classical QSAR approaches predict therapeutic response, environmental fate or toxicity from structure/property descriptors quantifying hydrophobicity, topological descriptors, electronic descriptors and steric effects;
b) SPARC (SPARC Performs Automated Reasoning in Chemistry), an expert system written in Prolog, models chemistry at the level of physical organic chemistry in terms of mechanism of interaction that contribute to the phenomena of interest;
c) SPARC uses algorithms based on fundamental chemical structure theory to estimate parameters such as acid dissociation constants (pK$_a$s), hydrolysis rate constants, UV, visible and IR absorption spectra, and other properties;
d) The information required to predict input data for broad classes of compounds is dispersed throughout the entire IR spectrum and can be extracted using Fourier transforms;
e) The accuracy of SPARC algorithm was demonstrated on the close match of calculated and experimental pK$_a$ values of 20 carboxylic acid derivatives near to the noise level of measurement.

1990/344

Title: Retention prediction of analytes in reversed-phase high-performance liquid chromatography based on molecular structure.

![Software](image3)
CRIPES (Chromatographic Retention Index Prediction Expert System).

Authors: Smith*, R.M.; Burr, C.M.
Department of Chemistry, Loughborough University of Technology
Loughborough, Leicestershire LE11 3TU, England.
Source: J. Chromatogr. 1989, 485(Dec), 325–340.
Compounds: 41 Benzene derivatives substituted with various groups on aromatic ring and on the alkyl side chain.

Data determined:
k_E, k_C (experimental and calculated capacity factors, respectively, determined by RP-HPLC);
I_E (experimental RP-HPLC retention index);
I_C (retention index calculated by the summation of the retention index of the parent compound (I_P, benzene) and contribution for aliphatic carbons (I_{R}, R, for substituents on aromatic ring (I_{AR}, x) and aliphatic carbons (I_{AR}, x) and for interactions between the substituents (I_{IY-Z}) (I = I_P + I_{R} + I_{AR} + \Sigma I_{AR}, x + I_{IY-Z}, where each term in the equation is expressed as the coefficients of a quadratic equation: I = a x^2 + bx + c, where x is the percentage of the modifier)).

Results: An expert system (CRIPES) has been developed for the prediction of RP-HPLC retention indices from molecular structure by combining a set of rules with retention coefficients stored in a database. The method underlying the system is based on the "alkyl arylation index scale" and aims to improve the reproducibility of prediction and compatibility between various instruments and column materials. The VP-Expert system shell from Microsoft was used for the development. The performance of CRIPES was demonstrated on several subtypes of substituted benzenes (phenacyl halides, substituted arylamines, arylamides and other types). In general the calculated and measured retention indices agreed well but relatively large deviations were observed between the I_E and I_C values for phenacyl bromides and chlorides, o- and p-bromo anilines, N-methylbenzamide and N,N-dimethylbenzamide and phthalate esters. The extension of the database with further interaction values was regarded as necessary for a consistently high accuracy at prediction.

1990/345

Title: Validation of the general purpose Tripos 5.2 force field.
Authors: Clark*, M.; Cramer III, R.D.; Van Opdenbosch, N.
Tripos Associates
1699 S. Hanley Road, St. Louis MO 63144, USA.
Source: J. Computat. Chem. 1989, 10(8), 982–1012.

Compounds:
a) Three cyclic hexapeptides;
b) Crambin;
c) 76 Diverse organic compounds.

Data taken from the literature:
Crystal structure (X-ray crystal structures of the peptides and compounds were obtained from the data banks).

Data determined:
| rms | [root mean square deviation (Å) of the position of corresponding atoms of two superimposed molecular structures]; |
| E_{ij} | [bond stretching energy (kcal/mol·Å^2) associated with bond stretching and compression given by the expression E_{ij} = k_{ij}(d_{ij} - d_0)^2, where d_{ij} is the actual bond length, d_0^2 is standard bond length and k_{ij} is a force constant]; |
| E_{b} | [angle bending term (kcal/mol·deg^2) given by the formula E_{b} = k_{ijk}(\text{actual angle} - \theta_0)^2, where k_{ijk} is a force constant]; |
| E_{p} | [out-of-plane bending energy (kcal/mol) given by the formula E_{p} = k d^2, where d is the distance from the atom to the plane defined by its three attached atoms and k is a force constant]; |
| E_{t} | [torsional energy (kcal/mol·deg^2) associated with four consecutive bonded atoms i,j,k,l given by the formula E_{t} = k_{ijkl}(1 + \sin(\theta)|\cos(\theta)|), where B is the torsion angle between atoms i,j,k and l, s, i, k are constants]; |
| E_{p} | [potential energy (kcal/mol) (nonbonded energy term) associated with any pair of atoms which are neither directly bonded to a common atom or belong to substructures more than a specified cutoff distance away given by the formula E_{p} = k_{ij}(1.04^{13} - 2.04^6), where a is the distance between the two atoms divided by the sum of their radii, and k is the geometric mean of the k constants associated with each atom]. |

Results: Model geometries produced by the Tripos 5.2 force field have been assessed by minimizing the crystalline structures of three cyclic hexapeptides, crambin and 76 diverse complex organic compounds. Comparative force field studies of the Tripos 5.2, Amber and Amber/OPLS force fields were carried out by energy minimization of three cyclic hexapeptides starting from the crystal structures showed the Tripos 5.2 force field superior to the others with the exception of Amber, ECEP2 and LEVB force fields as published by other workers. A direct comparison between the performance of Tripos 5.2 and Amber using isolated crambin showed that the bond and torsion angles of Tripos 5.2 averaged closer to the crystal structure than the angles calculated by Amber (rms = 0.025 Å, 2.97 deg and 13.0 deg for bond lengths, angles, and torsions, respectively, and rms = 0.42 Å for heavy atoms). Fig. 1 shows the superimposed structures of crambin before and after energy minimization.

1990/346

Title: New software weds molecular modeling, NMR.
Author: Krieger, J.
C&EN
1155 Sixteenth St., N.W., Washington DC 20036, USA.
Source: C&EN 1990, 68(13), 16.
**1990/347**

Title: Electronic structure calculations on workstation computers.

The program system TURBOMOLE.

Authors: Ahlrichs*, R.; Bar, M.; Häser, M.; Horn, H.; Kölmel, C.

Theoretical Chemistry, Institute of Physical Chemistry, University of Karlsruhe

D-7500 Karlsruhe, Federal Republic of Germany.

Source: Chem. Phys. Letters 1989, 162(3), 165 – 169.

Results: The main features of the program system TURBOMOLE for large-scale calculation of SCF molecular electronic structure on workstation computers is described:

a) The program system allows for SCF level treatments of energy, first- and second-order derivatives with respect to nuclear coordinates, and an evaluation of the MP2 correlation energy approximation;

b) The most important modules of TURBOMOLE are (i) DSCF performing closed and open shell RHF calculations; (ii) EGRAD used for analytical SCF gradient evaluations; (iii) KOBA calculating direct two-electron integral transformation (iv) FORCE for the computation and processing of integral derivatives and the solution of CPHF equations;

c) Comparison and evaluation of timings of representative applications of TURBOMOLE on various workstations showed that APOLLO DS 10.000 and IRIS 4D/210 were the fastest and comparable to the CONVEX C210 in scalar mode.

**1990/350**

Title: Computer program designed to predict and plot the secondary structure of proteins.

Authors: Sikaris*, K.; Minasian, E.; Leach, S.J.; Flegg, R.

Russell Grimwade School of Biochemistry, University of Melbourne Parkville, Victoria 3052, Australia.

Source: CABIOS 1989, 5(4), 323.

Results: A program has been developed for the prediction and display of the secondary structure of proteins using the primary amino acid sequence as database.

a) The program calculates and graphically demonstrates four predictive profiles of the proteins allowing interpretation and comparison with the results of other programs;

b) As a demonstration the sliding averages of n sequential amino acids were calculated and plotted for four properties of human interleukin 6: (i) plot of the probabilities of $\alpha$-helix, $\beta$-structure and $\beta$-turns according to Chou and Fasman; (ii) $\beta$-turn index of Chou and Fasman; (iii) plot of the hydrophobicity index of Hopp and Woods; (iv) flexibility index of Karplus and Schultz;

c) The regions of primary structure having properties which usually go together agreed reasonably well with each other, i.e. loops and turns with bend probability and hydrophilicity with flexibility.

**1990/351**

Title: 3DSEARCH, a system for three-dimensional substructure searching.
a) Representation of atom types consists of five fields: (i) element (He-U); (ii) number of non hydrogen neighbors (bonded atoms); (iii) number of π electrons; (iv) expected number of attached hydrogens; (v) formal charge; (vi) four type of dummy atoms are also used to define geometric points in space (e.g. centroid of a ring);

b) Definition of queries: (i) definition of spatial relationship between atoms; (ii) matches in atom type; (iii) preparation of keys (constituent descriptors); (iv) execution of key search; (v) geometric search including the handling of angle/dihedral constrains and takes into account "excluded volume";

c) Time tests showed that a search of 3D structures with 3 to 5 atoms in large databases with more than 200,000 entries took only a few minutes (22 - 492 s).

1990/352

Title: Using CONCORD to construct a large database of three-dimensional coordinates from connection tables.
Authors: Rusinko III, A.; Sheridan, R.P.; Nilakantan, R.; Haraki, K.S.; Bauman, N.; Venkataraman*, R.
Medical Research Division, Lederle Laboratories, American Cyanamid Company
Pearl River NY 10965, USA.

Source: J. Chem. Inf. Comput. Sci. 1989, 29(4), 251 - 255.
Results: A database containing about 265,000 compounds in connection tables and 30,000 experimentally determined structures from the Cambridge Structural Database has been transformed into a database of low energy 3D molecular structures using the program CONCORD. The strategy for building the 3D database consisted of the following four steps:

a) Generation of approximate 3D coordinates from connection tables (hydrogens were omitted);

b) Assignment of atom types from connection table information characterized by five descriptors: (i) element type (He-U); (ii) number of attached non-hydrogen neighbors (0–8); (iii) number of π electrons (0–2); (iv) calculated number of attached hydrogens (0–4); (v) formal charge (−1, 0, 1);

c) Addition of three types of chemically meaningful dummy atoms for purposes of 3D substructure searching: (i) centroids of planar 5- and 6-membered rings; (ii) dummy atoms representing the lone electron pairs; (iii) ring perpendiculars positioned orthogonal to and 0.5 Å above and below each planar ring;

d) Efficient storage of the resultant coordinate database indexing the compounds with identification number;

e) The database can be used among others to deduce pharmacophores essential for biological activity and to search for compounds containing a given pharmacophore.

1990/353

Title: Problems on molecular design and computer: XIV. Systematic analysis of organic processes on base of linear-cycle redistribution of topological connections.
Authors: Trach*, S.S.; Baskin, I.I.; Zefirov, N.S.
State University of Moscow
Moscow, USSR.

Source: Zh. Org. Khim. 1989, 25(8), 1585 - 1606 (Russ.).
Results: New organic reaction paths were elaborated by the help of the formal logical and topological methods proposed earlier by the authors (Trach, S.S.et al. Zh. Org. Khim. 1988, 24(6), 1121-1133) for a systematic topological description of the organic reactions with 4-6 vertices, where the vertices correspond to the reaction centres. A combinatoric algorithm generated all the possible topological connections. The algorithm permuted linear and cyclic part of graphs. Each of the 158 reactions are formally described and discussed. Dozens of them evaluated for practical feasibility.

1990/354

Title: New database for carbohydrate structure, information searches.
Author: Borman, S.
C&EN
1155 Sixteenth St., N.W., Washington DC 20036, USA.

Source: C&EN 1989, 67(43), 18 - 20.
Results: A Complex Carbohydrate Structure Database (CCSD) has been developed by the Complex Carbohydrate Research Center at the University of Georgia, having more than 2000 structures and related text files, with about 3000 more records to be added over the next two years. The following are the most important features of CCSD:

a) In CCSD, database records include full primary structures for each complex carbohydrate, citations to papers in which sequences were published, and supplementary information such as spectroscopic analysis, biological activity, information about binding studies, etc.

b) Structural display format visualize branching, points of attachment between glycosyl residues and substituents, anomeric configuration of glycosyl linkages, absolute configuration of glycosyl residues, ring size, identity of proteins or lipids to which carbohydrates are attached and other data;

c) It is planned that CCSD will provide three-dimensional coordinates, to visualize and rotate the structures in stereo and study their interaction with proteins or other biopolimers.