Chemoinformatics in Europe: Achievements and Perspectives

Johann Gasteiger
Computer-Chemie-Centrum
University of Erlangen-Nürnberg
D-91052 Erlangen, Germany

www2.chemie.uni-erlangen.de/

Overview

- chemoinformatics - definition
- the bright past
- the grim presence
- the bright future ?

Synthesis of Properties

The most fundamental and lasting objective of synthesis is not

production of new compounds

but

production of properties

George S. Hammond
Norris Award Lecture, 1968

Fundamental Questions in Chemistry

What structure do I need for a certain property?
structure-activity relationships

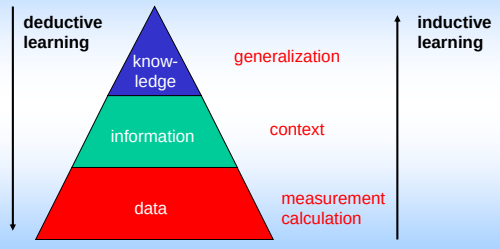
How do I make this structure?
synthesis design

What is the product of my reaction?
reaction prediction
structure elucidation

Chemoinformatics - Why?

- complex relationships
  - structure - biological activity
  - chemical reactivity
- amount of information
  - many millions of compounds and reactions
  - many millions of publications

From Data to Knowledge



Chemoinformatics: Definition

The application of

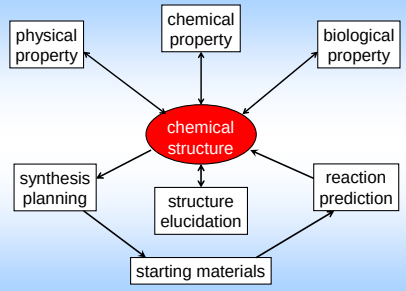
informatics methods

to solve

chemical problems

Application Areas for Chemoinformatics

- drug design
- analytical chemistry
- chemical engineering
- inorganic chemistry
- medicinal chemistry
- organic chemistry
- physical chemistry
- theoretical chemistry



The Scope of Chemoinformatrics

- structure representation and searching
- data analysis and chemometrics
- molecular modeling
- spectra analysis and structure elucidation
- reaction representation and searching
- reaction modeling and synthesis design

Chemoinformatics – An Old Discipline

- structure representation
  1965, Morgan
- structure elucidation
  1965, Sasaki, Munk, DENDRAL
- synthesis design
  1970, Corey & Wipke, Ugi, Gelernter, Hendrickson
- molecular modeling
  1970, Langridge, Marshall
- data analysis / chemometrics
  1970, Kowalski, Wold

Common Topics: Structure Representation

- data storage and retrieval
- property prediction
- drug design
- synthesis design
- spectra analysis and prediction

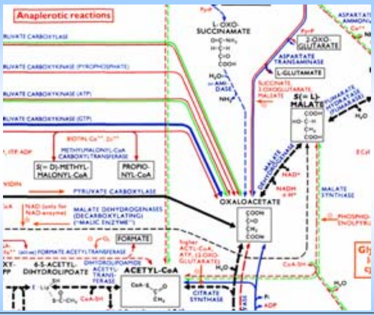
Common Topics: Data Analysis Methods

- property prediction
- drug design
- analytical chemistry
- spectra analysis and prediction

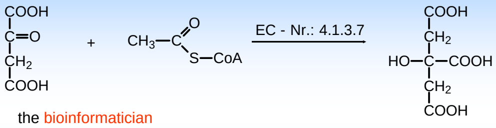
Bioinformatics **Chemoinformatics**

gene ←→ protein ←→ drug ←→ lead

Biochemical Pathways



What is a Chemical Reaction?

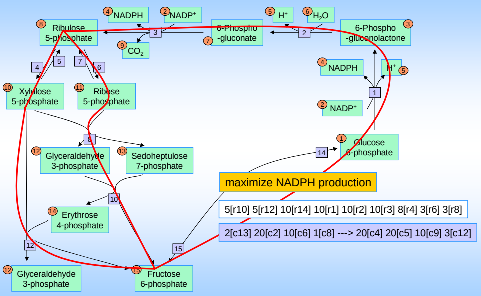


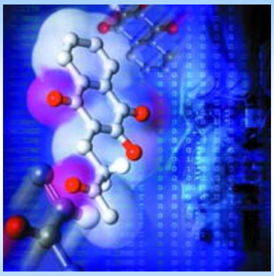
the bioinformatician
an event influenced by a gene, a protein

the computer scientist
a context sensitive graph rewriting rule

the chemist
an event breaking and making bonds

Pathway Searching





Handbook of Chemoinformatics

From Data to Knowledge

J. Gasteiger (Editor)

65 authors
73 contributions
4 volumes
1900 pages
Wiley-VCH, Weinheim
(August 2003)

Textbooks on Chemoinformatics

- V. Gillet, A. Leach
- J. Gasteiger, T. Engel
- J. Bajorath

Major Contributions from Europe

- structure representation
- data analysis methods
- databases
- research centers
- funding
- industry
- teaching

Structure Representation (Europe)

European industry: BASF, Hoechst, ICI, Thomae, BASIC, IDC

Sheffield: M. Lynch, P. Willett

Munich: I. Ugi, J. Gasteiger, C. Jochum

Data Analysis Methods (Europe)

PLS: S. Wold

Self-organizing neural network: Kohonen

Databases (Europe)

Cambridge CSD

Inorganic Structures Database

Beilstein

Gmelin

ChemInformRX

SpecInfo

Institutions and Research Centers (Europe)

FIZ Karlsruhe, D

FIZ Chemie, D

CAOS/CAMM Center, Nijmegen, NL

Computer-Chemie-Centrum, Erlangen, D

Center for Molecular Informatics, Cambridge, UK

Funding (Germany)

Fachinformationsprogramm 1981 - 1994
German Federal Minister of Research and Technology (BMFT)
(Dr. Riesenhuber, chemist !)
institutions
- FIZ Karlsruhe
- FIZ Chemie

databases
- Beilstein
- Gmelin
- ChemInformRX
- SpecInfo

Industry (Europe)

- early work on databases
- positions in chemoinformatics

Teaching (Europe)

Sheffield
UMIST
Strasbourg
Erlangen

Databases (Europe)

Cambridge CSD		*however*:
Inorganic Structures Database		
Beilstein	-	distributed by
Gmelin	-	MDL Elsevier
ChemInformRX	-	
SpecInfo		➡ expensive; academia cut off?

Institutions and Research Centers (Europe)

FIZ Karlsruhe, D	*however:* no research
FIZ Chemie, D	no research
CAOS/CAMM Center, Nijmegen, NL	no long-term commitment
Computer-Chemie-Centrum, Erlangen, D	successor to J. Gasteiger ?
Center for Molecular Informatics, Cambridge, UK	no long-term commitment

Funding (Germany)

Fachinformationsprogramm 1981 - 1994
German Federal Minister of Research and Technology (BMFT)
(Dr. Riesenhuber, chemist !)
institutions
- FIZ Karlsruhe
- FIZ Chemie

databases
- Beilstein
- Gmelin
- ChemInformRX
- SpecInfo

however:
now all BMBF projects go into bioinformatics

Industry

- early work on databases
- positions in chemoinformatics

however:

- hardly any in-house work done anymore
- mergers lead to elimination of positions

Teaching (Europe)

	however:
Sheffield	
UMIST	discontinued?
Strasbourg	
Erlangen	only for Molecular Science students

Funding (Europe)

6th Framework program of the European Union:

several programs for Bioinformatics

however: no mention of Chemoinformatics!

What Can be Done?

- conferences
- teaching
- cooperation academia – industry
- new application areas
- funding
- organization

Germany Chemical Society (GDCh)

Division “Chemical Information” changed to “Chemical-Information-Computer (CIC)” in 1987

Workshop:
Software Development in Chemistry 1986-2004
German Conference on Chemoinformatics 2005

Teaching

- define curriculum in chemoinformatics
- what contents of chemoinformatics have to go into regular chemistry curricula

➡ Round Table + Committee

Bioinformatics in Germany

Fachgruppe 4 der Gesellschaft für Informatik
(Division 4 of the Society for Informatics)

- definiton of a curriculum on Bioinformatics in 1990
- was then put on the web

⇨all positions in bioinformatics at German universities were given to computer scientists

Application Areas for Chemoinformatics

- drug design
- analytical chemistry
- chemical engineering
- inorganic chemistry
- medicinal chemistry
- organic chemistry
- physical chemistry
- theoretical chemistry

Cooperation Industry - Academia

- industry: generate data
- academia: develop methods

➡ provide academia access to data

Funding

- increase awareness for importance of Chemoinformatics
- go into committees

Get Organized!

Chemometrics Society

QSAR Society

FECS Working Party: Computational Chemistry

➡ Chemoinformatics Society