Looking Back, But Not In Anger:
My view of the history and future of chemical information

Engelbert Zass (*retd.*)
Chemistry Biology Pharmacy
Information Center
ETH Zurich
8093 Zuerich, Switzerland
Chemical Information

about:

• -1970 *print* sources only

• 1970-* isolated electronic sources for information specialists

• 1985-* isolated electronic sources for chemists (“end-users”)

• 2000-* integrated electronic sources for chemists
Chemists and their Information

• Past: **printed** sources
  ⇒ Chemists search themselves

• Transition (~1980-2000): Databases powerful but complex and expensive ⇒ „High Time“ for Specialists/Intermediaries

• Present: Information via/in the **Web**
  ⇒ Chemists search themselves (again)
Chemical Information Landmarks I

• 1817: Gmelin Handbook
• 1830: Pharmaceutisches Central-Blatt
Chemical Abstracts print: 1907-2009
Searching Beilstein (Handbook): 1975
Personal Literature before IT

A. Gossauer und R.-P. Hinze
J. Org. Chem. 43, 283 (1978)
An Improved Chemical Synthesis of Racemic Phycocyanobilin Dimethyl Ester

K. M. Smith, G. M. F. Bisset & M. J. Bushell
Partial Synthesis of Optically Pure Methyl Bacteriopheophorbides c and d from Methyl Pheophorbide a

Partial Syntheses of Optically Pure Methyl Bacteriopheophorbides c and d from Methyl Pheophorbide a

Kevin M. Smith, Graham M. F. Bisset, and Michael J. Bushell
Department of Chemistry, University of California at Davis, Davis, California 95616
Received January 18, 1980

A partial synthesis of the diastereomeric mixture of methyl bacteriopheophorbide a from the chlorophyll degradation product chl a, trimethyl ester (4), is described. Compound 2 is related to band 6 of the bacteriopheophorbide a (Chlorobium thiosulfatophilus, “CHT”) series and also to bacteriopheophorbide a, recently isolated from Chloroflexus aurantiacus. Separation of the 2.5 diastereomeric mixture (the 2.5-bacteriopheophorbide a) is readily accomplished by using reverse-phase high-performance liquid chromatography (high-performance LC). Methylation of the 2.5-bacteriopheophorbide a (25) similarly gives the 2.5 diastereomeric mixture [18(19),36(38)] which can in turn be separated by high-performance LC. On account of Woodwell’s earlier total synthesis of optically pure chl a, trimethyl ester (4), this work constitutes a formal total synthesis of the optically pure methyl phæophorbides of band 6 of the bacteriopheophorbide a (equivalent to methyl bacteriopheophorbide a) and of band 6 of the bacteriopheophorbide a.
<table>
<thead>
<tr>
<th>Year</th>
<th>Chemical Information Landmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>1961</td>
<td>Chemical Titles (CAS)</td>
</tr>
<tr>
<td>1968</td>
<td>CA Condensates</td>
</tr>
<tr>
<td>1980/81</td>
<td>CAS Online, DARC</td>
</tr>
<tr>
<td>1982</td>
<td>REACCS (MDL)</td>
</tr>
<tr>
<td>1986</td>
<td>CJACS (e-journals)</td>
</tr>
<tr>
<td>1991</td>
<td>World Wide Web</td>
</tr>
<tr>
<td>1993</td>
<td>CrossFire</td>
</tr>
</tbody>
</table>
Reaction Searching: 1974-

CHEMICAL REACTIONS DOCUMENTATION SERVICE

INSTRUCTION BULLETIN NO. 1, PT. 2

REACTION CODE

1ST EDITION - FEBRUARY 1974
ENTER COMMAND: SEARCH JONES1 DEMO
JONES1 DEFINED ON 26 AUG 80 AT 12:
SEARCHED LAST ON 26 AUG 80 AT 12:36:07
PRIMARY SEARCH: SC28 AND SC23 AND SC21
THE COMPONENT DEFINITIONS:
SC28 : 1312 AND 1099 AND 401 AND
SC23 : 212
SC21 : 1170 OR 1174
MODIFY SEARCH QUALIFICATION? (Y,N,*END): N
CREATE ANSWER FILE? (Y,N, OR NUMBER): Y
ANSFILE 1 ASSIGNED
ENTER SEARCH LIMITS (SAMPLE, FULL
SEARCH CONTROL -
ENTER SRCHSTART OR *END: SRCHSTART
12:54:48 SEARCH INITIATED (12:54:48)
SEARCH CONTROL -
ENTER STATUS, OUTPUT, STOP OR *END:
NORMAL TERM. OF FULL FILE SEARCH - 4

Guide to CAS ONLINE Commands

CHEMICAL ABSTRACTS SERVICE
A DIVISION OF THE AMERICAN CHEMICAL SOCIETY

© E. Zass, 244th ACS Natl.Meet., Philadelphia
The CAS Pre-1965 Registration Project

A Proposal for Expanding the CAS Chemical Registry System
By Chemical Abstracts Service,
A Division of the American Chemical Society

January 1, 1983
1989: in the ETH Chemistry Library

CHF 30’000 !
1989: at the Bench

Database Access in the ETHZ Organic Chemistry Laboratory

ETHZ Computer Center

DEC VAX 8600 (V/MS)
- REACCS
- ORAC
- SYNLIB
- CAMEO
- CHIRON
- ICSD
- PDB

CDC Cyber 180-855 (NOS/VE)
- CSD

ETHZ Main Library

ETHICS

KOMETH LAN

"T-Box"

© E. Zass, 244th ACS Natl.Meet., Philadelphia
1998

ETH Chemistry Information Center
Access to Chemical Information

Printed Sources
"Electronic Library"
CD-ROMs
In-house Databases

"at the bench"
End-Users

License Control

Public Online Databases
(Data-Star, STN, DIALOG,
ORBIT, CIS, Questel)

Library Staff (Chemists)

at InfoCenter

© E. Zass, 244th ACS Natl.Meet., Philadelphia
2012

CD Library on WWW

First steps

Please note our system requirements and access restrictions before first use of our CD Library.

Please ask the IT support of ICBP in case of any problems when using the CD Library.

CD Library

Biologie / Biology

Biochemical Interactions 3.1 Kinemages 2002
Biochemical Interactions 3.1 2002 (Voet & Voet, Lehrbuch der Biochemie)
Campbell BioIolgy 7th Edition 2005

Biology: 17 titles
Chemistry: 13 titles
Dictionaries: 8 titles
Pharmacy: 19 titles
Safety: 13 titles
Past, Present - and Future?
The Past: "Must Use" Brands
Present: Many "Equal" Brands

Google Schweiz

WIKIPEDIA Die freie Enzyklopädie

facebook

SciFinder® Essential content. Proven results™

reaxys® Chemistry Starts Here

WEB OF KNOWLEDGE™ DISCOVERY STARTS HERE

Springer Materials The Landolt-Börnstein Database

Integrity™
### Search for Element Systems

Select elements by clicking on the symbols. Deselect elements by clicking a second time.

Your Selection

**B-F-O**

<table>
<thead>
<tr>
<th>Period</th>
<th>Group</th>
<th>Element</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>IIA</td>
<td>H</td>
</tr>
<tr>
<td>2</td>
<td>IIB</td>
<td>Li, Bc</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>Na, Mg</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>K, Ca</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>Rb, Sr</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>Cs, Ba</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>Fr, Ra</td>
</tr>
</tbody>
</table>

Other Elements:
- B
- F
- O
<table>
<thead>
<tr>
<th>Query</th>
<th>Temporary result description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>49 substances</td>
</tr>
<tr>
<td></td>
<td>Substances: IDE.ELS = 'B' AND IDE.ELS = 'F' AND IDE.ELS = 'O' AND IDE.ELN=3'</td>
</tr>
<tr>
<td></td>
<td>130 citations</td>
</tr>
<tr>
<td></td>
<td>56635 substances</td>
</tr>
<tr>
<td></td>
<td>Substances: IDE.ELS = 'B' AND IDE.ELS = 'F' AND IDE.ELS = 'O'</td>
</tr>
<tr>
<td></td>
<td>31280 citations</td>
</tr>
<tr>
<td>BH₃</td>
<td>30 substances</td>
</tr>
<tr>
<td>HF</td>
<td>Substances: IDE.ELN=3', Substructure: on all atoms</td>
</tr>
<tr>
<td>H₂O</td>
<td>125 citations</td>
</tr>
<tr>
<td>BH₃</td>
<td>56129 substances</td>
</tr>
<tr>
<td>HF</td>
<td>Substances: Substructure: on all atoms</td>
</tr>
<tr>
<td>H₂O</td>
<td>31094 citations</td>
</tr>
</tbody>
</table>

Reaxys: BₓFᵧOₜ
PROPERTY VALUES TAGGED WITH IC ARE FROM THE ZIC/VINITI DATA FILE PROVIDED BY INFOCHEM.

STRUCTURE FILE UPDATES: 14 MAR 2011 HIGHEST RN 1268445-99-3
DICTIONARY FILE UPDATES: 14 MAR 2011 HIGHEST RN 1268445-99-3

CAS INFORMATION USE POLICIES APPLY AND ARE AVAILABLE AT:

http://www.cas.org/legal/infopolicy.html

TSCA INFORMATION NOW CURRENT THROUGH JANUARY 14, 2011.

Please note that search-term pricing does apply when conducting SmartSELECT searches.

REGISTRY INCLUDES NUMERICALLY SEARCHABLE DATA FOR EXPERIMENTAL AND PREDICTED PROPERTIES AS WELL AS TAGS INDICATING AVAILABILITY OF EXPERIMENTAL PROPERTY DATA IN THE ORIGINAL DOCUMENT. FOR INFORMATION ON PROPERTY SEARCHING IN REGISTRY, REFER TO:

http://www.cas.org/support/stngen/stndoc/properties.html

=> S (B AND F AND O)/ELS
   524327 B/ELS
   11976795 F/ELS
   50616266 O/ELS
L1  81621 (B AND F AND O)/ELS

=> S L1 AND 3/ELC.SUB
   4814319 3/ELC.SUB
L2  41 L1 AND 3/ELC.SUB

=> D SCAN

cf. SCIINDER: 7 !
„dot.disconnect“ + acidic H + normalization = TROUBLE!

\[ \text{Be}_x \text{H}_y (\text{PO}_4)_z \]

\[ \text{Be}_3 (\text{PO}_4)_2 \]

\[ \text{Be}(\text{H}_2\text{PO}_4)_2 \]

\[ \text{BeHPO}_4 \]
SciFinder: with nmr (3/2011)

- Registry: Refine by Property… 36 Subs. (28)
  19 refs. total (rel.: 19 / excl. 17)

- CAplus: Get References
  - Limit results to: Spectral Properties
    7 refs. total (rel.: 7 / excl. 0)
  - Categorize – Physical Chemistry – Spectra …
    9 refs. total (rel.: 8 / excl. 0)
  - Refine by Topic: nmr 25 Subs. (17)
    31 refs. total (rel.: 21 / excl. 9)

Reaxys: 328 substances
KnowItAll: 2 substances
SciFinder: Preparation of a Compound

• Search via:
  1. Get Reaction: Product (CASREACT)
  2. Additional Reactions (CASREACT ⇒ CAplus)
  3. Registry Role Matrix (Patents, Nonpatents) (Registry ⇒ CAplus)
  4. Get References – Limit results to: Preparation (CAplus)
Preparation of Lidocaine (27.6.2009)

- Reaxys
  23 refs. 1946-2008

- SciFinder Web
  - CASREACT: 4 refs. 1984-2009
  - CAplus: 109 refs. 1948-2009
    only 45 relevant!
21. Extraction of local anesthetics from horse urine

By: Hagedorn, Heinz Werner; Rose, Heidrun

The solid phase extn. was examd. of some local anesthetics from horse urine, including an internal std. Recovery and precision of standardized extn. were detd. to compare the efficiency of self-filled RP-18 columns with ready-to-use LiChrolut RP-18 columns. Depending on the compds. and column tests, recoveries varied from 44-108% with a maximal coeff. of variation of 11%. Recoveries of the drugs were lower when using self-filled RP-18 columns. The precision of the test was not influenced by the column type.

Indexing

- Biochemical Methods (Section 9-9)
- Section cross-reference(s): 1, 4

Concepts

- Anesthetics
- Urine analysis
- extn. of local anesthetics from horse urine
- Extraction
- solid-phase; extn. of local anesthetics from horse urine

Substances

- 59-46-1P Procaine
- 83-98-7P Orphenadrine
- 86-43-1P Propoxycaine
- 94-24-6P Tetracaine
- 96-88-8P Mepivacaine
- 137-58-6P Lidocaine

- extn. of local anesthetics from horse urine

- Analyte; Biological study, unclassified; Purification or recovery; Analytical study; Biological study; Preparation
# Lidocaine Syntheses (6/2009)

<table>
<thead>
<tr>
<th>Lidocaine</th>
<th>SFS CASREACT</th>
<th>SFS CAPIplus</th>
<th>Reaxys</th>
<th>Integrity</th>
<th>Kleemann-Engel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Astra DE 968561 (1943)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Name</td>
</tr>
<tr>
<td>Astra US 2441498 (1945)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Name</td>
</tr>
<tr>
<td>Ark.Kem. 1946</td>
<td></td>
<td></td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>US 2797241 (1953)</td>
<td></td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
</tr>
<tr>
<td>GB 706409 (1954)</td>
<td></td>
<td></td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB 758224 (1956)</td>
<td></td>
<td></td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Astra GB 771151 (1957)</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zh.Obshch.Khim. 1959</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zh.Obshch.Khim. 1960</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB 810212 (1960)</td>
<td></td>
<td></td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zh.Obshch.Khim. 1961</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J.Med.Pharm.Chem. 1961</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ugi DE 1103337 (1961)</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Astra GB 941195 (1963)</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J.Labelled Compd. 1966</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chugai JP 47024547 (1972)</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yakugaku Zasshi 1979</td>
<td></td>
<td></td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RO 85396 (1984)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Name</td>
</tr>
<tr>
<td>OBRBP (1986)</td>
<td></td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
</tr>
<tr>
<td>SU 1557964 (1992)</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J.Chem.Educ. 1999</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bull.Georg.Acad.Sci. 2001</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IT 1307824 (2001)</td>
<td></td>
<td></td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J.Comb.Chem. 2003</td>
<td></td>
<td></td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pharm.Chem J. 2008</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bioorg.Međ.Chem.Lett. 2008</td>
<td>Prepn + Categorize</td>
<td>Name</td>
<td>Name</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Comparison Reaxys / SciFinder

- exact reaction (*FS 2010*)

  ![Chemical Structure]  

  – **Reaxys:** 4 reactions / 16 references  
  exclusive: 10

  – **SciFinder:** 21 reactions / 14 references  
  exclusive: 8
CASREACT: no Stereosearch!
Stereosearch via the back door …

1. Search (sub)structure of chiral reaction partner(s)
   Explore Substances: Chemical Structure

2. Get Reactions: Product (Reactant)

3. Refine: by Reaction Structure
   (substructure: use „atom mapping“ if needed)
Databases: Problems and Dangers

2° Literature (A&I, Handbooks) = endangered species?

• (Relatively) complex interfaces
• Too many sources needed
• Too many differences between sources
• Problematic legacies

(Cost) Effort/Utility ratio not obvious enough
Information Processing Chain

Web: Google, Wikipedia, …

• 1° Literature (Fulltext)

• 2° Literature (A & I)

• 3° Literature
  - Selection
  - Concentration
  - Transformation
  - Structuring (meta-Data)
Dendrobine Total Syntheses (4/2012)

- Total refs.: 44 (1965-2012)
  - SciFinder 38 (10 excl.)
    - CASREACT 4
    - CASREACT addl. Reaction 9
    - CPlus 38
  - Reaxys 14 (1 excl.)
  - Web of Knowledge 20 (none excl.)
  - Scopus 18 (none excl.)
  - Google Scholar 18 (1 excl.)
CHEMICAL INFORMATION RETRIEVAL – A SHORT DISCUSSION ABOUT THE STATE OF THE ART, PROGRESS, AND PITFALLS

Engelbert Zass

Chemistry Biology Pharmacy Information Center, ETH Zurich, Wolfgang-Pauli-Strasse 10, CH-8093 Zurich, Switzerland. E-Mail: zass@chem.ethz.ch

Abstract – Examples of author, keyword, structure, and reaction searches related to the scientific achievements of A. Eschenmoser were analyzed to illustrate the power, but also the limitations of modern database systems like SciFinder, Reaxys, and Web of Knowledge.

# Gratefully dedicated to Prof. Dr. Albert Eschenmoser on the occasion of his 85th birthday.