

Chemical Structure and Text Hyperlinking in a Web-Based E-library

ACS Fall National Meeting
29th August 2001

James Weeks
ChemWeb, Inc.
84 Theobalds Road
Holborn, London
United Kingdom WC1X 8RR




Outline of Presentation

- **Introduction**
 - Chemical information sources
 - Organisation of disparate chemical information
 - ChemWeb.com
 - Gathering online information together
- **Searching & Linking in ChemWeb.com**
 - Text
 - Structures
 - Meta-data (Chemistry Preprint Server)
- **Conclusions**
- **Discussion**

Chemical Information Sources

- Information Sources for chemistry

- Vast amount of information
- CAS currently indexes over 18 million abstracts 
- Hundreds of scientific journals in chemistry
- Many databases also contain a wealth of information
- Many different publishers



Organisation of Chemical Information

- Organisation
 - Facilitated by the Internet
- Portals
 - Varied: TheScientificWorld, PubMedCentral, ChemPort & SciFinder, Scirus
 - Added community features: BioMedNet, ChemWeb.com
 - Centralised online resources of scientific information
 - Access many different journals & databases

ChemWeb.com 

BioMedNet
bmn.com 

SCIRUS
for scientific information only 

PubMed
central 

TheScientificWorld 

ChemPort
CONNECTION 

- Online resource for the worldwide chemical community
- Free registration and membership
- Essential services & information sources
 - Publisher-neutral
 - 230 journals, 25 databases
 - *the alchemist* Magazine, the Chemistry Preprint Server, Shopping Mall, Conference Centre, Careers Centre
- alchemist.chemweb.com
- preprint.chemweb.com



ChemWeb.com (2)

- Offer free services
 - Temporary & permanent
 - Access to journals
 - AutoNom and Beilstein Abstracts
- Encourage member interaction
 - e.g. forums
- Exploit web technology
 - e.g. structure searching
- Over 300,000 members
- Over 2.5m page impressions per month



How Can You Find Relevant Information?

- **Problem**
 - Scientific information is disparate
- **Solution**
 - Create centralised libraries of journals and databases on the Internet
- **Enhance**
 - Ability to search and inter-link adds value to a large library of online information
- **ChemWeb.com**
 - Supports searching & linking within all articles and records of hosted journals & databases



ChemWeb.com Technology: Introduction

- **Text searching (all of ChemWeb)**
 - Articles and records are tagged and indexed (multi-level) in a centralised database
 - Quick and advanced searching through all information
- **Text linking (journals)**
 - Inter article links enhance navigation
- **Structure searching (databases)**
 - Search by structure in records of many databases
- **Improvements**
 - Dymond: Link structures in journals & databases
 - Search history to be implemented soon

Text Searching:

All Information in ChemWeb: Sitewide

- **Restrict by field**
 - Full text, title, author names, abstract, keywords, references
- **Restrict by section**
 - Journals; databases; *the alchemist*; preprints; conferences; jobs; shopping mall; links
- **Advanced**
 - AND, NOT, OR
 - Search for phrases

The screenshot shows the 'advanced' search interface. It features three search boxes labeled 'a', 'b', and 'c', each with a dropdown menu set to 'ALL of:'. Between the boxes are 'AND' operators. To the right of each box is a dropdown menu for search fields, with 'Full Text' selected. Below the search boxes is a section titled 'Select section(s): (if all unchecked, all sections will be searched)'. This section contains several checkboxes: Journals, Databases*, Alchemist, Conferences, Preprints, Jobs, Shops, Members, and Links. A note below the checkboxes states '*Structure databases are not included in site searches'. At the bottom, there is a 'results per page:' dropdown set to '10', and buttons for 'hide tips', 'help', 'clear', and 'search'. Three red arrows point from text tips to the search interface: one to the search boxes, one to the 'AND' operators, and one to the 'Select section(s):' section.

tip
type your terms in boxes
a and/or b and/or c
[HIDE TIPS](#)

tip
select AND or NOT or OR operators to refine your search
[HIDE TIPS](#)

tip
restrict your search to one or more sections
[HIDE TIPS](#)

search advanced history results

a ALL of: [] in Full Text

b ALL of: [] in Full Text

c ALL of: [] in Full Text

AND

AND

AND

Select section(s): (if all unchecked, all sections will be searched)

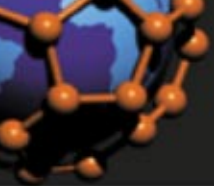
Journals Databases* Alchemist Conferences

Preprints Jobs Shops Members Links

*Structure databases are not included in site searches

results per page: 10

hide tips help clear search



Text Linking: Hyperlinks in Journal Articles

- **Article links**
 - *Author*: articles with the same author in the author list
 - *Author email*: principal author
 - *Related*: similar articles
- **Reference section**
 - *Full text*: articles on ChemWeb
 - *Cited-by*: articles which cite a particular reference
 - *LitLink*: external articles

INFO

[PRICE LIST](#) [HOW TO BUY](#) [EDIT ALERTS](#) [ADD TO MY ALERTS](#) [HELP](#)

E.V. Gribanovaa, A.N. Zhukova, I.E. Antonyuka, C. Benndorfba & E.N. Baskovaa
DIAMOND AND RELATED MATERIALS 2000, 9:1:1-6
Department of Colloid Chemistry, St. Petersburg University, Petrodvoretz, 198904, St. Petersburg, Russia.
Department of Physical Chemistry, University of Hamburg, Bundesstrasse 45, D-20146 Hamburg, Germany.
Correspondence Tel: +4940-4123-34-31 Fax: +4940-4123-34-52
*Received 7 January 1999
Accepted 23 August 1999
Available online 1 February 2000*

OTHER ARTICLES BY

E.V. Gribanovaa
A.N. Zhukova
I.E. Antonyuka
C. Benndorfba
E.N. Baskovaa

RELATED ARTICLES CITED BY..

[VIEW PDF](#)
requires Acrobat 3.0+

E-MAIL

benndorf@chemie.uni-hamburg.de

The contact angle between the solid/liquid/gas interface was measured in dependence on the pH of the liquid on different carbon surfaces (CVD-diamond films, graphite and pyrocarbon films). As function of the pH, the values show significant changes with maxima and minima. This behavior is interpreted as being due to the presence of surface acidic groups with different acidic strength. The appearance of maxima and minima in the (pH) curves are caused by the reduction of the solid surface interaction with a minimum dissociation of the individual functional groups.

RNA molecules are key players in all steps of gene expression. Beyond their function as messengers, RNAs are involved in regulatory processes and constitute essential parts of the cellular machinery responsible for mRNA splicing, transport and translation (reviewed in [1]). Several catalytic RNA molecules (ribozymes) have been discovered that function in the absence of

Tetrahedron Lett., 2000, 41, 6347. [[LitLink](#)] [[Full text](#)] [[Cited by](#)]

Wiley and Sons [[LitLink](#)]

Letf., 1988, 39, 2715. [[LitLink](#)] [[Cited by](#)]

Text Linking: Hyperlinks to Other Articles


- External links

- LitLink hyperlinks next to references (tagged)
- LitLink finds where the article may be downloaded
- Various selection rules
- Similar technologies include CrossRef, ChemPort
- Access all Elsevier Science articles via ScienceDirect



Get this document from [ACS PUBLICATIONS](#)

ChemWeb.com™
The World Wide Club for the Chemical Community

Get this abstract from 

[Send an e-mail document request](#)

Search Parameters
YEAR = 2000
_ORIGIN = ChemWeb
AUTHOR = N. Shibata
JOURNAL = J. AM. CHEM. SOC.
VOL = 122
PG = 10728



View the Full Text version of this article

Full Text via ScienceDirect:

 [\[ScienceDirect\]](#)



Access to full text articles from Analytica Chimica Acta via ScienceDirect is available to ChemWeb.com members that -

- belong to a library that has a current subscription to Analytica Chimica Acta via [ScienceDirect](#)

or

- belong to a library that has a current print subscription to Analytica Chimica Acta *and* has registered for [ScienceDirect Web Editions](#)

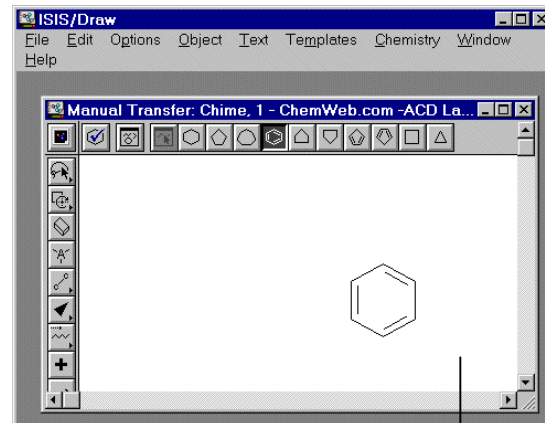
Full Text via ChemWeb.com:

 [\[HTML\]](#)
 [\[PDF 128K\]](#)

Access to full text articles from Analytica Chimica Acta via ChemWeb.com is available to members that have a subscription key for a current Associated Personal Online Subscription (no longer available for purchase)

Structure Searching: Database Records

- **ISIS/Draw**
 - Input structure
- **CHIME plug-in**
 - Structure is converted
 - Image is rendered
 - CHIME data submitted
- **Search database**
 - Compares similar data
 - Links to matches (exact, 60%, sub-structure)
 - Databases include ACD/Labs NMR, Brethericks



ACD Labs/HNMR Database Search

2D structure query box	Chemical Structure Searching exact structure match ▾ No. of hits 10 ▾ Search Clear
------------------------	--

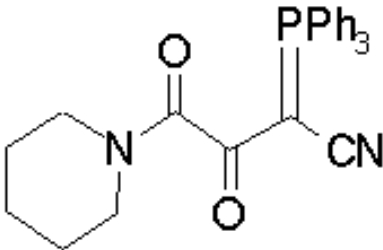


- **Dymond**

- Technology on ChemWeb.com, soon to be implemented on Elsevier Science sites
- Links chemical structures from journal articles to structure-searchable databases
- Links are keyed into the journal articles, where mentioned, using the CHIME plug-in
- Hyperlink sends CHIME data to a number of structure databases in ChemWeb.com
- Combined Chemical Dictionary, AutoNom, ACD/Labs NMR databases
- Beilstein Crossfire

Structure Linking:

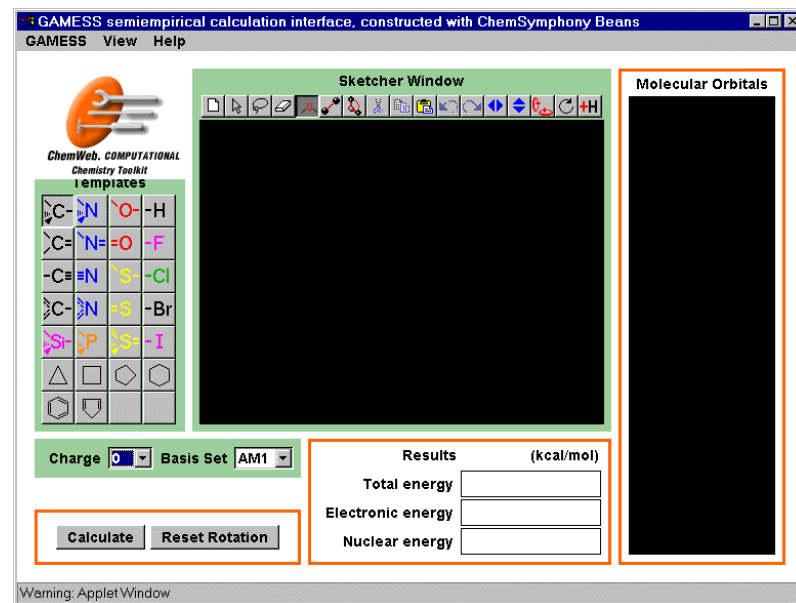
DYMOND Initiative in Journal Articles (2)

 <p>2</p> <p>Close Window</p>	Find this compound in Beilstein CrossFire <input type="button" value="Go"/> Help Note: users will need to have Commander 2000 installed and have access to CrossFire for this search to work
	Find related compounds in Combined Chemical Dictionary <input type="radio"/> Exact <input checked="" type="radio"/> 60% similar <input type="button" value="Go"/> Help <input type="radio"/> Substructure
	Generate IUPAC Name for this compound using AutoNom Standard <input type="button" value="Go"/> Help
	Search for NMR spectral details in the ACD/Labs databases <input type="radio"/> Exact <input checked="" type="radio"/> 60% similar <input type="button" value="Go"/> Help <input type="radio"/> Substructure <input checked="" type="radio"/> H NMR <input type="radio"/> C NMR <input type="radio"/> F NMR <input type="radio"/> P NMR
	Search for properties in the ACD/Labs databases <input type="button" value="Go"/> Help <input type="radio"/> Exact <input checked="" type="radio"/> 60% similar <input type="radio"/> Substructure <input checked="" type="radio"/> logP <input type="radio"/> pKa

<http://www.chemweb.com/dymond>

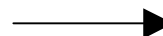
Structure Linking: Computational Toolkit

- New feature on ChemWeb.com
 - Semi-empirical calculations
- Possible future development
 - Send chemical structures to the toolkit
 - Mechanism similar to DYMOND initiative
 - Further enhance retrieval of chemical information



Improvement: Search History

- Search history:
 - To be implemented soon
 - Perform advanced search
 - Search by text or structure
 - “Save” search in user profile
 - “View” search results later...
 - Original search is re-run
- Future enhancement
 - Request re-run of search at a defined frequency
 - Search results are sent to the user in an email digest



The screenshot shows a search history interface with three entries. At the top, there are three tabs: 'search', 'advanced', and 'history', with 'history' being the active tab. Each entry consists of a search query, a 'view' button, a 'delete' button, the number of results, and the search date.

SEARCH	view	delete	RESULTS	DATE
'organic synthesis' AND 'organic synthesis' AND 'organic synthesis'			20,999,999	25-JAN-02
'organic synthesis' in Full text AND synthesis in Abstracts AND organic in Keywords, in Journals & Databases & The Alchemist			20,999,999	25-JAN-02
DATABASE NAME IS carbon OR NAME BEGINS WITH carbon AND [STRUCTURE]			20,999,999	25-JAN-02

Meta-Data Linking: Chemistry Preprint Server (CPS)

- Preprint
 - Publicly-available research article prior to formal publication
- Preprint server
 - Freely available and permanent archive and distribution medium
 - Rapid dissemination
 - Multimedia & supporting files
- CPS
 - First for chemistry, launched 21 August 2000 at ACS meeting
 - 300 preprints from 50 countries

The screenshot shows the homepage of the Chemistry Preprint Server. At the top, there is a navigation bar with links for Home, Browse, Search, Submit, Help, and Demo. The main content area includes a welcome message, a user login status, and a 'What's hot!' section with links to News and Logos. A 'New:' section lists recent updates. On the right side, there are several widgets: 'Communities' with a 'What's your speciality?' prompt, 'the Alchemist' with 'Daily news and reviews', 'last updated' showing the date 20 August 2001, 'no. of articles' showing 283, and a 'quick find' section with links for most viewed, most discussed, highest ranked, and most recent. A 'quick search' box is also present.

the chemistry preprint server

Home Browse Search Submit Help Demo ChemWeb.com

Welcome to The Chemistry Preprint Server

logged in ▶ Dr James Weeks - 21 August 2001, 13:49 (GMT)

What's hot!

News

cps Logos

The Chemistry Preprint Server (CPS) is a major new initiative for the chemistry community, powered by [ChemWeb.com](#). It is a freely available and permanent web archive and distribution medium for research articles in the field of chemistry.

New:

- [CPS logos](#) - link from your website to the CPS!
- We have now added a [conference proceedings section](#).
- Eprints.org invites you to participate in an important survey!

Communities
What's your speciality?

the Alchemist
Daily news and reviews

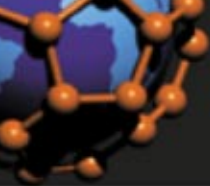
last updated
▶ 20 August 2001/08:48

no. of articles
▶ 283

quick find
▶ [most viewed](#)
▶ [most discussed](#)
▶ [highest ranked](#)
▶ [most recent](#)

quick search

Go



Meta-Data Linking: What is Meta-Data?

- Upload article to CPS
 - Every article has specific identifying information...
- Meta-data
 - Author names
 - Affiliation
 - Email address
 - Article title
 - Article abstract
 - Chemistry classification

Please enter your name, affiliation and email address

First Name: James

Last Name: Weeks

Add additional authors:

Example: J C Lewis, S A Henry. We regret we cannot accept individual affiliations for additional authors.

Affiliation:

Please enter the full name and address of your affiliation (institution).

Email: james.weeks@chemweb.com

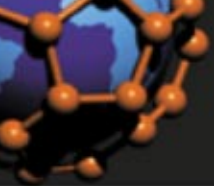
Please enter details about the article

Article Title:

Conference Proceeding:

Please tick the box to the right **only** if your preprint is directly relevant to a particular conference (for example, the paper and/or slides from a presented talk, a list of abstracts, a review). In this case, please also include the conference title in the Author Comments box.

Article Abstract:



Meta-Data Linking: Open Archives Initiative



- **Open Archives Initiative (OAI)**
 - Sets standards for the transfer of information between different web servers. The CPS will be compliant soon.
 - <http://www.openarchives.org>
- **“Data provision”**
 - CPS meta-data is made available in XML format
 - Meta-data can be harvested by remote preprint servers
 - Users of remote preprint servers can search CPS meta-data and link directly to the article information
- **“Service provision”**
 - CPS harvests information on remote preprint servers
 - Users of CPS can search meta-data on remote servers

Meta-Data Linking: CPS - Future Directions

• After submission...

- The author can revise the original text any number of times
- If published, access can be redirected to the publisher's site
- The author enters the journal ISSN number, volume, issue etc.
- Users of the CPS can find the article via LitLink

[pdf of latest version]

- [Version 1](#)

[Redirect to Published Article](#)

[Revise PDF](#)

Uploaded supplementary files

- [Fig1.pdf](#)
- [Fig2a.gif](#)
- [Fig2b.gif](#)
- [Fig2c.gif](#)
- [Fig3.pdf](#)

[Add more files](#)

ISSN:

Year:

Volume:

Starting page number:

• Future direction?

- CPS could act as a centralised library, inter-linking all research from preprint to publication

[Published full text](#)

Uploaded supplementary files
[None]

[Disclaimer](#)



Conclusions

- **Chemical information**
 - Vast amount of scientific information available
 - Need to bring different information sources together
- **ChemWeb.com**
 - Wide range of searching and linking capabilities
 - Members can seamlessly find relevant information
 - Choose from a wide variety of journals and databases
- **CPS will soon be compliant with the OAI**
 - Information is exchanged between the CPS and remote preprint servers
 - Search CPS and remote servers independent of location



Thank you for your attention...

James Weeks
james.weeks@chemweb.com

ChemWeb, Inc.
84 Theobalds Road
Holborn, London
United Kingdom WC1X 8RR