

**PubChem:
An Open Information Resource
Linking Chemistry and Biology**

Stephen H. Bryant

**Allen Press Emerging
Trends in Scholarly
Publishing Seminar**

National Press Club

April 12, 2007



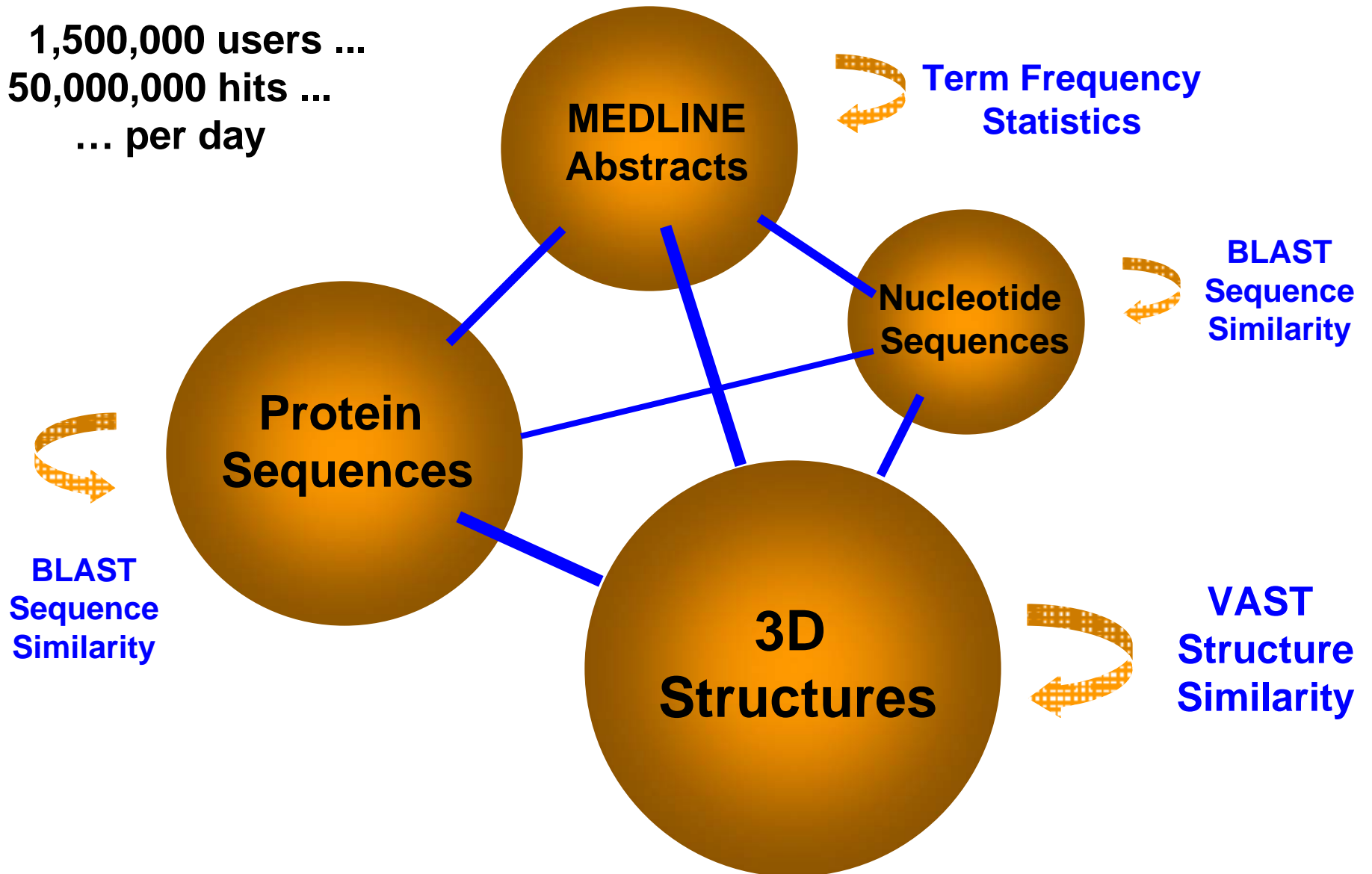
NCBI's Entrez Search Engine ...

The screenshot shows a Microsoft Internet Explorer browser window displaying the NCBI Entrez cross-database search page. The address bar shows the URL <http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi>. The page features the NCBI logo and the Entrez logo with the tagline "Entrez, The Life Sciences Search Engine". A navigation menu includes links for HOME, SEARCH, SITE MAP, PubMed, All Databases, Human Genome, GenBank, Map Viewer, and BLAST. A search bar is present with "GO" and "Clear" buttons. The main content area is titled "Welcome to the Entrez cross-database search page" and lists various databases with their descriptions and icons:

- PubMed:** biomedical literature citations and abstracts
- PubMed Central:** free, full text journal articles
- Books:** online books
- OMIM:** online Mendelian Inheritance in Man
- Site Search:** NCBI web and FTP sites
- Nucleotide:** sequence database (GenBank)
- Protein:** sequence database
- Genome:** whole genome sequences
- Structure:** three-dimensional macromolecular structures
- Taxonomy:** organisms in GenBank
- SNP:** single nucleotide polymorphism
- Gene:** gene-centered information
- UniGene:** gene-oriented clusters of transcript sequences
- CDD:** conserved protein domain database
- 3D Domains:** domains from Entrez Structure
- UniSTS:** markers and mapping data
- PopSet:** population study data sets
- GEO Profiles:** expression and molecular abundance profiles
- GEO DataSets:** experimental sets of GEO data

Entrez Retrieval System ...

1,500,000 users ...
50,000,000 hits ...
... per day



Search “fab Saul Poljak” ...

Entrez cross-database search - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi

Entrez cross-database search

NCBI Entrez, The Life Sciences Search Engine

HOME SEARCH SITE MAP PubMed All Databases Human Genome GenBank Map Viewer BLAST

Search across databases [GO] [CLEAR] Help

12	PubMed: biomedical literature citations and abstracts	none	Books: online books
66	PubMed Central: free, full text journal articles	none	OMIM: online Mendelian Inheritance in Man
	Site Search: NCBI web and FTP sites	none	OMIA: Online Mendelian Inheritance in Animals
none	Nucleotide: sequence database (includes GenBank)	36	UniGene: gene-oriented clusters of transcript sequences
17	Protein: sequence database		CDD: conserved protein domain database
6	Genome: whole genome sequences	18	3D Domains: domains from Entrez Structure
3	Structure: three-dimensional macromolecular structures	none	UniSTS: markers and mapping data
none	Taxonomy: organisms in GenBank	none	PopSet: population study data sets
74	SNP: single nucleotide polymorphisms	191	GEO Profiles: expression and molecular abundance profiles

Internet 100%

PubMed Article Retrieved ...

The screenshot shows a web browser window titled "Entrez PubMed - Windows Internet Explorer". The address bar contains the URL: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=>. The browser's menu bar includes File, Edit, View, Favorites, Tools, and Help. The search bar is set to "PubMed" and the search results show "All: 1" and "Review: 0".

The PubMed logo is prominently displayed, along with the text "A service of the National Library of Medicine and the National Institutes of Health" and the website address "www.pubmed.gov". Navigation tabs for "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "OMIM", "PMC", "Journals", and "Books" are visible. The search results display the following information:

1: [J Biol Chem](#). 1978 Jan 25;253(2):585-97. Full Text FREE Links
J Biol Chem

Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin new at 2.0 A resolution.

[Saul FA](#), [Amzel LM](#), [Poljak RJ](#).

PMID: 618887 [PubMed - indexed for MEDLINE]

Related Links

- The three-dimensional structure of the fab' fragm [Proc Natl Acad Sci U S A, 1974]
- Crystal structure of human immunoglobulin fragme [Proteins, 1992]
- Three-dimensional structure of the Fab' fragr [Proc Natl Acad Sci U S A, 1973]
- [Physico-chemical characterisation and crvs. [Biochem Soc Trans \(Lond\)](#), 1976]

The status bar at the bottom shows the address "http://www.nlm.nih.gov/" and the zoom level "100%".

Link to JBC Full Text Article ...

The screenshot shows a Windows Internet Explorer browser window. The title bar reads "Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin ne - Windows Internet Ex...". The address bar contains the URL "http://www.jbc.org/cgi/reprint/253/2/585". The browser's menu bar includes "File", "Edit", "View", "Favorites", "Tools", and "Help". The toolbar shows various navigation and utility icons, including a search bar with "Search Web" and a "65%" zoom level. The main content area displays the article page for "The Journal of Biological Chemistry", Vol. 253, No. 2, Issue of January 25, pp. 585-597, 1978. The article title is "Preliminary Refinement and Structural Analysis of the Fab Fragment from Human Immunoglobulin New at 2.0 Å Resolution*". The authors are Frederick A. Saul, L. Mario Amzel, and Roberto J. Poljak. The article is from the Department of Biophysics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205. The abstract discusses the three-dimensional structure of the Fab fragment from human psycloma IgG New, refined using "model building" and "real space" procedures. The text mentions the correlation between amino acid sequences and the 2.0 Å resolution multiple isomorphous replacement Fourier map, the average shift of all atoms during real space refinement (0.62 Å), and the distribution of ϕ , ψ angular values. The refined atomic coordinates for the 440 amino acid residues are provided, along with Ramachandran plots and the predominant β -planted sheet conformation. The structures of the homology subunits V_H , V_L , C_H1 , and C_L are also mentioned. The article is 1 of 13 pages. The right sidebar contains the "jbc ONLINE" logo and navigation links: HOME, HELP, FEEDBACK, SUBSCRIPTIONS, ARCHIVE, SEARCH, TABLE OF CONTENTS, and This Article. Under "This Article", there is a PDF version link for "Saul et al. 253 (2): 585. (1978)", along with links for "Purchase Article", "View Shopping Cart", "Alert me when this article is cited", and "Alert me if a correction is posted". A "Services" section includes "Email this article to a friend" and "Similar articles in this". The status bar at the bottom shows "Internet" and "100%" zoom.

Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin ne - Windows Internet Ex...

File Edit View Favorites Tools Help

http://www.jbc.org/cgi/reprint/253/2/585

Google

Preliminary refinement and structural anal...

Search Web

Pages

Attachments

Comments

ical Chemistry

The Journal of Biological Chemistry
Vol. 253, No. 2, Issue of January 25, pp. 585-597, 1978
Printed in U.S.A.

Preliminary Refinement and Structural Analysis of the Fab Fragment from Human Immunoglobulin New at 2.0 Å Resolution*

(Received for publication, July 7, 1977)

FREDERICK A. SAUL, L. MARIO AMZEL, AND ROBERTO J. POLJAK†

From the Department of Biophysics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

The three-dimensional structure of the Fab fragment from human psycloma IgG New has been refined using "model building" and "real space" procedures. By these techniques, the correlation between the amino acid sequences and the 2.0 Å resolution multiple isomorphous replacement Fourier map has been optimized. The average shift of all atoms during real space refinement was 0.62 Å. A list of the refined atomic coordinates for the 440 amino acid residues in the structure is given. Ramachandran plots prepared using the refined coordinates show a distribution of ϕ , ψ angular values which corresponds to the predominant β -planted sheet conformation present in the structure. The structures of the homology subunits V_H , V_L , C_H1 , and

Fab New.* Some features of the refined structure are discussed in relation to the genetic control and physiological function of immunoglobulins.

It is generally accepted (see review in Ref. 1) that electron density maps calculated by multiple isomorphous replacement techniques contain significant errors which may lead to imprecise determination of structural details such as the location of amino acid side chain atoms, bond angles, ϕ and ψ values, *cis* or *trans* character of proline residues, etc. This refinement project was undertaken with the aim of obtaining more accurate coordinates which can be applied to structural studies of other immunoglobulins and Fab-hapten complexes (2). The starting atomic coordinates were those of the structure previ-

Downloaded from www.jbc.org at National

jbc ONLINE

HOME HELP FEEDBACK

SUBSCRIPTIONS ARCHIVE SEARCH

TABLE OF CONTENTS

This Article

PDF version of:
Saul et al. 253 (2): 585.
(1978)

Purchase Article

View Shopping Cart

Alert me when this article is cited

Alert me if a correction is posted

Services

Email this article to a friend

Similar articles in this

1 of 13

Internet 100%

Link to JBC Full Text Article ...

Preliminary refinement and structural analysis of the Fab fragment from human immunoglobulin ne - Windows Internet Ex...

File Edit View Favorites Tools Help

http://www.jbc.org/cgi/reprint/253/2/585

Google

Preliminary refinement and structural anal...

65%

Search Web

Pages

580

Structural Refinement of Fab New

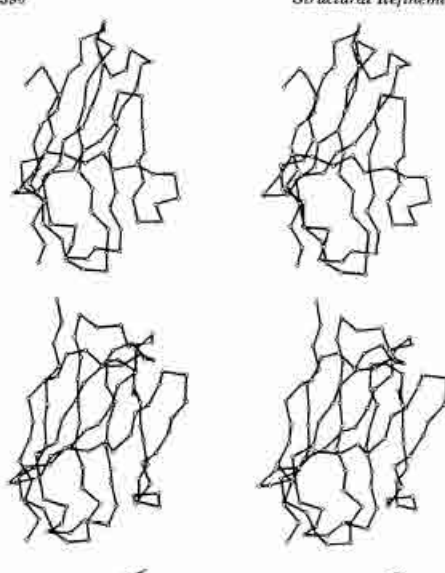


Fig. 6. Stereo pair drawings of the α carbon backbones of the V_h (top) and V_h (bottom) subunits. The subunits are viewed here in similar orientations.

Downloaded from www.jbc.org at Nank

Comments Attachments

Journal of Chemistry

HOME HELP FEEDBACK

SUBSCRIPTIONS ARCHIVE SEARCH

TABLE OF CONTENTS

This Article

PDF version of:
Saul et al. 253 (2): 585.
(1978)

[Purchase Article](#)

[View Shopping Cart](#)

[Alert me when this article is cited](#)

[Alert me if a correction is posted](#)

Services

[Email this article to a friend](#)

[Similar articles in this](#)

6 of 13

Internet 100%

Search “fab Saul Poljak” ...

Entrez cross-database search - Windows Internet Explorer

http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi

Entrez cross-database search

NCBI Entrez, The Life Sciences Search Engine

HOME SEARCH SITE MAP PubMed All Databases Human Genome GenBank Map Viewer BLAST

Search across databases [GO] [CLEAR] Help

12	PubMed: biomedical literature citations and abstracts	none	Books: online books
66	PubMed Central: free, full text journal articles	none	OMIM: online Mendelian Inheritance in Man
	Site Search: NCBI web and FTP sites	none	OMIA: Online Mendelian Inheritance in Animals
none	Nucleotide: sequence database (includes GenBank)	36	UniGene: gene-oriented clusters of transcript sequences
17	Protein: sequence database		CDD: conserved protein domain database
6	Genome: whole genome sequences	18	3D Domains: domains from Entrez Structure
3	Structure: three-dimensional macromolecular structures	none	UniSTS: markers and mapping data
none	Taxonomy: organisms in GenBank	none	PopSet: population study data sets
74	SNP: single nucleotide polymorphisms	191	GEO Profiles: expression and molecular abundance profiles

Internet 100%

Protein 3D Structure Record ...

Structure Summary, 7FAB, 3367 - Windows Internet Explorer

File Edit View Favorites Tools Help

http://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cgi?form=6&db=t&Dopt= Google

Structure Summary, 7FAB, 3367

Reference: Saul FA, Poljak RJ [Crystal structure of human immunoglobulin fragment Fab New refined at 2.0 A resolution](#) *Proteins* v14, p.363-371
[All References](#)

Description: Immunoglobulin Fab' New (Lambda Light Chain).

Deposition: 1991/11/20

Taxonomy: [Homo sapiens](#)

MMDB: [3367](#) **PDB:** [7FAB](#) **Structure Neighbors:** [VAST](#)

View 3D Structure of All Atom Model Cn3D Display [Download Cn3D!](#)

Molecular components in the MMDB structure are listed below. The icons indicate macromolecular chains, 3D domains, protein classifications and ligands. Please hold the mouse over each icon for more information on the component.

Protein Chain L

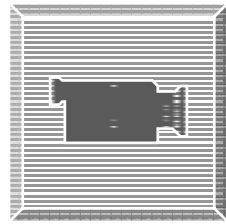
3D Domains 1 2

Domain Family IGv IGc

V-set

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=pubmed&c Internet 100%

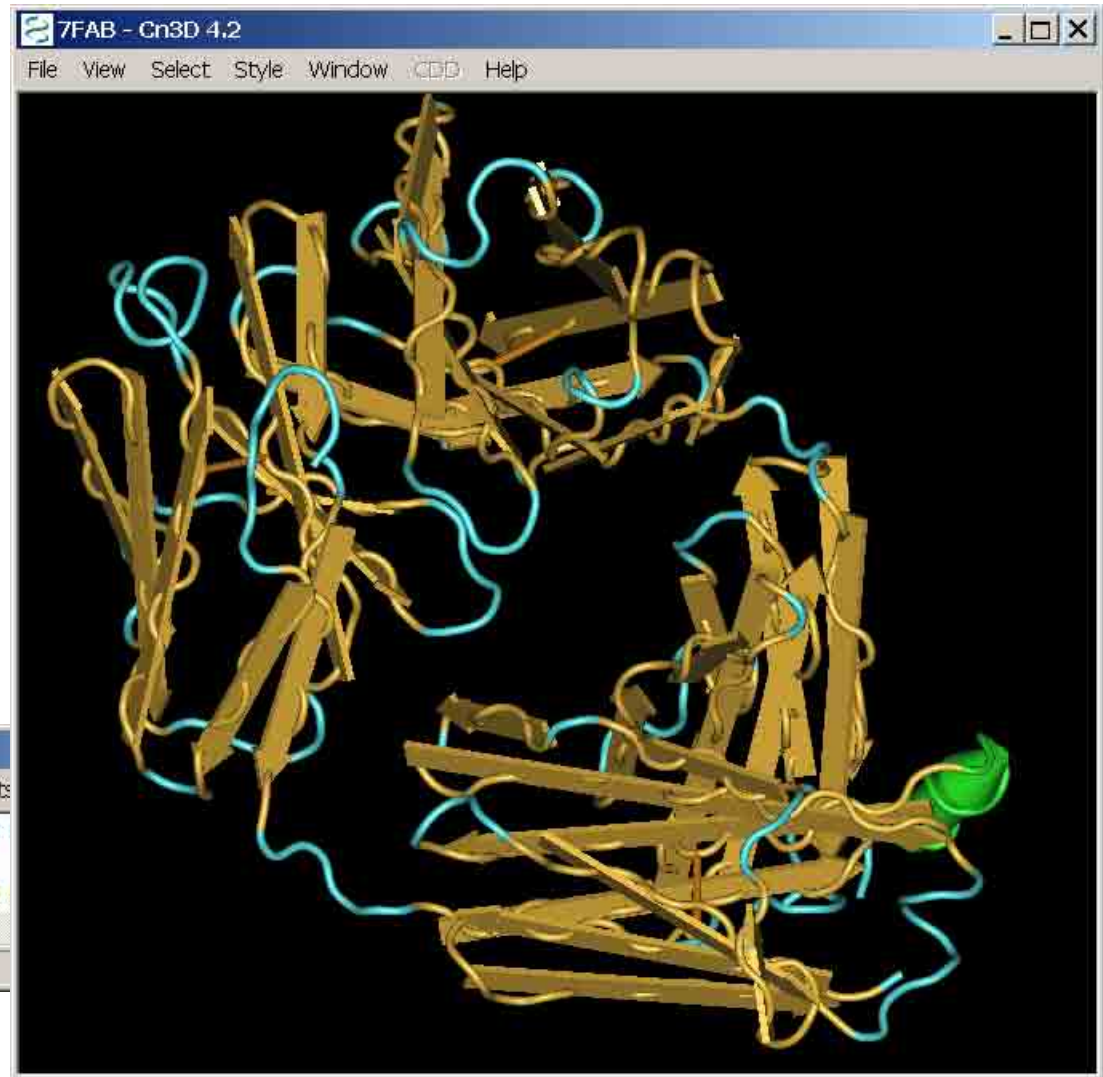
View Protein 3D Structure ...



7FAB - Sequence/Alignment Viewer

View Edit Mouse Mode Unaligned Justification Imports

7FAB_L	a s v l t q p p s y s g a p g q r v t i s o t e s
7FAB_H	a v q l e q s g p g l v r p s q t l i s l e t v s



3D Structure Neighbors ...

Vast Neighbor Summary - Windows Internet Explorer

File Edit View Favorites Tools Help

http://www.ncbi.nlm.nih.gov/Structure/vast/vastsrv.cgi?reqid=&sdid=8319&allbfc

Vast Neighbor Summary

View 3D Alignment of All Atoms with Cn3D Display [Download Cn3D!](#)

View Sequence Alignment using Hypertext for Selected VAST neighbors

List Non-identical seq. subset, sorted by Aligned Length in Graphics

Advanced neighbor search

Move the mouse over the red alignment footprints in the graphics below and click, you will obtain a structure-based sequence alignment.

Total neighbors: 8963; 1 - 60 of 3644 representatives from the [Non-identical seq.](#) subset displayed. Page: 1

Click to: [Check All](#) [Uncheck All](#)

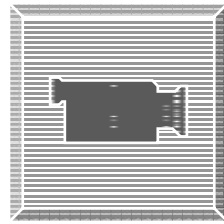
[7FAB L](#)
[3d Dom.](#)
[Protein Family](#)

[1MCS A 1](#) 131

[1A0K L](#) 131

Internet 100%

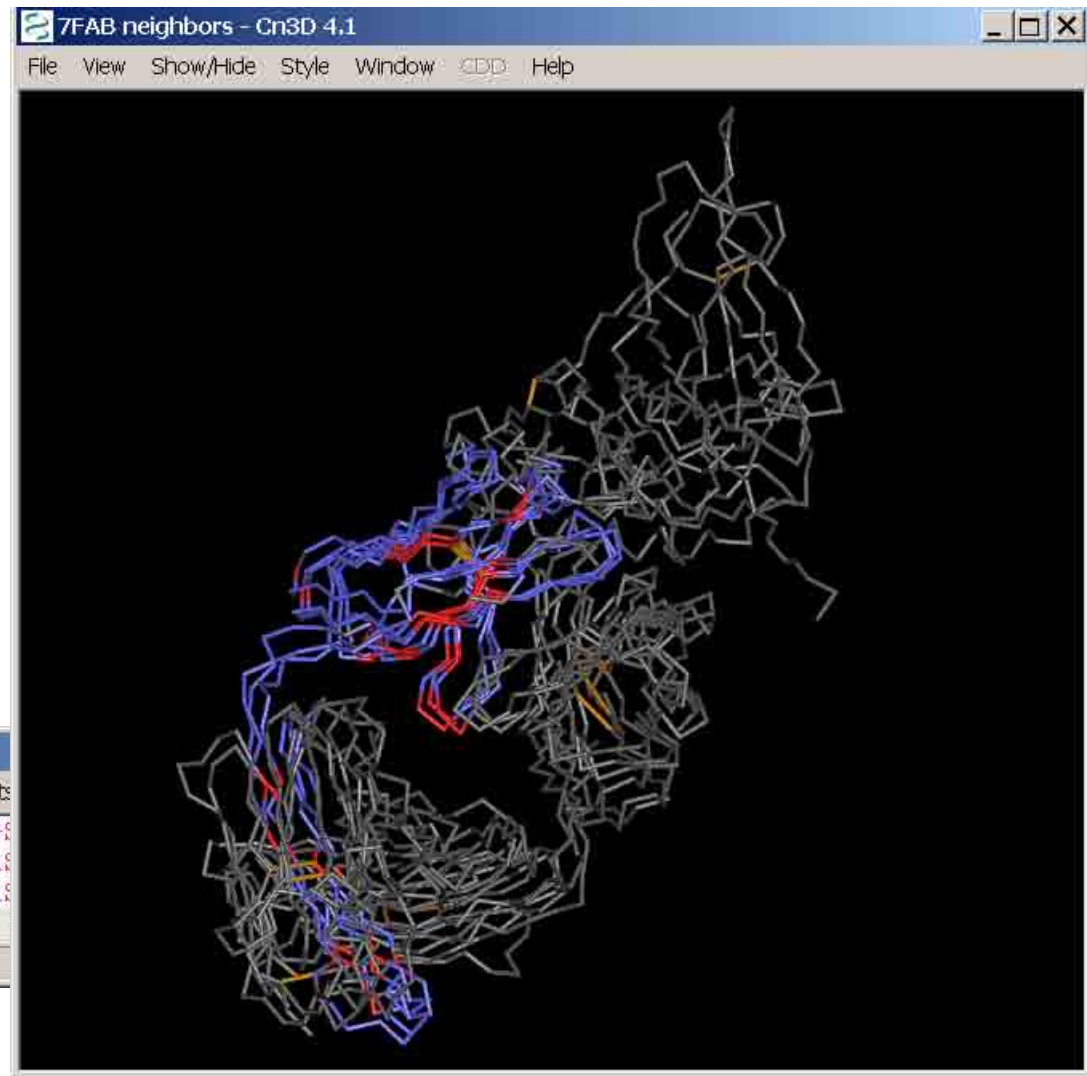
Link to Another Structure ...



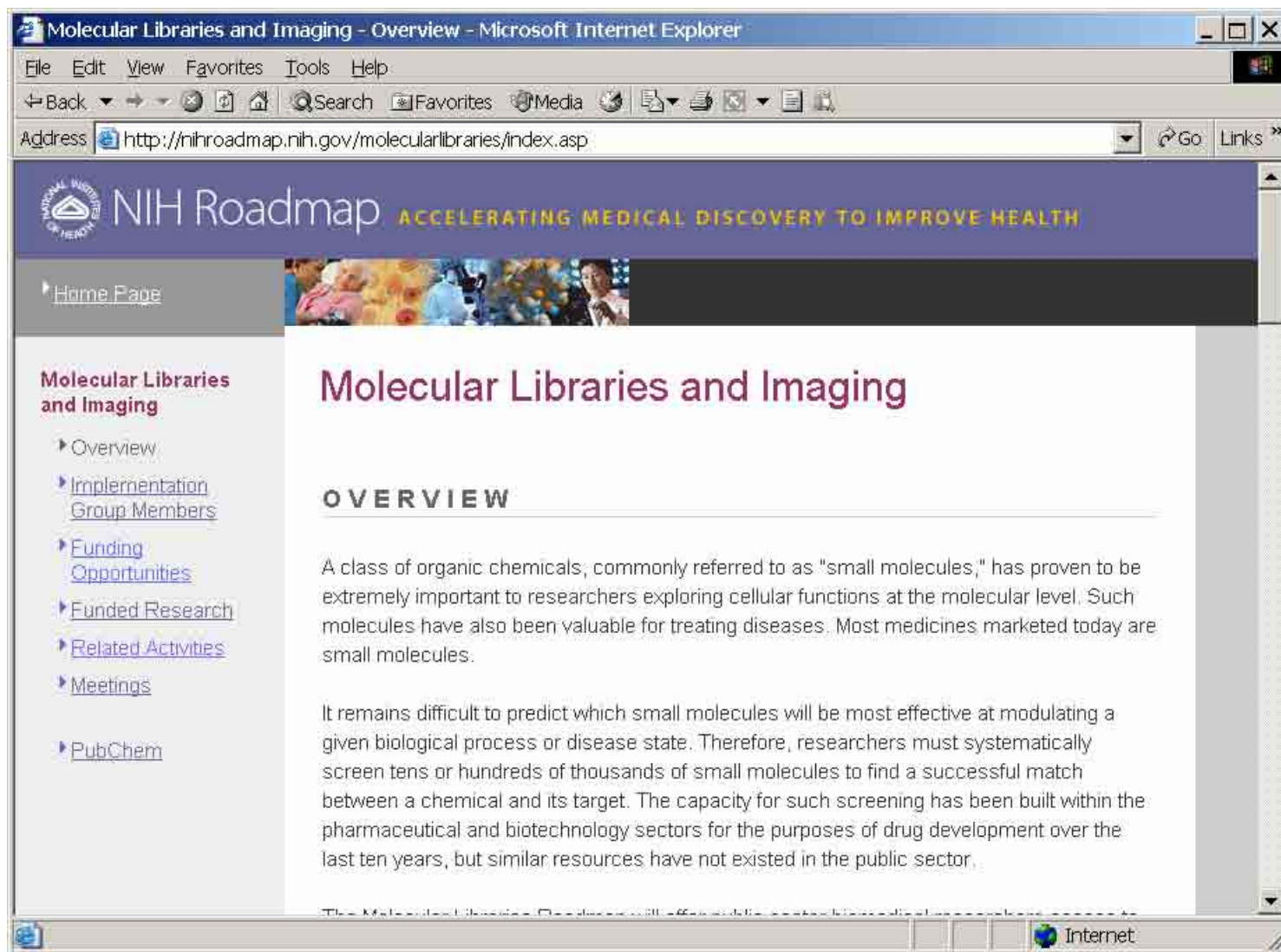
7FAB neighbors - Sequence/Alignment Viewer

View Edit Mouse Mode Unaligned Justification Imports

```
7FAB_L  a s VLTQP~PSVSGAPGQRVTISCTG$
1UJ3_B  q v QLLES~GAVLARPGTSVKISCKA$
1NSN_L  d i VLTQS p SSLAVSLGQRATISCRAS
```



Molecular Libraries Roadmap ...



The screenshot shows a Microsoft Internet Explorer browser window displaying the NIH Roadmap Molecular Libraries and Imaging Overview page. The browser's address bar shows the URL: <http://nihroadmap.nih.gov/molecularlibraries/index.asp>. The page header features the NIH Roadmap logo and the tagline "ACCELERATING MEDICAL DISCOVERY TO IMPROVE HEALTH". A navigation menu on the left includes links for Home Page, Overview, Implementation, Group Members, Funding Opportunities, Funded Research, Related Activities, Meetings, and PubChem. The main content area is titled "Molecular Libraries and Imaging" and contains an "OVERVIEW" section. The overview text discusses the importance of small molecules in research and drug development, noting that while the pharmaceutical and biotechnology sectors have built screening capacity, similar resources are lacking in the public sector.

Molecular Libraries and Imaging

- ▶ [Overview](#)
- ▶ [Implementation](#)
- ▶ [Group Members](#)
- ▶ [Funding Opportunities](#)
- ▶ [Funded Research](#)
- ▶ [Related Activities](#)
- ▶ [Meetings](#)

- ▶ [PubChem](#)

Molecular Libraries and Imaging

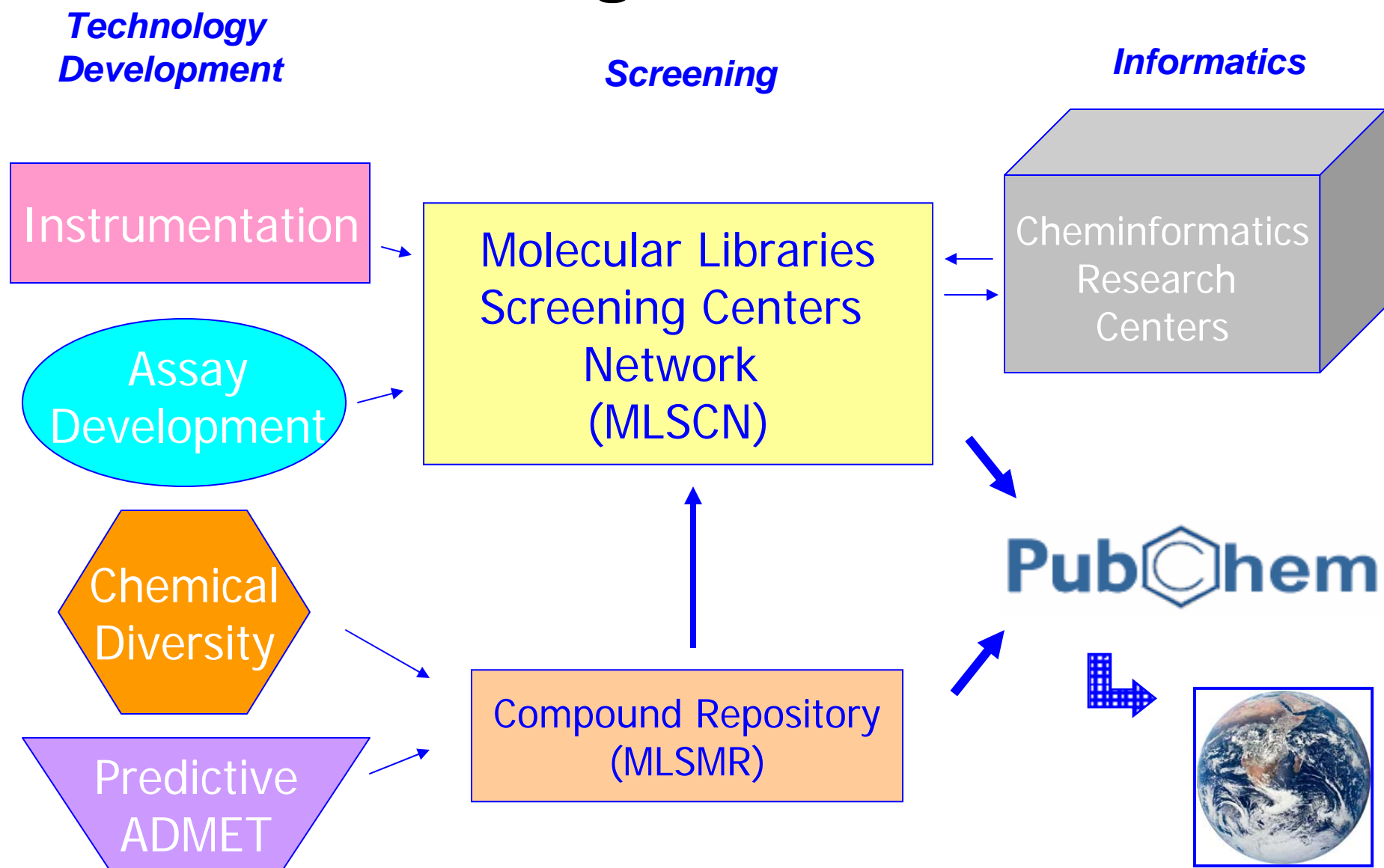
OVERVIEW

A class of organic chemicals, commonly referred to as "small molecules," has proven to be extremely important to researchers exploring cellular functions at the molecular level. Such molecules have also been valuable for treating diseases. Most medicines marketed today are small molecules.

It remains difficult to predict which small molecules will be most effective at modulating a given biological process or disease state. Therefore, researchers must systematically screen tens or hundreds of thousands of small molecules to find a successful match between a chemical and its target. The capacity for such screening has been built within the pharmaceutical and biotechnology sectors for the purposes of drug development over the last ten years, but similar resources have not existed in the public sector.

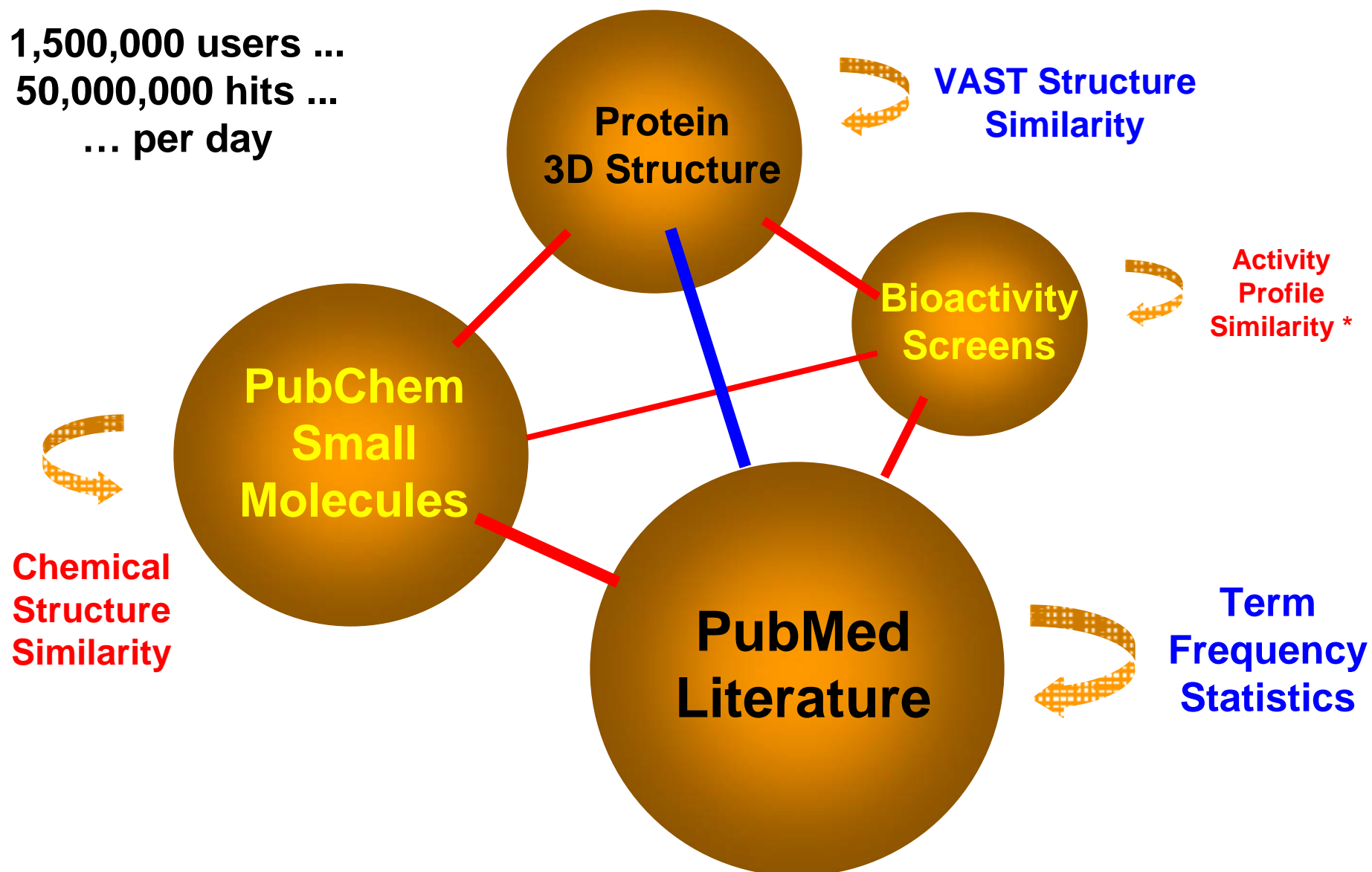
The Molecular Libraries Roadmap will offer public sector biomedical researchers

The Molecular Libraries Roadmap: An Integrated Initiative



Entrez Links and Neighbors ...

1,500,000 users ...
50,000,000 hits ...
... per day



PubChem Goals ...

... Archive molecular structure and bioassay data from the Molecular Libraries Screening Center Network

... Provide search, retrieval and data analysis tools to optimize utility of these results

PubChem Goals ...

... Further optimize research utility by including other public sources of chemical structure and bioactivity information

... and by integration with other NIH Biomedical information resources whenever possible

PubChem Approach ...

... “GenBank model”

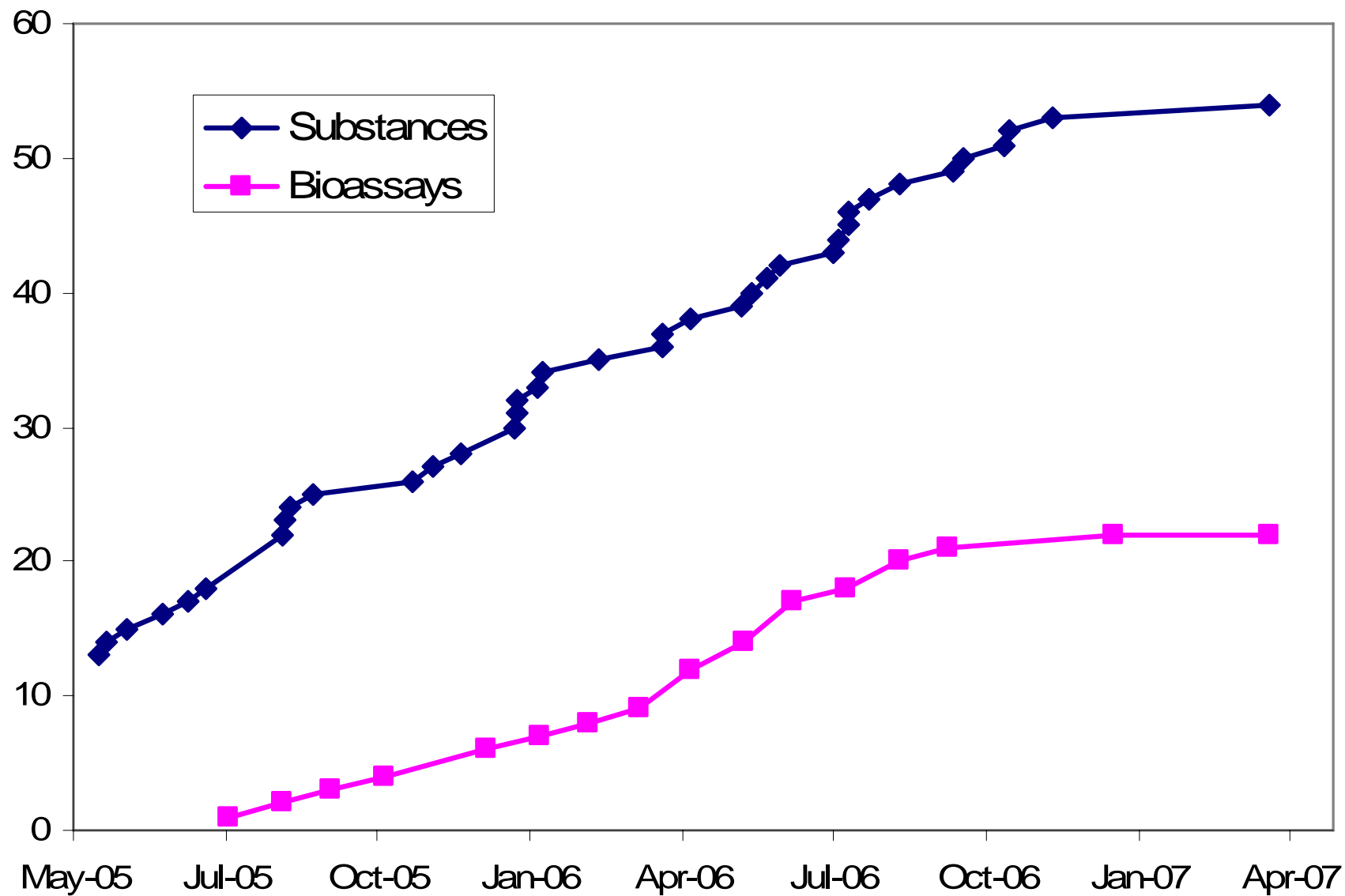
... direct depositions by investigators

... highly automated (low database cost)

... 25 year precedents in biology

... less precedent in chemistry

Growth In PubChem Contributing Organizations



PubChem Contents ...

... Contributed substance records

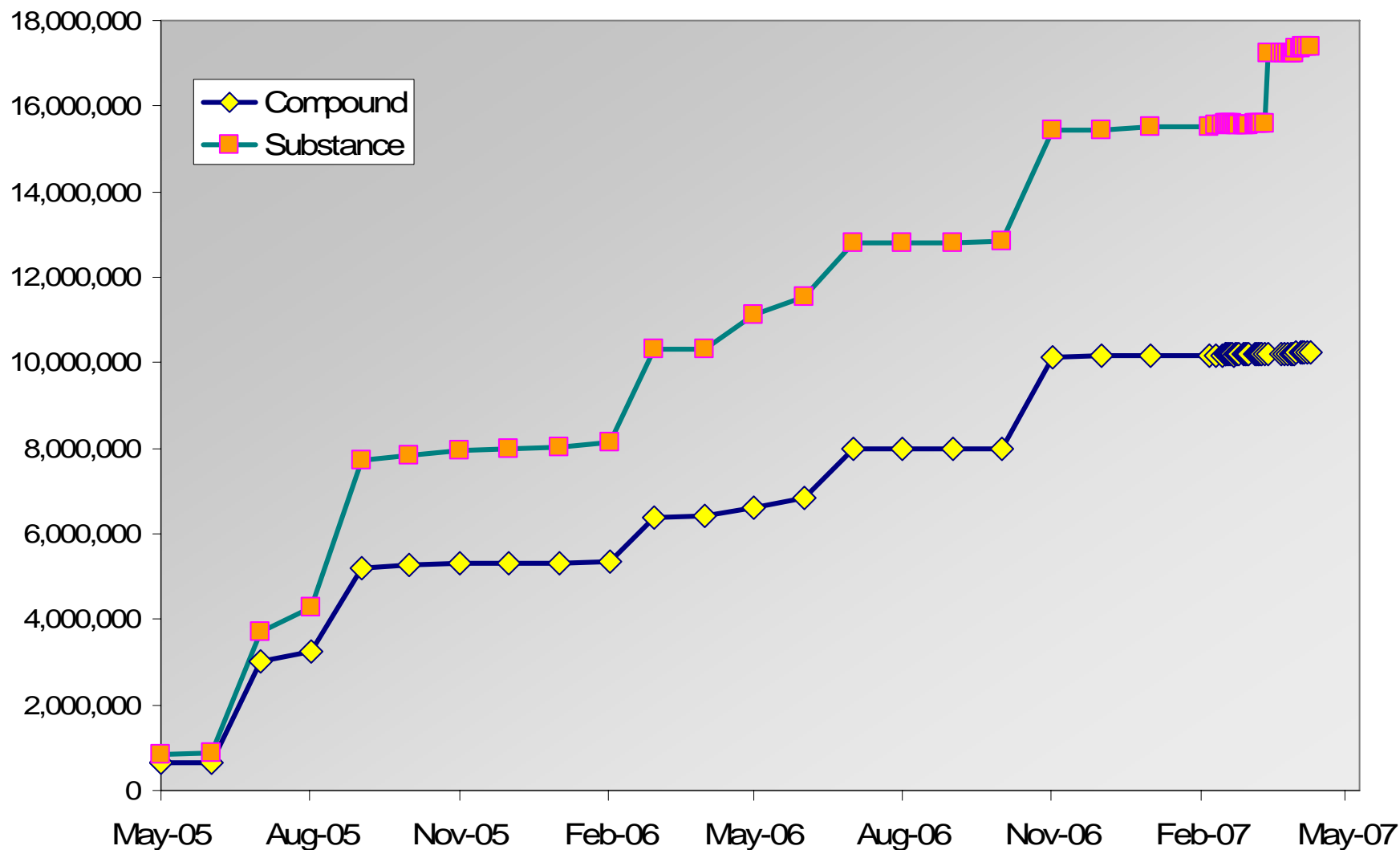
... with chemical structure

... chemical names and comments

... links to contributor web sites

**... contributed links to other NCBI
biomedical databases**

Growth In PubChem Substances / Compounds



PubChem Contents ...

... Contributed bioassay records

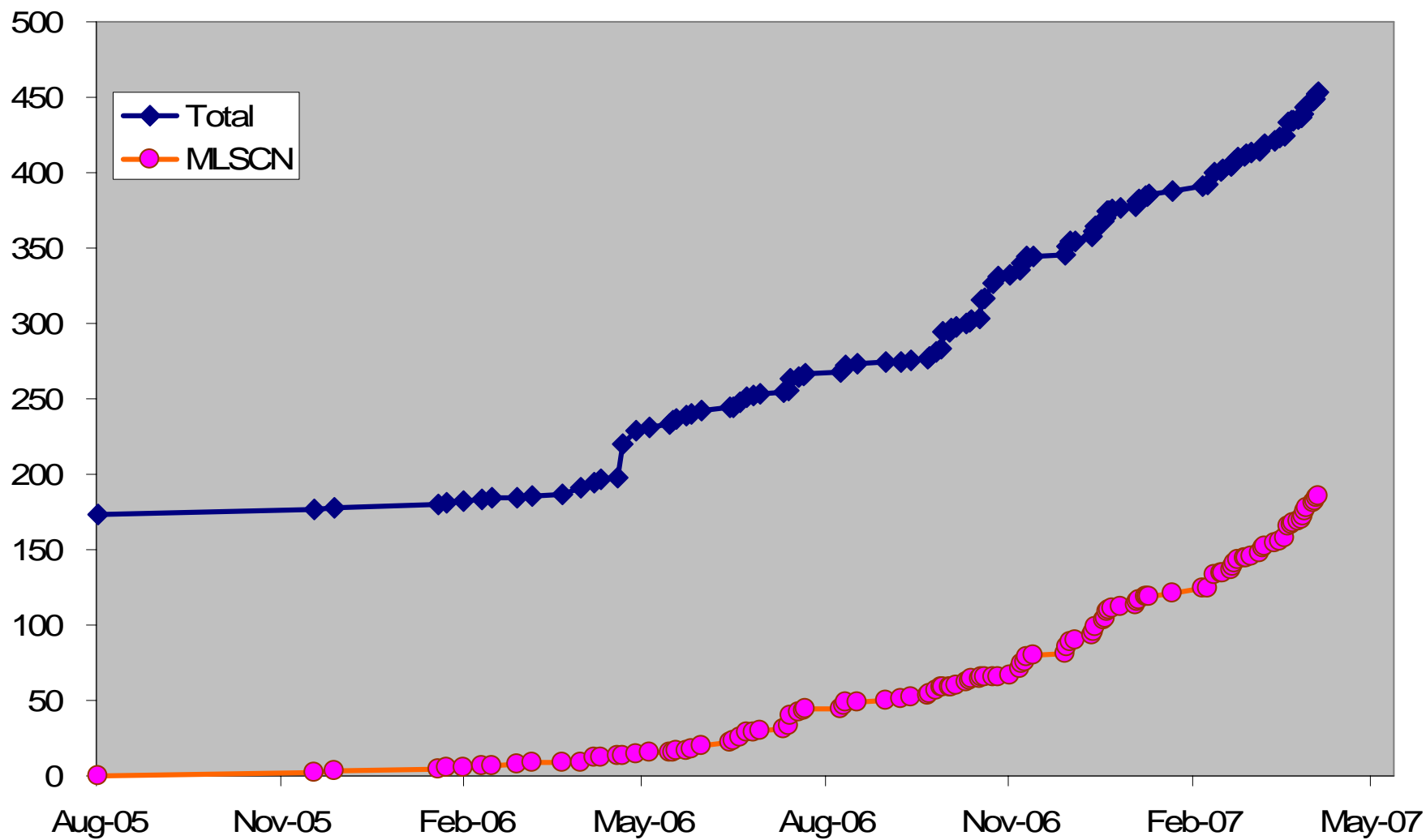
... with assay description / protocol

... links to tested substances

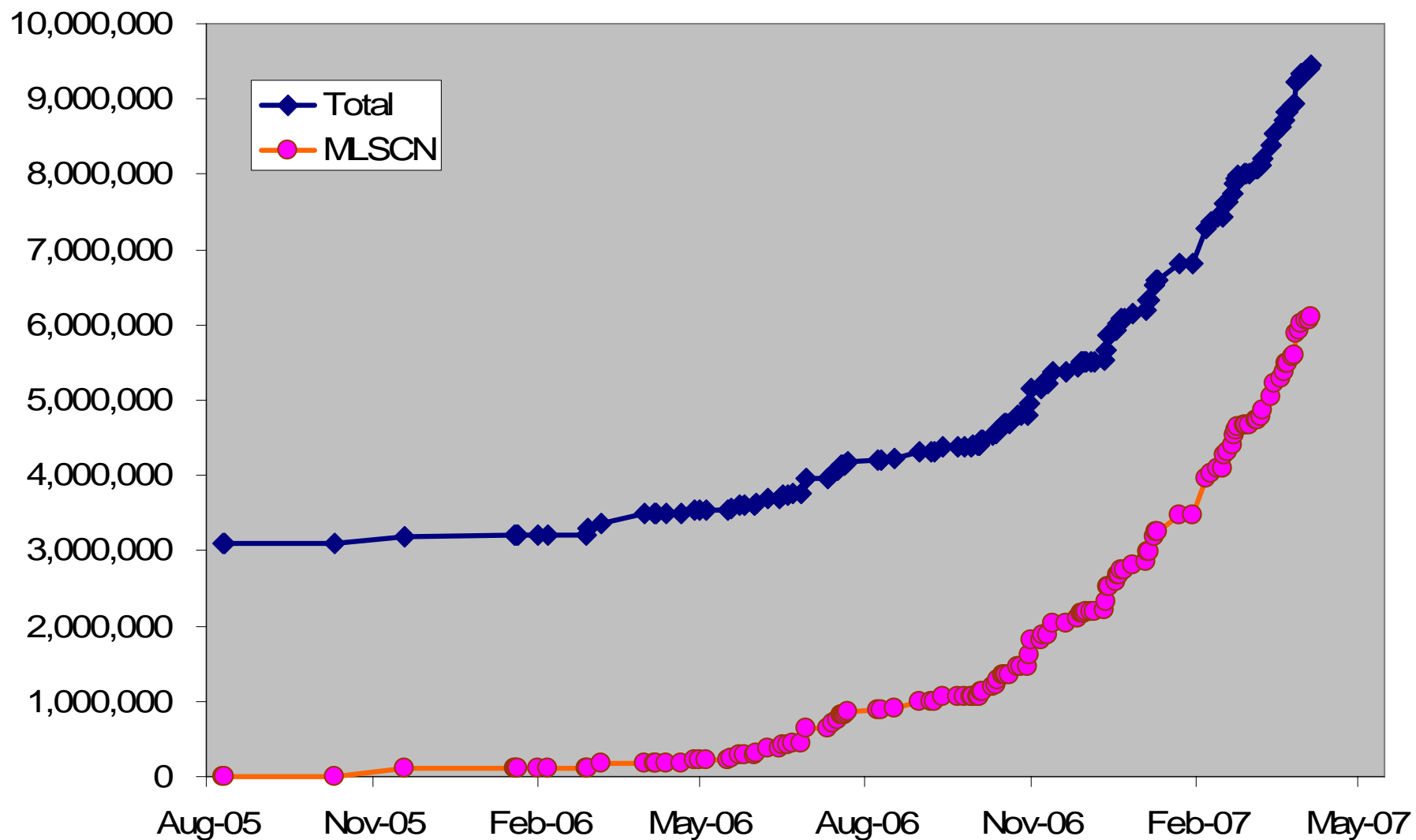
... summary and detailed test results

**... links to contributor web sites and other
NCBI databases**

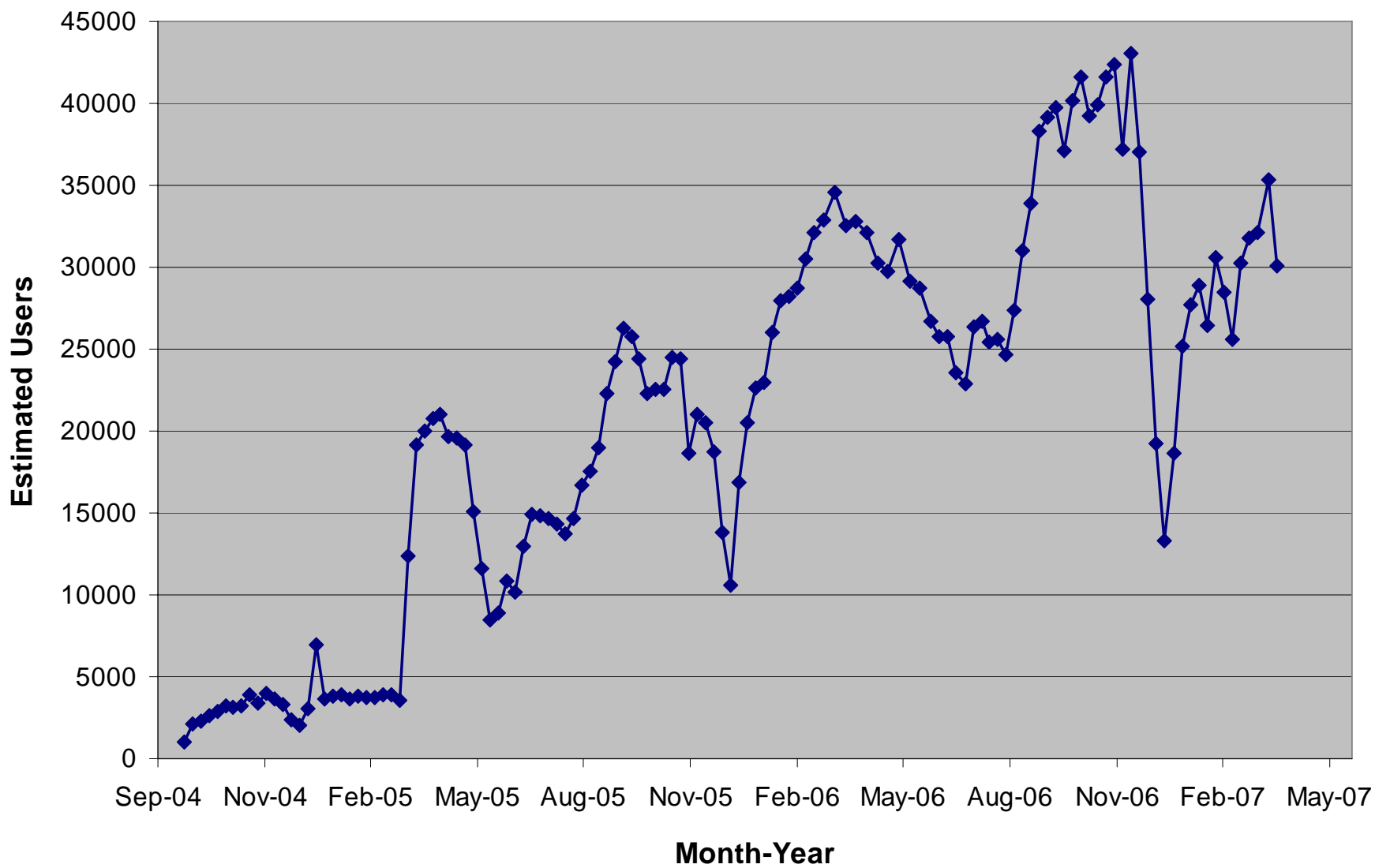
Growth In PubChem BioAssays



Growth In PubChem Tested Substances



Growth in PubChem Users per Day



PubChem Retrieval System ...

... Optimize “discoverability” for molecular biologists by integrating PubChem into NCBI’s Entrez / PubMed Search Engine

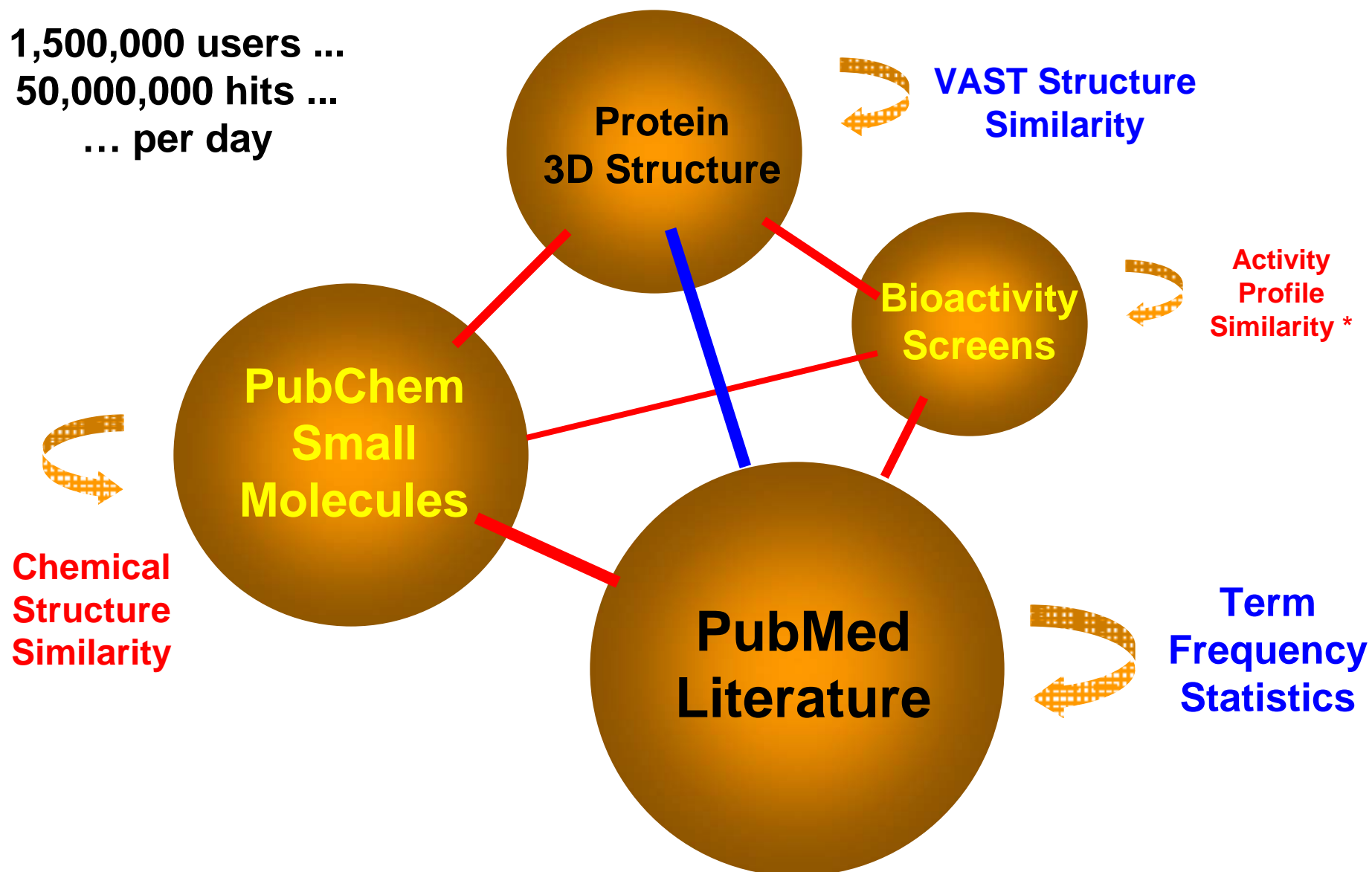
... Chemical structure search

... Bioassay result search

... Exploratory structure-activity tools

Entrez Links and Neighbors ...

1,500,000 users ...
50,000,000 hits ...
... per day



Search for “Shoichet inhibitors” ...

The screenshot shows a Windows Internet Explorer browser window titled "Entrez cross-database search - Windows Internet Explorer". The address bar contains the URL "http://www.ncbi.nlm.nih.gov/gquery/gquery.fcgi". The search bar on the page contains the text "shoichet inhibitors". The search results are displayed in a grid format, showing the number of hits for various databases. The results are as follows:

Database	Number of Hits
PubMed: biomedical literature citations and abstracts	51
PubMed Central: free, full text journal articles	33
Site Search: NCBI web and FTP sites	19
Nucleotide: sequence database (includes GenBank)	3303
Protein: sequence database	35
Genome: whole genome sequences	115
Structure: three-dimensional macromolecular structures	11
Taxonomy: organisms in GenBank	none
SNP: single nucleotide polymorphism	none
Books: online books	none
OMIM: online Mendelian Inheritance in Man	none
OMIA: Online Mendelian Inheritance in Animals	none
UniGene: gene-oriented clusters of transcript sequences	34
CDD: conserved protein domain database	112
3D Domains: domains from Entrez Structure	41
UniSTS: markers and mapping data	26
PopSet: population study data sets	63
GEO Profiles: expression and molecular abundance profiles	234943

PubMed Article Retrieved ...

The screenshot shows a web browser window titled "Entrez PubMed - Windows Internet Explorer". The address bar contains the URL: <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=>. The browser's menu bar includes File, Edit, View, Favorites, Tools, and Help. The address bar also features a search box with "Google" and a "Page" dropdown menu.

The main content area displays the NCBI PubMed logo and navigation tabs for "All Databases", "PubMed", "Nucleotide", "Protein", "Genome", "Structure", "OMIM", "PMC", "Journals", and "Books". A search bar is present with "PubMed" selected in the dropdown and a "Go" button. Below the search bar are buttons for "Limits", "Preview/Index", "History", "Clipboard", and "Details". The display settings show "AbstractPlus" selected, "Show 20" items, and a "Sort by" dropdown. A status bar indicates "All: 1" and "Review: 0".

The article information is displayed as follows:

- 1: [Nat Chem Biol.](#) 2005 Aug;1(3):146-8. Epub 2005 Jul 3.
- Comment in: [Nat Chem Biol.](#) 2005 Aug;1(3):125.
- High-throughput assays for promiscuous inhibitors.**
- Feng BY, Shelat A, Doman TN, Guy RK, Shoichet BK.**
- Department of Pharmaceutical Chemistry & Graduate Group in Chemistry and Chemical Biology, 1700 4th St., University of California San Francisco, San Francisco, California 94143-2550, USA.

On the right side, there is a "Links" dropdown menu with options: BioAssay, Substance via MeSH, Substance (Publisher), Cited in PMC, and LinkOut. Below this is a "Related Links" section with several entries:

- A common mechanism for promiscuous inhibition [J Med Chem. 2002]
- Identification and prediction of promiscuous aggregations [J Med Chem. 2003]
- Virtual screening to enrich hit lists from high-throughput screening [Proteins. 2003]
- Kinase inhibitors: not just for kinases anymore. [J Med Chem. 2003]

The browser's status bar at the bottom shows "Internet" and "100%" zoom level.

Link to PubChem Records ...

PubChem Substance - Windows Internet Explorer

File Edit View Favorites Tools Help

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?tool=pubmed_DocSum&db=pubm Google

PubChem Substance

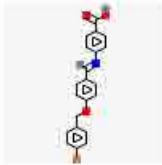
Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

All: 29 BioAssay: 0 Protein3D: 0 Rule of 5: 24


Items 1 - 20 of 29 Page 1 of 2 Next

1: SID: [855510](#) [Related Structures, Other Links](#)



CID: [657179](#), nchembio718-comp29, (E)-4-(4-(4-bromobenzoyloxy)benzylideneamino)benzoic acid
Source: [Nature Chemical Biology\(nchembio718-comp29\)](#)
IUPAC: 4-[[4-[(4-bromophenyl)methoxy]phenyl]methylideneamino]benzoic acid
MW: 410.261 | MF: C21H16BrNO3

2: SID: [855509](#) [Related Structures, Other Links](#)



CID: [5388956](#), nchembio718-comp28, (Z)-4-((4,6-dioxo-2-thioxo-tetrahydropyrimidin-5(6H)-ylidene)methyl)phenyl 3-phenylacrylate
Source: [Nature Chemical Biology\(nchembio718-comp28\)](#)
IUPAC: [4-[(4,6-dioxo-2-sulfanylidene-1,3-diazinan-5-ylidene)methyl]phenyl] (Z)-3-phenylprop-2-enoate
MW: 378.402 | MF: C20H14N2O4S

Internet 100%

“Kaempferol” in PubChem ...

CID 576 -- PubChem Compound Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print

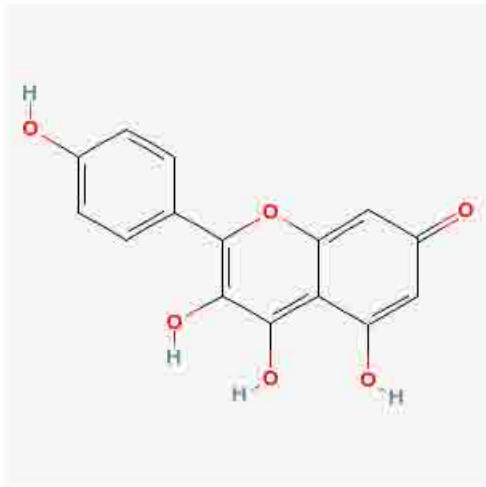
Address <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=576> Go Links

NCBI PubChem National Library of Medicine NLM









WOMI SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

Search PubChem Compound GO

Compound Summary:



The image shows the chemical structure of Kaempferol, a flavonoid. It consists of a central chromone ring system with a 4-hydroxyphenyl group at position 7, a 3,5-dihydroxyphenyl group at position 2, and a 4-hydroxyphenyl group at position 4. The structure is shown in a 2D representation with hydrogen atoms explicitly drawn on the hydroxyl groups.

-  CID: 576 [?](#)
-  Substances: [?](#)
 - All: 14 Links
 - Same: 12 Links
 - Mixture: 2 Links
-  BioActivity: 7 Links [?](#)
-  PubMed: 16 Links [?](#)
-  Protein Structure: 1 Link [?](#)
-  NLM Toxicology: Link [?](#)
-  Similar Compounds: 73 Links [?](#)
-  Structure Search [?](#)

MeSH Synonyms Properties Descriptors Category Exports

Internet

Link to Contributor Site ...

High-throughput assays for promiscuous inhibitors : Nature Chemical Biology - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print

Address <http://www.nature.com/nchembio/journal/v1/n3/full/nchembio718.html> Go Links

Accessibility statement Jump to main content Jump to main navigation Jump to search Jump to login

NATURE.COM NEWS@NATURE.COM NATUREJOBS NATUREEVENTS ABOUT NPG Help Nature.com site index

nature chemical biology

E-MAIL: PASSWORD

Save password: ? | [Forgotten password?](#)

login to nature.com

Search journal Advanced search MY ACCOUNT E-ALERTS SUBSCRIBE REGISTER

Journal home > Archive > Letter > Full Text

Journal home
Advance online publication
Current issue
Archive
Press releases
For authors
Online submission
For referees
Free online issue
About the journal

Letter

Nature Chemical Biology **1**, 146-148 (2005)
doi: 10.1038/nchembio718

High-throughput assays for promiscuous inhibitors

Brian Y Feng^{1,3}, Anang Shelat^{1,3}, Thompson N Doman², R Kip Guy¹ and Brian K Shoichet¹

High-throughput screening (HTS) searches large libraries of chemical compounds for those that can modulate the activity

nature chemical biology
Subscribe

FULL TEXT

Previous article | Next article

Table of contents

Download PDF

Send to a friend

Export citation

Export references

Done Internet

Similar Compounds in PubChem ...

PubChem Compound - Windows Internet Explorer

File Edit View Favorites Tools Help

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=DisplayFiltered&DB=pccomp

PubChem Compound

About Entrez
Entrez Help

PubChem
Help | FAQ

PubChem
Substance
Structures
supplied by
depositors


PubChem
Compound
Unique structures
with computed
properties

PubChem
BioAssay
Bioactivity assay
results supplied by
depositors

All 1481 BioAssay 163 Protein3D: 12 Rule of 5 1257


Items 1 - 12 of 12 One page.

1: CID: [5280863](#) Related Structures, Assays, Literature, Other Entrez Databases



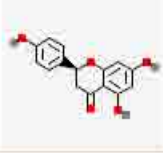
kaempferol, Kaempherol ...
IUPAC: 3,5,7-trihydroxy-2-(4-hydroxyphenyl)chromen-4-one
MW: 286.236 | MF: C15H10O6

2: CID: [5280343](#) Related Structures, Assays, Literature, Other Entrez Databases



quercetin, Sophoretin ...
IUPAC: 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-chromen-4-one
MW: 302.236 | MF: C15H10O7

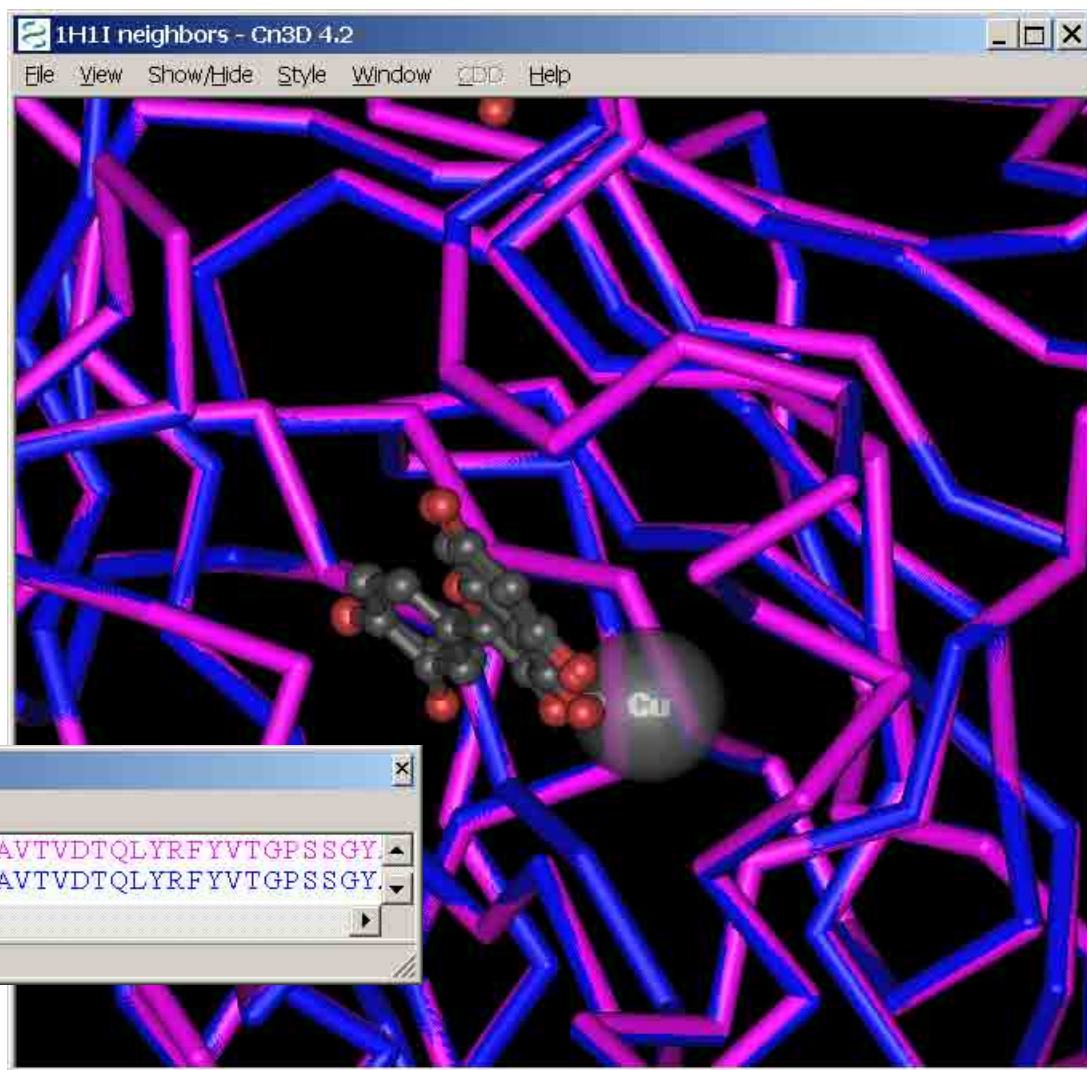
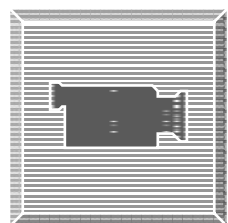
3: CID: [439246](#) Related Structures, Assays, Literature, Other Entrez Databases



naringenin, naringetol ...
IUPAC: (2S)-5,7-dihydroxy-2-(4-hydroxyphenyl)chroman-4-one
MW: 272.253 | MF: C15H12O5

Done Internet 100%

Compare Protein / Ligand Complexes ...



“Quercetin” in PubChem ...

CID 480 -- PubChem Compound Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print

Address <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=480> Go Links »

- CID: 480 [?](#)
- Substances: [?](#)
 - All: 21 Links
 - Same: 16 Links
 - Mixture: 5 Links
- BioActivity: 74 Links [?](#)
- PubMed: 27 Links [?](#)
- Protein Structures: 4 Links [?](#)
- NLM Toxicology: Link [?](#)
- Related Compounds: [?](#)
 - Same, Any Tautomers: 2 Links
- Similar Compounds: 109 Links [?](#)
- Structure Search [?](#)

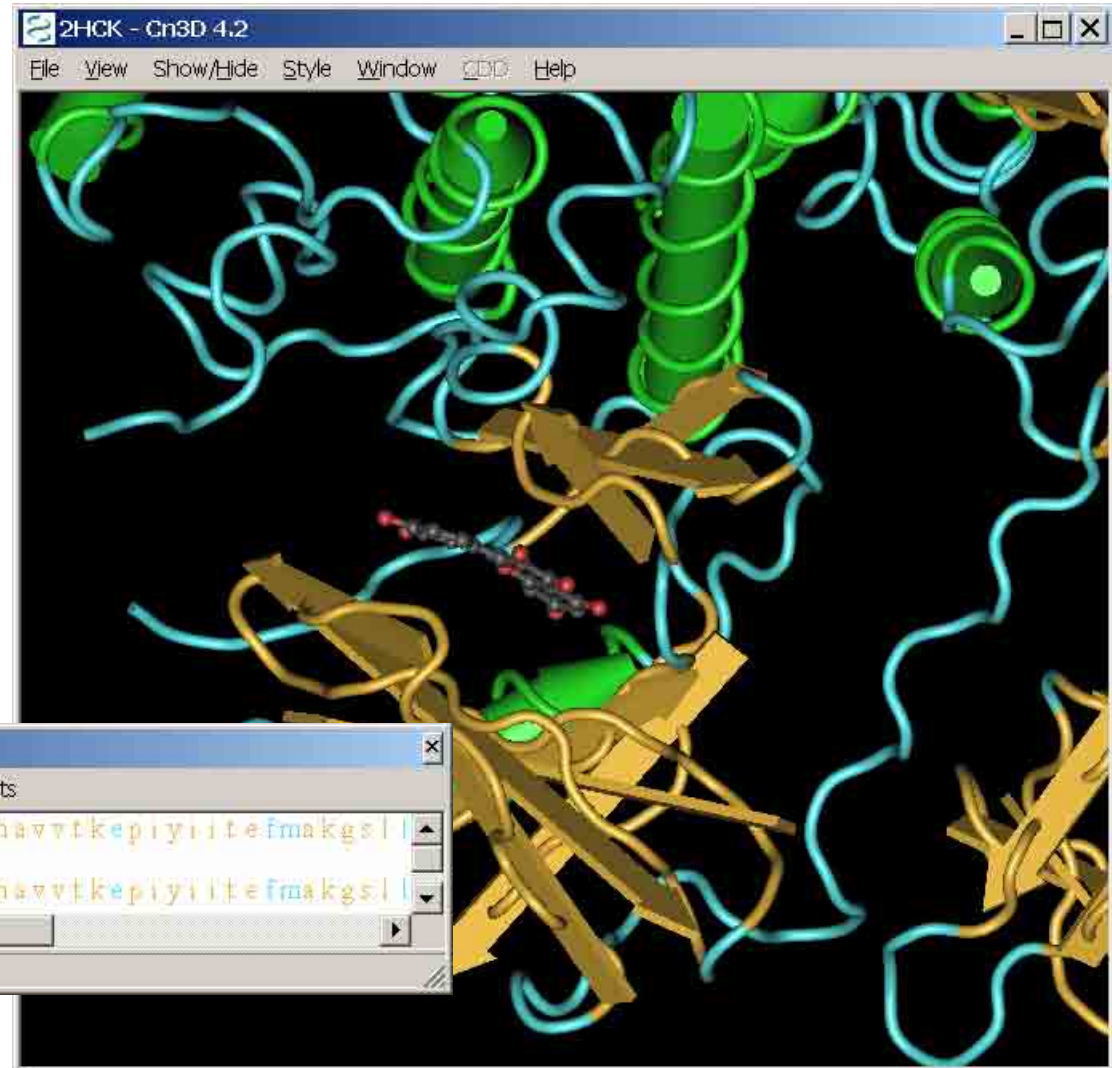
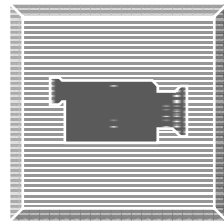
[MeSH](#) [Synonyms](#) [Properties](#) [Descriptors](#) [Category](#) [Exports](#)

Medical Subject Annotations: (Total: 2) [?](#) Display: Next 1 | All

Quercetin
A flavonol widely distributed in plants. It is an antioxidant, like many other phenolic

Internet

Link to Another Structure ...



Tyrosine Kinase Family Member ...

TyrKc - Cn3D 4.2
File View Show/Hide Style Win

CDD Descriptive Items
Name: TyrKc
Tyrosine kinase, catalytic domain. Phosphotransferases; tyrosine-specific kinase subfamily. Enzymes with TyrKc domains belong to an extensive family of proteins which share a conserved catalytic core common to both serine/threonine and tyrosine protein kinases. Enzymatic activity of tyrosine protein kinases is controlled by phosphorylation
Show Annotations Panel Show References Panel Dismiss

TyrKc - Sequence/Alignment Viewer
View Edit Mouse Mode Unaligned Justification Imports

<i>IFGI_A</i>	l p e d p r W E L P R D R L V L g ~ ~ ~ k P L
<i>IBYG_A</i>	e f y r s g W A L N M K E L K L l ~ ~ ~ q T I
<i>IIR3_A</i>	v f v p d e W E V S R E K I T L l ~ ~ ~ r E L
<i>IIRK</i>	v f v p d e W E V S R E K I T L l ~ ~ ~ r E L
<i>2HCK_A</i>	p w e k d a W E I P R E S L K L e ~ ~ ~ k K L
gi136238	k p d t y v Q H I K R R D I V L k ~ ~ ~ r E L
gi422541	n a k l l s L E Y P R N N I E Y v ~ ~ ~ r D I
gi232109	

Links from Compound to PubMed ...

PubChem Compound Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print

Address <http://pubchem.ncbi.nlm.nih.gov/summary/summary.cgi?cid=480> Go Links

MeSH Synonyms Properties Descriptors Category Exports

Medical Subject Annotations: (Total: 2) [?](#) Display: Next 1 | All

Quercetin
A flavonol widely distributed in plants. It is an antioxidant, like many other phenolic heterocyclic compounds. Glycosylated forms include RUTIN and quercetrin.

[Show MeSH Tree Structure](#)

PubMed via MeSH Choose by Subheadings:

administration and dosage	adverse effects	analogs and derivatives
analysis	antagonists and inhibitors	biosynthesis
blood	chemical synthesis	chemistry
diagnostic use	genetics	immunology
isolation and purification	metabolism	pharmacokinetics
pharmacology	physiology	radiation effects
secretion	standards	therapeutic use
toxicity	urine	

Depositor-Supplied Synonyms: (Total: 74) [?](#)

Internet

Links from Compounds to PubMed ...

Entrez PubMed - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=pubmed&term=%22Quercetin%2fantagonists%20and%20> Go Links »

PubMed Services Journals Database MeSH Database Single Citation Matcher Batch Citation Matcher Clinical Queries Special Queries LinkOut My NCBI (Cubby) Related Resources Order Documents NLM Mobile NLM Catalog NLM Gateway TOXNET Consumer Health Clinical Alerts ClinicalTrials.gov PubMed Central

2: [Lin HC, Cheng TH, Chen YC, Juan SH](#) Related Articles, Links
Mechanism of heme oxygenase-1 gene induction by quercetin in rat aortic smooth muscle cells.
Pharmacology. 2004 Jun;71(2):107-12.
PMID: 15118350 [PubMed - indexed for MEDLINE]

3: [Ko WC, Chen MC, Wang SH, Lai YH, Chen JH, Lin CN](#) Related Articles, Links
3-O-methylquercetin more selectively inhibits phosphodiesterase subtype 3.
Planta Med. 2003 Apr;69(4):310-5.
PMID: 12709896 [PubMed - indexed for MEDLINE]

4: [Revueita MP, Cantabrana B, Hidalgo A](#) Related Articles, Links
Mechanisms involved in kaempferol-induced relaxation in rat uterine smooth muscle.
Life Sci. 2000 Jun 8;67(3):251-9.
PMID: 10983869 [PubMed - indexed for MEDLINE]

5: [Ciolino HP, Daschner PJ, Yeh GC](#) Related Articles, Links
Dietary flavonols quercetin and kaempferol are ligands of the aryl hydrocarbon receptor that affect CYP1A1 transcription differentially.
Biochem J. 1999 Jun 15;340 (Pt 3):715-22.
PMID: 10359656 [PubMed - indexed for MEDLINE]

6: [Lagarrigue S, Chaumontet C, Heberden C, Martel P, Gaillard-Sanchez I](#) Related Articles, Links
Suppression of oncogene-induced transformation by quercetin and retinoic acid in rat liver epithelial cells.
Cell Mol Biol Res. 1995;41(6):551-60.
PMID: 8551124 [PubMed - indexed for MEDLINE]

Internet

Link to BioAssays where “Active” ...

PubChem BioAssay - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Refresh Print

Address: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pccompound&cmd=Display&dopt=pccompound_pcassay_ Go Links »

Items 1 - 9 of 9 One page. ▾

<p>PubChem Help FAQ</p> <p>PubChem Substance Structures supplied by depositors</p> <p>PubChem Compound Unique structures with computed properties</p> <p>PubChem BioAssay Bioactivity assay results supplied by depositors</p> <p>PubChem</p>	<p><input type="checkbox"/> 1: AID: 396 Compound Screen Assay, Human YWHAB Source: SGCoxCompounds Total substances tested: 1; Active: 1 Links</p> <p><input type="checkbox"/> 2: AID: 395 Compound Screen Assay, Human STK16 Source: SGCoxCompounds Total substances tested: 35; Active: 35 Links</p> <p><input type="checkbox"/> 3: AID: 393 Compound Screen Assay, Human PIM1 Source: SGCoxCompounds Total substances tested: 91; Active: 91 Links</p> <p><input type="checkbox"/> 4: AID: 386 Compound Screen Assay, Human DIRAS Source: SGCoxCompounds Total substances tested: 2; Active: 2 Links</p> <p><input type="checkbox"/> 5: AID: 383 Compound Screen Assay, Human CLK3 Links</p>
--	--

Local intranet

BioAssay where “Active” ...

AID 367 - PubChem BioAssay Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=367&> Go

NCBI PubChem National Library of Medicine NLM

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

Select BioAssay Results

Test Results:

BioAssay Display Option

Group Results By:

Duplicate Test Option:

AID: 367

Name: HIV-2 RNase H Inhibition
Data Source: [MTDP \(RNAI\)](#)
BioAssay Version: 1.2

Other Filters:

Substance Filter

Internet

BioAssay where “Active” ...

AID 367 - PubChem BioAssay Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=367&> Go


AID: 367
Name: HIV-2 RNase H Inhibition
Data Source: [MTDP \(RNAI\)](#)
BioAssay Version: 1.2

Test Results: [Re-Show](#) [Select](#) [Summary](#)

[Results](#) [Download](#)

BioAssay Display Option
Group Results By:
Duplicate Test Option:

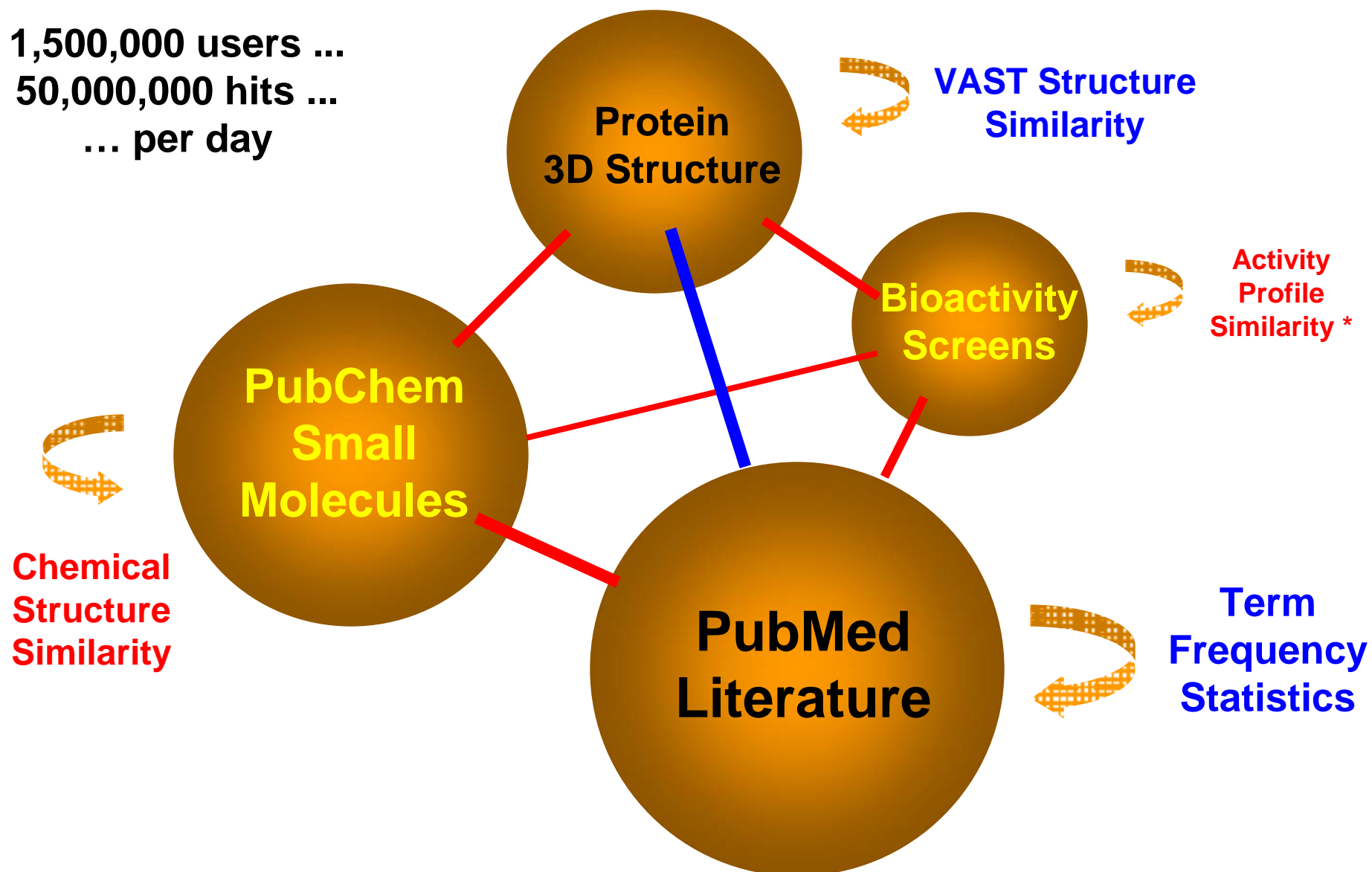
Selected BioAssay Result Count: 1 out of 206 Page: 1 of 1

#	Structure	CID	Score	Outcome	Links	IC-50 uM
1		5280343	26	Active		9

Done Internet

Entrez Links and Neighbors ...

1,500,000 users ...
50,000,000 hits ...
... per day



PubChem Retrieval System ...

... Optimize “discoverability” for molecular biologists by integrating PubChem into NCBI’s Entrez / PubMed Search Engine

... Chemical structure search

... Bioassay result search

... Exploratory structure-activity tools

Quercetin Structure-Activity ...

PubChem Compound - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Select+> Go

NCBI PubChem Compound National Library of Medicine NLM My NCBI [Sign In] [Register]

All Databases PubMed Nucleotide Protein Genome Structure PMC PubChem Books

Search PubChem Compound for #4 #11 Go Clear Save Search

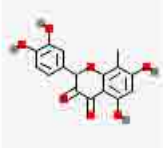
Limits Preview/Index History Clipboard Details

Display PubChem BioActivity Summary Show 20 Sort by Send to


All: 49 BioAssay: 49 Protein3D: 4 Rule of 5: 44

Items 1 - 20 of 49 Page 1 of 3 Next

1: CID: [6419774](#) Related Structures, Assays


CNC-244128, S00072
IUPAC: 2-(3,4-dihydroxyphenyl)-5,7-dihydroxy-8-methyl-chroman-3,4-dione
MW: 316.262 | MF: C16H12O7

2: CID: [6419772](#) Related Structures, Assays


CNC-364524, S00070
IUPAC: 5,7-dihydroxy-2-(3,4,5-trihydroxyphenyl)chroman-3,4-dione
MW: 318.235 | MF: C15H10O8




Local intranet

PubChem Bioactivity Summary ...

PubChem BioAssay Summary - Microsoft Internet Explorer

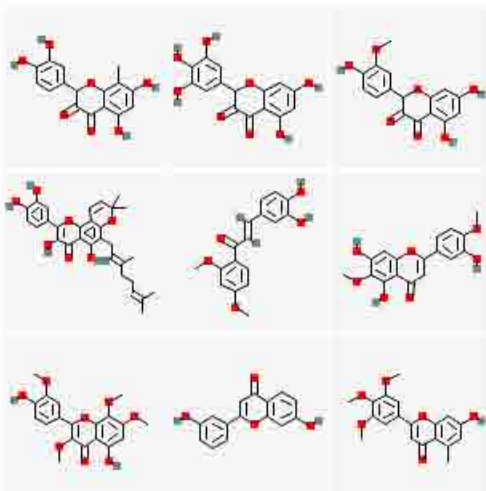
File Edit View Favorites Tools Help





Back Forward Stop Home Links Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?q=cids&q=...> Go

HOME SEARCH SITE MAP PubMed Entrez Structure GenBank PubChem Help

BioActivity Summary:



-  **Total Compound count: 49**
(9 structures shown)
Total tested: 49
Active: 49
Inactive: 43
-  **Structure Activity Analysis**
-  **Structure Clustering**
-  **Selected BioAssays to Entrez**

Total BioAssays: 193

javascript:AssayQueryHandler('heat'); Local intranet

PubChem Bioactivity Summary ...

PubChem BioAssay Summary - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?q=cids&qur> Go

Total BioAssays: 193
Total Pages: : 2

Display: 100 Go To Page 1

#	<input type="checkbox"/> AID	Active	Inactive	Total Tested	Data	Name
1	<input checked="" type="checkbox"/>	410	22	1	23	Data p450-cyp1a2
2	<input checked="" type="checkbox"/>	382	11		11	Data Compound Screen Assay, Human CLK1
3	<input checked="" type="checkbox"/>	505	9		9	Data Compound Screen Assay, Human PIM2
4	<input checked="" type="checkbox"/>	395	8		8	Data Compound Screen Assay, Human STK16
5	<input checked="" type="checkbox"/>	354	7	2	9	Data NCTR Estrogen Receptor Binding Database (NCTER)
6	<input checked="" type="checkbox"/>	383	5		5	Data Compound Screen Assay, Human CLK3
7	<input checked="" type="checkbox"/>	115	5	12	17	Data NCI human tumor cell line growth inhibition assay. Data for the SR Leukemia cell line
8	<input checked="" type="checkbox"/>	384	4		4	Data Compound Screen Assay, Human CSNK1G2
9	<input checked="" type="checkbox"/>	367	4		4	Data HIV-2 RNase H Inhibition
10	<input checked="" type="checkbox"/>	179	4	15	19	Data NCI AIDS Antiviral Assay
11	<input checked="" type="checkbox"/>	125	4	16	20	Data NCI human tumor cell line growth inhibition assay. Data for the HL-60(TB) Leukemia cell line

Local intranet

Active Compound Cluster ...

PubChem Compound - Microsoft Internet Explorer

File Edit View Favorites Tools Help

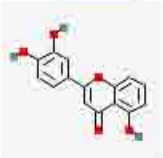
Back Forward Stop Home Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?WebEnvRq=1> Go

About Entrez: Entrez Help


All: 6 BioAssay: 6 Protein3D: 3 Rule of 5: 4

Items 1 - 6 of 6 One page.

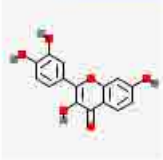
1: CID: [88281](#) [Related Structures, Assays](#)


TNP00056, ZINC00039317 ...
IUPAC: 2-(3,4-dihydroxyphenyl)-5-hydroxy-chromen-4-one
MW: 270.237 | MF: C15H10O5

2: CID: [5280343](#) [Related Structures, Assays, Literature, Other Entrez Databases](#)


quercetin, Sophoretin ...
IUPAC: 2-(3,4-dihydroxyphenyl)-3,5,7-trihydroxy-chromen-4-one
MW: 302.236 | MF: C15H10O7

3: CID: [5281614](#) [Related Structures, Assays, Literature, Other Entrez Databases](#)


Fisetin, Cotinin ...
IUPAC: 2-(3,4-dihydroxyphenyl)-3,7-dihydroxy-chromen-4-one
MW: 286.236 | MF: C15H10O6

4: CID: [5281677](#) [Related Structures, Assays, Literature, Other Entrez Databases](#)

PubChem
Help | FAQ

PubChem
Substance
Structures
supplied by
depositors

PubChem
Compound
Unique structures
with computed
properties

PubChem
BioAssay
Bioactivity assay
results supplied by
depositors

Local intranet

BioAssay Cluster ...

PubChem BioAssay - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?WebEnvRq=1> Go

PubChem
Help | FAQ

PubChem
Substance
Structures
supplied by
depositors

PubChem
Compound
Unique structures
with computed
properties

PubChem
BioAssay
Bioactivity assay
results supplied by
depositors

PubChem
Structure Search

1: AID: [384](#) Links

Compound Screen Assay, Human CSNK1G2
Source: SGCoxCompounds
Total substances tested:29; Active:29

2: AID: [383](#) Links

Compound Screen Assay, Human CLK3
Source: SGCoxCompounds
Total substances tested:55; Active:55

3: AID: [395](#) Links

Compound Screen Assay, Human STK16
Source: SGCoxCompounds
Total substances tested:55; Active:55

4: AID: [505](#) Links

Compound Screen Assay, Human PIM2
Source: SGCoxCompounds
Total substances tested:67; Active:67

5: AID: [382](#) Links

Compound Screen Assay, Human CLK1
Source: SGCoxCompounds
Total substances tested:121; Active:121

<http://pubchem.ncbi.nlm.nih.gov/assay/assay.cgi?aid=505> Local intranet

Another BioAssay Cluster ...

PubChem BioAssay - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Links Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?WebEnvRq=1> Go

About Entrez
Entrez Help

All: 4
Items 1 - 4 of 4 One page.

1: AID: [115](#) Links
NCI human tumor cell line growth inhibition assay. Data for the SR Leukemia cell line
Source: DTP/NCI
Total substances tested:35247; Active:3567

2: AID: [125](#) Links
NCI human tumor cell line growth inhibition assay. Data for the HL-60(TB) Leukemia cell line
Source: DTP/NCI
Total substances tested:38933; Active:3681

3: AID: [121](#) Links
NCI human tumor cell line growth inhibition assay. Data for the K-562 Leukemia cell line
Source: DTP/NCI
Total substances tested:41721; Active:3228

4: AID: [15](#) Links
NCI human tumor cell line growth inhibition assay. Data for the NCI-H522 Non-Small Cell Lung cell line
Source: DTP/NCI
Total substances tested:38588; Active:3026

PubChem
Help | FAQ

PubChem
Substance
Structures
supplied by
depositors

PubChem
Compound
Unique structures
with computed
properties

PubChem
BioAssay
Bioactivity assay
results supplied by
depositors

Local intranet

PubMed Connection ...

Entrez PubMed - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=Displayf> Go

NCBI PubMed A service of the National Library of Medicine and the National Institutes of Health [My NCBI](#) [\[Sign In\]](#) [\[Register\]](#)

All Databases PubMed Nucleotide Protein Genome Structure OMIM PMC Journals Books

Search PubMed for leukemia kinase inhibitor Go Clear [Save Search](#)

Limits Preview/Index History Clipboard Details

Display Summary Show 20 Sort by Send to

All 3124 **Review: 281** ✕

Items 1 - 20 of 281 Page 1 of 15 [Next](#)

1: [Larsen TS, Pallisgaard N, Christensen JH, Gram-Hansen P, Kerndrup GB, Moller MB, Hasselbalch HC.](#) [Related Articles, Links](#)

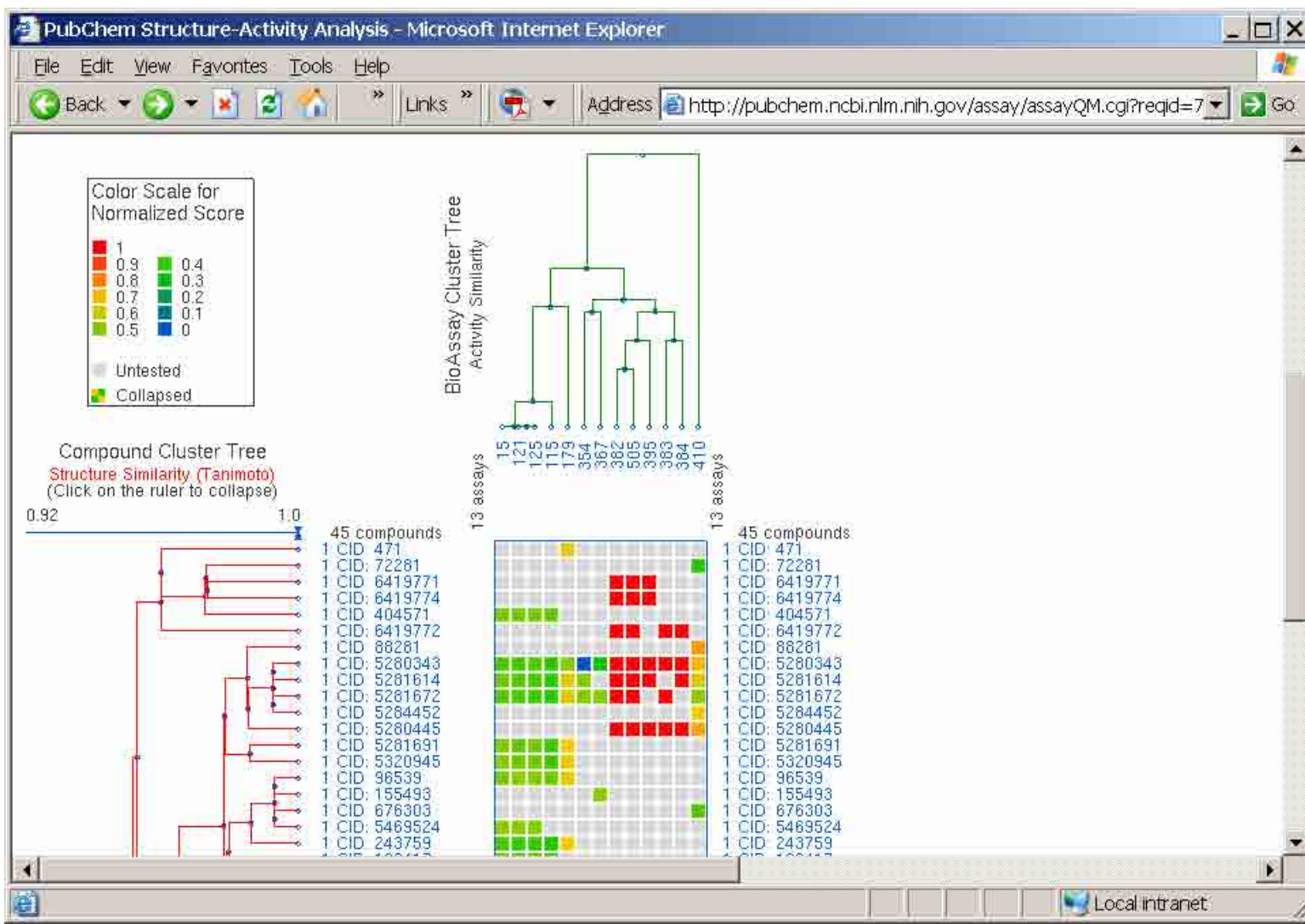
[New molecular markers within the chronic myeloproliferative disorders. II: the JAK2 mutation]
Ugeskr Laeger 2006 Sep 25;168(39):3299-303 Review Danish.
PMID: 17032592 [PubMed - indexed for MEDLINE]

2: [Witzig TE, Kaufmann SH.](#) [Related Articles, Links](#)

Inhibition of the phosphatidylinositol 3-kinase/mammalian target of rapamycin pathway in hematologic malignancies.
Curr Treat Options Oncol. 2006 Jul;7(4):285-94 Review.
PMID: 16916489 [PubMed - indexed for MEDLINE]

Local intranet

PubChem Structure-Activity ...



BioAssay Neighbor Philosophy ...

We have seen, then, that the effect of the *division of labour*, both in mechanical and in mental operations, is, that it enables us to purchase and apply to each process precisely that quantity of skill and knowledge which is required for it: we ... avoid the loss arising from the employment of an accomplished mathematician in performing the lowest processes of arithmetic.

Charles Babbage, *On the Economy of Machinery and Manufactures*, 1835

<http://pubchem.ncbi.nlm.nih.gov>

Evan Bolton

Yulia Borodina

Jie Chen

Svetlana Dracheva

Fabien Fontaine

Lewis Geer

Jane He

Siqian He

Wolf-Dietrich Ihlenfeldt

Karen Karapetian

Vahan Simonyan

Ben Shoemaker

Wenyao Shi

Tugba Suzek

Paul Thiessen

Valery Tkachenko

Jiyao Wang

Yanli Wang

Jewen Xiao

Jian Zhang

