

Finding unusual peptides on the Internet using plain three letter sequence codes.

Dr. Hans-Juergen Himmler, contact@hjhimmler.de, Dr. Alexander Kos, globalsearch@akosgmbh.de, AKos GmbH, Austr. 26, 79585 Steinen, Germany

ECCB 2012 Basel, Tech Track session

CWM Global Search provides a single User Interface that allow to perform a true federated search over all major free accessible chemistry databases and drug discovery data sources on the Internet using

- Chemical structure
- Chemical name
- CAS Registry numbers

Major characteristics of CWM Global Search :

- As easy to use as Google (Single structure box, single text box and one button to start the search)
- Other than Google only science relevant data sources are searched
- Like in Google the result of the search are links
- Unlike Google result links can be sorted, filtered and grouped
- Allows search for multiple independent search terms with one single click
- Queries are automatically extended if possible

Simple search

- ☒ Search all datasources
- ☐ Search profiles

☐ Include isomers
☐ Include Substructures
☐ Include similar compounds
Similarity coefficient (1-100)

90,00

Double click to edit structure

Search for CAS Registry number or synonym name or systematic name or free text.

AVP

Search All Data Sources

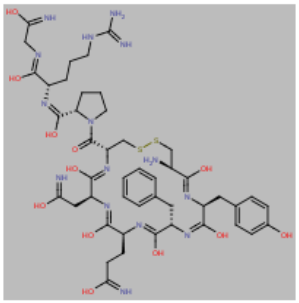
CWM Global Search

Home Simple search Advanced search Results

Results

Structure

Double click to zoom structure



Remove all filter

RECORD_1 27

Drag a column header and drop it here to group by that column

Rz		Datasource	Link	Category	Keywords	Searchtype	Query
1	<input type="checkbox"/>	PubChem	Hit(s) found	Bioactivity		SYNONYMNAME	"Arginine antidiuretic hormone"
1	<input type="checkbox"/>	PubChem	Hit(s) found	Bioactivity		SYNONYMNAME	Rindervasopressin
11	<input type="checkbox"/>	ChEBI	100 ChEBI ID(s) found.	Bioactivity		SYNONYMNAME	"Arginine antidiuretic hormone"
1	<input type="checkbox"/>	PubChem	Hit(s) found	Bioactivity		FULLSTRUCTURE	
63	<input type="checkbox"/>	EBI	Total hit(s) 0	Bioactivity		SYNONYMNAME	"Arginine antidiuretic hormone"
47	<input type="checkbox"/>	ChEMBL	CHEMBL373742	Bioactivity		FULLSTRUCTURE	
60	<input type="checkbox"/>	ChemBank	CHEMBANK ID found	Bioactivity		MOLSIMILARITY	
6	<input type="checkbox"/>	Chemicaland21	2 Hit(s)	Chemistry		SYNONYMNAME	"Arginine antidiuretic hormone"
2	<input type="checkbox"/>	ChemSpider	Hit(s) found	Chemistry		SYNONYMNAME	"Arginine antidiuretic hormone"
2	<input type="checkbox"/>	ChemSpider	Hit(s) found	Chemistry		SYNONYMNAME	Rindervasopressin
2	<input type="checkbox"/>	ChemSpider	Hit(s) found	Chemistry		FULLSTRUCTURE	
2	<input type="checkbox"/>	ChemSpider	Hit(s) found	Chemistry		FULLSTRUCTURE	
37	<input type="checkbox"/>	NextBio	PubChem Substance NextBio link found	Drugs		FULLSTRUCTURE	
10	<input type="checkbox"/>	Quertle	(35) relationships	Literature		SYNONYMNAME	"Arginine antidiuretic hormone"
31	<input type="checkbox"/>	Google Books	Books 1 - 10 of 10	Literature		SYNONYMNAME	"Arginine antidiuretic hormone"
31	<input type="checkbox"/>	Google Books	Books 1 - 10 of 23	Literature		SYNONYMNAME	Rindervasopressin
33	<input type="checkbox"/>	PubMed	Results: 2	Literature		SYNONYMNAME	"Arginine antidiuretic hormone"
41	<input type="checkbox"/>	PubMedCentral	Results: 1 to 20 of 1822	Literature		SYNONYMNAME	"Arginine antidiuretic hormone"
41	<input type="checkbox"/>	PubMedCentral	Results: 4	Literature		SYNONYMNAME	Rindervasopressin

Right click for menu

Home
Simple search
Advanced search
Results

Advanced search

9/7/2012 12:56:22 PM
 9/7/2012 1:00:12 PM
[Read only query page](#)

Double click to edit structure

Structure Search

☐ Include isomers

☐ Include Substructures

☐ Include similar compounds

Similarity coefficient (1-100)

Use the rightmost 'Pin' button to show/hide this pane

Free text

07.09.2012 13:00:12

Free text

Click here to add new item

- ☐ [8-Arginine]vasopressin
- ☐ [Arg8]-Vasopressin
- ☐ "1,2-Dithia-5,8,11,14,17-pentaazacycloicosane-10-propionamide"
- ☐ 3-(Phenylalanine)-8-arginineoxytocin
- ☐ 8-L-Arginine-vasopressin
- ☐ Arg8-vasopressin
- ☒ "Arginine antidiuretic hormone"
- ☐ Arginine-8-vasopressin
- ☐ Arginine-vasopressin
- ☐ "Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-NH2[Disulfide Bridge]"
- ☐ "Glycinamide, L-cysteinyl-L-tyrosyl-L-phenylalanyl-L-glutamate"
- ☐ NCGC00166306-01
- ☐ "Oxytocin, 3-(L-phenylalanine)-8-L-arginine-"
- ☒ Rindervasopressin
- ☐ "roxybenzyl-6,9,12,15,18-pentaoxo- (6CI)"
- ☐ "Vasopressin, 8-L-arginine- (7CI,8CI,9CI)"
- ☐ (2S)-2-[[[(2S)-1-[(4R,7S,10S,13S,16S,19R)-19-amino-13-oxo-9-oxa-8-azabicyclo[5.5.1]undec-5-ene-2-carboxamide]-2-carboxamide]-2-carboxamide]-2-carboxamide]-2-carboxamide]

Use the rightmost 'Pin' button to show/hide this pane

Filter Remove all filter

Filter operations

07.09.2012 13:00:12

Drag a column header and drop it here to group by that column

Rank		Name
40	<input checked="" type="checkbox"/>	ACS Publications
5	<input checked="" type="checkbox"/>	AKOSSAMPLES
40	<input checked="" type="checkbox"/>	Bielefeld Academic Search Engine
42	<input checked="" type="checkbox"/>	BindingDB
25	<input checked="" type="checkbox"/>	BioMed Central
61	<input checked="" type="checkbox"/>	BMRB
15	<input checked="" type="checkbox"/>	Buyersguide
26	<input checked="" type="checkbox"/>	CCRIS
11	<input checked="" type="checkbox"/>	ChEBI
24	<input checked="" type="checkbox"/>	ChemAxon Chemicalize Search
60	<input checked="" type="checkbox"/>	ChemBank
47	<input checked="" type="checkbox"/>	ChEMBL
62	<input checked="" type="checkbox"/>	Chemo
7	<input checked="" type="checkbox"/>	ChemExper
4	<input checked="" type="checkbox"/>	ChemicalBook
18	<input checked="" type="checkbox"/>	CHEMICALDATABASE
6	<input checked="" type="checkbox"/>	Chemicaland21
27	<input checked="" type="checkbox"/>	ChemIDPlus
2	<input checked="" type="checkbox"/>	ChemSpider
21	<input checked="" type="checkbox"/>	ChemSynthesis

Simple search
 Advanced search
 Results

Finding unusual peptides on the Internet using plain three letter sequence codes.

Scenario 1

The Effects of N-Terminal Part Modification of Arginine Vasopressin Analogues with 2-Aminoindane-2-carboxylic Acid: A Highly Potent V₂ Agonist

Wioleta Kowalczyk,^{*,†} Dariusz Sobolewski,[†] Adam Prahl,[†] Izabela Derdowska,[†] Lenka Borovičková,[‡] Jiřina Slaninová,[‡] and Bernard Lammek[†]

Faculty of Chemistry, University of Gdańsk, Sobieskiego 18, 80-952 Gdańsk, Poland, and Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo sq. 2, 16610 Prague, Czech Republic

Received February 15, 2007

In this study we present the synthesis and some pharmacological properties of nine new analogues of arginine vasopressin modified in the N-terminal part of the molecule with 2-aminoindane-2-carboxylic acid (Aic). The peptides were tested for their in vitro uterotonic and in vivo pressor and antidiuretic activities. One of the new peptides, [Mpa¹,Aic²,Val⁴,D-Arg⁸]VP, exhibited an antidiuretic activity similar to that of [Mpa¹,D-Arg⁸]VP, thus being one of the most potent antidiuretic vasopressin analogues reported to date.

Introduction

Arginine vasopressin (AVP^a), a neuropeptide, exerts its action through at least three types of G-protein-coupled receptors with seven transmembrane spanning domains. V₁ vascular receptors

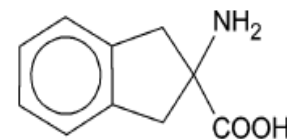


Figure 1. Structure of 2-aminoindane-2-carboxylic acid (Aic).

The task:

How do I find Arginine Vasopressin analogues with unusual amino acid residues in position 2 of AVP that show antidiuretic activity ?

A possible approach :

Search through the ChEBI database, since this database contains biological active peptides with associated assay results

The problem :

In the ChEBI database you can not do BLAST searches such as Cys-X-Phe-Gln-Asn-Cys-Pro-Arg-Gly instead you need to perform a structure search

The solution:

CWM Global Search professional with Proteax option !

CWM Global Search

Home Simple search Advanced search Results

Progress Start Global Search

Modify query New query Override query

Add/Remove custom profile

Clear structure box

Import

Open Save

Load/Save queries

Proteax

Session manager

Help

Advanced search

9/9/2012 11:31:26 A

Double click to edit structure

Structure Search

☐ Include isomers

☐ Include Substructures

☐ Include similar compounds

Similarity coefficient (1-100) 90,00

se the rightmost 'Pin' button to show/hide this pane

CAS number Free text

09.09.2012 11:31:26 09.09.2012 11:31:26

Page 1

CAS number Free text

Click here to add new item Click here to add new item

Use the rightmost 'Pin' button to show/hide this pane

Filter Remove all filter

Filter operations

09.09.2012 11:31:26

Drag a column header and drop it here to group by that column

Rank		Name	Category
40	<input type="checkbox"/>	ACS Publications	Literature
5	<input type="checkbox"/>	AKOSSAMPLES	Suppliers
40	<input type="checkbox"/>	Bielefeld Academic Search Engine	Open Access
42	<input type="checkbox"/>	BindingDB	Proteins
25	<input type="checkbox"/>	BioMed Central	Open Access
61	<input type="checkbox"/>	BMRB	Spectra
15	<input type="checkbox"/>	Buyersguide	Suppliers
26	<input type="checkbox"/>	CCRIS	Toxicity
11	<input checked="" type="checkbox"/>	ChEBI	Bioactivity
24	<input type="checkbox"/>	ChemAxon Chemicalize Search	Search engine
60	<input type="checkbox"/>	ChemBank	Bioactivity
47	<input type="checkbox"/>	ChEMBL	Bioactivity
62	<input type="checkbox"/>	Chemo	Chemical prop
7	<input type="checkbox"/>	ChemExper	Suppliers
3	<input type="checkbox"/>	Chemical Structure Lookup Service	Search engine
4	<input type="checkbox"/>	ChemicalBook	Suppliers

Simple search

Advanced search

Results

Home

Simple search

Advanced

Progress

Start Global Search

Modify

9/9/2012 11:31:26 A

Str

Sim

Clean

3Letter

Configure

Turn On Off

Format 1 3

Global Search

PLN

Hydrogens

Editor

H

Cys(1)

X

Phe

Gln

Asn

Cys(1)

Pro

Arg

Gly

[NH2]

Protein text - PLN format

H-Cys (1)-X-Phe-Gln-Asn-Cys (1)-Pro-Arg-Gly-[NH2]

N-terminal

C-terminal

Ala

Arg

Asn

Asp

Cys

Glu

Gln

Gly

His

Ile

Leu

Lys

Met

Phe

Pro

Ser

Thr

Trp

Tyr

Val

Sec

Pyl

Xaa

Sequence

Molecule

Sum formula: C38 H57 N15 O11 S2 X Avg. MW: 964,08308

1:1

S

S

H-Cys-Xaa-Phe-Gln-Asn-Cys-Pro-Arg-NH2

Simple search

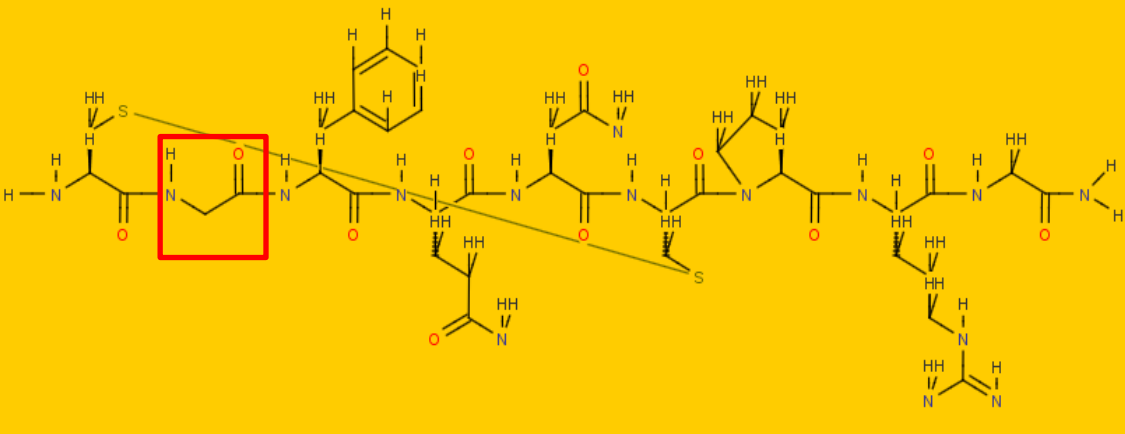
Advanced search

Results

MarvinSketch

File Edit View Insert Atom Bond Structure Tools Help

100%



Home Progress Advance

Category

- Literature
- Suppliers
- Open Access
- Proteins
- Open Access

61 ☐ BMRB Spectra

15 ☐ Buyersguide Suppliers

26 ☐ CCRIS Toxicity

11 ☒ ChEBI Bioactivity

24 ☐ ChemAxon Chemicalize Search Search engine

60 ☐ ChemBank Bioactivity

47 ☐ ChEMBL Bioactivity

62 ☐ Chemo Chemical prop

7 ☐ ChemExper Suppliers

3 ☐ Chemical Structure Lookup Service Search engine

4 ☐ ChemicalBook Suppliers

Include isomers

☒ Include Substructures

☐ Include similar compounds

Similarity coefficient (1-100) 90,00

Simple search

Advanced search

Results

The screenshot displays the CWM Global Search application interface. The top navigation bar includes tabs for 'Home', 'Simple search', 'Advanced search', and 'Results'. Below this, a toolbar contains icons for 'Export', 'Open selected links in Browser', 'Cancel search', 'Session manager', 'Proteax', and 'Help'. The main workspace is divided into three primary sections. On the left, the 'Results' panel shows a search history entry for '9/9/2012 11:31:26 A'. The central 'Structure' panel features a chemical structure diagram of a complex organic molecule, with a prompt 'Double click to zoom structure'. To the right, the 'Globalsearch links' panel contains a 'Remove all filter' button, a date/time filter set to '09.09.2012 11:31:26', and a table with columns 'Extreg' and 'Number of hits'. The table contains one row: 'RECORD_1' with '1' hit. Below the table is a section for grouping by column, with a prompt 'Drag a column header and drop it here to group by that column'. A row of column headers is visible: 'Rank', 'Link', 'D:', 'Link', 'Categ', 'Keywords', 'Searchtype', 'Query', 'Proteax', and 'Errormessage'. The 'Keywords' column is highlighted with a red box, and its value 'SUBSTRUCTURE' is visible. Below the column headers is a table with one row: '11', a checkbox, 'ChEBI', '8 ChEBI ID(s) found.', 'Bioactivity', and 'SUBSTRUCTURE'. The 'SUBSTRUCTURE' cell is also highlighted with a red box. At the bottom of the interface, there are buttons for 'Simple search', 'Advanced search', and 'Results'.

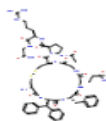
Another problem :

Peptides in ChEBI are not represented in 3 or 1 letter sequence codes and therefore make the interpretation of the results rather difficult

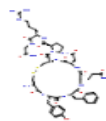
The results in ChEBI :

8 entries found, displaying 1 to 8.

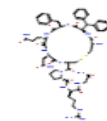
<< < 1 > >>



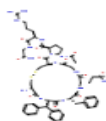
[CHEBI:655167](#) (ChEMBL) (S)-N-((R)-1-(2-amino-2-oxoethylamino)-5-guanidino-1-oxopentan-2-yl)-1-((4R,7S,10S,13S,16R,19R)-19-amino-7-(2-amino-2-oxoethyl)-10-(3-amino-3-oxopropyl)-16-benzhydryl-13-benzyl-6,9,12,15,18-pentaoxo-1,2-dithia-5,8,11,14,17-pentaazacycloicosane-4-carbonyl)pyrrolidine-2-carboxamide



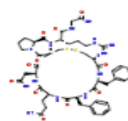
[CHEBI:655164](#) (ChEMBL) (S)-N-((R)-1-(2-amino-2-oxoethylamino)-5-guanidino-1-oxopentan-2-yl)-1-((4R,7S,10S,13S,16S,19R)-19-amino-7-(2-amino-2-oxoethyl)-10-(3-amino-3-oxopropyl)-13-benzyl-16-(4-hydroxybenzyl)-6,9,12,15,18-pentaoxo-1,2-dithia-5,8,11,14,17-pentaazacycloicosane-4-carbonyl)pyrrolidine-2-carboxamide



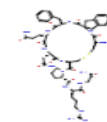
[CHEBI:658626](#) (ChEMBL) (S)-N-((R)-1-(2-amino-2-oxoethylamino)-5-guanidino-1-oxopentan-2-yl)-1-((4R,7S,10S,13S,16S,19R)-19-amino-7-(2-amino-2-oxoethyl)-10-(3-amino-3-oxopropyl)-16-benzhydryl-13-benzyl-6,9,12,15,18-pentaoxo-1,2-dithia-5,8,11,14,17-pentaazacycloicosane-4-carbonyl)pyrrolidine-2-carboxamide



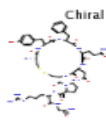
[CHEBI:656873](#) (ChEMBL) (S)-N-((S)-1-(2-amino-2-oxoethylamino)-5-guanidino-1-oxopentan-2-yl)-1-((4R,7S,10S,13S,16R,19R)-19-amino-7-(2-amino-2-oxoethyl)-10-(3-amino-3-oxopropyl)-16-benzhydryl-13-benzyl-6,9,12,15,18-pentaoxo-1,2-dithia-5,8,11,14,17-pentaazacycloicosane-4-carbonyl)pyrrolidine-2-carboxamide



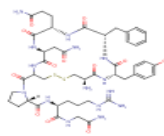
[CHEBI:1271649](#) (ChEMBL) (S)-N-((S)-1-(2-amino-2-oxoethylamino)-5-guanidino-1-oxopentan-2-yl)-1-((4R,7S,10S,13S,16S,19R)-19-amino-7-(2-amino-2-oxoethyl)-10-(3-amino-3-oxopropyl)-13,16-dibenzyl-6,9,12,15,18-pentaoxo-1,2-dithia-5,8,11,14,17-pentaazacycloicosane-4-carbonyl)pyrrolidine-2-carboxamide



[CHEBI:658624](#) (ChEMBL) (S)-N-((S)-1-(2-amino-2-oxoethylamino)-5-guanidino-1-oxopentan-2-yl)-1-((4R,7S,10S,13S,19R)-19-amino-7-(2-amino-2-oxoethyl)-10-(3-amino-3-oxopropyl)-13-benzyl-6,9,12,15,18-pentaoxo-1',3'-dihydro-1,2-dithia-5,8,11,14,17-pentaazaspiro[[1,2]cycloicosane-16,2'-indene]-4-ylcarbonyl)pyrrolidine-2-carboxamide



[CHEBI:466860](#) (ChEMBL) CHEBI:466860



[CHEBI:34543](#) argipressin

Stars: ★★★★★

The solution :

The structure -> 1(3) letter sequence code converter in Proteax

CWM Global Search

Home Simple search Advanced search Results

Export Open selected links in Browser Cancel search Session manager Proteax Help

Results 9/9/2012 11:31:26 A

Structure Double click to zoom structure

Chemical structure image

Globalssearch links

Remove all filter

09.09.2012 11:31:26

Extreg	Number of hits
RECORD_1	1

Page 1 of 1

Drag a column header and drop it here to group by that column

Rank	Link	Categ	Keywords	Searchtype	Query	Proteax	Errormessage
11	ChEBI 8 ChEBI ID(s) found.	Bioactivity		SUBSTRUCTURE		Proteax	

The same ChEBI result in CWM Global Search with integrated Proteax

Drag a column header and drop it here to group by that column

	ID	Link	PLN	Dernot_Diff	
<input type="radio"/>	CHEBI:655167	CHEBI:655167	H-C(1){d}[Dip]FQNC(1)P{d}RG-[NH2] name=CHEBI:655167		
<input type="radio"/>	CHEBI:656873	CHEBI:656873	H-C(1){d}[Dip]FQNC(1)PRG-[NH2] name=CHEBI:656873		
<input type="radio"/>	CHEBI:1271649	CHEBI:1271649	H-C(1)FFQNC(1)PRG-[NH2] name=CHEBI:1271649		
<input type="radio"/>	CHEBI:658624	CHEBI:658624	H-C(1)[Aic]FQNC(1)PRG-[NH2] name=CHEBI:658624		
<input type="radio"/>	CHEBI:655164	CHEBI:655164	H-C(1)YFQNC(1)P{d}RG-[NH2] name=CHEBI:655164		

Compare Usual peptides Unusual peptides Show all

Show only ,unusual' peptides and open ChEBI record for specific peptide

Drag a column header and drop it here to group by that column

	ID	Link	PLN	Demot_Diff	
>	CHEBI:658626	CHEBI:658626	H-C(1)[Dip]FQNC(1)P{d}RG-[NH2] name=CHEBI:658626		
	CHEBI:655167	CHEBI:655167	H-C(1){d}[Dip]FQNC(1)P{d}RG-[NH2] name=CHEBI:655167		
	CHEBI:656873	CHEBI:656873	H-C(1){d}[Dip]FQNC(1)PRG-[NH2] name=CHEBI:656873		
	CHEBI:658624	CHEBI:658624	H-C(1)[Aic]FQNC(1)PRG-[NH2] name=CHEBI:658624		

Compare

Usual peptides

Unusual peptides

Show all

Page 1 of 1

Errormessage

teax

ChEMBL

ChEMBLdb

Malaria Data

ChEMBL-NTD

Kinase SARfari

GPCR SARfari

DrugEBility

ChEMBL Group

Downloads

Web Services

FAQ

ChEMBLdb Statistics

- DB: ChEMBL_14
- Targets: 9,003
- Compound records: 1,376,469
- Distinct compounds: 1,213,239
- Activities: 10,129,256
- Publications: 46,133

ChEMBL Blog

- [New Drug Approvals 2012 - Pt. XVII - Linacotide \(LinzessTM\)](#)
- [Antibody Drugs To Have Reached Clinical](#)

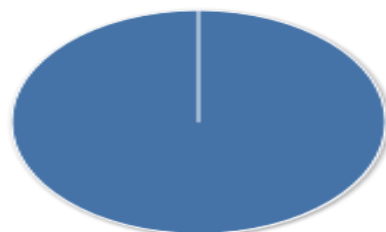
EBI > Databases > Small Molecules > ChEMBL Database

Document Report Card

Doc ID	CHEMBL1152921
Journal	Eur. J. Med. Chem. (2009) 44:2862-2867
Title	Arginine vasopressin and its analogues--the influence of position 2 modification with 3,3-diphenylalanine enantiomers. Highly potent V2 agonists.
Authors	Kwiatkowska A, Sobolewski D, Prahl A, Borovickova L, Slaninova J, Lammek B
Abstract	Eleven new analogues of arginine vasopressin (AVP) modified in position 2 by 3,3-diphenyl-L-alanine or its D-enantiomer (Dip or D-Dip) were synthesized and pharmacologically evaluated for their pressor, antidiuretic and in vitro uterotonic activities. Both the Dip and D-Dip modifications at position 2 of AVP are sufficient to completely change the pharmacological profile of the peptides. They preserve or increase antidiuretic activity, cause its prolongation, transform uterotonic property in antagonistic one and cancel the effect on blood pressure. Four of the new peptides ([Mpa1,D-Dip2]AVP, [Mpa1,D-Dip2,Val4]AVP, [Mpa1,D-Dip2,D-Arg8]VP, [Mpa1,D-Dip2,Val4,D-Arg8]VP) are exceptionally potent antidiuretic agents with significantly prolonged activities.
CiteXplore	19418631
DOI	

Protein Target Summary

ChEMBL Protein Target Classes for Doc CHEMBL1152921



■ Membrane receptor (2)

Total: 2



Assay Summary

Finding unusual peptides on the Internet using plain three letter sequence codes.

Scenario 2

The task:

How do I find specifically NON CYCLIC nona peptides with two cysteine residues in position 1 and 6 in the ChEBI database and in PubChem ?

Clean

3Letter

Configure

Turn On Off

Format 1 3

Global Search

PLN

Hydrogens

Editor

H

Cys

X

X

X

X

Cys

X

X

X

NH2

Protein text - PLN format

H-Cys-X-X-X-X-Cys-X-X-X-[NH2]

N-terminal

C-terminal

Ala

Arg

Asn

Asp

Cys

Glu

Gln

Gly

His

Ile

Leu

Lys

Met

Phe

Pro

Ser

Thr

Trp

Tyr

Val

Sec

Pyl

Xaa

Sequence

Molecule

Sum formula: C13 H20 N10 O9 S2 X7 Avg. MW: 524,4895

+

-

1:1

MarvinSketch

File Edit View Insert Atom Bond Structure Tools Help

100%

Proteas

Session manager

Help

10001101

10001101

10001101

10001101

10001101

Use the rightmost 'Pin' button to show/hide this pane

Remove all filter

operations

2:21:51

header and drop it here to group by that column

	Name	Category	S
<input checked="" type="checkbox"/>	ChEBI	Bioactivity	
<input checked="" type="checkbox"/>	PubChem	Bioactivity	

☐ Include isomers
☒ Include Substructures
☐ Include similar compounds
Similarity coefficient (1-100) 90,00

Simple search

Advanced search

Results

CWM Global Search

Home Simple search Advanced search **Results**

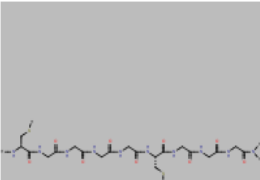
Export Open selected links in Browser Cancel search Session manager Proteax Help

Results

9/9/2012 11:31:26 AM
9/9/2012 12:21:51 PM

Structure

Double click to zoom structure



Globalsearch links

Remove all filter

09.09.2012 12:21:51

Extreg	Number of hits
RECORD_1	2

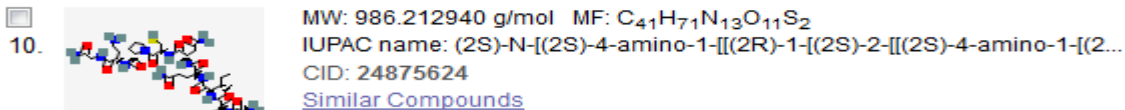
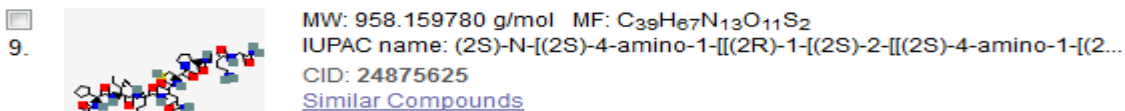
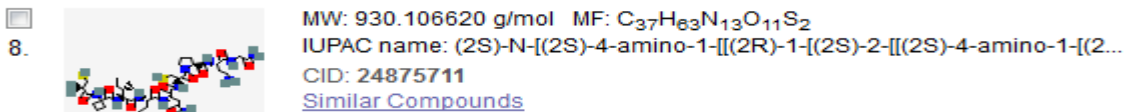
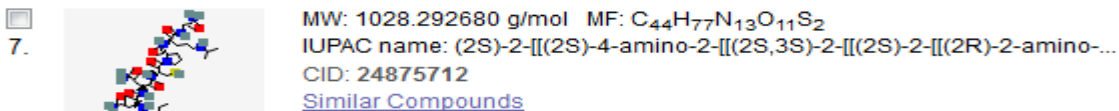
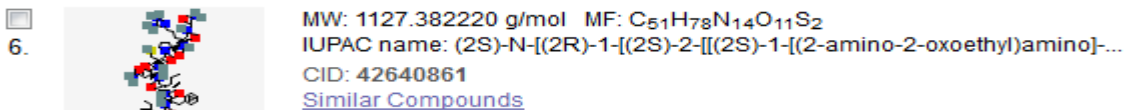
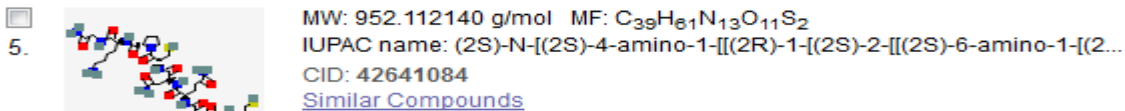
Page 1

Drag a column header and drop it here to group by that column

Rank	Datas	Link	Categ	Keywords	Searchtype	Query	Protea	Errormess
1	<input type="checkbox"/> PubChem	100 or more PubChem CIDs found.	Bioactivity		SUBSTRUCTURE		Protea	
11	<input type="checkbox"/> ChEBI	12 ChEBI ID(s) found.	Bioactivity		SUBSTRUCTURE		Protea	

Simple search
Advanced search
Results

The result in the PubChem database



The result in CWM Global Search

Drag a column header and drop it here to group by that column					
	ID	Link	PLN	Demot_Diff	
<input type="radio"/>	42640861	42640861	H-CIFINCPRG-[NH2] name=42640861		
<input type="radio"/>	42641084	42641084	H-CGFQNCPKG-[NH2] name=42641084		
<input type="radio"/>	44277798	44277798	H-FCSDYSCYLD-[NH2] name=44277798		
<input type="radio"/>	44308796	44308796	H-CYI{d}[Glu[N(CH3)2]]NC{d}PLG-[NH2] name=44308796		
<input type="radio"/>	44308989	44308989	H-CYI{d}[Glu[N(C3H7)2]]NC{d}PLG-[NH2] name=44308989		

Scenario 3 :

I have a PubChem hit list containing both small molecules and peptides and I only want to look at peptides and compare them !

CWM Global Search

Home Simple search Advanced search Results

Progress Start Global Search New search

Modify query New query Override query Add/Remove custom profile Clear structure box Import Load/Save queries Proteas Session manager Help

Advanced search

- 9/9/2012 11:31:26 AM
[Read only query page](#)
- 9/9/2012 12:21:51 PM
[Read only query page](#)
- 9/9/2012 12:57:37 PM
[Read only query page](#)
- 9/9/2012 1:01:33 PM

Double click to edit structure

Structure Search

- ☐ Include isomers
- ☐ Include Substructures
- ☐ Include similar compounds

Similarity coefficient (1-100) 90,00

Use the rightmost 'Pin' button to show/hide this pane

Free text

09.09.2012 12:57:37

Page 1 of 1

Click here to add new item

<input checked="" type="checkbox"/>	"vasoconstrictor agents"
<input checked="" type="checkbox"/>	vasoconstrictor agents

Use the rightmost 'Pin' button to show/hide this pane

Filter Remove all filter Filter operations

09.09.2012 12:57:37

Drag a column header and drop it here to group by that column

Rank	Name	Category
44	IBRIDGE	Innovations
57	IPCS	Safety
39	KEGG	Drugs
8	MOLPORT	Suppliers
20	MPBIO	Supplier
55	Nature chemical biology	Literature
23	NCIDATABASEROWSER	Chemistry
37	NextBio	Drugs
34	NIAID	Drugs
29	NISTWebbook	Spectra
53	PharmGKB	Drugs
32	PLOSONe	Open Access
58	PNAS	Literature
43	Protein Data Bank	Proteins
1	<input checked="" type="checkbox"/> PubChem	Bioactivity
33	PubMed	Literature
41	PubMedCentral	Literature

Simple search

Advanced search

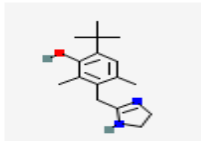
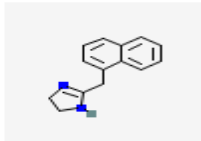
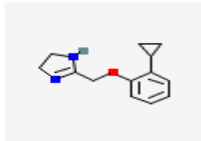
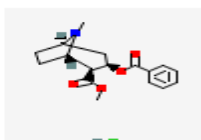
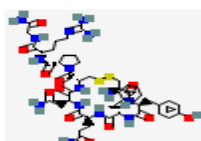
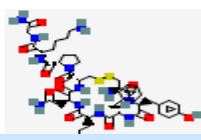
Results

Right click for menu

The result list in PubChem

Results: 21 to 40 of 166

<< First < Prev Page 2 of 9 Next > Last >>

21.  [oxymetazoline; Oxymethazoline; Oxyazine ...](#)
 MW: 260.374560 g/mol MF: C₁₆H₂₄N₂O
 IUPAC name: 6-tert-butyl-3-(4,5-dihydro-1H-imidazol-2-ylmethyl)-2,4-dime...
[Active in 44 BioAssays](#) [Tested in 277 BioAssays](#)
 CID: 4636
[Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [BioAssays, activity ≤ 1 μM](#) [PubMed \(MeSH Keyword\)](#)
22.  [naphazoline; Naphthizine; Clearine ...](#)
 MW: 210.274360 g/mol MF: C₁₄H₁₄N₂
 IUPAC name: 2-(naphthalen-1-ylmethyl)-4,5-dihydro-1H-imidazole
[Active in 8 BioAssays](#) [Tested in 144 BioAssays](#)
 CID: 4436
[Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [BioAssays, activity ≤ 1 μM](#) [PubMed \(MeSH Keyword\)](#)
23.  [cirazoline; Cirazolinum \[INN-Latin\]; Cirazolina \[INN-Spanish\] ...](#)
 MW: 216.278940 g/mol MF: C₁₃H₁₆N₂O
 IUPAC name: 2-[(2-cyclopropylphenoxy)methyl]-4,5-dihydro-1H-imidazole
[Active in 11 BioAssays](#) [Tested in 235 BioAssays](#)
 CID: 2765
[Similar Compounds](#) [Mixture/Component Compounds](#) [BioAssays, activity ≤ 1 μM](#) [PubMed \(MeSH Keyword\)](#)
24.  [Cocaine muriate; Cocaine chloride; Sal de merck ...](#)
 MW: 339.813880 g/mol MF: C₁₇H₂₂ClNO₄
 IUPAC name: methyl (1R,3R,4S,5S)-3-benzoyloxy-8-methyl-8-azabicyclo[3.2....
 CID: 644270
[Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [PubMed \(MeSH Keyword\)](#)
25.  [ARGIPRESSIN; Arginine vasopressin; \[Arg8\]-Vasopressin ...](#)
 MW: 1084.231600 g/mol MF: C₄₆H₆₅N₁₅O₁₂S₂
 IUPAC name: (2S)-1-[(4R,7S,10S,13S,16S,19R)-19-amino-7-(2-amino-2-oxoeth...
[Active in 2 BioAssays](#) [Tested in 44 BioAssays](#)
 CID: 644077
[Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [BioAssays, activity ≤ 1 μM](#) [PubMed \(MeSH Keyword\)](#)
26.  [Lysipressin; Lysopressin; Syntopressin ...](#)
 MW: 1056.218200 g/mol MF: C₄₆H₆₅N₁₃O₁₂S₂
 IUPAC name: (2S)-N-[(2S)-6-amino-1-[(2-amino-2-oxoethyl)amino]-1-oxohexa...
 CID: 644076
[Similar Compounds](#) [Same Parent, Connectivity](#) [Mixture/Component Compounds](#) [PubMed \(MeSH Keyword\)](#)

The filtered list only containing peptides in CWM Global Search

Drag a column header and drop it here to group by that column				
	ID	Link	PLN	Demot_Diff
<input type="radio"/>	53477739	53477739	H-C(1)YFQNC(1)PKG-[NH2] name=53477739	
<input type="radio"/>	53395954	53395954	H-C(1)YFQNC(1)PRG-[NH2] name=53395954	
<input checked="" type="radio"/>	44428105	44428105	H-C(1)YFQNC(1)PRG-[NH2] name=44428105	
<input type="radio"/>	44285245	44285245	H-{d}C(1){d}YF{d}QNC(1){d}PRG-[NH2] name=44285245	
<input type="radio"/>	23725099	23725099	H-GGGC(1)YFQNC(1)PKG-[NH2] name=23725099	
<input type="radio"/>	16132429	16132429	H-C(1)TC(2)NDMTDEEC(2)LNFC(1)HQDVIW-OH name=16132429	
<input type="radio"/>	11979316	11979316	H-C(1)YFQNC(1)PRG-[NH2].H-C(2)YFQNC(2)PKG-[NH2] name=11979316	

The filtered list compared to AVP

Drag a column header and drop it here to group by that column

PLN	Dernot_Diff	
H-C(1)YFQNC(1)P[Orn]G-[NH2] name=533960	[Orn](8) 44428105	
H-C(1)YFQNC(1)PKG-[NH2] name=53477739	K(8) 44428105	
H-C(1)YFQNC(1)PRG-[NH2] name=53395954	44428105	
H-C(1)YFQNC(1)PRG-[NH2] name=44428105	44428105	
H-{d}C(1){d}YF{d}QNC(1){d}PRG-[NH2] name {d}C{d}Y(1-2){d}Q(4){d}P(7) 44428105	{d}C{d}Y(1-2){d}Q(4){d}P(7) 44428105	
H-GGGC(1)YFQNC(1)PKG-[NH2] name=237250	-GGG- K(8) 44428105	
H-C(1)TC(2)NDMTDEEC(2)LNFC(1)HQDVIW-OH	endo-DMTDEE(5) des-(4) TC(2-3)LNFC(7-9) 44428105 -CHQDVIW-OH	

Compare Usual peptides Unusual peptides Show all

Summary :

- CWM Global Search provides an user interface that is as easy to use as Google, but only searches science relevant data sources on the Internet.
- CWM Global Search extend your original query, thus automatically broaden your search.
- Resulting links can, other than Google, be sorted/filtered and grouped, thus making the interpretation of the search results easier for less experienced users.
- Try the FREE version of CWM Global Search at <http://cwmglobalsearch.com/gsweb>

Summary :

- CWM Global Search professional with Proteax option allows to perform ,BLAST like' searches in structure centric databases such as ChEBI or PubChem.
- You can use ,ANY' residues or specific structure fragments to search for unusual peptides.
- ChEBI/PubChem hit lists are ,mapped back' to three (one) letter sequence codes, thus making the interpretation of results much easier compared with the original database.
- You can filter an arbitrary list of structures to only show peptides and proteins.
- You can easily compare peptides and proteins.

Where can I find more information about CWM Global Search and Proteax ?

- Visit the CWM Global Search home page at <http://www.akosgmbh.de/globalsearch>
- Visit the Proteax home page at <http://www.biochemfusion.com>
- Contact us via email globalsearch@akosgmbh.de