



Consiglio Nazionale delle Ricerche



# Bioinformatics resources to study PPIs

Allegra Via

This presentation is adapted from the materials kindly provided by **Luana Licata**, MINT, University of Rome Tor Vergata

<https://goo.gl/forms/wt9aPbQodTUZe31F3>

Shortened URL: bit.ly/2IGhVKx

# Protein-Protein Interactions (PPIs)

- PPIs are physical contacts due to biochemical events and/or electrostatic forces between two or more proteins.
- PPIs can occur only in a specific biological context and between specific molecular regions
- Proteins rarely act alone
- Most molecular processes within a cell are carried out by molecular machines made up of a large number of interacting proteins (complexes)

# Protein-Protein Interactions (PPIs)

- **INTERACTOME:** A protein-protein interaction network
- An interactome provides information about the organization of a cell
- Interactome models can be used to predict poorly characterised proteins into their functional context
- If the function of one protein is known, the function of its binding partners is likely to be related ("guilty by association")

# Methods to detect protein-protein interactions

A variety of techniques and methods have been developed to generate PPI data.

- **High throughput techniques**
- **Low throughput techniques**

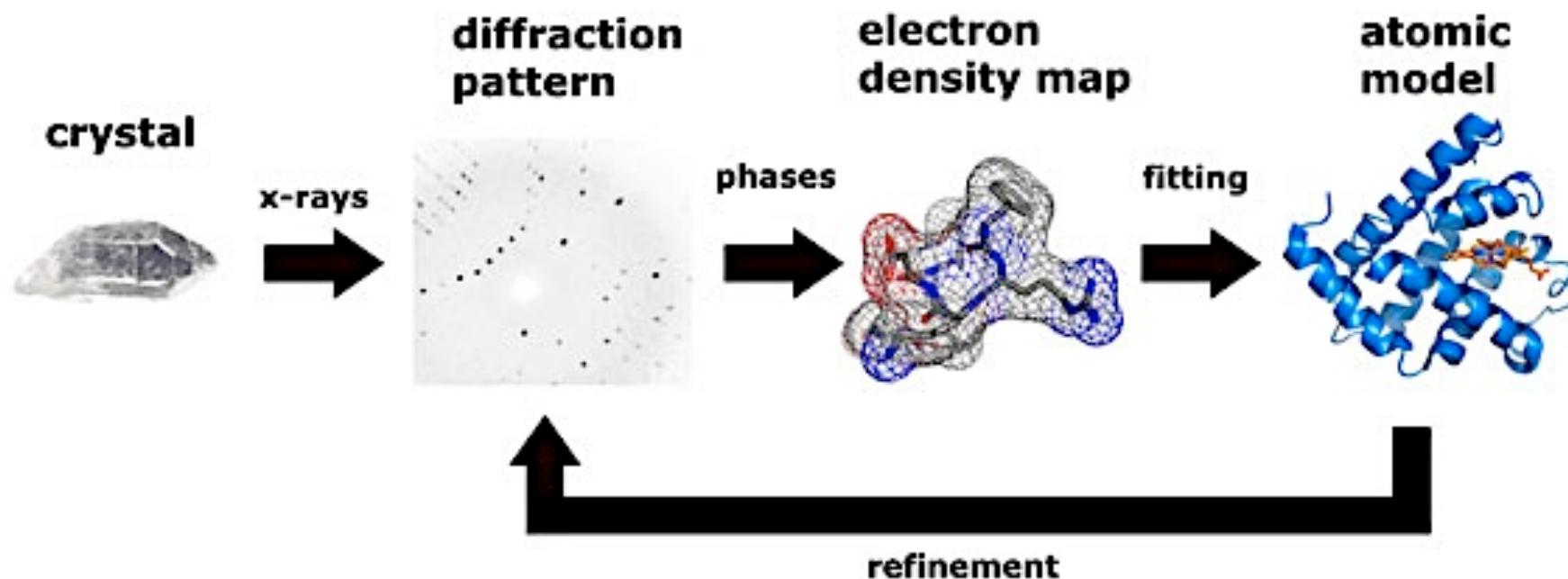
These techniques can be further divided in:

- **Binary methods**: techniques detecting **direct physical associations** between two proteins
- **Co-complex methods**: techniques detecting interactions among groups of proteins that may not be in physical contact

# Methods to detect protein-protein interactions

Some low throughput techniques, such as FRET, SPR, ITC, NMR and X-ray crystallography, provide deeper insight certain characteristic of an interaction.

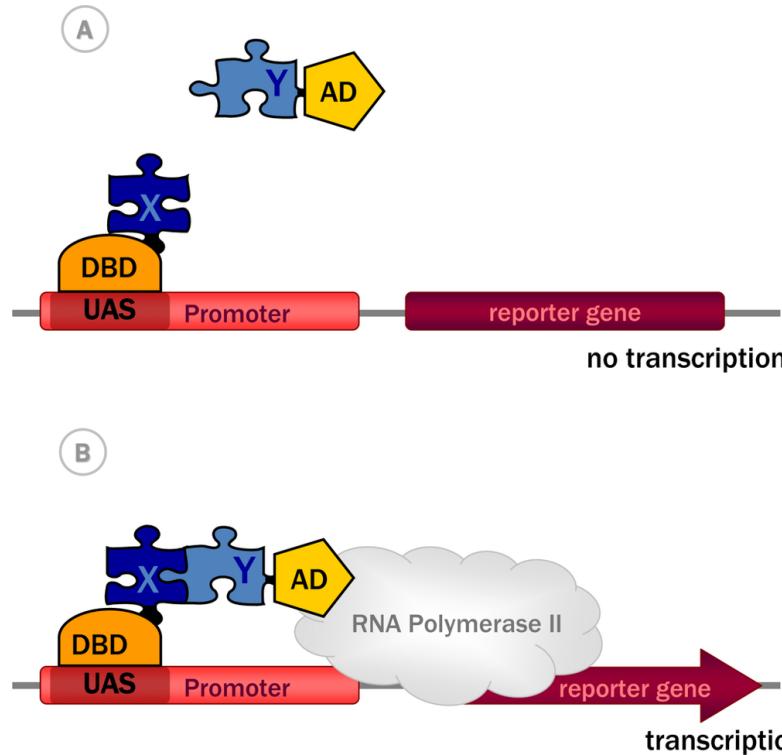
X-ray crystallography is considered the **gold standard** for PPI, since it provides high quality data for binding surfaces to the level of individual atoms and binding sites.



# Methods to detect protein-protein interactions

The main binary methods to detect direct physical interactions between protein pairs is

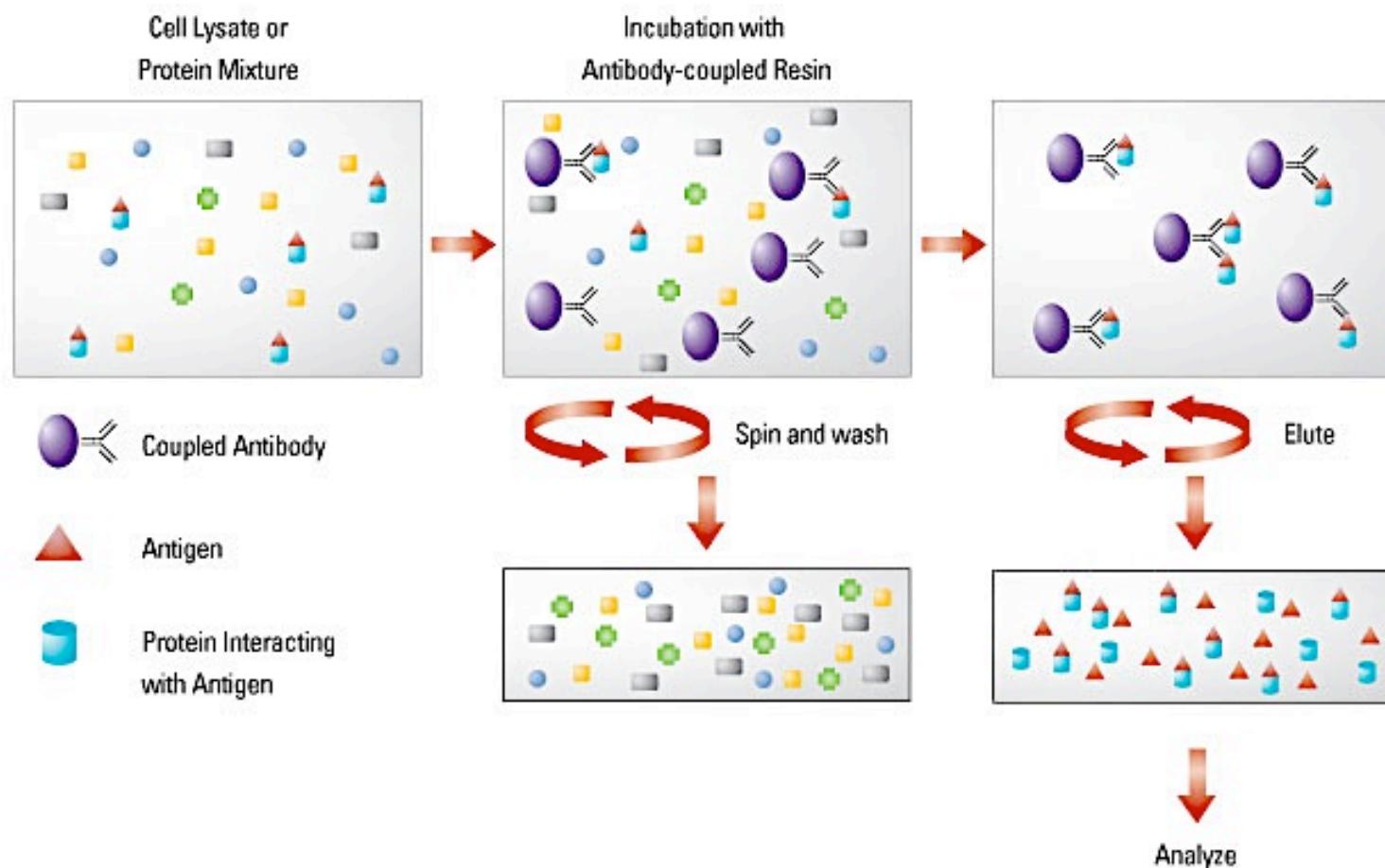
**Yeast two-hybrid (Y2H)** - The strategy interrogates two proteins, called **bait** and **prey**, coupled to two halves of a transcription factor and expressed in yeast. If the proteins make a contact, they reconstitute the transcription factor that activates a reporter gene.



From Yeast Two-Hybrid, a Powerful Tool for Systems Biology, Anna Brückner , Cécile Polge, Nicolas Lentze, Daniel Auerbach and Uwe Schlattner, *Int. J. Mol. Sci.* **2009**

# Methods to detect protein-protein interactions

The most common co-complex method is **co-immunoprecipitation (co-IP) coupled with mass spectrometry (MS)**. In this approach, the bait protein, usually expressed in the cell *in vivo* conditions, is affinity purified and the interacting partners are detected by mass spectrometry.

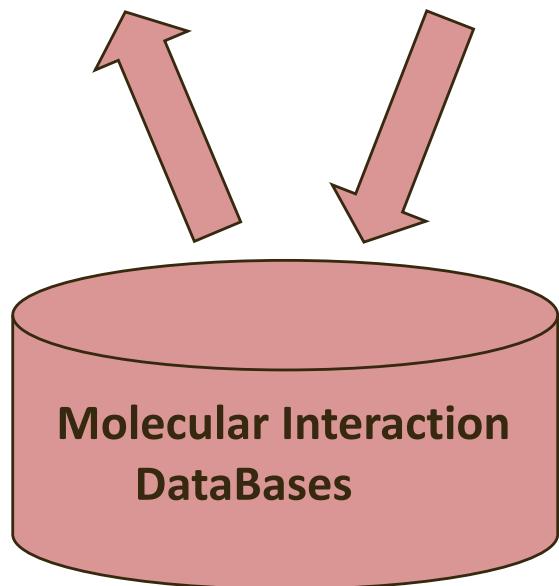


# Databases and curation

Wet Lab Scientists



Scientific Curators



# Molecular Interaction Databases

**Molecular interaction databases** have been established to **collect, archive** and **disseminate** molecular interaction data in a **structured format** to perform searches and bioinformatics analyses

# Molecular Interaction Databases

Molecular interaction databases can be divided in:

**Primary databases:** experimentally determined protein interactions coming from either small- or large-scale published studies that have been manually curated

**Meta databases:** experimentally determined PPIs obtained by consistent integration of several primary databases

**Prediction databases:** mainly predicted PPIs derived using different approaches, combined with experimentally determined PPIs.

See <http://www.pathguide.org>

## Molecular Interaction Databases

---

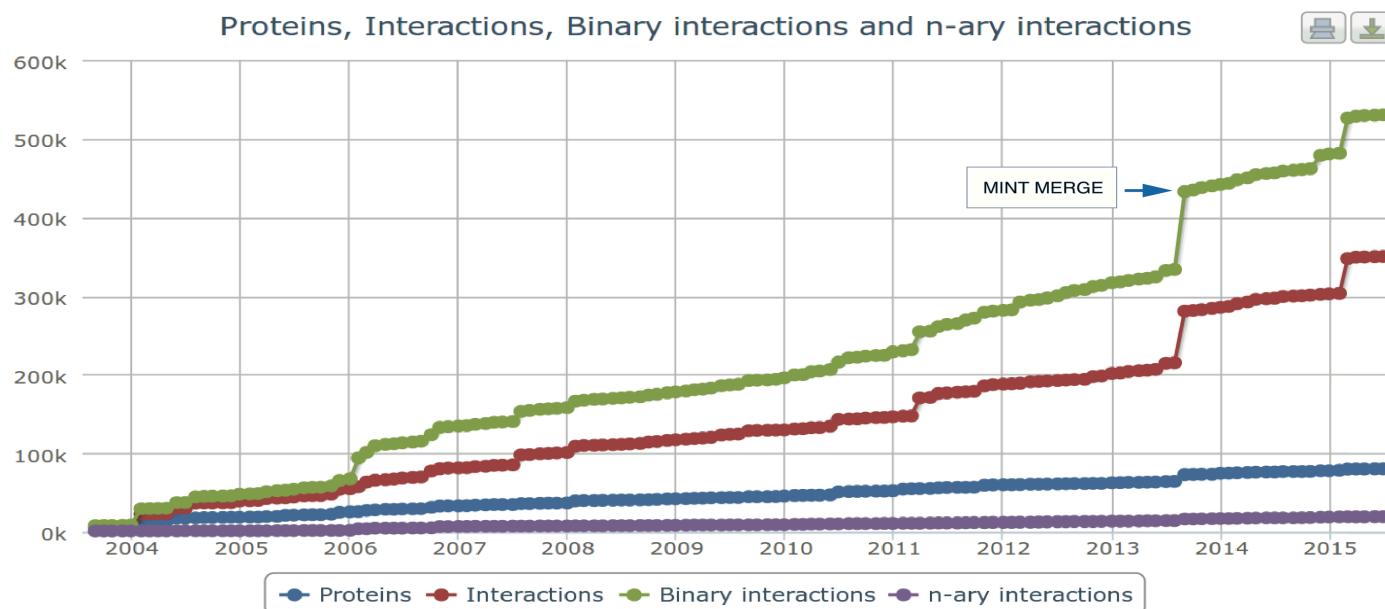
- In the early 2000s, DIP and BIND were the first protein-protein interaction (PPI) repositories to contain freely available, manually curated interaction data.
- MINT and IntAct databases are two of the largest databases (number of manuscripts curated and the number of non-redundant interactions).
- Both adopted the highest possible data quality standards.

## MintAct Project

IntAct and MINT were founder members of the IMEx Consortium

IntAct and MINT joined forces to create a single resource to improve curation and software development efforts.

Data maintenance, and the PSICQUIC and IMEx web services are the responsibility of the IntAct team, while the curation effort is undertaken by both IntAct and MINT curators. .



# PPI databases categorisation and qualification

## Type of data captured:

- Only PPI information as in MINT and DIP.
- Interactions between proteins and other molecules (DNA, RNA, small molecules) as in IntAct and MatrixDB.
- PPIs and genetic interactions as in BIOGRID.
- Only PPIs related to a specific scientific topic such as : InnateDB (PPIs in the immune system), MPIDB (PPIs in microbes) and MatrixDB (extracellular PPIs).

## Type of curation Policy:

- Databases describing PPIs with low level of curation detail and quality control procedures
- Databases describing PPIs with high level of curation details and high accuracy of quality control procedures such as IMEx-complying databases.

# The IMEx Consortium

- An international collaboration between a group of major public interaction data providers who have agreed to share curation effort ([www.imexconsortium.org](http://www.imexconsortium.org))
- 12 active molecular interaction databases dedicated to producing high quality, annotated data, curated to the same standards and following the same curation rules
- Data is curated once at a single centre then exchanged between partners
- Users need only go to a single site to obtain all data

## Imex Central

The web service IMEx Central (<https://imexcentral.org/icentralbeta/>) is a central resource to assign IMEx IDs to the publications curated by IMEx members (version BETA-0.93 has been recently released).

Curators can check by using the NCBI PubMed identifier (PMID) if other IMEx members have curated an already published paper and therefore it allows avoiding work duplication.

# IMEx Consortium

The screenshot shows the IMEx Consortium website. At the top, there's a navigation bar with icons for back, forward, search, and download. The URL 'imexconsortium.org' is in the address bar. Below the header, the IMEx logo is displayed with the text 'The International Molecular Exchange Consortium'. A search bar says 'Search the IMEx data resource' with a placeholder 'Use as input: UniProtKB Accs, Gene names, Publication Ids'. Below the search bar is a navigation menu with links to 'Home', 'About', 'Curation', 'Submit Your Data', and 'Contact us'. The main content area has a dark background with a circular graphic.

## IMEx data

- A non-redundant set of physical molecular interaction data from a broad taxonomic range of organisms.
- Expertly curated from direct submissions or peer-reviewed journals to a consistent high standard.
- Available in standard formats [MITAB](#) or [PSI-MI XML 2.5](#).
- Provided by a network of participating major public domain databases.

The screenshot shows two columns of content. On the left, under 'News', there's a section about access to a single unified IMEx dataset, mentioning a collaboration with PSICQUIC and providing a download link. On the right, under 'Citing IMEx', there's a citation for a paper by Orchard et al. in Nat Methods.

### IMEx Partners



### IMEx Observers



# PPI (web) resources

# Main Molecular Interaction Databases

Database name	Data types	Main Taxonomies	Archival/thematic	Curation depth	IMEx Member	PSICQUIC service	Ref.
IntAct	All	Full	Archival	IMEx/MIMIx	Full	Yes	[6]
MINT	PPIs	Full	Archival	IMEx/MIMIx	Full	Yes	[7]
InnateDB	PPIs	Human and mouse	Proteins involved in innate immunity	IMEx/MIMIx	Full	Yes	[10]
MPIDB	PPIs	Bacteria and archaea	Microbial proteins	IMEx/MIMIx	Full	Yes	[9]
I2D	PPIs	Model organisms	Cancer related proteins	IMEx/MIMIx	Full	Yes	
DIP	PPIs	Full	Archival	IMEx	Full	Yes	[1]
MatrixDB	PPIs; PSMIs	Human and mouse	Extracellular matrix	IMEx	Full	Yes	[8]
BioGRID	PPIs	Model organisms	Archival	Limited	Observer	Yes	[13]
HPRD	PPIs	Human	Human	Limited	No	No	[38]
ChEMBL	Drug-target PSMIs	Targets mainly human or pathogens	Drug-target	MIABE [39]/MIMIx	No	Yes	[16]
BindingDB	Drug-target PSMIs	All	Drug-target	MIABE/MIMIx	No	Yes	[40]
PubChem BioAssay	Drug-target PSMIs	Targets mainly human or pathogens	Drug-target	MIABE/MIMIx	No	No	[19]
PrimesDB	PPIs	Human and mouse	EGFR network	Limited	Observer	No	[34]
HPIDB	PPIs	Model organisms and pathogens	Host-pathogen systems	IMEx	Full	Application pending	

IMEx/MIMIx – the database contains both IMEx and MIMIx standards data.  
PPIs – Protein-Protein Interactions; PSMIs –Protein-Small Molecule Interactions.

A more complete list of molecular interaction databases is available at:  
<http://www.pathguide.org>

<b>Database</b>	<b>Website</b>
2P2I	<a href="http://2p2fdb.cnrs-mrs.fr/">http://2p2fdb.cnrs-mrs.fr/</a>
PrePPI	<a href="http://bhapp.c2b2.columbia.edu/PrePPI">http://bhapp.c2b2.columbia.edu/PrePPI</a>
STRING	<a href="http://string-db.org/">http://string-db.org/</a>
IBIS	<a href="http://www.ncbi.nlm.nih.gov/Structure/ibis/ibis_help.shtml#whatisIBIS">http://www.ncbi.nlm.nih.gov/Structure/ibis/ibis_help.shtml#whatisIBIS</a>
PIPS	<a href="http://www.compbio.dundee.ac.uk/www-pips/">http://www.compbio.dundee.ac.uk/www-pips/</a>
PredUS	<a href="https://bhapp.c2b2.columbia.edu/PredUs/">https://bhapp.c2b2.columbia.edu/PredUs/</a>
DIP, LiveDIP	<a href="dip.doe-mbi.ucla.edu/ldipc/tmpl/livedip.cgi">dip.doe-mbi.ucla.edu/ldipc/tmpl/livedip.cgi</a>
BIND	<a href="http://www.bindingdb.org/bind/index.jsp">http://www.bindingdb.org/bind/index.jsp</a>
MPact/MIPS	<a href="http://mips.helmholtz-muenchen.de/proj/ppi/">http://mips.helmholtz-muenchen.de/proj/ppi/</a>
YPD and WormPD	<a href="https://portal.biobase-international.com/build_ghpywl/idb/1.0/html/bkldoc/source/bkl/proteome/proteome_wormpd_intro.html">https://portal.biobase-international.com/build_ghpywl/idb/1.0/html/bkldoc/source/bkl/proteome/proteome_wormpd_intro.html</a>
MINT	<a href="http://mint.bio.uniroma2.it/mint/Welcome.do">http://mint.bio.uniroma2.it/mint/Welcome.do</a>
IntAct	<a href="http://www.ebi.ac.uk/intact/">http://www.ebi.ac.uk/intact/</a>
BioGRID	<a href="http://thebiogrid.org/">http://thebiogrid.org/</a>
HPRD	<a href="http://www.hprd.org/">http://www.hprd.org/</a>
ProtCom	<a href="http://www.ces.clemson.edu/compbio/protcom">http://www.ces.clemson.edu/compbio/protcom</a>
3did, Interprets	<a href="http://3did.irbbarcelona.org/">http://3did.irbbarcelona.org/</a>
Pibase, ModBase	<a href="http://modbase.compbio.ucsf.edu/pibase/introduction.html">http://modbase.compbio.ucsf.edu/pibase/introduction.html</a>
CBM	<a href="http://www.cazy.org/Carbohydrate-Binding-Modules.html">http://www.cazy.org/Carbohydrate-Binding-Modules.html</a>
SCOPPI	<a href="http://scoppi.biotech.tu-dresden.de/scoppi/">http://scoppi.biotech.tu-dresden.de/scoppi/</a>
iPfam	<a href="http://www.ipfam.org/">http://www.ipfam.org/</a>
InterDom	<a href="http://interdom.i2r.a-star.edu.sg/">http://interdom.i2r.a-star.edu.sg/</a>
DIMA	<a href="http://webclu.bio.wzw.tum.de/dima/">http://webclu.bio.wzw.tum.de/dima/</a>
Prolinks	<a href="http://prl.mbi.ucla.edu/prlbeta/">http://prl.mbi.ucla.edu/prlbeta/</a>
Sable&Jois, Molecules, 2015	

# A close look into a molecular interaction database

- How to search a molecular interaction database?
- Which information is it possible to retrieve?
- Which kind of data are collected in an interaction database?
- How is it possible to download data?
- How can data be visualised?

# Web resources describing PPI data

Six working groups (or three, depending on your background knowledge):

- UNIPROT
- PDB
- GO
- INTACT/MintIntact
- PSICQUIC
- MENTHA and VIRUS MENTHA
- Discuss the properties of each type of resource/tool,
- Make a list of the main characteristics of each resource/tool examined
- Show one example interaction for each resource/tool
- All groups report to the class
- Groups start writing individual reports

# UniProt



**Uniprot** is the highest quality and most accurately annotated protein sequence database.

- A non redundant protein database, with maximal coverage including splicing isoforms, disease variants and PTMs.
- Easy protein identification, stable identifiers and consistent nomenclature/controlled vocabularies.
- Detailed information on protein function, biological processes, molecular interactions and pathways cross-referenced to external source

# UniProt Search

UniProtKB▼ ataxin 1 Advanced ▾ 

BLAST Align Retrieve/ID mapping Help Contact

UniProtKB results  ? About UniProtKB 

Filter by:

Reviewed (26) Swiss-Prot

Unreviewed (536) TrEMBL

Popular organisms

Human (22)

Mouse (20)

Zebrafish (9)

Rat (5)

Bovine (4)

Other organisms

BLAST Align Download Add to basket Columns > 1 to 25 of 562 Show 25

	Entry	Entry name		Protein names	Gene names	Organism	Length	
2 result(s) selected. (Clear selection)								
<input type="checkbox"/>	Q9NRR5	UBQL4_HUMAN		Ubiquilin-4	UBQLN4 C1orf6,CIP75,UBIN	Homo sapiens (Human)	601	
<input type="checkbox"/>	Q99NB8	UBQL4_MOUSE		Ubiquilin-4	Ubqln4 Cip75,Ubin	Mus musculus (Mouse)	596	
<input checked="" type="checkbox"/>	P54253	ATX1_HUMAN		Ataxin-1	ATXN1 ATX1,SCA1	Homo sapiens (Human)	815	
<input type="checkbox"/>	P54254	ATX1_MOUSE		Ataxin-1	Atxn1 Sca1	Mus musculus (Mouse)	791	
<input checked="" type="checkbox"/>	P0C7T5	ATX1L_HUMAN		Ataxin-1-like	ATXN1L BOAT,BOAT1	Homo sapiens (Human)	689	
<input type="checkbox"/>	P0C7T6	ATX1L_MOUSE		Ataxin-1-like	Atxn1l Boat	Mus musculus (Mouse)	687	
<input type="checkbox"/>	Q63540	ATX1_RAT		Ataxin-1	Atxn1 Sca1	Rattus norvegicus (Rat)	789	

Go

# PROTEIN INFORMATION

- Function
- Names & Taxonomy
- Subcell. location
- Pathol./Biotech
- PTM / Processing
- Expression
- Interaction
- Structure
- Family & Domains
- Sequence
- Cross-references
- Publications
- Entry information
- Miscellaneous
- Similar proteins

## 3 (ATX1\_HUMAN)

**taxin-1**

**TXN1**

*Homo sapiens (Human)*

Show only features (sites, domains, PTMs ...)

Reviewed - Annotation score:  - Experimental evidence at protein level<sup>i</sup>

 BLAST  Align  Format  Add to basket  History

unction<sup>i</sup>

chromatin-binding factor that repress Notch signaling in the absence of Notch intracellular domain by acting as a CBF1 corepressor. Binds RNA in vitro. May be involved in RNA metabolism.  1 Publication ▾

**O - Molecular function<sup>i</sup>**

- chromatin binding  Source: Ensembl
- DNA binding  Source: UniProtKB-KW
- identical protein binding  Source: IntAct ▾
- poly(G) binding  Source: UniProtKB ▾
- poly(U) RNA binding  Source: UniProtKB ▾
- protein C-terminus binding  Source: UniProtKB ▾
- protein self-association  Source: UniProtKB ▾

**O - Biological process<sup>i</sup>**

- adult locomotory behavior  Source: Ensembl
- excitatory postsynaptic potential  Source: Ensembl
- lung alveolus development  Source: Ensembl
- negative regulation of insulin-like growth factor receptor signaling pathway  Source: Ensembl
- negative regulation of phosphorylation  Source: Ensembl
- negative regulation of transcription, DNA-templated  Source: UniProtKB ▾
- negative regulation of transcription from RNA polymerase II promoter  Source: Ensembl
- nuclear export  Source: UniProtKB ▾

# GOING IN DETAILS...

Display None Family & Domains<sup>i</sup>

Display None

Function  
 Names & Taxonomy  
 Subcellular location  
 Pathology & Biotech  
 PTM / Processing  
 Expression  
 Interaction  
 Structure  
 Family & Domains  
 Sequence  
 Cross-references  
 Publications  
 Entry information  
 Miscellaneous  
 Similar proteins  
▲ Top

Sequence<sup>i</sup>

Sequence status<sup>i</sup>: Complete.  
This entry describes 1 isoform<sup>i</sup> produced by alternative splicing. [Align](#) [Add to basket](#)

Note: At least 2 isoforms are produced.

**Isoform 1** (identifier: P54253-1) [UniParc] [FASTA](#) [Add to basket](#)

This isoform has been chosen as the 'canonical' sequence. All positional information in this entry refers to it. This is also the sequence that appears in the downloadable versions of the entry.  
« Hide

Length: 815  
Mass (Da): 86,923  
Last modified: September 23, 2008 - v2  
Checksum: i 657876F8FD19ECB2

BLAST [GO](#)

10	20	30	40	50
MKSQNQRSNE	CLPPKKREIP	ATSRSEEKA	PTLPSDNHRV	EGTAWLPGNP
60	70	80	90	100
GGRGHGGGRH	GPAGTSVELG	LQQGIGLHK	LSTGLDYSPP	SAPRSVPVAT
110	120	130	140	150
TLPAAYATPQ	PGTPVSPVQY	AHLPHTFQFI	GSSQYSGTYA	SFIPSQQLIPP
160	170	180	190	200
TANPVTSAVA	SAAGATTPSQ	RSQLEAYSTL	LANMGSLSQ	PGHKAEQQQQ
210	220	230	240	250
QQQQQQQQHQ	HQQQQQQQQQ	QQQQQHLSRA	PGLITPGSPP	PAQQNQYVHI
260	270	280	290	300
SSSPQNTGRT	ASPPAIPVHL	HPHQTMIPHT	LTLGPPSQVV	MQYADSGSHF
310	320	330	340	350
VPREATKAE	SSRLQQAIQA	KEVLNGEMEK	SRRY GAPSSA	DLGLGKAGGK
360	370	380	390	400
SVPHPYESRH	VVHPSPSDY	SSRDPSGVRA	SVMVLPNSNT	PAADLEVQQA
410	420	430	440	450

# STRUCTURAL INFORMATION IN UNIPROT

Display None

Structure <sup>i</sup>

**Secondary structure**  
1 815 Legend:

[Helix](#) [Turn](#) [Beta strand](#)

[Show more details](#)

**3D structure databases**

Select the link destinations:

	Entry	Method	Resolution (Å)	Chain	Positions	PDBsum
<input type="radio"/> PDB <i>b</i>	1OA8	X-ray	1.70	A/B/C/D	562-693	[*]
<input checked="" type="radio"/> RCSB PDB <i>b</i>	2M41	NMR	-	B	566-688	[*]
<input type="radio"/> PDB <i>b</i>	4APT	X-ray	2.50	A/B/C/D	566-688	[*]
	4AQP	X-ray	2.45	A/B/C/D	566-688	[*]
	4I21	X-ray	2.50	A/B/C	562-688	[*]
	4J2L	X-ray	3.15	A/B	562-688	[*]

[ProteinModelPortal<sup>i</sup>](#) P54253.

[SMR<sup>i</sup>](#) P54253. Positions 573-693.

[ModBase<sup>i</sup>](#) Search...

[MobiDB<sup>i</sup>](#) Search...

# PROTEIN DATA BANK (PDB)

The screenshot shows the RCSB PDB homepage. At the top, there is a navigation bar with links for Deposit, Search, Visualize, Analyze, Download, Learn, More, and MyPDB Login. Below the navigation bar, the RCSB PDB logo is displayed, along with a banner stating "An Information Portal to 114217 Biological Macromolecular Structures". A search bar allows users to search by PDB ID, author, macromolecule, sequence, or ligands, with a "Go" button. Below the search bar are links for Advanced Search and Browse by Annotations. The main content area features a large image of a protein structure. On the left, there is a sidebar with links for Welcome, Search, Visualize, Analyze, Download, and Learn. The central content area includes sections for "A Structural View of Biology" and "December Molecule of the Month". The "A Structural View of Biology" section describes the resource as powered by the Protein Data Bank archive, which contains information about the 3D shapes of proteins, nucleic acids, and complex assemblies. It highlights the resource's role in understanding all aspects of life, from protein synthesis to health and disease. The "December Molecule of the Month" section features an image of the Vancomycin molecule.

**PDB is an portal promoting exploration in the world of proteins and nucleic acids.**

**Along with our Worldwide PDB collaborators, RCSB PDB curates, annotates, and makes publicly available the PDB data deposited by scientists around the globe.**

**The RCSB PDB provides a window to these data through a rich online resource with powerful searching, reporting, and visualization tools for researchers.**

# PDB:1OA8

RCSB PDB Deposit Search Visualize Analyze Download Learn More MyPDB Login

Structure Summary 3D View Annotations Sequence Sequence Similarity Structure Similarity Experiment Literature

Biological Assembly 1 ? Display Files Download Files

# 1OA8

AXH DOMAIN OF HUMAN SPINOCEREBELLAR ATAXIN-1

DOI: 10.2210/pdb1oa8/pdb

Classification: RNA BINDING

Deposited: 2003-01-02 Released: 2003-11-06

Deposition author(s): Allen, M.D., Chen, Y.W., Bycroft, M.

Organism: Homo sapiens

Expression System: ESCHERICHIA COLI

Structural Biology Knowledgebase: 1OA8 (1 model >14 annotations) SBKB.org

Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
Resolution: 1.7 Å  
R-Value Free: 0.252  
R-Value Work: 0.210

wwPDB Validation

Metric	Percentile Ranks	Value
Clashscore	5	5
Ramachandran outliers	0.4%	0.4%
Sidechain outliers	6.7%	6.7%

Full Report

Literature Download Primary Citation

The Structure of the Axh Domain of Spinocerebellar Ataxin-1

Chen, Y.W., Allen, M.D., Veprintsev, D., Lowe, J., Bycroft, M.

(2004) J.Biol.Chem. 279: 3758

PubMed: 14583607 Search on PubMed

View in 3D: JSmol or PV (in Browser)

Standalone Viewers

Simple Viewer Protein Workshop  
Ligand Explorer Kiosk Viewer

Protein Symmetry: Asymmetric (View in 3D)

Protein Stoichiometry: Homo 4-mer - A4

Biological assembly 1 assigned by authors and generated by PQS (software)



EMBL-EBI

## Protein Data Bank in Europe

Bringing Structure to Biology

Search Examples: hemoglobin, BRCA1\_HUMAN Search EMDB

Share Feedback

### PDBe > 1oa8

AXH DOMAIN OF HUMAN SPINOCEREBELLAR ATAXIN-1

Source organism: *Homo sapiens*

Primary publication:

The structure of the AXH domain of spinocerebellar ataxin-1.  
Chen YW, Allen MD, Vepriksev DB, Löwe J, Bycroft M  
*J. Biol. Chem.* **279** 3758-65 (2004)  
PMID: 14583607

X-ray diffraction  
1.7 Å resolution

Released: 06 Nov 2003

Model geometry Fit model/data Not available

Quick links

- 1oa8 overview
- Citations
- Structure analysis
- Function and Biology
- Ligands and Environments
- Experiments and Validation
- View
- Downloads
- 3D Visualisation

Function and Biology

Details

Biochemical function: RNA binding

Biological process: not assigned

Cellular component: not assigned

Sequence domains:

- Ataxin-1/HBP1 module (AXH) [IPR013723]
- Ataxin, AXH domain [IPR003652]

Structure domain:

- AXH domain

Ligands and Environments

1 bound ligand:  
Na<sup>+</sup>  
3 x NA

No modified residues

Citations

7 review citations

Kaleidoscopic protein-protein interactions in the life and death of ataxin-1: new strategies against protein aggregation.  
de Chiara et al. (2014)

6 more

Structure analysis

Details

Assembly composition: homo tetramer (preferred)

Entry contents: 1 distinct polypeptide molecule

Macromolecule:

- Ataxin-1

Experiments and Validation

Details

Metric Percentile Ranks Value

Metric	Percentile Ranks	Value
Clashscore	5	5
Ramachandran outliers	0.4%	0.4%
Sidechain outliers	6.7%	6.7%

Percentile relative to X-ray structures

## Map proteins into UniProt reference sequence database

UniProtKB is the protein sequence reference database chosen by the majority of the interaction databases.

Choosing UniProtKB allow the curators to annotate:

- the specific isoform utilized in an experiment
- to describe all isoforms simultaneously (using the canonical sequence)
- To specify a peptide, resulting from a post-translational cleavage.

Uniprot mapping tools allow users to convert several proteins identifiers to UniProt ACs



# Uniprot Retrieve/ID mapping

UniProt

UniProtKB yourlist:M2015120965JB6AN7UC Advanced Search

BLAST Align Retrieve/ID mapping Help Contact

UniProtKB results

8 out of 8 Ensembl identifiers were successfully mapped to 36 UniProtKB IDs in the table below.

Filter by:

- Reviewed (8) Swiss-Prot
- Unreviewed (28) TrEMBL

Popular organisms

Human (36)

View by

Taxonomy

Keywords

Gene Ontology

Enzyme class

Pathway

UniRef

	Your list:...6AN7U	Entry	Entry name	Protein names	Gene names	Organism	Length
<input type="checkbox"/>	ENSG00000107796	F6QUT6	F6QUT6_HUMAN	Actin, aortic smooth muscle	ACTA2	Homo sapiens (Human)	151
<input type="checkbox"/>	ENSG00000107796	F6UVQ4	F6UVQ4_HUMAN	Actin, aortic smooth muscle	ACTA2	Homo sapiens (Human)	151
<input type="checkbox"/>	ENSG00000107796	P62736	ACTA_HUMAN	Actin, aortic smooth muscle	ACTA2 ACTSA,ACTVS,GIG46	Homo sapiens (Human)	377
<input type="checkbox"/>	ENSG00000075624	C9JTX5	C9JTX5_HUMAN	Actin, cytoplasmic 1	ACTB	Homo sapiens (Human)	80
<input type="checkbox"/>	ENSG00000075624	C9JUM1	C9JUM1_HUMAN	Actin, cytoplasmic 1	ACTB	Homo sapiens (Human)	99
<input type="checkbox"/>	ENSG00000075624	C9JZR7	C9JZR7_HUMAN	Actin, cytoplasmic 1	ACTB	Homo sapiens (Human)	102
<input type="checkbox"/>	ENSG00000075624	E7EVS6	E7EVS6_HUMAN	Actin, cytoplasmic 1	ACTB	Homo sapiens (Human)	163
<input type="checkbox"/>	ENSG00000075624	G5E9R0	G5E9R0_HUMAN	Actin, cytoplasmic 1	ACTB hCG_15971	Homo sapiens	125



# PSICQUIC (PSI Common Query Interface)

The screenshot shows the PSICQUIC View interface. At the top, there's a navigation bar with links for Services, Research, Training, About us, and a search bar. Below the header, a teal banner displays "PSICQUIC View". Underneath, a sub-navigation bar has "Input Form" selected, followed by "Browse" and "Help". A breadcrumb trail shows "Input Form > Browse". The main content area displays a large number of search results for the term "126.036.189". Each result is a link to a specific database, accompanied by its name, a small icon, and a numerical value. To the right, a sidebar titled "Status of the service" lists four categories: ONLINE (green), OFFLINE (grey), WARNING: Time out (orange), and ERROR: Unexpected Error (red). Below the sidebar, a message indicates "126.036.189 selected interactions" and a note about reducing the number of interactions.

Service	Count
APID	104.854
BindingDB	102.153
DrugBank	
I2D	817.915
InnateDB-IMEx	855
MatrixDB	
MolCon	495
Spike	36.248
VirHostNet	49.390
BAR	2.126
BioGrid	795.886
EBI-GOA-nonIntAct	54.610
I2D-IMEx	1.118
IntAct	423.652
MBInfo	638
MPIDB	1.759
STRING	
InnateDB	
Interporc	208.558
mentha	920.980
Reactome	141.996
TopFind	
UniProt	13.171
Bind	192.961
DIP	107.619
HPIDb	2.557
InnateDB-All	443.375
iReflindex	
MINT	130.601
Reactome-Fis	209.988

PSICQUIC is an effort from the HUPO Proteomics Standard Initiative (HUPO-PSI) to standardize the access to molecular interaction databases programmatically.

It is available at:

<http://www.ebi.ac.uk/Tools/webservices/psicquic/view/main.xhtml>

and users can obtain results both in MITAB 2.5, 2.6 and 2.7 or PSI-XML formats using the MIQL 2.7 query language

## MENTHA and VIRUS MENTHA

The screenshot shows the main interface of the MENTHA interactome browser. At the top, there's a green header bar with the MENTHA logo, which features a stylized green leaf shape containing a white molecular network icon. Below the logo, the word "mentha" is written in a large, green, lowercase sans-serif font, with "the interactome browser" in smaller text underneath. To the left of the search bar, there are links for "Example", "Advanced Tools", and "Help". The search bar itself has a light blue background and contains the placeholder text "Genes, uniprot IDs, keywords". To the right of the search bar is a green "search" button. Below the search bar, there's a dropdown menu labeled "Organism: All (including those not listed)". At the bottom of the header, a message indicates the last update: "Last update: 6th December 2015" followed by statistics: "84023 proteins, 569081 interactions, 45310 publications".

An interactome browser that offers protein-protein physical/enzymatic interaction information integrated from various sources. Its data comes from manually curated protein-protein interaction databases

This screenshot shows the VIRUS MENTHA interactome browser. It has a similar layout to the MENTHA version, with a blue header bar and a blue footer bar. The central part features the "virus mentha" logo, where "virus" is in blue and "mentha" is in a larger, blue, stylized font. Below the logo, it says "the interactome browser". The search bar at the top has a light blue background and contains the placeholder text "Genes, uniprot IDs, keywords". To the right is a blue "search" button. Below the search bar, there's a "Search by:" section with radio buttons for "Entire database" (selected), "Families", and "Hosts". At the bottom, a message indicates the last update: "Last update: 6th December 2015" followed by statistics: "4125 proteins, 8475 interactions, 9271 publications".

virus mentha archives evidence about viral protein-protein interactions Viruses vs Hosts and viceversa - evidence collected from different databases that have adhered to the IMEx consortium.



# Course on Protein Networks and Systems Biology

## References

- Taylor CF, Hermjakob H, Julian RK, Garavelli JS, Aebersold R, Apweiler R (2006) The work of the Human Proteome Organisation's Proteomics Standards Initiative (HUPO PSI). OMICS 10:145–51. doi: 10.1089/omi.2006.10.145
- Hermjakob H, Montecchi-Palazzi L, Bader G, Wojcik J, Salwinski L, Ceol A, Moore S, Orchard S, Sarkans U, von Mering C, Roechert B, Poux S, Jung E, Mersch H, Kersey P, Lappe M, Li Y, Zeng R, Rana D, Nikolski M, Husi H, Brun C, Shanker K, Grant SGN, Sander C, Bork P, Zhu W, Pandey A, Brazma A, Jacq B, Vidal M, Sherman D, Legrain P, Cesareni G, Xenarios I, Eisenberg D, Steipe B, Hogue C, Apweiler R (2004) The HUPO PSI's molecular interaction format--a community standard for the representation of protein interaction data. Nat Biotechnol 22:177–83. doi: 10.1038/nbt926
- Orchard S, Kerrien S, Abbani S, Aranda B, Bhate J, Bidwell S, Bridge A, Briganti L, Brinkman FSL, Brinkman F, Cesareni G, Chatr-aryamontri A, Chautard E, Chen C, Dumousseau M, Goll J, Hancock REW, Hancock R, Hannick LI, Jurisica I, Khadake J, Lynn DJ, Mahadevan U, Perfetto L, Raghunath A, Ricard-Blum S, Roechert B, Salwinski L, Stümpflen V, Tyers M, Uetz P, Xenarios I, Hermjakob H (2012) Protein interaction data curation: the International Molecular Exchange (IMEx) consortium. Nat Methods 9:345–50. doi: 10.1038/nmeth.1931
- The Universal Protein Resource (UniProt) 2009. Nucleic Acids Res 37:D169–74. doi: 10.1093/nar/gkn664
- Kerrien S, Orchard S, Montecchi-Palazzi L, Aranda B, Quinn AF, Vinod N, Bader GD, Xenarios I, Wojcik J, Sherman D, Tyers M, Salama JJ, Moore S, Ceol A, Chatr-Aryamontri A, Oesterheld M, Stümpflen V, Salwinski L, Nerothin J, Cerami E, Cusick ME, Vidal M, Gilson M, Armstrong J, Woppard P, Hogue C, Eisenberg D, Cesareni G, Apweiler R, Hermjakob H (2007) Broadening the horizon--level 2.5 of the HUPO-PSI format for molecular interactions. BMC Biol 5:44. doi: 10.1186/1741-7007-5-44
- Aranda B, Blankenburg H, Kerrien S, Brinkman FSL, Ceol A, Chautard E, Dana JM, De Las Rivas J, Dumousseau M, Galeota E, Gaulton A, Goll J, Hancock REW, Isserlin R, Jimenez RC, Kerssemakers J, Khadake J, Lynn DJ, Michaut M, O'Kelly G, Ono K, Orchard S, Prieto C, Razick S, Rigna O, Salwinski L, Simonovic M, Velankar S, Winter A, Wu G, Bader GD, Cesareni G, Donaldson IM, Eisenberg D, Kleywegt GJ, Overington J, Ricard-Blum S, Tyers M, Albrecht M, Hermjakob H (2011) PSICQUIC and PSISCORE: accessing and scoring molecular interactions. Nat Methods 8:528–9. doi: 10.1038/nmeth.1637
- del-Toro N, Dumousseau M, Orchard S, Jimenez RC, Galeota E, Launay G, Goll J, Breuer K, Ono K, Salwinski L, Hermjakob H (2013) A new reference implementation of the PSICQUIC web service. Nucleic Acids Res 41:W601–6. doi: 10.1093/nar/gkt392
- Calderone A, Castagnoli L, Cesareni G (2013) mentha: a resource for browsing integrated protein-interaction networks. Nat Methods 10:690–1. doi: 10.1038/nmeth.2561
- Calderone A, Licata L, Cesareni G (2015) VirusMentha: a new resource for virus-host protein interactions. Nucleic Acids Res 43:D588–92. doi: 10.1093/nar/gku830
- Mayer G, Jones AR, Binz PA, Deutsch EW, Orchard S, Montecchi-Palazzi L, Vizcaíno JA, Hermjakob H, Oveillero D, Julian R, Stephan C, Meyer HE, Eisenacher M., Controlled vocabularies and ontologies in proteomics: overview, principles and practice. Biochim Biophys Acta. 2014 Jan;1844(1 Pt A):98-107. doi: 10.1016/j.bbapap.2013.02.017. Epub 2013 Feb 19. Review.
- Orchard S., Data standardization and sharing-the work of the HUPO-PSI. Biochim Biophys Acta. 2014 Jan;1844(1 Pt A):82-7. doi: 10.1016/j.bbapap.2013.03.011. Epub 2013 Mar 20. Review.

# The IntAct database

IntAct Home Page (<http://www.ebi.ac.uk/intact/>)

EMBL-EBI  Go to the EMBL-EBI homepage

**IntAct**

Home Advanced Search About Resources Download Feedback

## IntAct Molecular Interaction Database

IntAct provides a freely available, open source database system and analysis tools for molecular interaction data. All interactions are derived from literature curation or direct user submissions and are freely available. The IntAct Team also produce the [Complex Portal](#).

Search in IntAct  
Enter search term(s)  
**Creb1**  
Search Search Tips

**Examples**

- Gene, Protein, RNA or Chemical name: [BRCA2](#), [Staurosporine](#)
- UniProtKB or ChEBI AC: [Q06609](#), [CHEBI:15996](#)
- UniProtKB ID: [LCK\\_HUMAN](#)
- RNACentral ID: [URS00004C95F4\\_559292](#)
- PMID: [25416956](#)
- IMEx ID: [IM-23318](#)

**Dataset of the month: December**  
**Systematic profiling of the human centrosome-cilium proximity interaction landscape.**

- Gupta et al. [IntAct](#) [PSI-MI 2.5](#) [PSI-MI TAB](#)
- [Go to Archive](#)

**Contributors**  
Manually curated content is added to IntAct by curators at the EMBL-EBI and the following organisations:

**Data Content**

- Publications: **13952**
- Interactions: **577297**
- Interactors: **89716**

**News** Follow

 **EMBL-EBI Training** @EBItraining 10 Dec Finding out more about molecular interactions for #proteomics with Pablo Porras @intact\_project pic.twitter.com/SgsPaROs6U

 **IntAct at EBI** @intact\_project 2 Dec IntAct release 194 out! 577297 binary interactions from 13952 curated publications and 1378 biological complexes. [ebi.ac.uk/intact/about/statistics](http://ebi.ac.uk/intact/about/statistics)

 **EMBL-EBI Training** @EBItraining 18 Nov Hello @Cambridge\_Uni students! Today we're teaching network analysis using @cytoscape and @PSICQUIC: ow.ly/Ub8ku

 **EMBL-EBI Training** @EBItraining 3 Nov Practical #networkanalysis with Pablo Porras @intact\_project @emblebi pic.twitter.com/oASdjV1QR

[Show Photo](#) [Expand](#) [Retweeted by IntAct at EBI](#)

[Tweet to @intact\\_project](#)

# The IntAct database: the result page

156 binary interactions found for search term *creb1*

Interactions (156) Interactors Interaction Details Graph

Filter out the spoke expanded co-complexes (54)

Your query also matches 2 biological complexes in IntAct.

Your query also matches 1,553 interaction evidences from 7 other databases.

Customize view Select format to Download Download

(2 of 8) 1 2 3 4 5 6 7 8 >> >>

Dts	Molecule 'A'	Links 'A'	Molecule 'B'	Links 'B'	Interaction Detection Method	Interaction AC	Source Database
•	CREB1	P16220 EBI-711855	CREB1	P16220 EBI-711855	peptide array	EBI-10890938	HPIDb
•					tandem affinity purification	EBI-11317087 imex : IM-24178-98	MINT
•					tandem affinity purification	EBI-11320514 imex : IM-24178-153	MINT
•	CREB1	P16220 EBI-711855	NFIL3	Q16649 EBI-3951858	peptide array	EBI-10890678	HPIDb
•					peptide array	EBI-10890943	HPIDb
•					tandem affinity purification	EBI-11317087 imex : IM-24178-98	MINT
•	CREB1	P16220 EBI-711855	CREBBP	Q92793 EBI-81215	anti tag coimmunoprecipitation	EBI-2880014 imex : IM-14346-3	IntAct
•					pull down	EBI-7397274 MINT-4979977 imex : IM-11271-2	MINT
•	CREB1	P16220 EBI-711855	NFATC1	O95644 EBI-6907210	tandem affinity purification	EBI-11319967 imex : IM-24178-143	MINT
•					tandem affinity purification	EBI-10690378 imex : IM-24178-23	MINT
•					pull down	EBI-10711187 imex : IM-24178-17	MINT
•	Creb1	Q01147 EBI-2291098	Crebl2	Q32M00 EBI-5314489	anti tag coimmunoprecipitation	EBI-5314488	UniProt
•					anti bait coimmunoprecipitation	EBI-5314548	UniProt
•					confocal microscopy	EBI-5315256	UniProt
•					fluorescence microscopy	EBI-5315419	UniProt
•	CREB1	P16220 EBI-711855	PASK	Q96RG2 EBI-1042651	protein kinase assay	EBI-8613283 MINT-8145385 imex : IM-15853-7	MINT
•					peptide array	EBI-8613680 MINT-8148634 imex : IM-15853-25	MINT
•	CREB1	P16220 EBI-711855	p03259-2	P03259-2 EBI-7225021	pull down	EBI-7225044 MINT-2790592	MINT
•					anti bait coimmunoprecipitation	EBI-7225250 MINT-2790667	MINT
•	Creb1	Q01147 EBI-2291098	Myod1	P10085 EBI-4405734	anti bait coimmunoprecipitation	EBI-8629205 MINT-50074	MINT

# The IntAct database: interaction details

IntAct > Interaction Details

Interactions (156) Interactors Interaction Details Graph

## Publication

**PubMed Id:** 17476304      **Title:** Cooperative interactions between CBP and TORC2 confer selectivity to CREB target gene expression.

**Journal:** EMBO J. (0261-4189)      **Author List:** Ravnskjær K., Kester H., Liu Y., Zhang X., Lee D., Yates JR., Montminy M.

**Year of Publication:** 2007

**Cross References:**

Database	Identifier	Secondary identifier	Qualifier
doi	<a href="https://doi.org/10.1038/sj.emboj.7601715">10.1038/sj.emboj.7601715</a>	-	secondary-ac
mint	MINT-5219899	-	primary-reference
pubmed	17476304	-	primary-reference
mint	MINT-4979961	-	identity
imex	IM-11271	-	imex-primary

**Annotations:**

Topic	Text
dataset	Virus - Publications including interactions involving viral proteins

**Experiment (1 interaction)**

**Accession:** EBI-7397264      **Host organism:** In vitro

**Name:** ravnskjær-2007-2      **Interaction Detection Method:** pull down      **Participant Identification Method:** predetermined

**Cross References:**

Database	Identifier	Secondary identifier	Qualifier
mint	MINT-4979997	-	identity
pubmed	17476304	-	secondary-ac
imex	IM-11271	-	imex-primary
doi	<a href="https://doi.org/10.1038/sj.emboj.7601715">10.1038/sj.emboj.7601715</a>	-	secondary-ac
pubmed	17476304	-	primary-reference

**Annotations:**

Topic	Text
imex-exported	2009-01-08
curation depth	imex curation
full coverage	Only protein-protein interactions
imex curation	-

# The IntAct database: interaction details

## Interaction

Accession: EBI-7397274      Description: -

Name: creb1-crebbp      Type: physical association

Cross References:

Database	Identifier	Secondary identifier	Qualifier
mint	MINT-4979977	-	identity
imex	IM-11271-2	-	imex-primary

Annotations:

Topic	Text
figure legend	F2E
comment	homomint

Participant: Cyclic AMP-responsive element-binding protein 1

Participant: Cyclic AMP-responsive element-binding protein 1

## Participants (2)

Legend: A Annotation and Cross Reference

#	Name	Links	Primary Identifier	Aliases	Description
1	EBI-711855	<a href="#">UnProt</a>	P16220	CREB1	Cyclic AMP-responsive element
2	EBI-81215	<a href="#">UnProt</a>	Q92793	CREBBP CBP	CREB-binding protein CBP

Accession: EBI-7397283

Name: EBI-711855

### Features:

Legend: A Annotation and Cross Reference

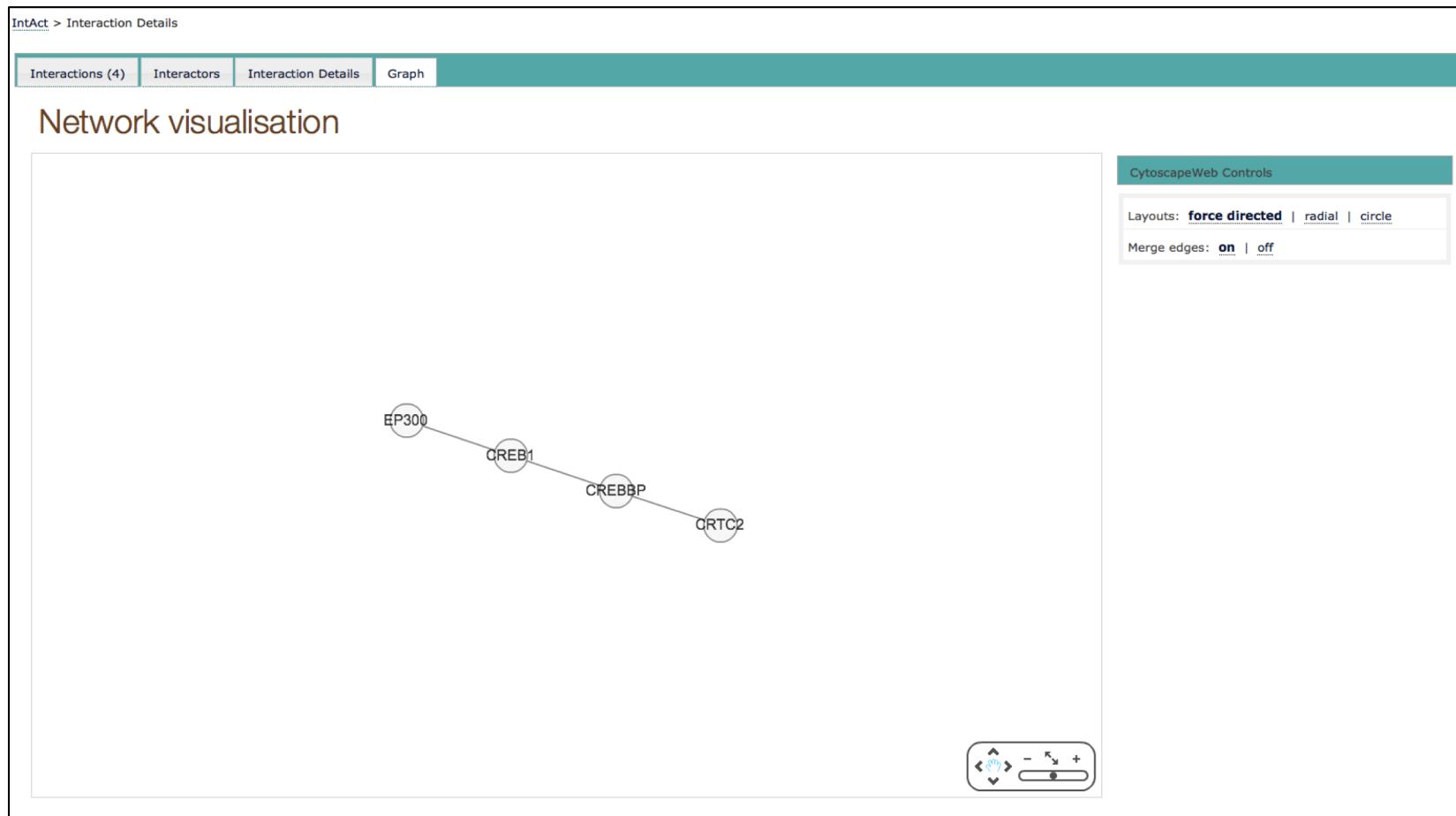
Feature Name	Feature Type	Detection Method	Range positions	More...
binding site	binding region		1-283	<a href="#">A</a>
mutation increasing	mutation increasing	mutation analysis	142-142	<a href="#">A</a>
gst tagged	gst tag		?-?	<a href="#">A</a>

## Graphical Representation of Experiment



# The IntAct database: network visualisation

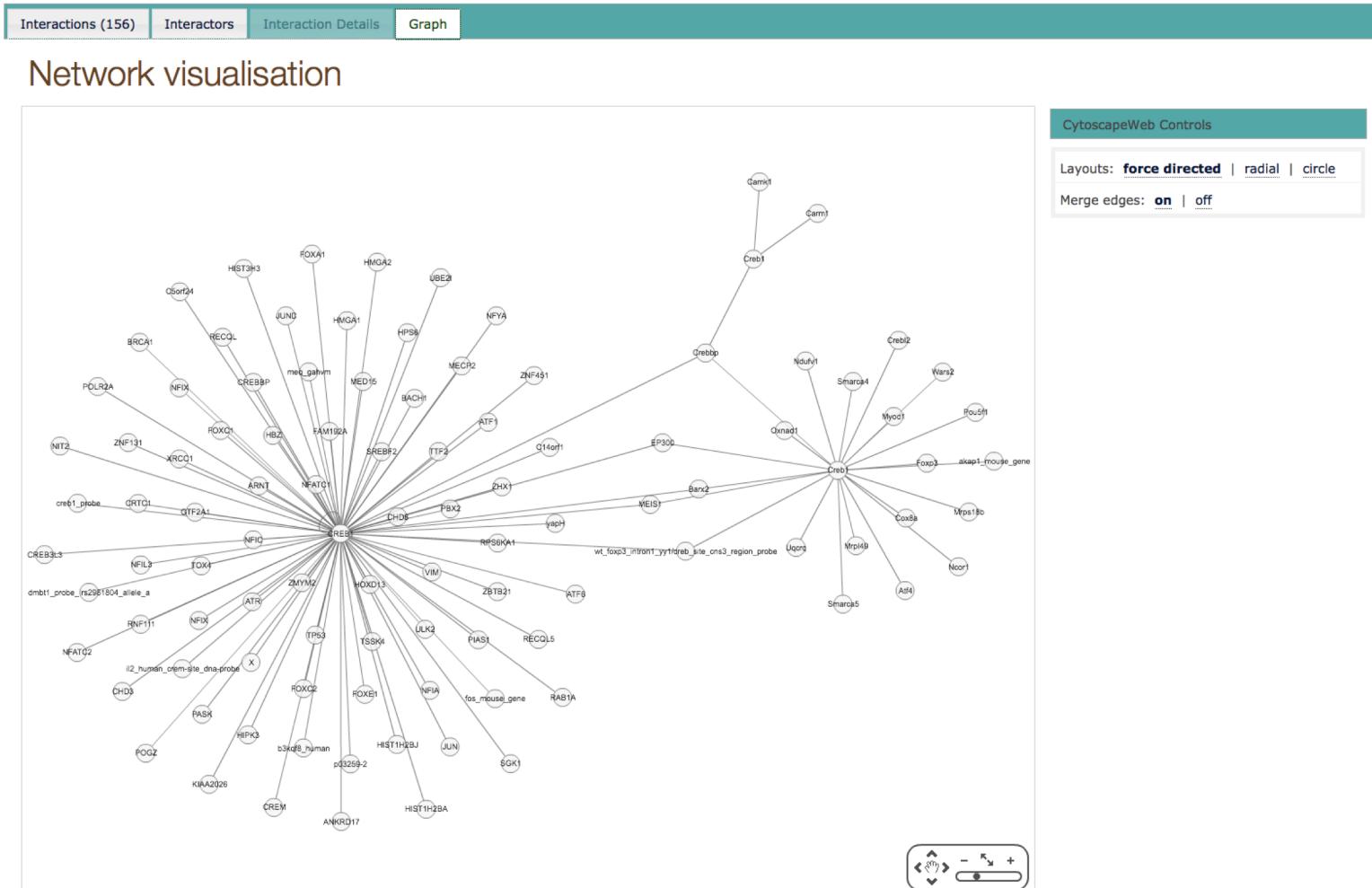
Network visualisation of the interactions present in a specific paper



# The IntAct database

Network visualisation of all the “Creb1” interactions present in the database

156 binary interactions found for search term *creb1*

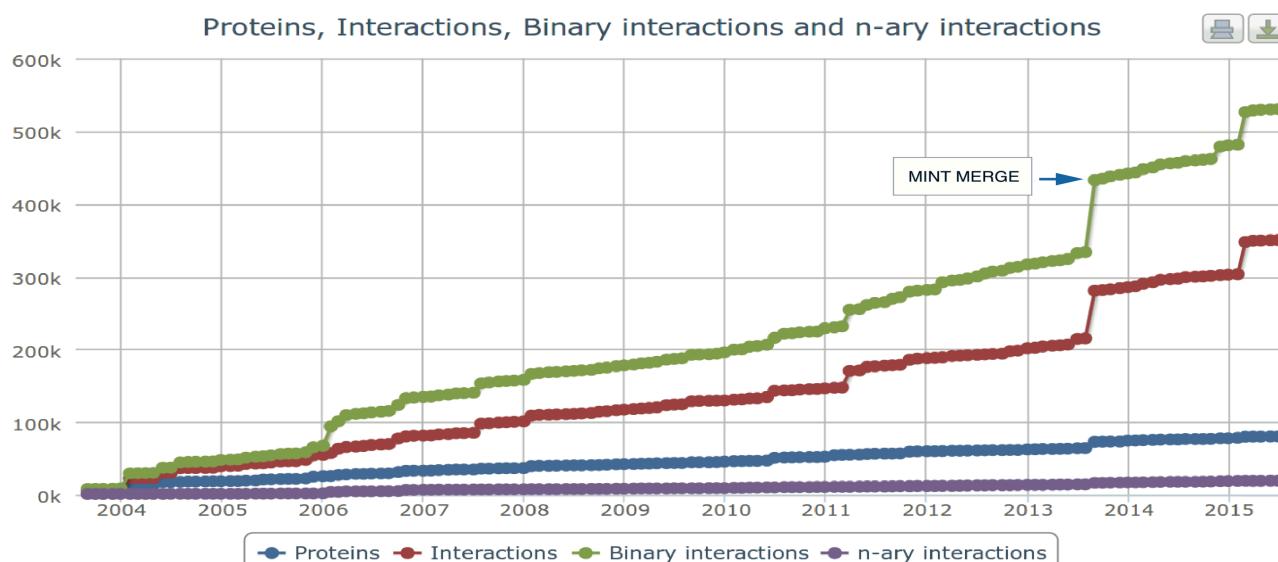


# MintAct Project

- MINT and IntAct databases were two of the largest databases (number of manuscripts curated + number of non-redundant interactions).
- Both adopted the highest possible data quality standards.
- Both were founder members of the IMEx Consortium.

IntAct and MINT joined their forces to create a single resource to improve curation and software development efforts.

Data maintenance, and the PSICQUIC and IMEx web services are the responsibility of the IntAct team, while the curation effort is undertaken by both IntAct and MINT curators.



# MintAct goals & achievements

Freely available open-source database of molecular interactions (mainly PPIs). ~560K binary interactions taken from >13892 publications (November 2015)

The screenshot shows a web page with a teal header bar containing the text "News Press releases". Below the header is a navigation menu with links: Overview, Leadership, Funding, Background, Collaborations, Jobs, People & groups, and News. Under the "News" link, there is a breadcrumb trail: About us > News > Press releases > 500,000 binary interactions... and growing. The main content features a large teal heading "500,000 binary interactions... and growing". To the left of the text is a circular visualization composed of numerous small lines connecting yellow and blue dots, representing protein interactions. To the right of the visualization is a block of text describing the database's purpose and data source. Below this is another block of text with a bolded "Image" label.

About us > News > Press releases > 500,000 binary interactions... and growing

## 500,000 binary interactions... and growing

More than half a million experimentally determined protein interactions are held in EMBL-EBI's IntAct database, providing a means to build and visualize interactions at play in living things. This public data resource contains experimental data submitted directly by researchers, and is bolstered by the scientific literature.

**Image:** Visualisation of all the human interactome data held in the proteins present in *A proteome-scale map of the human interactome* al., 2015. White: interactions observed in the same dataset. [Image]

Data is standard-compliant and available via the IntAct website, for download at the ftp site or via PSICQUIC:

<http://www.ebi.ac.uk/intact>

<ftp://ftp.ebi.ac.uk/pub/databases/intact>

[www.ebi.ac.uk/Tools/webservices/psicquic/view/main.xhtml](http://www.ebi.ac.uk/Tools/webservices/psicquic/view/main.xhtml)

## References

- Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U, Jandrasits C, Jimenez RC, Khadake J, Mahadevan U, Masson P, Pedrucci I, Pfeiffenberger E, Porras P, Raghunath A, Roechert B, Orchard S, Hermjakob H (2012) The IntAct molecular interaction database in 2012. *Nucleic Acids Res* 40:D841–6. doi: 10.1093/nar/gkr1088
- Licata L, Briganti L, Peluso D, Perfetto L, Iannuccelli M, Galeota E, Sacco F, Palma A, Nardozza AP, Santonico E, Castagnoli L, Cesareni G (2012) MINT, the molecular interaction database: 2012 update. *Nucleic Acids Res* 40:D857–61. doi: 10.1093/nar/gkr930
- Orchard S, Ammari M, Aranda B, Breuza L, Briganti L, Broackes-Carter F, Campbell NH, Chavali G, Chen C, del-Toro N, Duesbury M, Dumousseau M, Galeota E, Hinz U, Iannuccelli M, Jagannathan S, Jimenez R, Khadake J, Lagreid A, Licata L, Lovering RC, Meldal B, Melidoni AN, Milagros M, Peluso D, Perfetto L, Porras P, Raghunath A, Ricard-Blum S, Roechert B, Stutz A, Tognolli M, van Roey K, Cesareni G, Hermjakob H (2014) The MIntAct project--IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res* 42:D358–63. doi: 10.1093/nar/gkt1115
- Orchard S, Salwinski L, Kerrien S, Montecchi-Palazzi L, Oesterheld M, Stümpflen V, Ceol A, Chatr-aryamontri A, Armstrong J, Woollard P, Salama JJ, Moore S, Wojcik J, Bader GD, Vidal M, Cusick ME, Gerstein M, Gavin A-C, Superti-Furga G, Greenblatt J, Bader J, Uetz P, Tyers M, Legrain P, Fields S, Mulder N, Gilson M, Niepmann M, Burgoon L, De Las Rivas J, Prieto C, Perreau VM, Hogue C, Mewes H-W, Apweiler R, Xenarios I, Eisenberg D, Cesareni G, Hermjakob H (2007) The minimum information required for reporting a molecular interaction experiment (MIMIx). *Nat Biotechnol* 25:894–8. doi: 10.1038/nbt1324
- Koh GC, **Porras P**, Aranda B, Hermjakob H, Orchard SE., Analyzing protein-protein interaction networks. *J Proteome Res.* 2012 Apr 6;11(4):2014-31. doi: 10.1021/pr201211w. Epub 2012 Mar 2. Review.
- Orchard S., Molecular\_interaction\_databases. *Proteomics.* 2012 May;12(10):1656-62. doi: 10.1002/pmic.201100484. Review.

## Practical 1: Web resources describing results of PPI experiments

### **Training session: Searching UniProt,**

In this training session you will use UniProt database.

- perform different search
- check the cross-references
- look for the PDB information
- look for the PPI information

**Step2:** UniProt and PDB

**Step3:** UniProt and PPI

**Step4:** Protein isoforms

## Practical 2: Protein-Protein Interactions and Bioinformatics tools

### **Training session: Searching IntAct database**

In this training session you will use the IntAct database.

- perform different search
- check the number of interaction evidences
- look for the experimental details
- download data using different formats.

**Step1:** *Quick Search*

**Step2:** Refining a search using the Advanced Search

**Step3:** Download data