

PIR

A UniProt CONSORTIUM MEMBER

Protein Information Resource

LIGCLRNC SAVTAAAKQLAE
VTGF SN----AKTTAQHVKK



Introduction to the Protein Ontology (PRO)



**Immunology Ontologies and
Their Applications in Processing
Clinical Data Workshop**

Buffalo, NY · June 11-13, 2012

Cathy H. Wu, Ph.D.

Director, Protein Information Resource (PIR)

Edward G. Jefferson Chair and Director

Center for Bioinformatics & Computational Biology, University of Delaware

Professor of Biochemistry & Molecular Biology, Georgetown University

PRO in OBO Foundry



The Open Biological and Biomedical Ontologies

OBO Foundry

Establishing a set of principles to create a suite of orthogonal interoperable reference ontologies

RELATION TO TIME GRANULARITY	CONTINUANT				OCCURRENT
	INDEPENDENT		DEPENDENT		
ORGAN AND ORGANISM	Organism (NCBI Taxonomy)	Anatomical Entity (FMA, CARO)	Organ Function (FMP, CPRO)	Phenotypic Quality (PaTO)	Organism-Level Process (GO)
CELL AND CELLULAR COMPONENT	Cell (CL)	Cellular Component (FMA, GO)	Cellular Function (GO)		Cellular Process (GO)
MOLECULE	Molecule (ChEBI, SO, RNAO, PRO)		Molecular Function (GO)		Molecular Process (GO)

The Protein Ontology: a structured representation of protein forms and complexes




Natale DA, Arighi CN, Barker WC, Blake JA, Bult CJ, Caudy M, Drabkin HJ, D'Eustachio P, Evsikov AV, Huang H, Nchoutmboube J, Roberts NV, Smith B, Zhang J, Wu CH. (2011)

Nucleic Acids Res. 39, D539-545 [PMID:20935045]

PRO in OBO Foundry

Protein Ontology (PRO)

- One of the first set of six OBO Foundry ontologies
- Reference Ontology for Proteins

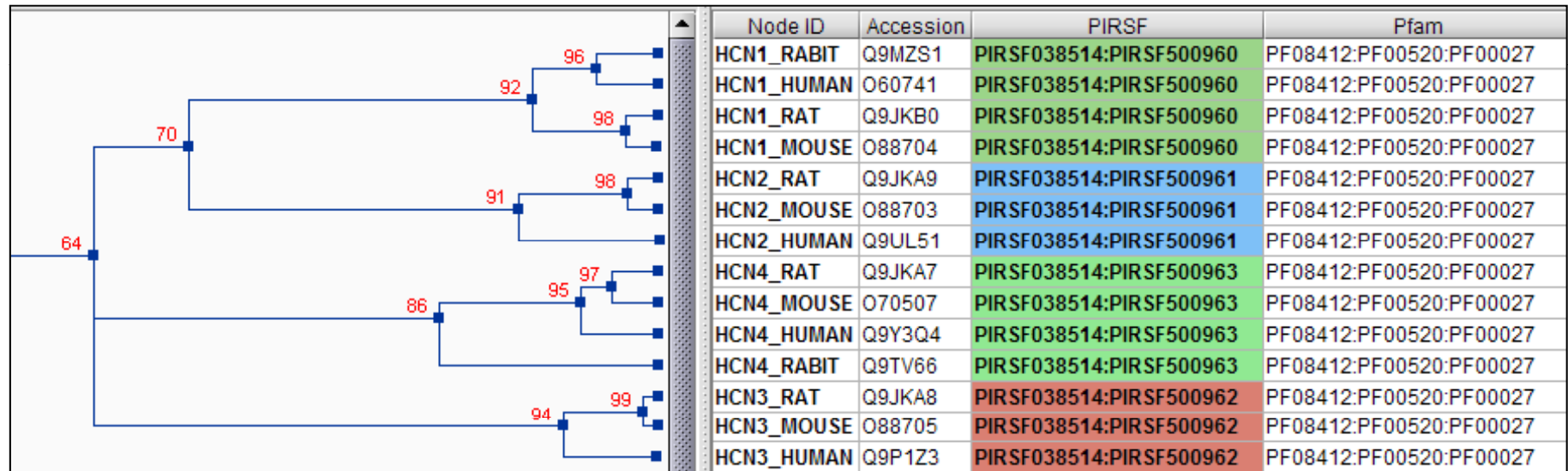
OBO Foundry ontologies				
<u>Title</u>	<u>Domain</u>	<u>Prefix</u>	<u>File</u>	
Biological process	biological process	GO	gene_ontology_edit.obo	
Cellular component	anatomy	GO	gene_ontology_edit.obo	
Chemical entities of biological interest	biochemistry	CHEBI	chebi.obo	
Molecular function	biological function	GO	gene_ontology_edit.obo	
Phenotypic quality	phenotype	PATO	quality.obo	
Protein Ontology (PRO)	proteins	PR	pro.obo	
Xenopus anatomy and development	anatomy	XAO	xenopus_anatomy.obo	
Zebrafish anatomy and development	anatomy	ZFA	zebrafish_anatomy.obo	

PRO Overview: ProEvo

PRO in OBO Foundry to represent protein entities

Three sub-ontologies to connect protein types necessary to model biology












- **Ontology for Protein Evolution (ProEvo):** Captures protein classes reflecting evolutionary relatedness of whole proteins



<input type="checkbox"/>	PR:000000676	potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel protein	family
<input type="checkbox"/>	PR:000000705	potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 1	gene
<input type="checkbox"/>	PR:000000706	potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 2	gene
<input type="checkbox"/>	PR:000000707	potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 3	gene
<input type="checkbox"/>	PR:000000708	potassium/sodium hyperpolarization-activated cyclic nucleotide-gated channel 4	gene

PRO Overview: ProForm

- **Ontology for Protein Forms (ProForm):** Captures different protein forms of a given gene locus from genetic variations, alternative splicing, proteolytic cleavage, post-translational modifications (PTMs)

	PR:000002184 <i>Bcl2 antagonist of cell death</i>	gene	
	PR:000002280 <i>Bcl2 antagonist of cell death isoform 1</i>	sequence	
	PR:000003084 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated form</i>	modification	Q35147 ; Q61337 ; Q92934
	PR:000003085 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 1</i>	modification	Q35147-1 ; Q61337-1
	PR:000003086 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 2</i>	modification	
	PR:000003087 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 3</i>	modification	Q61337-1 : pS112 / pS136
	PR:000003233 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 4</i>	modification	Q61337-1 : pS112 / pS136 / pS155
	PR:000003238 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 5</i>	modification	Q35147-1 : pS112
	PR:000003269 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 6</i>	modification	Q61337-1 : pT201
	PR:000025849 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 7</i>	modification	Q61337-1 : pS136
	PR:000025850 <i>Bcl2 antagonist of cell death isoform 1 phosphorylated 8</i>	modification	Q61337-1 : pS128 / pS136

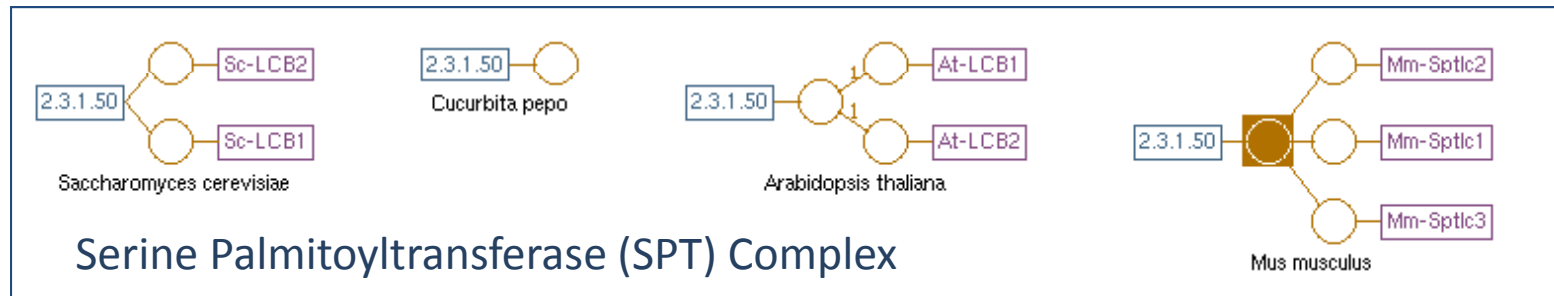
Need for Representing Proteins Forms

Alternative Splicing/Genetic Variation/PTM/Cleavage

Gene	Protein Form	Distinctive Attributes	
SMAD2	Long isoform phosphorylated (PR:000000468)	NOT has_function GO:00036677 DNA binding	Function
	Short isoform phosphorylated (PR:000000469)	has_function GO:00036677 DNA binding	
CUL1	Unmodified form (PR:000002507)	NOT part_of GO:0019005 SCF ubiquitin ligase complex	Association
	Needylated form (PR:000000542)	part_of GO:0019005 SCF ubiquitin ligase complex	
CD14	Membrane form (PR:000002149)	located_in GO:0005886 plasma membrane	Localization
	Soluble form (PR:000002147)	located_in GO:0005615 extracellular space	
ROCK1	Full length (PR:000002529)	has_function GO:0004674 protein serine/threonine kinase activity	Modification
	Cleaved form (PR:000000563)	Increased has_function GO:0004674 protein serine/threonine kinase activity	
CREBBP	Variant R -> P(1378) (PR:000000266)	agent_in MIM:180849, RUBINSTEIN-TAYBI SYNDROME SO:1000118, loss_of_function_of_polypeptide	Disease

PRO Overview: ProComp

- **Ontology for Protein Complexes (ProComp):** Captures **distinct complexes** as they exist in different species and defines complexes through **component proteins**

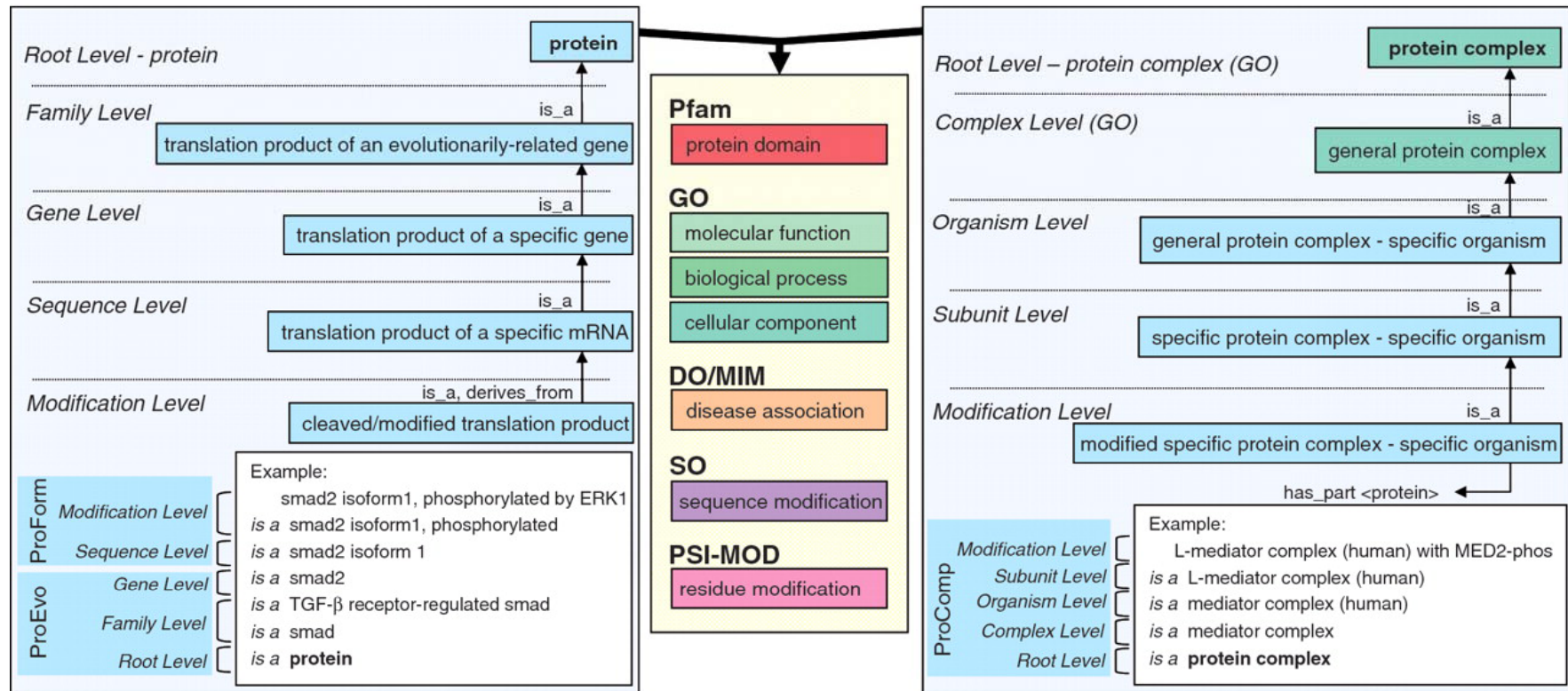


GO:0002178	palmitoyltransferase complex	
GO:0002179	homodimeric serine palmitoyltransferase complex	
PR:000026130	bacterial serine palmitoyltransferase complex (<i>Sphingomonas paucimobilis</i>)	organism-complex
PR:000026132	bacterial serine palmitoyltransferase complex (<i>Sphingomonas wittichii</i>)	organism-complex
PR:000026169	bacterial serine palmitoyltransferase complex (<i>Sphingobacterium multivorum</i>)	organism-complex
GO:0031211	endoplasmic reticulum palmitoyltransferase complex	
GO:0017059	serine C-palmitoyltransferase complex	
PR:000026146	serine palmitoyltransferase complex 3 (human)	organism-complex
PR:000026153	serine palmitoyltransferase complex 5 (human)	organism-complex
PR:000026145	serine palmitoyltransferase complex core 2 (human)	organism-complex

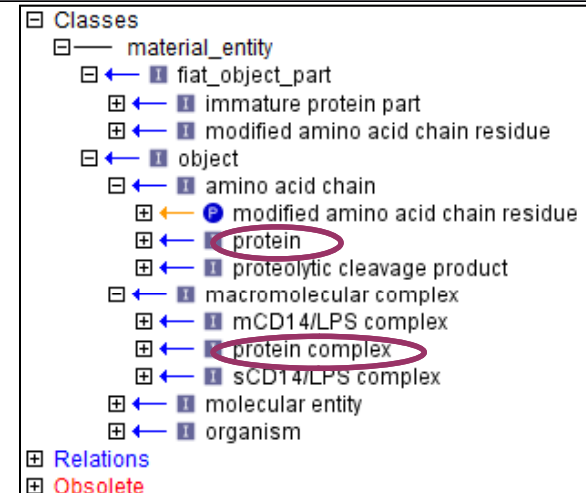
Why PRO

- Provides formalization to support **precise annotation** of specific protein classes/forms/ complexes, allowing accurate and consistent data mapping, integration and analysis
- Allows **specification of relationships** between PRO and other ontologies, such as GO, SO (Sequence Ontology), PSI-MOD, ChEBI, MIM/Disease Ontology, CL (Cell Ontology)
- Provides **stable unique identifiers** to distinct protein types
- Provides a formal structure to support computer-based **reasoning** based on homology and shared protein attributes, including “ortho-isoform,” “ortho-modified form”

PRO Framework



- PRO (ProForm, ProEvo, ProComp) is aligned with other OBO Foundry ontologies under the umbrella of the **Basic Formal Ontology (BFO)**
- PRO terms are defined/annotated using other ontologies and resources via definition of relations or mappings when appropriate



[illegible]

Toll-like receptors PRO Terms

- *Toll-like receptors (TLR)*: 24 gene-level terms with PANTHER mapping of 12 reference genomes
- 25 terms for protein forms and complexes; more forthcoming

PRO Home

search

PRO Name

OR

PRO Name

+ add input box

- del input box

Toll-like receptor

TLR

Display Options

52 entries | 2 pages | 50 / page | 1 | 2

Save

50 selected

show

click to show: selected

Hierarchy

selected

OBO / PAF

OR related

OBO / PAF

Cytoscape view

PRO ID	PRO Name	PRO Term Definition	Category	Parent
<input checked="" type="checkbox"/> GO:0035354	Toll-like receptor 1-Toll-like receptor 2 protein complex	A heterodimeric protein complex containing Toll-like receptor 1 (TLR1) and Toll-like receptor 2 (TLR2). [GOC:add, PMID: 17889651]		GO:0043234
<input checked="" type="checkbox"/> GO:0035355	Toll-like receptor 2-Toll-like receptor 6 protein complex	A heterodimeric protein complex containing Toll-like receptor 2 (TLR2) and Toll-like receptor 6 (TLR6). [GOC:add, PMID: 19931471]		GO:0043234
<input checked="" type="checkbox"/> PR:000001096	toll-like receptor	A protein with a core domain composition consisting of a signal peptide, an extracellular domain with multiple Leu-rich repeats (LRR), a cysteine-rich region, a single-pass transmembrane domain and a C-terminal cytoplasmic tail containing a TIR domain (PF01582). [PMID: 11022119]	family	PR:000000001
<input checked="" type="checkbox"/> PR:000001151	toll-like receptor 1	A toll-like receptor that is a translation product of the TLR1 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001152	toll-like receptor 10	A toll-like receptor that is a translation product of the TLR10 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001153	toll-like receptor 2	A toll-like receptor that is a translation product of the TLR2 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001154	toll-like receptor 3	A toll-like receptor that is a translation product of the TLR3 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001155	toll-like receptor 4	A toll-like receptor that is a translation product of the TLR4 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001156	toll-like receptor 5	A toll-like receptor that is a translation product of the TLR5 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001157	toll-like receptor 6	A toll-like receptor that is a translation product of the TLR6 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001158	toll-like receptor 7	A toll-like receptor that is a translation product of the TLR7 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001159	toll-like receptor 8	A toll-like receptor that is a translation product of the TLR8 gene or a 1:1 ortholog thereof. [PRO:CNA]	gene	PR:000001096
<input checked="" type="checkbox"/> PR:000001160	toll-like receptor 9	A toll-like receptor that is a translation product of the TLR9 gene or a 1:1 ortholog thereof. Non-methylated CpG motifs, present in viral and bacterial DNA, are one of many pathogen-associated molecular patterns (PAMP) recognized	gene	PR:000001096

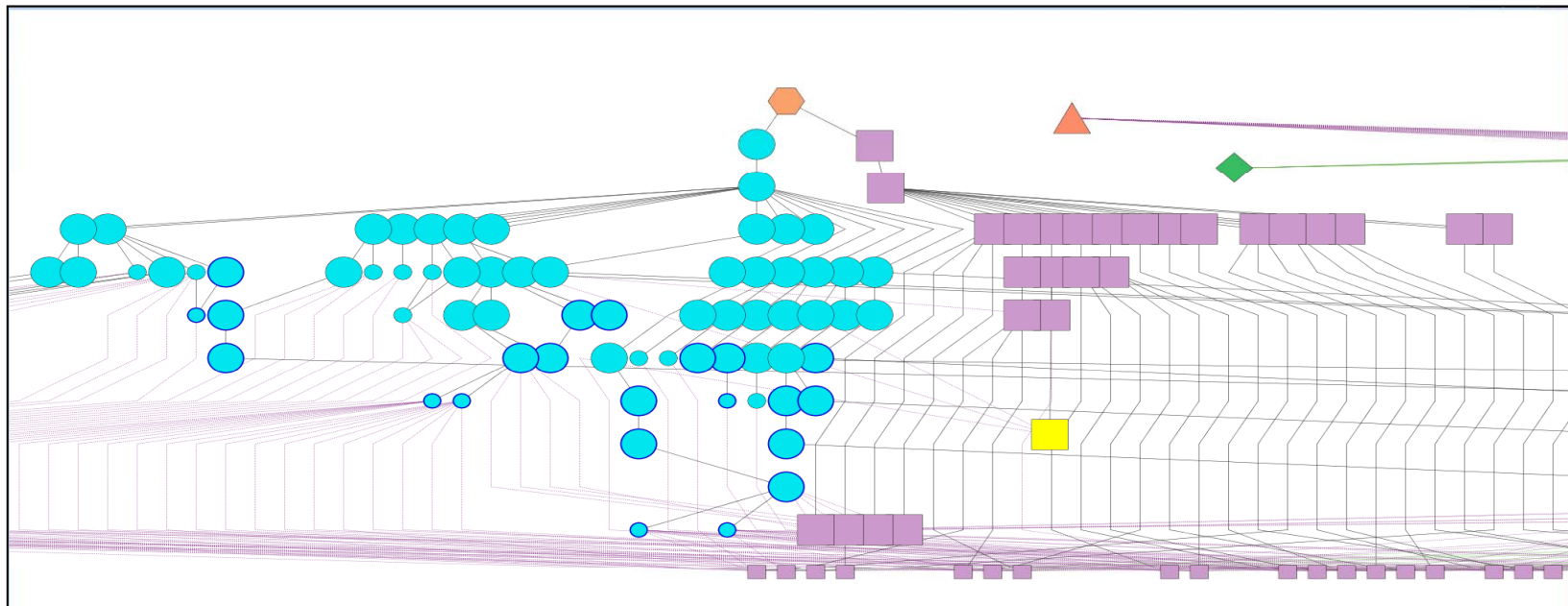
Hierarchical View: TLR4 Complex-Related Terms

PR:000000001	protein	
PR:000001096	toll-like receptor	family
PR:000001155	toll-like receptor 4	gene
UniProtKB:C4PCF3	Toll-like receptor (chicken)	organism-gene
UniProtKB:B8A691	Toll-like receptor (zebrafish)	organism-gene
UniProtKB:O00206	Toll-like receptor 4 (human)	organism-gene
UniProtKB:Q9QUK6	Toll-like receptor 4 (mouse)	organism-gene
UniProtKB:Q9QX05	Toll-like receptor 4 (rat)	organism-gene
PR:000025490	toll-like receptor 4 isoform 1	sequence
PR:000025492	toll-like receptor 4 isoform 1 cleaved and glycosylated 1	modification
PR:000025787	toll-like receptor 4 isoform 1 cleaved and glycosylated 1 (human)	organism-modification
PR:000027172	toll-like receptor 4 isoform 1 cleaved and glycosylated 1 (mouse)	organism-modification
PR:000026585	toll-like receptor 4 isoform 1, signal peptide removed form	modification
PR:000025491	toll-like receptor 4 isoform 2	sequence
PR:000018358	toll-like receptor 4, full-length form	modification
PR:000018359	toll-like receptor 4, signal peptide removed form	modification
GO:0032991	macromolecular complex	
GO:0043234	protein complex	
GO:0043235	receptor complex	
PR:000025774	MyD88:Mal:activated TLR complex	complex
PR:000025784	MyD88:Mal:activated TLR4 receptor	complex
PR:000025785	MyD88:Mal:activated TLR4 receptor (human)	organism-complex
PR:000027175	MyD88:Mal:activated TLR4 receptor (mouse)	organism-complex
PR:000027174	MyD88s:Mal:activated TLR4 receptor (mouse)	organism-complex
PR:000027193	IRAK4:MyD88:Mal:activated TLR complex	complex
PR:000027182	IRAK4:MyD88:Mal:activated TLR4	complex
PR:000027183	IRAK4:MyD88:Mal:activated TLR4 (human)	organism-complex
PR:000027187	IRAK4:MyD88:Mal:activated TLR4 (mouse)	organism-complex
PR:000027192	pIRAK4:MyD88:Mal:activated TLR4 (human)	organism-complex
PR:000027194	Mal:active TLR complex	complex
PR:000027195	Mal:activated TLR4 complex	complex
PR:000027196	Mal:activated TLR4 (mouse)	organism-complex
PR:000027217	pMal:activated TLR4 complex (mouse)	organism-complex
PR:000027218	pMal:activated TLR4 complex(human)	organism-complex
PR:000027202	ticam2:activated TLR4 complex	complex
PR:000028678	ticam2:activated TLR4 complex (human)	organism-complex
PR:000027204	ticam2:activated TLR4 complex (mouse)	organism-complex
PR:000027205	ticam1:ticam2:activated TLR4 complex	complex
PR:000027208	ticam1:ticam2:activated TLR4 complex (human)	organism-complex

Cytoscape Network View: TLR4 Complex-Related Terms

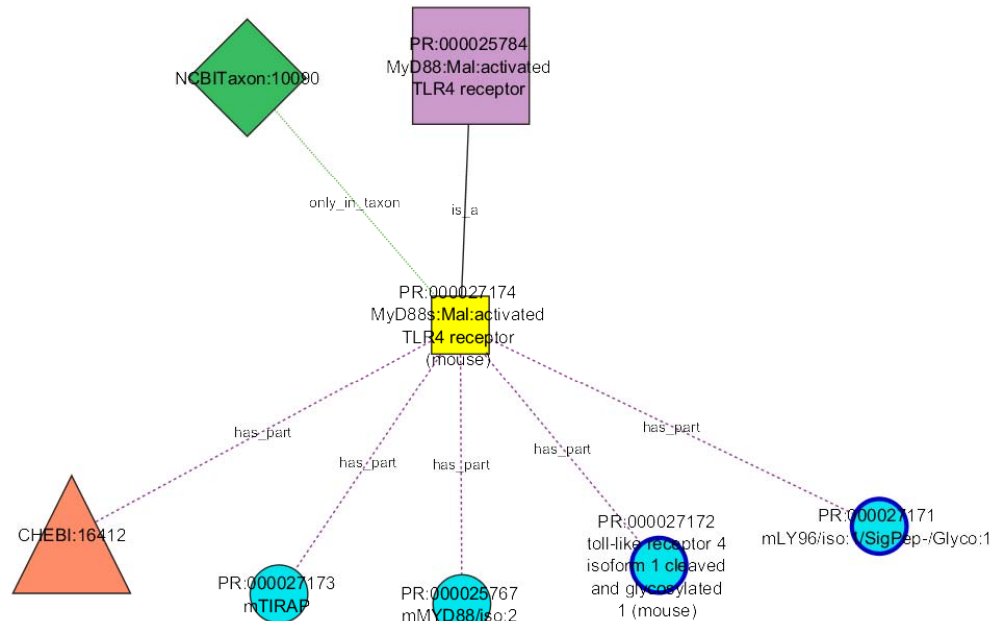
Connecting protein forms and complexes with annotation => Modeling biology

- Show full set of relationships between terms, including those from other ontologies
- Traverse hierarchical (parent-child) relationships
- Find all components of a complex, including modified forms
- Find all complexes that a protein is constituent of = first (complex) neighbors of a protein
- Find all complexes that bind small molecules = those that have relationship to ChEBI



Cytoscape & Entry Views

Connecting protein forms and complexes with annotation



Data Panel						
ID	Relation	Ontology_ID	Ontology_term	Evidence_source	Evidence_code	Taxon
PR:000027174	participates_in	GO:0034142	toll-like receptor 4 signaling pathway	PMID:12538665	EXP	10090

PRO Home Protein Ontology report for entry - PR:000027174 Show OBO stanza Retrieve related PRO nodes Save related nodes in OWL format	
Ontology Information	
PRO ID	PR:000027174
PRO name	MyD88s:Mal:activated TLR4 receptor (mouse)
Synonyms	
Definition	A MyD88:Mal:activated TLR4 receptor in mouse. This complex contains the short isoform of Myd88. [PMID:12538665, TLR:AMM]
Comment	Category=organism-complex.
Hierarchical relationship	Parent: PR:000025784 MyD88:Mal:activated TLR4 receptor Children: none has_part CHEBI:16412 {cardinality="2"} lipopolysaccharide has_part PR:000025767 {cardinality="1"} myeloid differentiation primary response protein MyD88 isoform 2 (mouse) has_part PR:000027171 {cardinality="2"} lymphocyte antigen 96 isoform 1, signal peptide removed glycosylated 1 (mouse) has_part PR:000027172 {cardinality="2"} toll-like receptor 4 isoform 1 cleaved and glycosylated 1 (mouse) has_part PR:000027173 {cardinality="1"} toll/interleukin-1 receptor domain-containing adapter protein (mouse) only_in_taxon NCBITaxon:10090 Mus musculus ⓘ (click to see DAG view.)

PRO Workflow

■ Data Sources

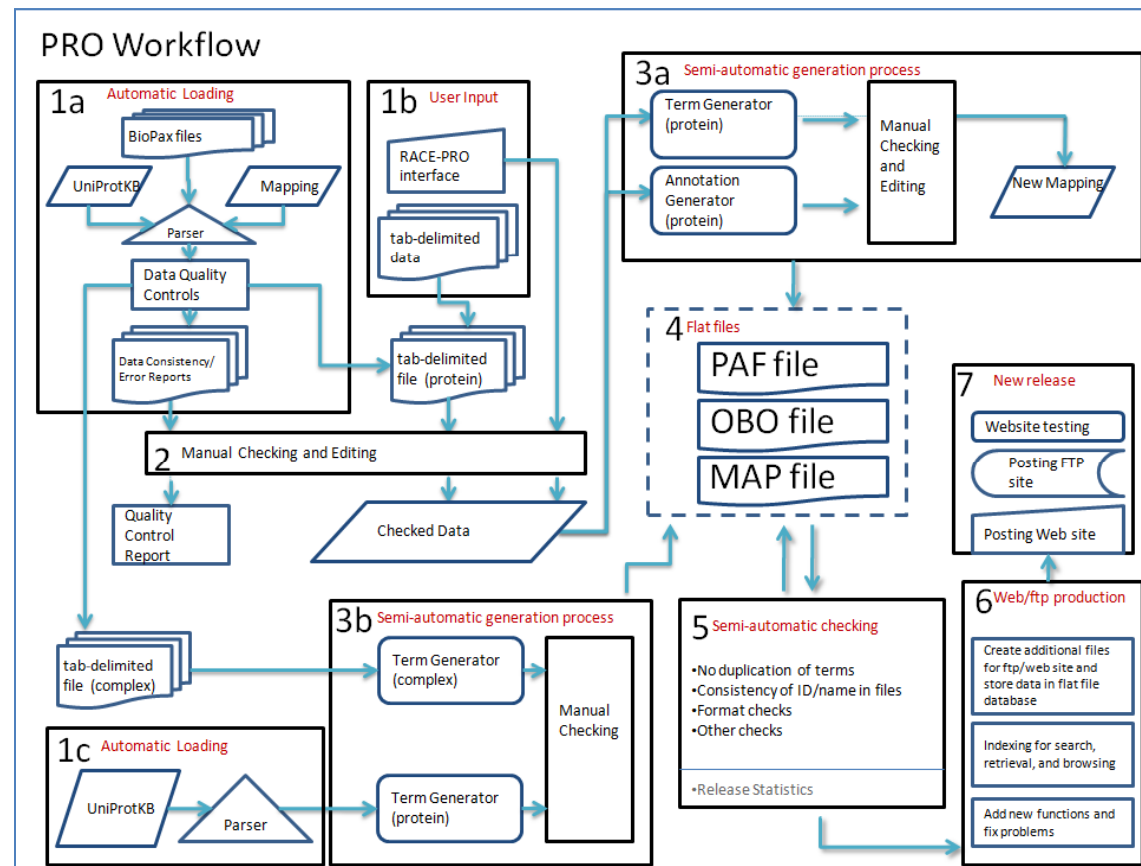
- Manual annotation (curator, collaborator, user): sourceforge tracker; Race-PRO
- Semi-automated processing of external databases (e.g., UniProtKB, Reactome, MouseCyc, EcoCyc); coverage of 12 reference genomes in progress

■ Integration with text mining tool: eFIP (*F*unctional *I*mpact of *P*hosphorylation)

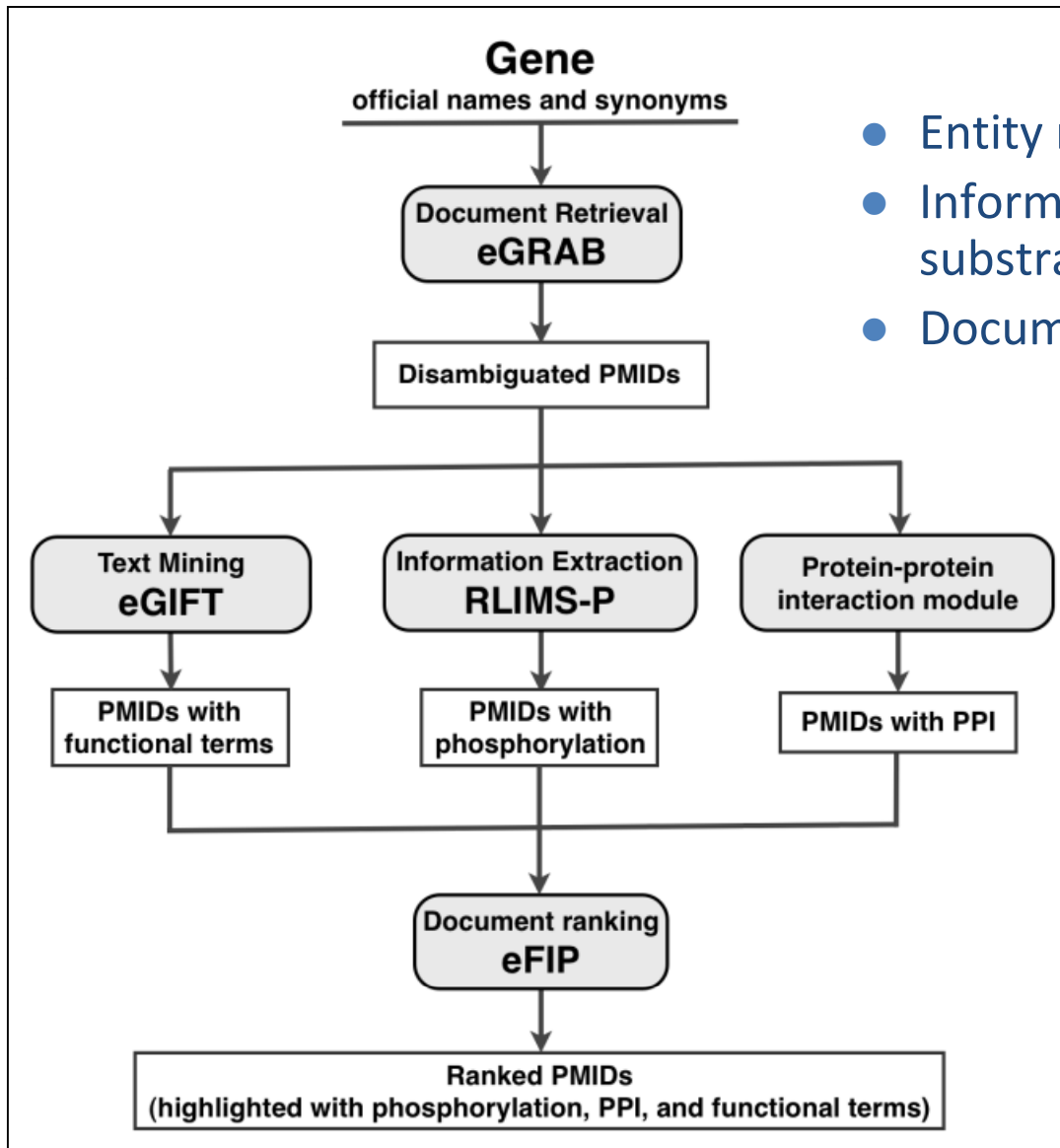
■ Distribution Files

- Ontology (OBO)
- Annotation (PAF)
- Mappings (exact; is_a)

sourceforge



eFIP: an integrated system for mining Functional Impact of Phosphorylation from literature



- Entity recognition and document retrieval
- Information extraction (RLIMS-P: kinase, substrate, phosphorylation site)
- Document ranking and evidence tagging

eFIP: a tool for mining functional impact of phosphorylation from literature

Arighi CN, Siu AY, Tudor CO, Nchoutmboube JA, Wu CH, Shanker VK. (2011)

Methods in Molecular Biology 694, 63-75
[PMID:21082428]

Discovery from Literature Mining

- Distinct phosphorylated forms of a protein may have different interacting proteins, leading to different subcellular locations, functions and pathways
- iPTMnet: PTM network of enzyme-substrate relationships and protein-protein interactions from text mining and data mining => Knowledge captured in PRO

PMID 10837486 for gene BAD - Bcl2-associated agonist of cell death

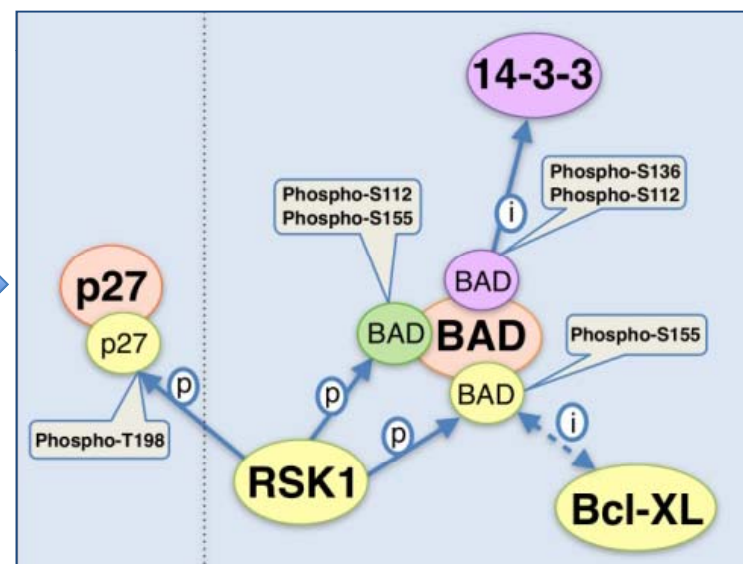
Predicted impact of phosphorylation:

Sentence #	Phosphorylation			Impact
	Substrate	Site	Kinase	
1	BAD	Ser-155	RSK1	regulates BAD/Bcl-XL interaction regulates cell survival
3,4	BAD	Ser-112 and Ser-136	N/A	promotes binding of BAD to 14-3-3 proteins
6	BAD	Ser-155	RSK1	blocking the binding of BAD to Bcl-XL
7	BAD	both Ser-112 and Ser-155	RSK1	rescues BAD -mediated cell death

- ☒ Tag substrate
- ☒ Tag kinase
- ☒ Tag phosphorylation site
- ☒ Tag protein-protein interaction
- ☒ Tag functional term

Text of title and abstract:

Sentence #	Sentence
1	TI - BAD Ser-155 PHOSphorylation regulates BAD/Bcl-XL interaction and cell survival .
2	AB - The BH3 domain of BAD mediates its death-promoting activities via heterodimerization to the Bcl-XL family of death regulators .
3	Growth and survival factors inhibit the death-promoting activity of BAD by stimulating PHOSphorylation at multiple sites including Ser-112 and Ser-136 .
4	PHOSphorylation at these sites promotes binding of BAD to 14-3-3 proteins , sequestering BAD away from the mitochondrial membrane where it dimerizes with Bcl-XL to exert its killing effects .
5	We report here that the phosphorylaTION of BAD at Ser-155 within the BH3 domain is a second PHOSphorylation -dependent mechanism that inhibits the death-promoting activity of BAD .
6	Protein kinase A , RSK1 and survival factor signaling stimulate PHOSphorylation of BAD at Ser-155 , blocking the binding of BAD to Bcl-XL .
7	RSK1 phosphorylaTES BAD at both Ser-112 and Ser-155 and rescues BAD -mediated cell death in a manner dependent upon PHOSphorylation at both sites .



ABI Development: Integrative Bioinformatics for Knowledge Discovery of PTM Networks (<http://www.nsf.gov/awardsearch/showAward.do?AwardNumber=1062520>)



User Annotation: RACE-PRO

Capture knowledge of protein forms/complexes of interest to support integrated analysis and computer-based reasoning

- ❖ Obtain a PRO ID for the protein objects of interest
- ❖ Define a protein object (based on literature, experimental data)
- ❖ Add annotation to that protein object
- ❖ How does it work?
 - ❖ Input your personal information (only for internal use)
 - ❖ Complete form with sequence information and annotation
 - ❖ Submit when ready (otherwise you can save for later)
 - ❖ PRO curation team will take the data, revise it, and create the corresponding PRO node in the ontology
 - ❖ User will be informed via email about the new PRO IDs and when they will be public

RACE-PRO Annotation Tool

Definition of the Protein Object

1. Enter a UniProtKB identifier (?)
OR, click [here](#) to insert a different sequence:

```

MTRDFKPGDL IFAKMKGYPH WPARVDEVDP GAVKPPFNKL PIFFFGTHET AFLGPKDIFP 60
YSENKEYGK PNKRKGFNEG LWEIDNNPKV KFSSQQAATK QSNASSDVEV EEKETSVSKK 120
D'TDHEEKASN EDVTKAVDIT TPKAARRGRK RKAERQVETE EAGVVTATA SVNLRVSPKR 180
GRPAATEVKI PKPRGRPKMV KQPCPSESDI ITEEDKSKKK GQEEKQPKKQ PKKDEEGQKE 240
EDKPRKEPDK KEGKKEVESK RKNLAKTGV TSTDSSEEGD DQEGEKRRKG GRNFQTAHRR 300
NMLKGQHEKE AADRKRKQEE QMETEHQTT NLQ
    
```

Organism:

2. Specify sequence region
☐ Full-length ☒ Region: from to
3. Indicate post-translational modifications (add amino acid number re
Amino acid number: --choose PTM--

4. Protein Object name (separate multiple names using ";")

5. Evidence Source (separate multiple IDs using
Db name: IDs:

A-Enter accession or
paste sequence

B-Define protein region
and/or PTMs

C-Enter protein form name

D- Data source

Annotation of the Protein Object

Domain

Modifier	Relation	Pfam ID	Pfam name	PMIDs
NOT	has_part	PF00855	PWWP domain	18708362

Functional Annotation

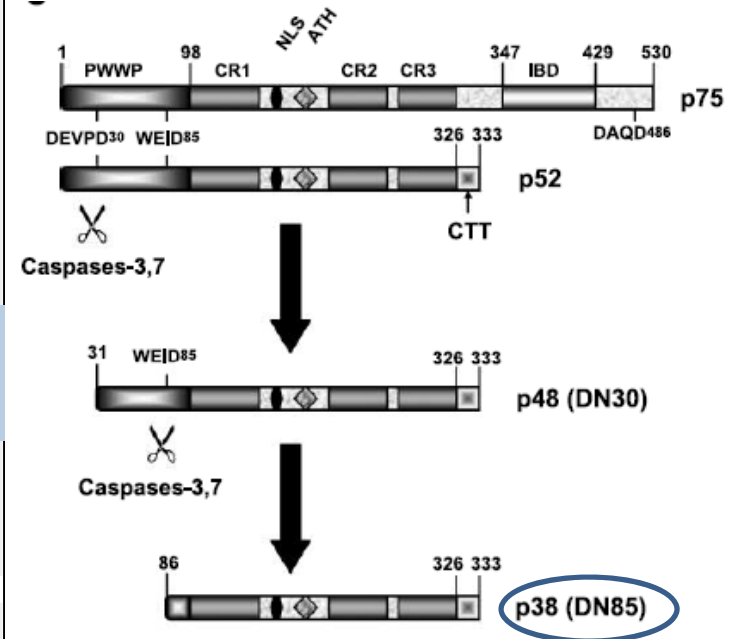
Modifier	Relation	GO ID	GO term	Interaction with
	located_in	GO:00056	nucleus	
	participates_in	GO:00164	negative regulation of tra	

Sequence Ontology

Disease

E- Annotation

PMID:18708362



Comments:

WB: anti-HA

C

HcRed-p75

HcRed-p52

p38/HA-DN85

D

Sticky Note

4/6/2009 3:41:37 PM

Options

arigic

localization:nucleus

HcRed-p75

HcRed-p52

p38/HA-DN85

Merged

Merged

PRO Dissemination

- PRO Website (<http://www.proconsortium.org>)
 - Searching, browsing, downloading
- PRO Views
 - Entry view
 - Table summary
 - OBO stanza, OWL
 - Ontology hierarchy
 - Cytoscape network
- PRO Link: Persistent URL: http://purl.obolibrary.org/obo/PR_xxxxxxxxx
- OBO Foundry (<http://www.obofoundry.org/>)
- NCBO Bioportal (<http://bioportal.bioontology.org/>)

PRO Communities

- Ontology Developers
 - GO ontology: Interfaces of GO/PRO complexes; GO definition (e.g., GO:0005109)
 - GO annotation: precise annotation of protein forms in PomBase
 - Dendritic Cell Ontology: Define cell types based on +/- protein types [PMID:19243617]
 - Annotation Ontology for annotating scientific documents on the web [PMID:21624159]
 - Brucellosis Ontology (IDOB RU), extension of the Infectious Disease Ontology (IDO) [PMID:22041276]
- Semantic Resources
 - Royal Society of Chemistry (RSC); Science Collaboration Framework; Semantic Web Applications in Neuromedicine (SWAN); Neuroscience Information Framework (NIF)
- Pathway/Process-Modeling Resources:
 - Reactome, MouseCyc, EcoCyc/BioCyc, Center for Molecular Immunology (Duke)
- Molecule-Modeling Resources: Int'l Union of Basic and Clinical Pharmacology (IUPhar)
- Pharma/Clinical Communities: Drug Discovery & Disease Biomarker
 - Alzforum
 - Salivaomics KB/SALO (Saliva Ontology): Saliva Biomarkers
 - Clinical flow cytometry, immunology, infectious disease community

AlzForum Driving Project

- Aids AlzForum as a comprehensive source of information about Alzheimer's disease (AD), linking to the wider biomedical knowledge via community-accepted ontologies
- Clinical driving project to address the needs of AD researchers: identify gaps in PRO for representing AD-related protein entities, especially those involved in etiology
- AD-related protein entities include various types of **proteolytic cleavage products** and **protein complexes**: Drives new terms/relationships: (i) **proteolytic cleavage product** (PR:000018264); (ii) **union_of** terms (PR:000025744 = set of 2 precursors)
- Coverage: (i) amyloid beta A4 protein (APP) (variants and cleavage products); (ii) microtubule-associated protein Tau (mutations and population variants); (iii) secretases (complexes)
- Alzforum/PRO Kick-Off Meeting: October 4-5, 2011, Buffalo, NY

Classes
material_entity
fiat_object_part
immature protein part
modified amino acid chain residue
object
amino acid chain
modified amino acid chain residue
protein
proteolytic cleavage product
macromolecular complex
mCD14/LPS complex
protein complex
sCD14/LPS complex
molecular entity
organism

expand	sort (ID)	sort (STR)	find	Category
PR:000018263				amino acid chain
PR:000000001				protein
PR:000004168				amyloid beta A4 protein
PR:000019036				amyloid beta A4 protein proteolytic cleavage product
PR:000025744				precursor of amyloid beta A4 protein gamma-secretase C-terminal fragment
PR:000025589				gamma-secretase C-terminal fragment 50
PR:000025585				gamma-secretase C-terminal fragment 57
PR:000025586				gamma-secretase C-terminal fragment 59

Flow Cytometry Driving Project

- How the use of clinical flow cytometry data can serve as a driver of ontology development in both the PRO and the Cell Ontology (CL) by assessing current standard clinical assays and recent approaches based on automated gating of multidimensional flow cytometry
- Representation of hematopoietic and other cell types in the CL to allow for the logical definition of cell types based on cellular attributes; PRO as a key component of these definitions
- Which protein isoforms, PTM forms, complex component proteins identified by flow cytometry typing reagents need to be represented in the PRO to enable cell types defined in their terms to be represented in the CL?
- How can use of the PRO and CL ontologies promote standardization in interpretation and integration of clinical flow cytometry data?
- Representation requirements for flow cytometry in PRO, CL, IEDB, OBI, ImmPort, and Immune System Modeling
- Initial PRO terms: IDOBRU (314), TLR (74), CL (176)

PRO Consortium Team (current)

Protein Information Resource (PIR) [Georgetown U & U Delaware]

Cathy Wu, Cecilia Arighi, Darren Natale, Natalia Roberts
Hongzhan Huang, Jian Zhang



The Jackson Lab – Mouse Genome Informatics (MGI)

Judith Blake, Carol Bult, Harold Drabkin, Alexei Evsikov



University at Buffalo-SUNY

Barry Smith, Alan Ruttenberg, Alexander Diehl

NYU School of Medicine – Reactome

Peter D'Eustachio, Veronica Shamovsky



AlzForum

Elizabeth Wu



1R01GM080646-01
3R01GM080646-04S2
2R01GM080646-06

PRO: A Protein Ontology in OBO Foundry for Integration of Biomedical Knowledge
(http://projectreporter.nih.gov/project_info_description.cfm?aid=8187900&icde=10696318)