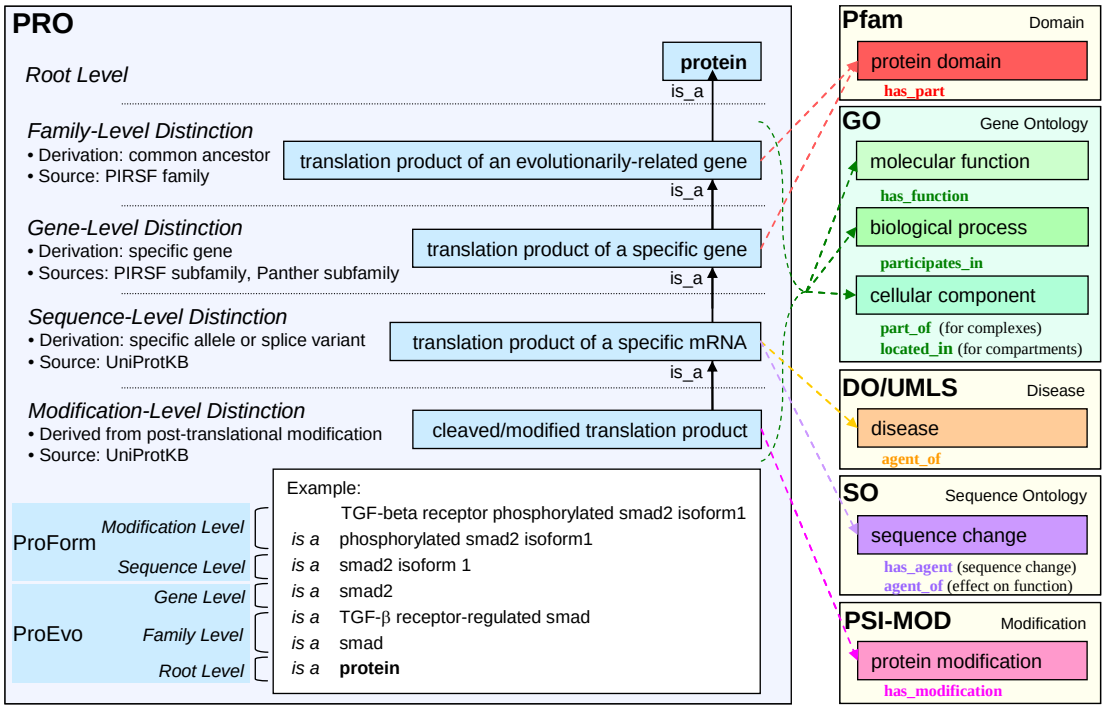
PIR
Protein Information Resource

PRO: *PR*otein *O*ntology

(http://pir.georgetown.edu/pro)

PRO framework

We have designed a PRotein Ontology (PRO) that provides a framework to define protein entities and the relations between them. The components of PRO extend from the classification of proteins on the basis of evolutionary relatedness (ProEvo) to the representation of the multiple protein forms of a gene (ProForm). PRO can be used to annotate the attributes of each entity, map objects in pathways, and model biological system dynamics and disease. Development of PRO is illustrated using transforming growth factor (TGF)-beta signaling proteins.

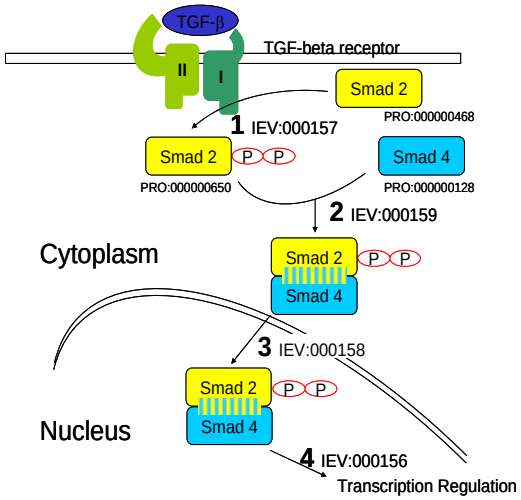


PRO and the TGF-beta signaling pathway

Smad proteins are essential components of serine/threonine kinase receptor signaling pathways that are regulated by phosphorylation. The figure on the right depicts the TGF-beta signaling pathway, focusing on the Smad 2 component. The steps shown here (described by INOH event ontology, IEV) are preceded by TGF-beta binding to the receptor, and receptor phosphorylation.

Step 1: Phosphorylation of Smad 2 by TGF beta receptor I.
Step 2: Complex formation of R-smad and Smad4.
Step 3: Nuclear import of R-smad:Smad4.
Step 4: Binding of R-smad:Smad 4 complex coactivator to responsive element.

The PRO ID describes the specific entity involved in the pathway (next page).



PRO Sample Report- Smad 2 protein

$**PRO:000000001** protein
%**PRO:000000027** Smad protein { PIRSF037286 }
has_part **Pfam:PF03165** MH1 domain
has_part **Pfam:PF03166** MH2 domain
participates_in **GO:0007165** signal transduction [ISS]
participates_in **GO:0006355** regulation of transcription, DNA-dependent [ISS]
has_function **GO:0005515** protein binding [ISS]
%**PRO:000000066** TGF-beta receptor-regulated Smad protein { PIRSF500455 }
participates_in **GO:0007179** transforming growth factor beta receptor signaling pathway [ISS]
participates_in **GO:0007183** SMAD protein complex assembly [ISS]
%**PRO:000000364** Smad2 { Q62432 (mouse), Q15796 (human) }
%**PRO:000000468** Smad2 isoform 1 (long form) { Q62432-1 (mouse), Q15796-1 (human) }
has_function **GO:0005102** receptor binding [PMID:8980228, TaxID:9606]
participates_in **GO:0007179** transforming growth factor beta receptor signaling pathway [PMID:8752209, TaxID:9606; PMID:11557747, TaxID:10090]
located_in **GO:0005737** cytoplasm [PMID:15280432, TaxID:9606; PMID:14701940, TaxID:10090]
>**PRO:000000574** Smad2 sequence 1 phosphorylated form
has_modification **MOD:00696** phosphorylated residue
% **PRO:000000650** Smad2 isoform 1 phosphorylated 1 (TGF-beta receptor I-phosphorylated) { Q15796-1, phosphorylated, Ser-465/Ser-467 }
has_modification **MOD:00046** O-phosphorylated L-serine
has_function **GO:0030618** transforming growth factor beta receptor, pathway-specific cytoplasmic mediator activity [PMID: 8980228, TaxID:9606; PMID:9346966, TaxID:9606]
has_function **GO:0046332** SMAD binding [PMID:9311995, TaxID:9606; PMID:11043574, TaxID:10090]
has_function **GO:0003713** transcription coactivator activity [PMID:9689110, TaxID:9609; PMID:15282343, TaxID:10090]
participates_in **GO:0007179** transforming growth factor beta receptor signaling pathway [PMID:9873005, TaxID:9606; PMID:15630024; TaxID:10090]
participates_in **GO:0007183** SMAD protein complex assembly[PMID:9346966, TaxID:9606]
participates_in **GO:0006355** regulation of transcription, DNA-dependent [PMID: 9689110, TaxID:9606; PMID:15690394 TaxID:9606]
located_in **GO:0005634** nucleus [PMID:9006934, TaxID:9606; PMID:9346966, TaxID:9606; PMID:11557747, TaxID:10090]
%**PRO:000000651** Smad2 isoform 1 phosphorylated 2 (TGF-beta receptor I and ERK1-phosphorylated) { Q15796-1, phosphorylated, Thr-8/Ser-465/Ser-467 }
has_modification **MOD:00046** O-phosphorylated L-serine
has_modification **MOD:00047** O-phosphorylated L-threonine
has_function **GO:0030618** transforming growth factor beta receptor, pathway-specific cytoplasmic mediator activity
has_function **GO:0046332** SMAD binding
has_function **GO:0003713** transcription coactivator activity [PMID: 12193595, TaxID:9606]
participates_in **GO:0007165** signal transduction [PMID: 12193595, TaxID:9606]
participates_in **GO:0007183** SMAD protein complex assembly[PMID: 12193595, TaxID:9606]
participates_in **GO:0006355** regulation of transcription, DNA-dependent [PMID: 12193595, TaxID:9606]
located_in **GO:0005634** nucleus [PMID: 12193595, TaxID:9606]
part_of **GO:0005667** transcription factor complex [PMID: 12193595, TaxID:9606]
%**PRO:000000652** Smad2 isoform 1 phosphorylated 3 (TGF-beta receptor I and CAMK2-phosphorylated) { Q15796-1, phosphorylated, Ser-240/Ser-465/Ser-467 }
has_modification **MOD:00046** O-phosphorylated L-serine
has_function **GO:0046332** SMAD binding [PMID:11027280, TaxID:9606]
NOT has_function **GO:0003713** transcription coactivator activity [PMID:11027280, TaxID:9606]
participates_in **GO:0007165** signal transduction [PMID:11027280, TaxID:9606]
participates_in **GO:0007183** SMAD protein complex assembly[PMID:11027280, TaxID:9606]
located_in **GO:0005737** cytoplasm [PMID:11027280, TaxID:9606]
located_in **GO:0005769** early endosome when bound to PRO:000000313 [PMID:12356868, TaxID:9606]
%**PRO:000000469** Smad2 isoform 2 (splice variant short form) { Q62432-2 (mouse), Q15796-2 (human) }
has_function **GO:0005102** receptor binding [PMID:9873005, TaxID:9606]
participates_in **GO:0007179** transforming growth factor beta receptor signaling pathway [PMID:9873005, TaxID:9606; PMID:14701940, TaxID:9606; PMID:15630024, TaxID:10090]
located_in **GO:0005737** cytoplasm [PMID:15280432, TaxID:10090]
>**PRO:000000576** Smad2 isoform 2 phosphorylated form
has_modification **MOD:00696** phosphorylated residue
%**PRO:000000656** Smad2 isoform 2 phosphorylated 1 (TGF-beta receptor I-phosphorylated) { Q15796-2, phosphorylated, Ser-436/Ser-438 }
has_modification **MOD:00046** O-phosphorylated L-serine
has_function **GO:0030618** transforming growth factor beta receptor, pathway-specific cytoplasmic mediator activity [PMID:9873005, TaxID:9606; PMID:14701940, TaxID:9606]
has_function **GO:0003677** DNA binding [PMID:9873005, TaxID:9606; PMID:14701940, TaxID:9606]
has_function **GO:0003713** transcription coactivator activity [PMID:9873005, TaxID:9606; PMID:15630024, TaxID:10090]
participates_in **GO:0007179** transforming growth factor beta receptor signaling pathway [PMID:9873005, TaxID:9606; PMID:15630024, TaxID:10090]
%**PRO:000000470** Smad2 sequence variant 1 (genetic variant related to carcinoma of the large intestine) { Q15796, R -> C (133) }
has_agent **SO:1000093** amino_acid_substitution
lacks_modification **MOD:00696** phosphorylated residue
NOT has_function **GO:0003713** transcription coactivator activity [PMID:8752209, TaxID:9606]
agent_of **UMLS:C0009402** carcinoma of the large intestine [PMID:8752209, TaxID:9606]

[a]The symbols preceding each PRO accession are as follows: $ root, % is_a, , > derives_from ,
[b]Text in curly braces indicates the PRO ID, typically derived from the source of the class
[c]Not all examples shown.

PRO can be downloaded from ftp://ftp.pir.georgetown.edu/databases/ontology/pro_obo/

PRO is funded by NIH grant 1 R01 GM080646-01.
Principal Investigators: Cathy Wu, Protein Information Resource, Georgetown University Medical Center; Judith Blake, The Jackson Laboratory; and Barry Smith, Department of Philosophy, State University of New York at Buffalo.

Protein Information Resource
Georgetown University Medical Center
3300 Whitehaven St., NW Suite 1200. Washington, DC 20057-1414, USA
pirmail@georgetown.edu