

○太田元規¹⁾，嘉戸裕美子¹⁾，坂本盛宇²⁾，細田和男³⁾，鹿間周子¹⁾，
大安裕美¹⁾，高木大輔²⁾，安保勲人³⁾，山口敦子^{4,5)}，畠中秀樹⁵⁾，
小池亮太郎¹⁾，廣明秀一⁶⁾，福地佐斗志³⁾

1) 名古屋大学大学院情報学研究科，2) (株) ホロニクス，3) 前橋工科大学工学部，
4) 東京都市大学総合理工学研究科，5) ライフサイエンス統合データベースセンター，
6) 名古屋大学大学院創薬科学研究科

<https://www.ideal-db.org/>

IDEALってなに？

- ・天然変性タンパク質のデータベース
 - 2009年活動開始 (2011年公開)
 - 995 エントリ, 11,643 領域：真核, 核内
 - **世界最大規模**のアノテーションデータベース
- ・天然変性領域中の相互作用領域 (ProS) を収集
 - 317 エントリにProSあり
- ・全データ, **マニュアルアノテーション**
- ・全コンテンツはXML, RDFでダウンロード可能
- ・**新バージョン近日公開** (1,111エントリ)

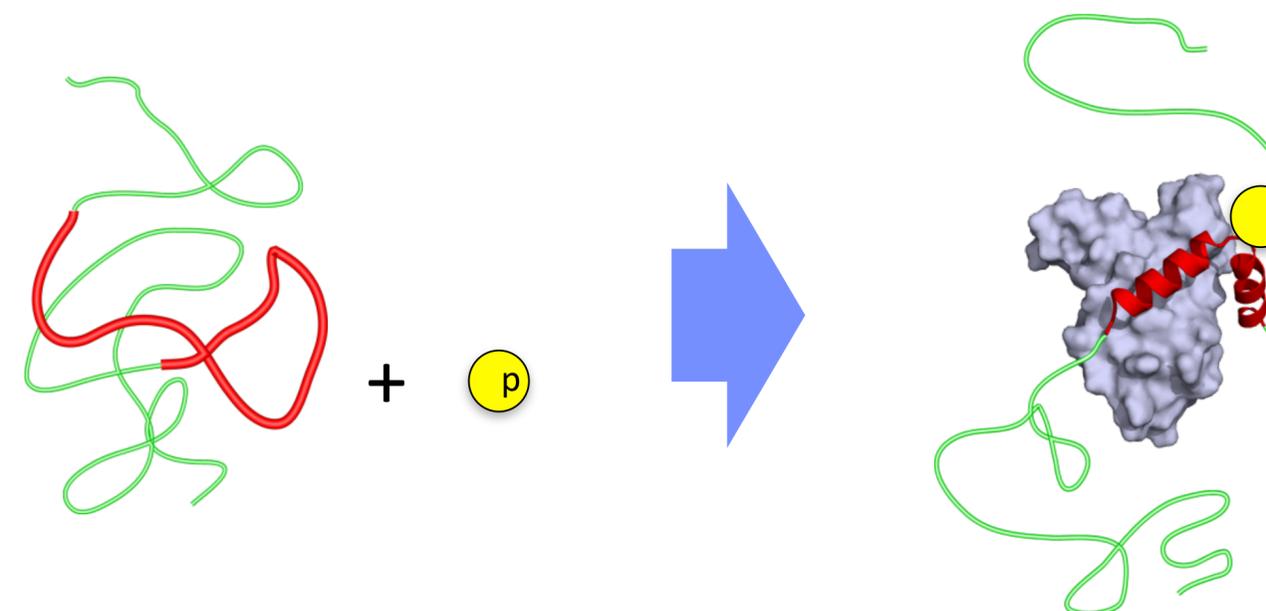


にも参画しています！

IDPcentral

天然変性タンパク質ってなに？

形をとらないタンパク質
 機能：シグナル伝達, 転写制御など
 多くの翻訳後修飾部位を持つ



天然変性領域にある機能部位
 構造を形成し結合

Protean Segment (**ProS**) として収集

IDEALをのぞいてみよう！

エッジ
EDGE

ネットワーク

エントリ

IID00070 Histone acetyltransferase p300 (*Homo sapiens*) Q09472

IID00179 Cop/p300-interacting transactivator 2 (*Homo sapiens*) Q99967

IID00015 Cellular tumor antigen p53 (*Homo sapiens*) P53

NODE

EDGE

クリックしていくうちに

相互作用, 機能が実感できる

SPARQL endpoint 設置

XML → RDF

SPARQL endpoint
クエリによる検索も可能

Virtuoso SPARQL Query Editor

Default Data Set Name (Graph IRI)

Query Text

```

PREFIX up:<http://purl.uniprot.org/core/>
PREFIX ideal:<http://www.ideal.force.cs.is.nagoya-u.ac.jp/IDEAL/>

SELECT ?protein ?idp ?region_start ?region_end ?order_disorder
WHERE
{
  SERVICE <https://integbio.jp/rdf/mirror/alt/uniprot/sparql>{
    ?protein a up:Protein .
    ?protein up:mnemonic 'MYC_HUMAN' .
  }

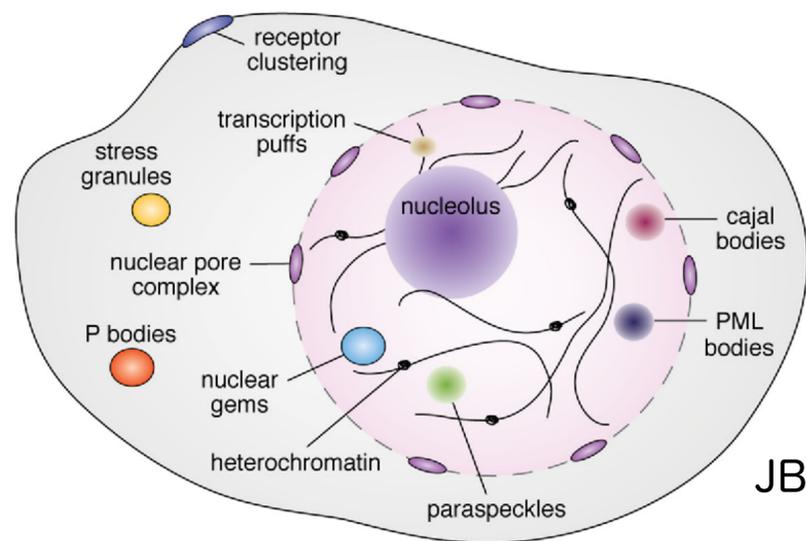
  ?idp ideal:uniprot ?protein .
  ?idp ideal:region ?rid.
  ?rid ideal:region_start ?region_start.
  ?rid ideal:region_end ?region_end.
  ?rid ideal:order_disorder ?order_disorder.
}

```

UniProt, PDBj からリンクされている



液滴を形成する天然変性タンパク質



JBC, 2019, Gomes

IID00012 Myc proto-oncogene protein (*Homo sapiens*) [P01106](#) LLPS

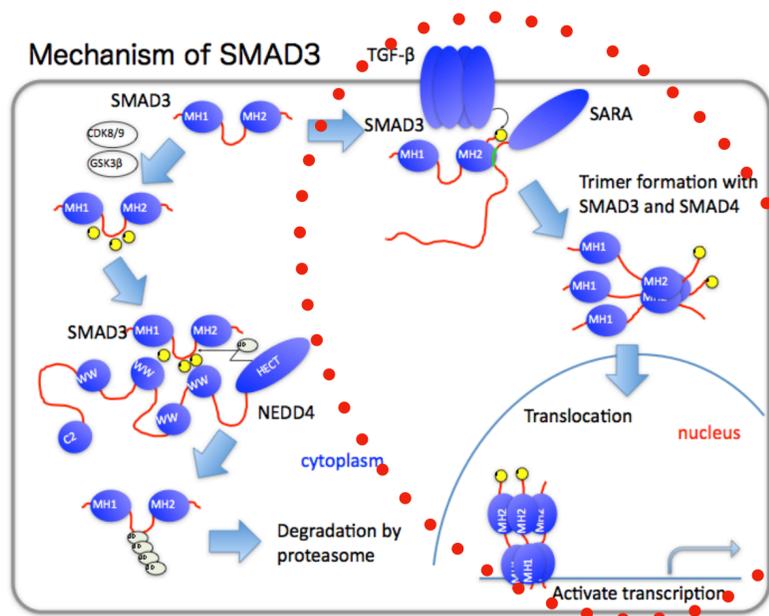


Subcellular location
Nucleus, Nucleolus.

液滴の局在情報

UniProtの局在情報から液滴をリスト化 2020年7月公開

リン酸化に伴うイベントの表現



リン酸化に起因する生命現象の記述

例) Smad3のリン酸化カスケード情報 エクセルで公開

id		Process				Additional Information		
protein_id	ideal_id	process_id	operand	operator	result	special object	location	property
P84022	IID00113	1	SMAD3 (obj_id 1)	TGF β receptor (obj_id 2)	SMAD3-P (obj_id 3)			
		2	SMAD3-P, SMAD3-P, SMAD4 (obj_id 4)		SMAD3-P/SMAD3-P/SMAD4 (obj_id 5)	SMAD3-P/SMAD3-P/SMAD4 (obj_id 5)	Cytosol	
		3	SMAD3-P/SMAD3-P/SMAD4 (obj_id 5)		SMAD3-P/SMAD3-P/SMAD4 (obj_id 6)	SMAD3-P/SMAD3-P/SMAD4 (obj_id 6)	Nucleus	positive regulation of transcription, DNA-templated



リン酸化に伴うイベントの表現：情報のXML・RDF・可視化計画

リン酸化情報入力エディター

Uniprot Accession : Q15797

Process List			
Operand	Operator	Result	Actions
SMAD1 obj_id1	BMP_receptor obj_id2	SMAD1(P) obj_id3	⊖
SMAD1(P),SMAD1(P),SMAD4 obj_id4		SMAD1(P)/SMAD1(P)/SMAD4 obj_id5	⊖
SMAD1(P)/SMAD1(P)/SMAD4 obj_id5		SMAD1(P)/SMAD1(P)/SMAD4 obj_id6	⊖
SMAD1 obj_id1	CDK8,CDK9 obj_id7	SMAD1(Ps) obj_id8	⊖
SMAD1(Ps),YAP obj_id9		SMAD1(Ps)/YAP obj_id10	⊖
SMAD1(Ps) obj_id8	GSK3_beta obj_id11	SMAD1(Ps) obj_id12	⊖

Object
SMAD1 obj_id1
BMP_receptor obj_id2
SMAD1(P) obj_id3
SMAD1(P),SMAD1(P),SMAD4 obj_id4
SMAD1(P)/SMAD1(P)/SMAD4 obj_id5
SMAD1(P)/SMAD1(P)/SMAD4 obj_id6
CDK8,CDK9 obj_id7

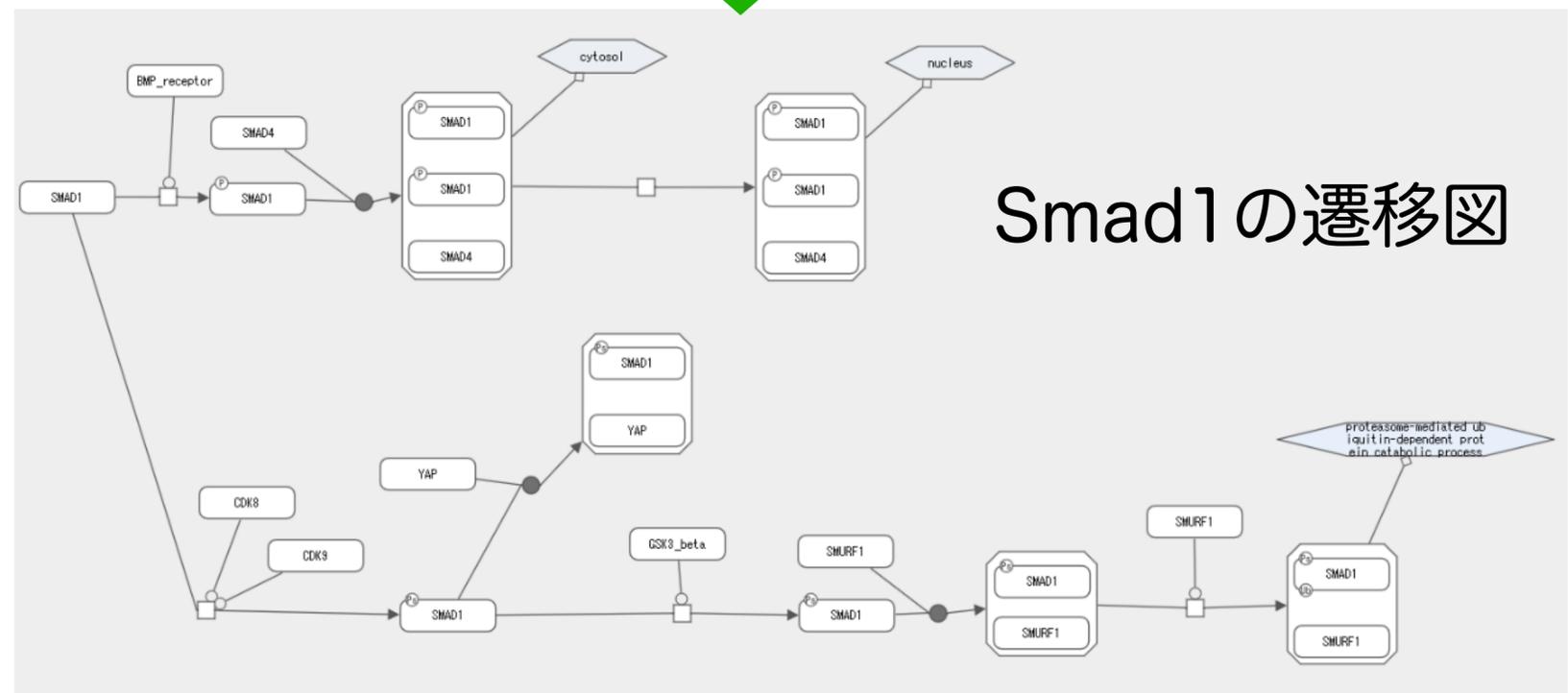
SBGN-ML



ChiSE

グラフツール

ChiSE is a library with an API based on [SBGNViz.js](#), which in turn is based on pathway models represented by process description (PD) and activity flow interaction format (SIF).



リン酸化に伴うイベントの **データ化・可視化・標準化**
 (同じデータからは同じ図が描かれる)



IDEAL

おわりに

Intrinsically Disordered proteins with Extensive Annotations and Literature

PDBが整備されてタンパク質の立体構造研究が飛躍的に進展したように、
まだ未知のことだらけの天然変性タンパク質研究の進展には、
質の高いデータベースが必要不可欠です。

IDEALは欧州のDisProtとともに、研究のインフラストラクチャとして
継続的、安定的な情報提供を行ってきたいと考えています。

IDEAL開発チームの活動をご理解いただけると幸いです。