

*Pathway Tools Schema and
Semantic Inference Layer*

Compounds, Reactions, Proteins and RNAs

**Alexander Shearer
Pallavi Kaipa**

SRI International

References

- **Pathway Tools User's Guide, Volume I**
 - Appendix A: Guide to the Pathway Tools Schema
- **Ontology Papers section of**
<http://biocyc.org/publications.shtml>
 - "An Evidence Ontology for use in Pathway/Genome Databases,"
 - "An ontology for biological function based on molecular interactions,"
 - "Representations of metabolic knowledge: Pathways,"
 - "Representations of metabolic knowledge,"

Use GKB Editor to Inspect the Pathway Tools Ontology

- **GKB Editor = Generic Knowledge Base Editor**
- **Type in Navigator window: (GKB) or**
- **[Right-Click] Edit->Ontology Editor**

- **View->Browse Class Hierarchy**
- **[Middle-Click] to expand hierarchy**
- **To view classes or instances, select them and:**
 - Frame -> List Frame Contents
 - Frame -> Edit Frame

Slots

- Describe an attribute or a property of the object that the frame represents.
 - Slots valid in only a particular set of classes
 - Slots valid in multiple classes
 - ◆ Common-Name
 - Primary name by which an on object is known
 - ◆ Synonyms
 - Names by which one may attempt to retrieve this object
 - ◆ Names
 - Values combined from all other name related slots
 - ◆ Comment
 - Stores the general comment about the object
 - ◆ Citations
 - Lists general citations pertaining to the object.
 - ◆ Database links
 - Links to variety of other databases

Compounds

Gkb Editor
 Application Knowledge-Base Frame View Preferences
 KB: METABASE Package: ECOCPK Class-Instance Hierarchy Viewer

```

graph TD
  CCO --> CCO-MEMBRANE
  CCO --> CCO-ORGANELLE
  CCO --> CCO-SPACE
  CCO-MEMBRANE --> CCO-ER-MEM
  CCO-MEMBRANE --> CCO-MBODY-MEM
  CCO-MEMBRANE --> CCO-MIT-MEM
  CCO-MEMBRANE --> CCO-HUC-MEM
  CCO-MEMBRANE --> 8_MORE[8 MORE]
  CCO-ORGANELLE --> CCO-MEM-ORG
  CCO-ORGANELLE --> CCO-HOH-MEM-ORG
  CCO-SPACE --> CCO-ENDO-LUM
  CCO-SPACE --> CCO-ER-LUM
  CCO-SPACE --> CCO-MICRO-LUM
  CCO-SPACE --> CCO-PERIPLASM
  CCO-SPACE --> CCO-PLAST-IM-SPC
  CCO-SPACE --> 12_MORE[12 MORE]
  Chemicals --> Compounds-And-Elements
  Chemicals --> Holder-Class
  Compounds-And-Elements --> Compounds
  Compounds --> All-Amines
  Compounds --> All-Amino-Acids
  Compounds --> All-Carbohydrates
  Compounds --> All-Carboxy-Acids
  Compounds --> All-Nucleosides
  Compounds --> Amides
  Compounds --> Aromatics
  Compounds --> Coenzymes
  Compounds --> Esters
  Compounds --> Groups
  Compounds --> Halides
  Compounds --> Hormones
  Compounds --> Inorganic-Minerals
  Compounds --> Inositols
  Compounds --> Ions
  Compounds --> Lipids
  Compounds --> Non-Metabolic-Compounds
  Compounds --> OLIGOPEPTIDES
  Compounds --> ORGANOSULFUR
  Compounds --> Porphyrins
  Compounds --> Prostaglandins
  Compounds --> Pseudo-Compounds
  Compounds --> Secondary-Metabolites
  Compounds --> Steroids
  Compounds --> Tautomers
  Compounds --> Unclassified-Compounds
  Compounds --> Vitamins
  Macromolecules --> Complexes
  Macromolecules --> Polynucleotides
  Macromolecules --> Proteins
  Databases --> CAS
  Databases --> CGSC
  Databases --> COLIBLAST
  Databases --> DBOET
  Databases --> ECHOBASE
  Databases --> 51_MORE[51 MORE]
  
```

Command: Deselect
 Command: Deselect
 Command: |

Compounds

- **Very few things come from within the compound editor**
 - MW, formula calculated from edited structure
- **Most traits defined in other editors**
 - “In pathway reactions” comes from reaction editing followed by pathway editing
 - activator, etc come from the enzymatic reaction editor

-- Instance TRP ---

Types: |Amino-Acid|, |Aromatic-Amino-Acids|, |Non-polar-amino-acids|

APPEARS-IN-LEFT-SIDE-OF: RXN0-287, TRANS-RXN-76, TRYPTOPHAN-RXN,
TRYPTOPHAN--TRNA-LIGASE-RXN

APPEARS-IN-RIGHT-SIDE-OF: RXN0-2382, RXN0-301, TRANS-RXN-76, TRYPSYN-RXN

CHEMICAL-FORMULA: (C 11), (H 12), (N 2), (O 2)

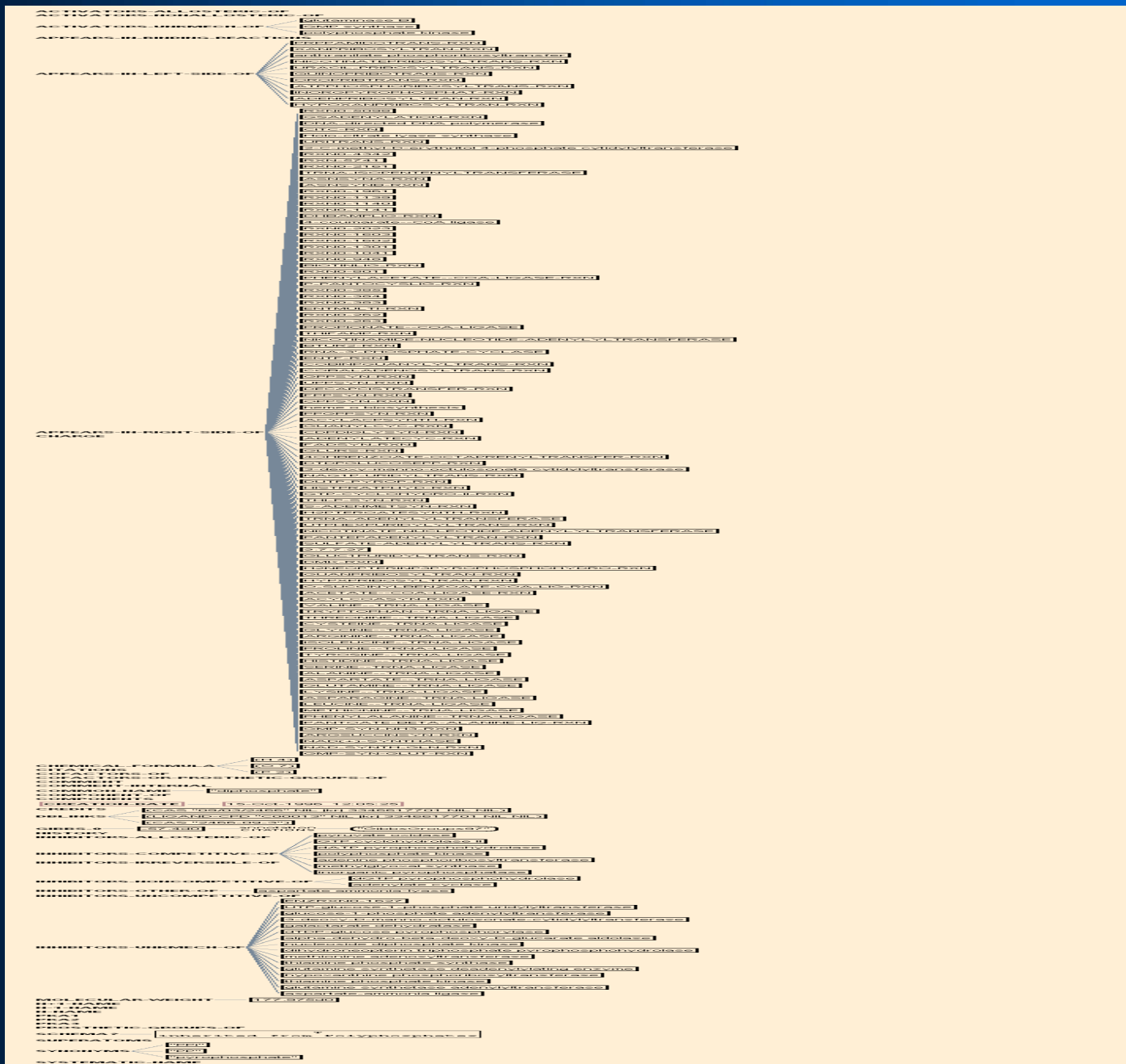
COMMON-NAME: "L-tryptophan"

DBLINKS: (LIGAND-CPD "C00078" NIL |kaipa| 3311532640 NIL NIL),
(CAS "6912-86-3"), (CAS "73-22-3")

NAMES: "L-tryptophan", "W", "tryptacin", "trofan", "trp", "tryptophan",
"2-amino-3-indolylpropanoic acid"

SMILES: "c1(c(CC(N)C(=O)O)c2(c([nH]1)cccc2))"

SYNONYMS: "W", "tryptacin", "trofan", "trp", "tryptophan",
"2-amino-3-indolylpropanoic acid"



ACTIVATORS-ALLOSTERIC-OF

ACTIVATORS-HOMIALLOSTERIC-OF

ACTIVATORS-UNIKMECH-OF

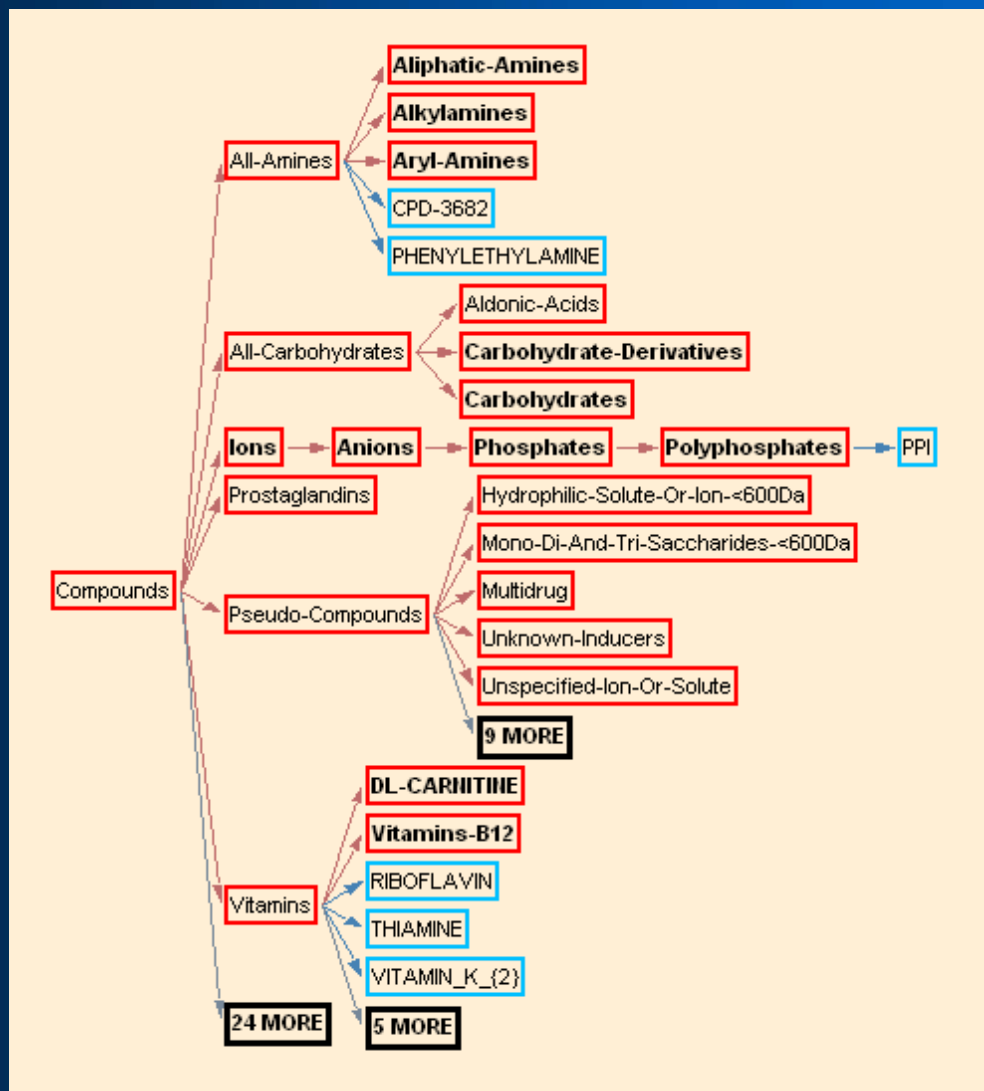
APPEARS-III-BINDING-REACTIONS

APPEARS-III-LEFT-SIDE-OF

- glutaminase B
- GMP synthase
- polyphosphate kinase
- PRPPAMIDOTRANS-RXN
- XANPRIBOSYLTRAN-RXN
- anthranilate phosphoribosyltransfer
- NICOTINATEPRIBOSYLTRANS-RXN
- URACIL-PRIBOSYLTRANS-RXN
- QUINOPRIBOTRANS-RXN
- OROPRIBTRANS-RXN
- ATPPHOSPHORIBOSYLTRANS-RXN
- INORGYPYROPHOSPHAT-RXN
- ADENPRIBOSYLTRAN-RXN
- HYPOXANPRIBOSYLTRAN-RXN
- RXN0-5098
- GSADENYLATION-RXN
- DNA-directed DNA polymerase
- CITC-RXN
- Holo-citrate lyase synthase
- URITRANS-RXN
- 2-C-methyl-D-erythritol 4-phosphate cytidyltransferase
- RXN0-4342
- RXN-5741
- RXN0-2161
- TRNA-ISOPENTENYLTRANSFERASE
- ASNSYNA-RXN
- ASNSYNB-RXN
- RXN0-1961
- RXN0-1139
- RXN0-1140
- RXN0-1141
- DHBAMPLIG-RXN
- 4-coumarate--CoA ligase
- RXN0-2023
- RXN0-1603
- RXN0-1602
- RXN0-1301
- RXN0-1041
- RXN0-948
- BIOTINLIG-RXN
- RXN0-801
- PHENYLACETATE--COA-LIGASE-RXN
- P-PANTOCYSLIG-RXN
- RXN0-385
- RXN0-384

COMMENT-INTERNAL
 COMMON-NAME — "diphosphate"
 COMPONENT-OF
 COMPONENTS
 CREATION-DATE — 15-Oct-1996 12:05:25
 CREDITS — [CAS "09/03/2466" NIL [kr] 3346617701 NIL NIL]
 DBLINKS — [LIGAND-CPD "C00013" NIL [kr] 3346617701 NIL NIL]
 [CAS "2466-09-3"]
 GIBBS-θ — -57.4d0 — annotation CITATIONS — "GibbsGroups97"
 HISTORY
 INHIBITORS-ALLOSTERIC-OF — pyruvate oxidase
 GTP cyclohydrolase II
 dATP pyrophosphohydrolase
 INHIBITORS-COMPETITIVE-OF — polyphosphate kinase
 INHIBITORS-IRREVERSIBLE-OF — adenine phosphoribosyltransferase
 methylglyoxal synthase
 inorganic pyrophosphatase
 INHIBITORS-NONCOMPETITIVE-OF — dGTP pyrophosphohydrolase
 adenylate cyclase
 INHIBITORS-OTHER-OF — aspartate ammonia-lyase
 INHIBITORS-UNCOMPETITIVE-OF
 ENZRXND-1E27
 UTP-glucose-1-phosphate uridylyltransferase
 glucose-1-phosphate adenyltransferase
 3-deoxy-D-manno-octulosonate-cytidylyltransferase
 galactarate dehydratase
 dTDP-glucose pyrophosphorylase
 alpha-dehydro-beta-deoxy-D-glucarate aldolase
 INHIBITORS-UNIMECH-OF — nucleoside diphosphate kinase
 dihydroneopterin triphosphate pyrophosphohydrolase
 methionine adenosyltransferase
 thiamine phosphate synthase
 glutamine synthetase deadenylating enzyme
 hypoxanthine phosphoribosyltransferase
 thiamine phosphate kinase
 glutamine synthetase adenyltransferase
 aspartate-ammonia ligase
 MOLECULAR-WEIGHT — 177.975d0
 H-1-NAME
 H-1-NAME
 H-NAME
 PKA1
 PKA2
 PKA3
 PROSTHETIC-GROUPS-OF
 SCHEMA? — T
 inherited from Polyphosphates
 SUPERATOMS — "PPi"
 SYNONYMS — "PPi"
 "pyrophosphate"
 SYSTEMATIC-NAME

Where is diphosphate in the ontology?



Semantic Inference Layer

- **Reactions-of-compound (cpd)**
- **Pathways-of-compound (cpd)**
- **is-substrate-an-autocatalytic-enzyme-p (cpd)**
- **Activated/inhibited by? (cpds slots)**
 - Returns a list of enrxtns for which a cpd in cpds is a modulator (example slots: activators-all, activators-allosteric)
- **All-substrates (rxns)**
 - All unique substrates specified in the given rxns
- **Has-structure-p? (cpd)**
- **Obtain-cpd-stats**
 - Returns two values:
 - ◆ Length of :all-cpds, cpds with structures

Miscellaneous things....

- **History List**

- Back/Forward and History buttons
- Default list is 50 items

- **Show frame**

- **(print-frame 'frame)**

Pathway Tools version 10.0

File Overview Pathway Reaction Protein RNA Gene Compound Chromosome Tools Help

Escherichia coli K-12

Home Back Forward History Next Answer Clone Save DB

***E. coli* K-12 Enzyme: DNA polymerase I, 3' → 5' polymerase, 5' → 3' and 3' → 5' exonuclease / 5' to 3' exonuclease / 3' to 5' proofreading exonuclease**

Protein Sequence

Synonyms: B3863, ResA, PolA

Comment:

DNA Polymerase I (Pol I) is a multifunctional enzyme that combines a DNA polymerase activity, a 5' to 3' exonuclease activity and a 3' to 5' proofreading exonuclease activity. It is required for several types of DNA repair and appears to be the primary enzyme responsible for stripping RNA primers from newly-synthesized DNA and replacing them with DNA.

Pol I is involved in several DNA repair pathways. It is required for excision repair, displacing the UvrABC nuclease and patching the gap it leaves behind [Sharon75, Sung03, Glickman75, Heyneker75, Orren92, Husain85, Matson81]. Sharon1975 Pol I is also required in MutHLS-mediated very short patch repair [Dzidic89]. Pol I can excise and replace pyrimidine dimers directly [Dorson78]. It also cleaves the faulty nucleotide from abasic lesion sites following nicking by endonuclease III [Mosbaugh82]. Finally, Pol I is generally involved in postreplication repair of DNA gaps and double-strand breaks [Sharma87].

Pol I primer removal and subsequent DNA gap filling has been shown directly in phiX174 phage DNA synthesis [Shlomai81]. A similar role for Pol I in *E. coli* is supported by the observations that Pol I can initiate synthesis at a DNA nick, that Okazaki fragment joining is only 10% of normal in mutants lacking *polA* and that normal replication depends on Pol I [Kelly70, Okazaki71, Olivera74].

Pol I consists of two domains. The larger domain, commonly known as the Klenow fragment when it is proteolytically separated, contains the polymerase and 3' to 5' exonuclease activities [Setlow72]. The smaller domain contains the 5' to 3' exonuclease activity [Setlow72a]. The Klenow domain itself has a large and a small subdomain, with its carboxy-terminal large domain containing the polymerase but not the 3' to 5' exonuclease function [Freemont86]. The Klenow domain also contains a "thumb" structure that is required for DNA binding, processivity and frameshifts and a J-helix region that regulates both the polymerase and 3' to 5' exonuclease functions [Minnick96, Tuske00, Singh05]. The Klenow portion undergoes conformational changes on binding template, then again on the subsequent binding of dNTPs [Dzantiev00].

Pol I and its subdomains have been crystallized several times. The initial crystallization was to 3.5 Å resolution [Steitz83]. Crystal structures have been determined for Klenow fragment bound to dNTP, pyrophosphate, ssDNA and dsDNA [Beese93, Freemont88]. Crystal structures of Pol I bound to dNMP and ssDNA have been determined to 2.6 Å and 3.1 Å resolution, respectively [Beese91].

Pol I binds DNA via hydrogen bonding between the minor groove and a hydrogen-bonding track on the protein [Spratt01, Singh03, Meyer04, Freemont88]. Pol I binds only one oligonucleotide at a time and only binds dsDNA at nicks or strand ends [Englund69]. It has a higher affinity for primers containing template mismatches or hairpin-like elements, and has a separate binding site for the 3'-hydroxyl end of substrates [Ljach92, Huberman70]. The binding of DNA by the Pol I 5' to 3' exonuclease function has been examined in detail [Xu01].

Polymerization by Pol I is processive, typically covering stretches of 20-40 nucleotides but potentially going up to hundreds of nucleotides [Bambara78, Uyemura75]. The kinetics of polymerization have been extensively evaluated [Travaglini75, McClure75, Mizrahi85, elDeiry88, Dahlberg91]. The nucleotide-dependence of polymerization termination has also been examined [Abbotts88].

Pol I polymerization is also specific; though error rates *in vitro* of 1 in 8,000-80,000 have been measured, the estimated rate on natural DNA is between 1 in 680,000 and 1 in 6.3 million [Agarwal79, Kunkel80]. The mechanisms behind specificity have been examined, as well as the role of differing metal cofactors in specificity [Astatke98, Astatke98a, Sirover79, Hillebrand84].

Two exonuclease activities of Pol I have also been evaluated. The 5' to 3' exonuclease activity requires a free 5' end at an ssDNA-dsDNA junction and is slower than the 3' to 5' exonuclease activity [Xu97, Deutscher69]. The 3' to 5' exonuclease, which is responsible for proofreading, does not identify base-pair mismatches. Instead, Pol I lingers when a mismatch occurs, allowing more time for the exonuclease to act on the mismatch [Kuchta88, Bailly84]. The transfer of DNA from the polymerase to the 3' to 5' exonuclease active site can occur either intra- or intermolecularly, with mismatches favoring the latter [Joyce89].

polA mutants are more vulnerable to UV and X-rays and experience more deletions, duplications and frameshifts [Billen85, Nagata02, Barfknecht78]. *polA mutU4* double mutants are inviable [Siegel73].

Gene: [polA](#)

Sequence Length: 928 AAs

Molecular Weight of Polypeptide (from nucleotide sequence): 103.12 kD

Queries with Multiple Answers

- **Navigator queries:**

- Example: Substring search for “pyruvate”
- Selected list is placed on the [Answer list](#)
- Use “Next Answer” button to view each one of them

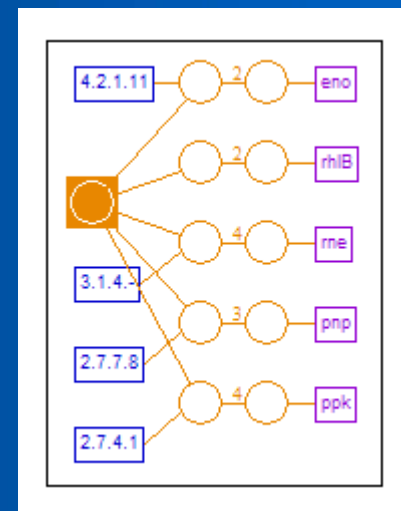
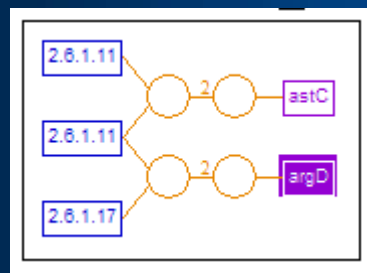
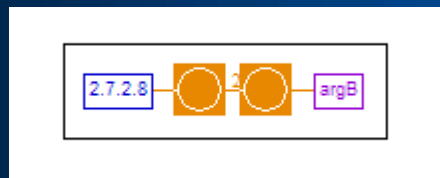
- **Lisp queries:**

Example : Find reactions involving pyruvate as a substrate

- ◆ (get-class-all-instances '|Compounds|)

```
(loop for rxn in (get-class-all-instances '|Reactions|)
      when (member 'pyruvate (get-slot-values rxn 'substrates)
            collect rxn)
      (replace-answer-list *))
```


Gene Reaction Schematic

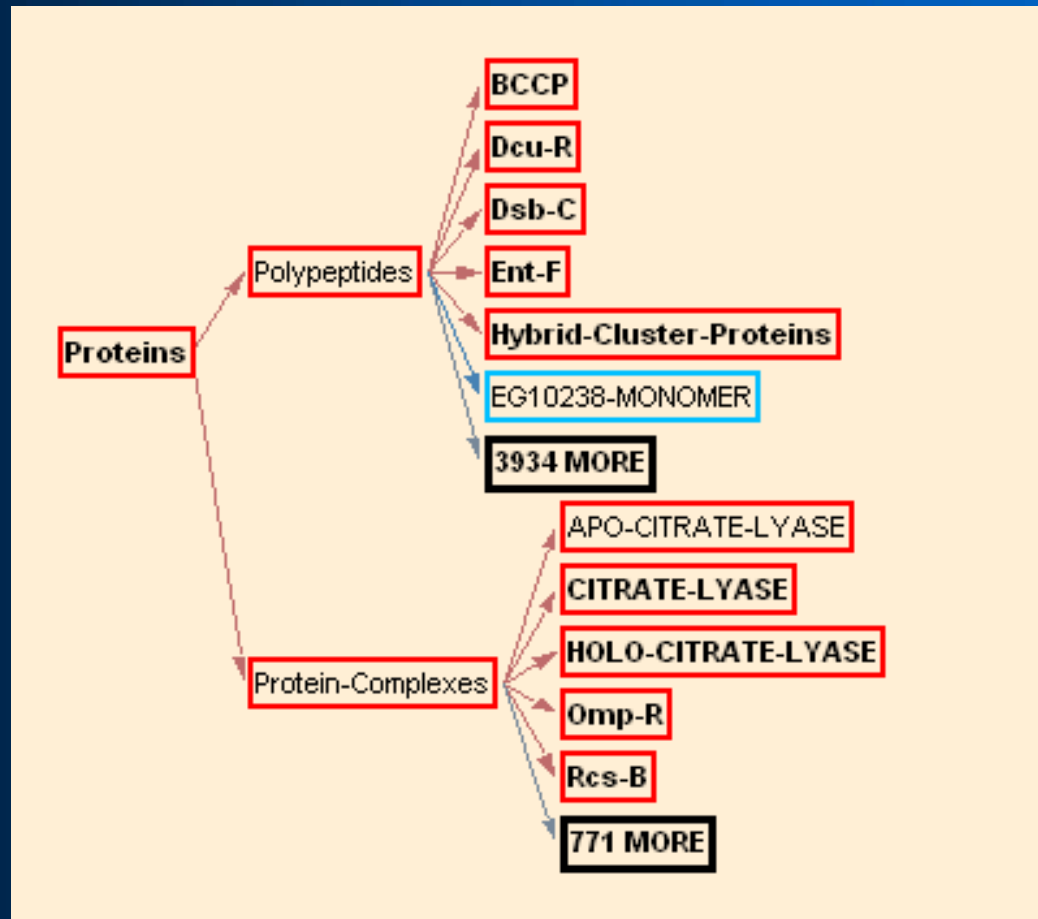


Proteins

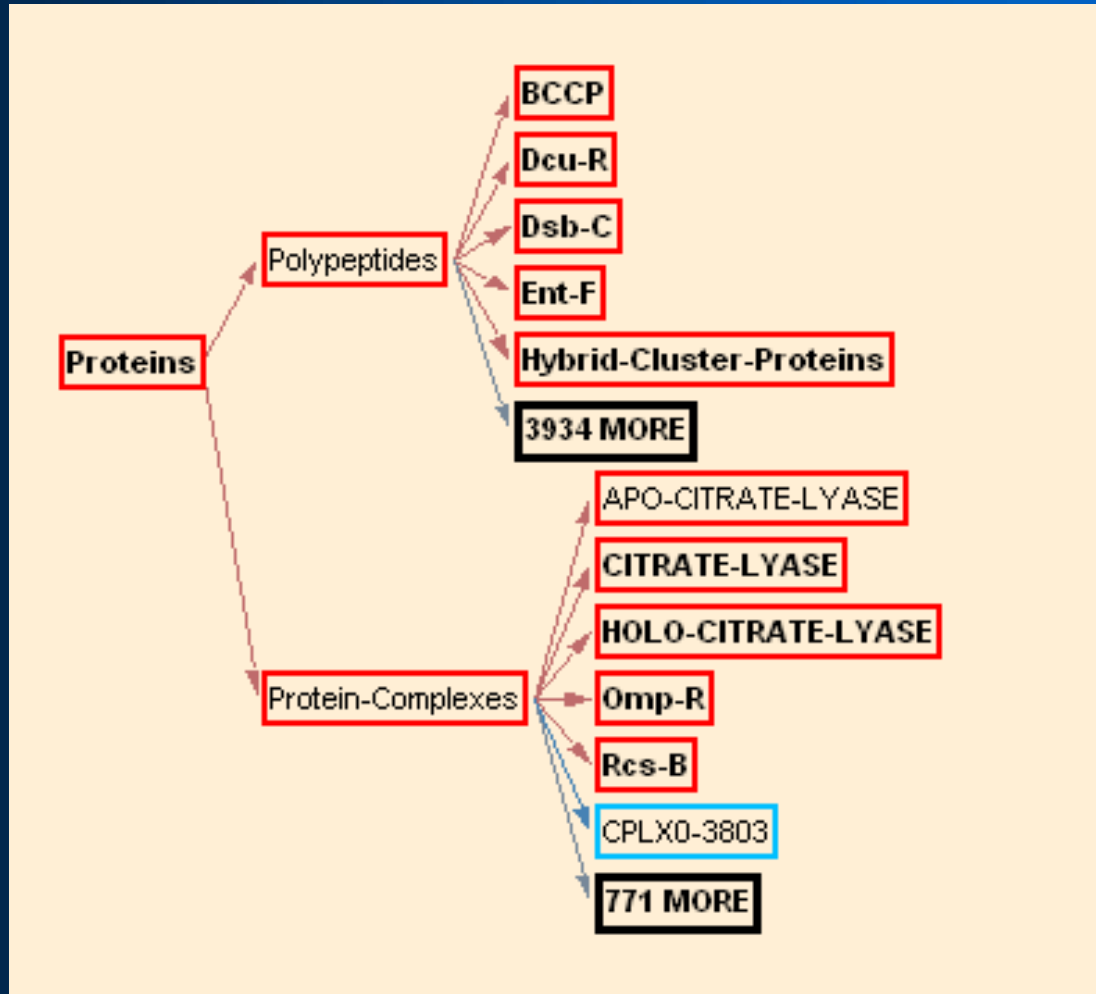
Proteins and Protein Complexes

- **Polypeptide: the monomer protein product of a given (may have multiple isoforms, as indicated at gene level)**
- **Protein complex: proteins consisting of multiple polypeptides or protein complexes**
- **Example: DNA pol III**
 - DnaE is a polypeptide
 - pol III core is DnaE and two other polypeptides
 - pol III holoenzymes is several protein complexes combined

Where is *DnaE* in the ontology?



Where is *pol III* in the ontology?



Features of a protein at the frame level (DnaE)

- catalyzes
- Is it an activator/reactant/etc?
- comments
- component-of
- dblinks
- features (edited in feature editor)

- Many other features possible

ACTIVATORS-ALLOSTERIC-OF
ACTIVATORS-HOMALLOSTERIC-OF
ACTIVATORS-UIKMECH-OF
APPEARS-III-BINDING-REACTIONS
APPEARS-III-LEFT-SIDE-OF
APPEARS-III-RIGHT-SIDE-OF

CATALYZES — ENZR_XNO-6081

CHEMICAL-FORMULA

CITATIONS — "6288664"

COFACTORS-OF

COFACTORS-OR-PROSTHETIC-GROUPS-OF

COMMENT —

"The alpha subunit of DNA polymerase III catalyzes the polymerase activity of the holoenzyme complex [CITS: [2997151]]. Replicative 5' to 3' polymerization of DNA requires dNTPs and template DNA with a bound RNA primer [CITS: [4560196][4589895]]. The newly polymerized DNA is covalently attached to the RNA primer [CITS: [1089643]]. The presence of the epsilon subunit increases the polymerase activity of the alpha subunit two-fold [CITS: [3037519]].

The alpha subunit is required for misincorporation and bypass during UV mutagenesis [CITS: [2184308][2184309]].

The middle portion of the alpha subunit (residues 542-991) is involved in binding to the polymerase III beta subunit. Deletion of the amino-terminal portion of alpha (residues 1-542) actually increases its affinity for beta [CITS: [8702820]]. The carboxy-terminus of alpha is required for binding to the polymerase III tau subunit [CITS: [8702819]]. The amino-terminal php domain of alpha is required for binding to the epsilon subunit [CITS: [16517598]].

Transcription of *<i>dnaE</i>* is induced by nalidixic acid, but not by mitomycin C, and induction does not require LexA [CITS: [11544210]].

Overproduction of alpha can compensate for an otherwise lethal deficiency in DNA polymerase I [CITS: [1597430]]."

COMMENT-INTERNAL

COMMON-NAME — "DNA polymerase III, α subunit"

COMPONENT-OF — DNA polymerase III, core enzyme

COMPONENTS

CREATION-DATE — 24-Jan-2000 10:28:54

CREATOR — pkarp

CREDITS

(MODBASE "P10443" NIL [pkarp] 3355444109 NIL NIL)

(SMSSMODEL "P10443" NIL [pkarp] 3355444109 NIL NIL)

DBLINKS

(PFAM "PF02231" IN-FAMILY [pkarp] 3346700315 NIL NIL)

(REFSEQ "NP_414726" NIL NIL NIL NIL NIL)

(UNIPROT "P10443" NIL [paley] 3169408120)

DNA-FOOTPRINT-SIZE

FEATURES

Protein-Binding-Region

Metal-Binding-Site

FUNCTIONAL-ASSIGNMENT-COMMENT

FUNCTIONAL-ASSIGNMENT-STATUS

A complex at the frame level (pol III)

- Same features as polypeptide frame, different use
- comment
- component-of and components
 - note coefficients

ACTIVATORS-ALLOSTERIC-OF
ACTIVATORS-NONALLOSTERIC-OF
ACTIVATORS-UNKMECH-OF
APPEARS-IN-BINDING-REACTIONS
APPEARS-IN-LEFT-SIDE-OF
APPEARS-IN-RIGHT-SIDE-OF
CATALYZES
CHEMICAL-FORMULA
CITATIONS
COFACTORS-OF
COFACTORS-OR-PROSTHETIC-GROUPS-OF

COMMENT

"The DNA polymerase III core enzyme contains one each of the alpha, epsilon and theta subunits and can carry out the basic polymerase and exonuclease activities of polymerase III [CITS: [368075]].
Based on yeast two-hybrid data, both alpha and theta interact with epsilon, but not each other [CITS: [9515927]].
The interaction between epsilon and theta has been examined via lanthanide-labeling NMR [CITS: [16536542]]."

COMMENT-INTERNAL

COMMON-NAME — "DNA polymerase III, core enzyme"

COMPONENT-OF — DNA polymerase III, holoenzyme

COMPONENTS — DNA polymerase III, alpha subunit — annotation COEFFICIENT 1
DNA polymerase III, epsilon subunit — annotation COEFFICIENT 1
DNA polymerase III, theta subunit — annotation COEFFICIENT 1

CREATION-DATE — 25-Jun-2004 15:54:36

CREATOR — keseler

CREDITS

DBLINKS

DNA-FOOTPRINT-SIZE

HISTORY

INHIBITORS-ALLOSTERIC-OF

INHIBITORS-COMPETITIVE-OF

INHIBITORS-IRREVERSIBLE-OF

INHIBITORS-NONCOMPETITIVE-OF

INHIBITORS-OTHER-OF

INHIBITORS-UNCOMPETITIVE-OF

INHIBITORS-UNKMECH-OF

ISOZYME-SEQUENCE-SIMILARITY

LOCATIONS — cytoplasm

MODIFIED-FORM

MOLECULAR-WEIGHT

MOLECULAR-WEIGHT-EXP

MOLECULAR-WEIGHT-SEQ

II+1-NAME

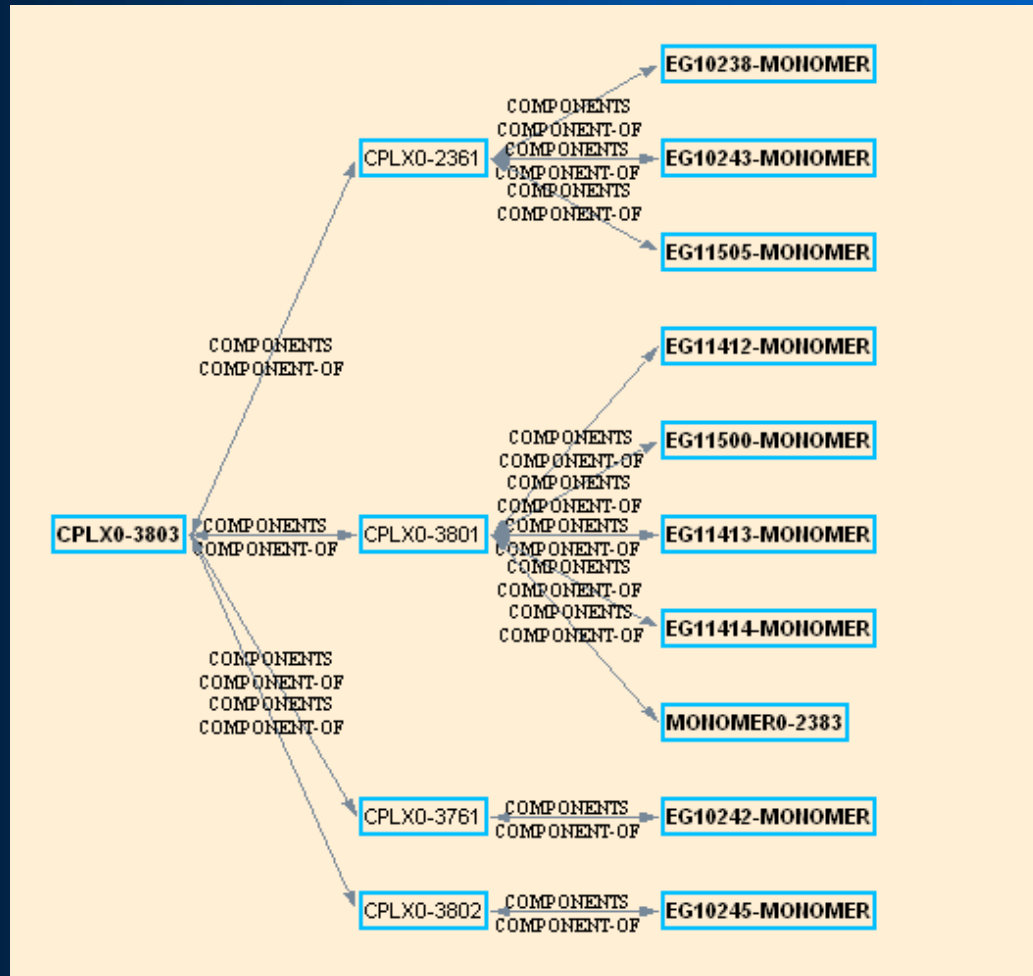
II-1-NAME

II-NAME

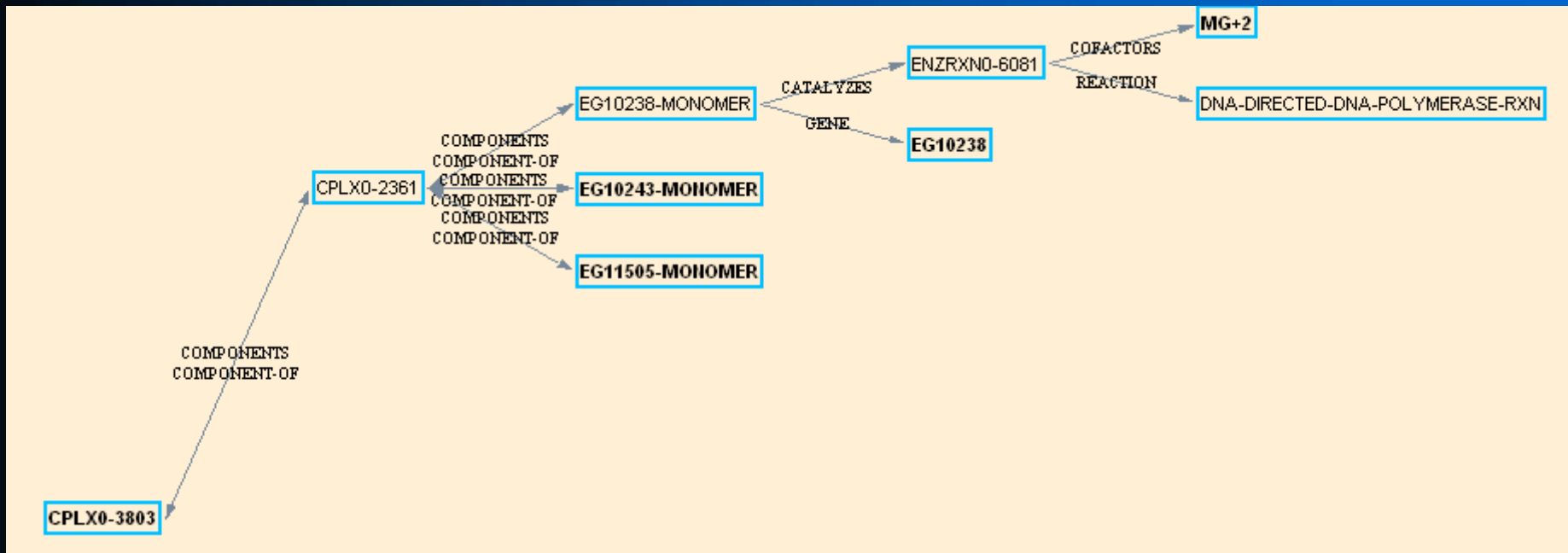
HEIDHARDT-SPOT-NUMBER

PI

Protein complex relationships



Protein complex relationships



Relationships are defined in many places

- **component-of** comes from creating a complex
- **appears-in-left-side-of** comes from defining a reaction (as do modified forms)
- **inhibitor-of** comes from an enzymatic reaction
- **can only edit dna-footprint** if protein has been associated with a TU

Semantic Inference Layer

- **Reactions-of-protein (pro)**
 - Returns a list of rxns this protein catalyzes
- **Transcription-units-of-proteins(pro)**
 - Returns a list of TU's activated/inhibited by the given protein
- **transporter? (pro)**
 - Is this protein a transporter
- **Polypeptide-or-homomultimer?**
- **Is this protein a transcription factor? (pro)**
- **Obtain-protein-stats**
 - Returns 5 values
 - ◆ Length of : all-polypeptides, complexes, transporters, enzymes, etc...

Sample

- **Find all enzymes that use pyridoxal phosphate as a cofactor or prosthetic group**

- (loop for protein in (get-class-all-instances '|Proteins|)

For enzrxn = (get-slot-value protein 'enzymatic-reaction)

when (and enzrxn

(or (member-slot-value-p enzrxn 'cofactors 'pyridoxal_phosphate)

(member-slot-value-p enzrxn 'prosthetic-groups

'pyridoxal_phosphate))

collect protein)

(member-slot-value-p frame slot value) : T if Value is one of the values of Slot of Frame.

Sample

- Find all proteins without a comment anywhere

```
(defun proteins-sans-comments-al ()
  (replace-answer-list
   (loop for x in (get-class-all-instances 'Proteins)
         for cmnts = (slot-has-value-p x 'comment)
         for rxn = (enz-rxn-comments x)
         for cplx = (complex-with-comments x)
         for cplx2 = (slot-has-value-p x 'components)
         for modform = (modified-form-comments x)
         for modform2 = (slot-has-value-p x 'unmodified-form)
         unless (or cmnts rxn cplx cplx2 modform modform2)
         collect x)
   )
  )

(defun enz-rxn-comments (subj-protein)
  (loop for x in (get-slot-values subj-protein 'catalyzes)
        for cmnts = (slot-has-value-p x 'comment)
        when cmnts
        collect x)
  )

(defun complex-with-comments (subj-protein)
  (loop for x in (get-slot-values subj-protein 'component-of)
        for rxn = (enz-rxn-comments x)
        for cmnts = (slot-has-value-p x 'comment)
        when (or cmnts rxn)
        collect x)
  )

(defun modified-form-comments (subj-protein)
  (loop for x in (get-slot-values subj-protein 'modified-form)
        for cmnts = (slot-has-value-p x 'comment)
        for rxn = (enz-rxn-comments x)
        for cplx = (complex-with-comments x)
        when (or cmnts rxn cplx)
        collect x)
  )
)
```

Reactions

Enzymatic reactions (DnaE and 2.7.7.7)

- **A necessary bridge between enzymes and “generic” versions of reactions**
- **Carries specific reaction features:**
 - activators
 - inhibitors
 - cofactors
 - alternative substrates
- **Frame generated when protein associated with reaction (protein or reaction editor)**

ACTIVATORS-ALLOSTERIC

ACTIVATORS-NONALLOSTERIC

ACTIVATORS-UNKMECH

ALTERNATIVE-COFACTORS

ALTERNATIVE-SUBSTRATES

BASIS-FOR-ASSIGNMENT

CITATIONS

"4560196"

"1089643:EV-EXP-IDA-PURIFIED-PROTEIN:3336919499:shearer"

COFACTOR-BINDING-COMMENT

COFACTORS

Mg2+

annotation
CITATIONS

"4560196"

COFACTORS-OR-PROSTHETIC-GROUPS

COMMENT

"This enzyme can also use divalent manganese ion as a cofactor [CITS: [788784]]."

COMMENT-INTERNAL

COMMON-NAME

COMPONENT-OF

COMPONENTS

CREATION-DATE

28-Sep-2005 10:55:45

CREATOR

shearer

CREDITS

DBLINKS

ENZYME

DNA polymerase III, alpha subunit

HISTORY

INHIBITORS-ALLOSTERIC

INHIBITORS-COMPETITIVE

INHIBITORS-IRREVERSIBLE

INHIBITORS-NONCOMPETITIVE

INHIBITORS-OTHER

INHIBITORS-UNCOMPETITIVE

INHIBITORS-UNKMECH

KM

PH-OPT

PHYSIOLOGICALLY-RELEVANT

PROSTHETIC-GROUPS

REACTION

DNA-directed DNA polymerase

REACTION-DIRECTION

REQUIRED-PROTEIN-COMPLEX

SCHEMA?

T

inherited from Enzymatic-Reactions

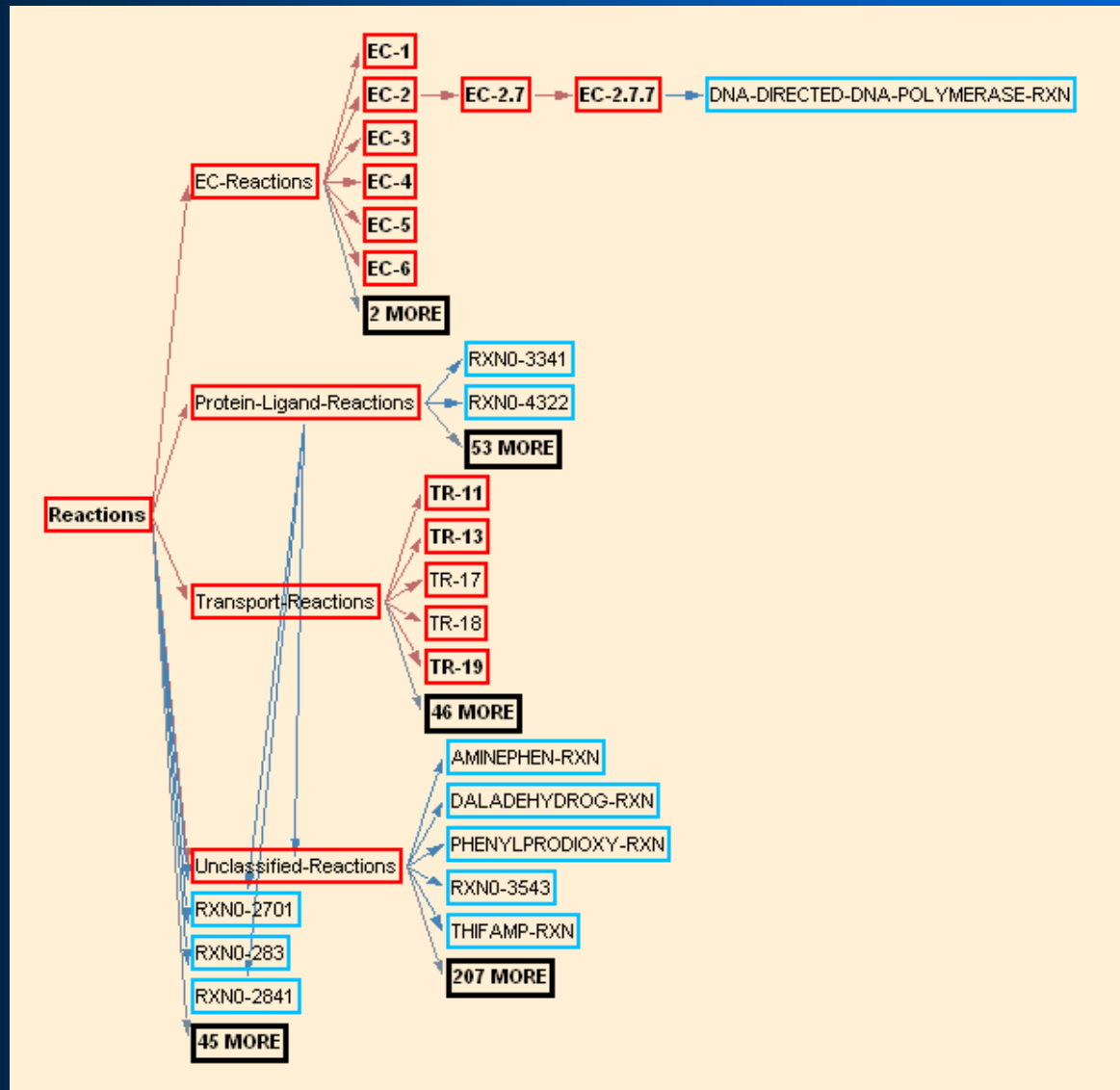
SYNONYMS

TEMPERATURE-OPT

Reactions

- Represent the “generic” form of the reaction
- Connected to proteins via enzymatic reaction frames
- Classified with EC system when possible
- Example: 2.7.7.7 – DNA-directed DNA polymerization
 - Carried out by five enzymes in *E. coli*

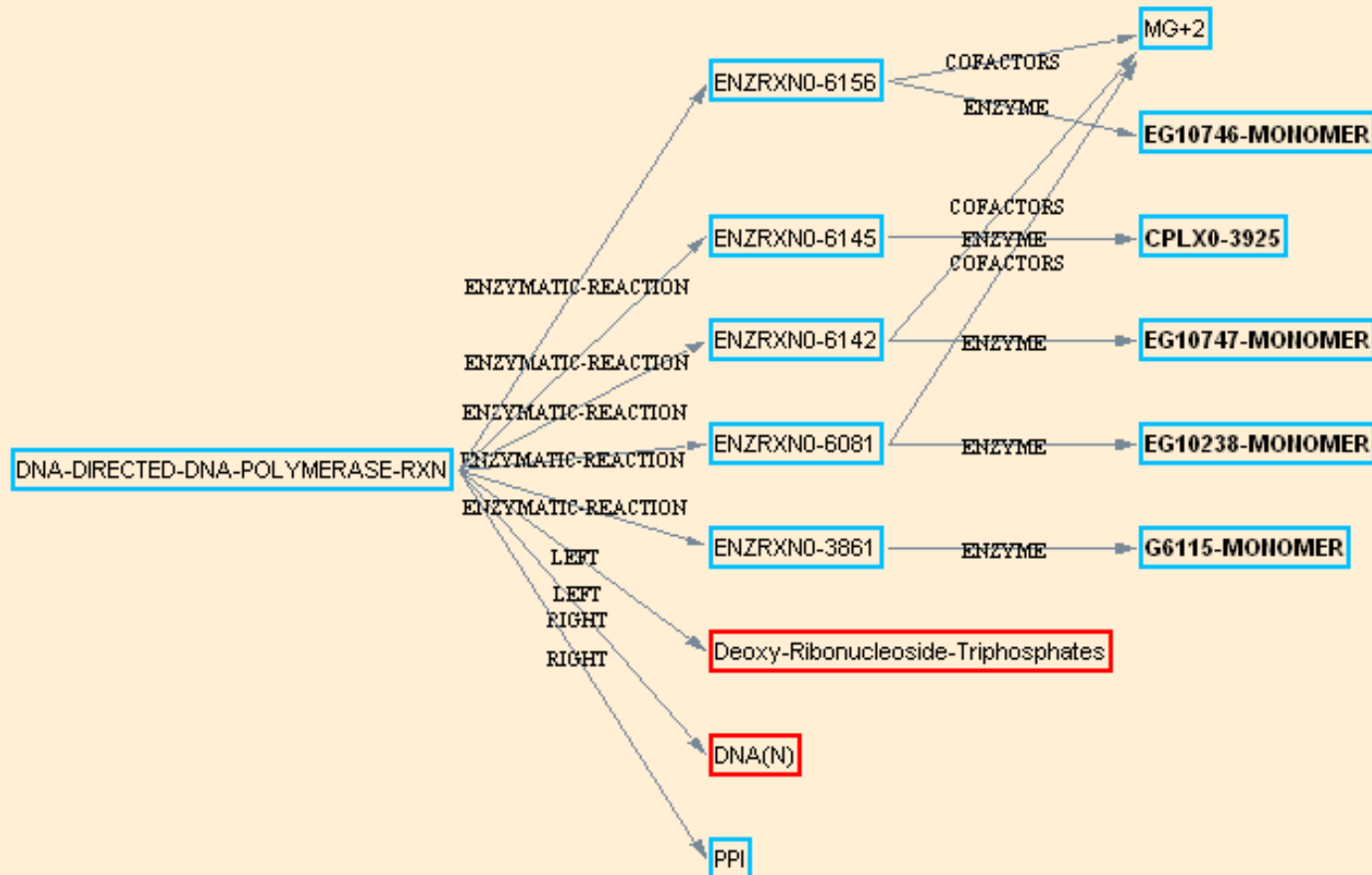
Where is 2.7.7.7 in the ontology?



Features of a reaction at the frame level

- **balance-state**
- **ec-number**
- **enzymatic-reaction**
 - Generated in protein or reaction editor
- **in-pathway**
 - Generated in pathway editor
- **left and right-side compounds**
 - Can make modified forms of proteins, RNAs, etc here
 - Not all reactants/products need to be frames

Reaction relationships



Semantic Inference Layer

- **Genes-of-reaction (rxn)**
- **Substrates-of-reaction (rxn)**
- **Enzymes-of-reaction (rxn)**
- **Lacking-ec-number (organism)**
 - Returns list of rxns with no ec numbers in that database
- **Get-reaction-direction-in-pathway (pwy rxn)**
- **Reaction type? (rxn)**
 - ◆ Small molecule rxn, transport rxn, protein-small-molecule rxn (one substrate is protein and one is a small molecule), protein rxn (all substrates are proteins), etc.
- **(all-rxns :type)**
 - Specify the type of reaction (see above for type)
- **Obtain-rxn-stats**
 - Returns six values
 - ◆ Length of : all-rxns, transport, non-transport, etc...

Orphan reactions in an organism:

```
(defun orphan-reactions (&optional (verbose? t))
```

```
  (loop for r in (all-rxns :small-molecule)
        when (and (not (slot-has-value-p r 'enzymatic-reaction))
                  (not (get-slot-value r 'spontaneous?)))
        collect r)
```

```
)
```

RNAs

RNAs

- **Currently only represent RNAs that are the “terminal gene product”**
 - tRNAs
 - rRNAs
 - miscellaneous small RNAs
- **Frame features similar to proteins**
- **tRNAs can have an anticodon**

ACTIVATORS-ALLOSTERIC-OF
ACTIVATORS-NONALLOSTERIC-OF
ACTIVATORS-UNKMECH-OF

ANTICODON — UUUU

APPEARS-III-BINDING-REACTIONS
APPEARS-III-LEFT-SIDE-OF
APPEARS-III-RIGHT-SIDE-OF

CHEMICAL-FORMULA

CITATIONS

CODOHS

COFACTORS-OR-PROSTHETIC-GROUPS-OF

"tRNA(lysT) is one of six lysine tRNAs.

tRNAs are the adapters that allow synthesis of proteins from mRNAs. Each tRNA carries a specific amino acid to the ribosome for protein synthesis. There, the tRNA recognizes an RNA codon with its own three-nucleotide anticodon, thus allowing synthesis of a specific peptide based on an mRNA template.

tRNAs are processed to their active, mature forms by RNA cleavage and by modification of their bases. RNA cleavage consists of removal of both 5' and 3' extensions in a multistep process involving many RNases [CITS: [11252717]]. RNases taking part in tRNA processing include [FRAME: CPLX0-3461], [FRAME: CPLX0-3601], [FRAME: EG10858-MONOMER], [FRAME: EG11620-MONOMER], [FRAME: EG11620-MONOMER], [FRAME: EG11620-MONOMER], and [FRAME: CPLX0-3602]. tRNAs are also subject to a wide variety of base modifications catalyzed by proteins such as

COMMENT — [FRAME: EG11620-MONOMER], [FRAME: EG11332-MONOMER], [FRAME: EG10595-MONOMER], [FRAME: G6364-MONOMER], [FRAME: EG11344-MONOMER], [FRAME: G7199-MONOMER], [FRAME: CPLX0-1101], [FRAME: EG11779-MONOMER], [FRAME: G7422-MONOMER], [FRAME: G7449-MONOMER], [FRAME: EG11177-MONOMER], [FRAME: EG10454-MONOMER], [FRAME: EG10967-MONOMER], and [FRAME: EG11022-MONOMER].

NMR analyses of the fully modified anticodon stem-loop domain of tRNA(lysT) show that base modifications play an important role in maintaining the structure of that domain [CITS: [11027137][15924427]].

Mature tRNAs are linked via a 3' CCA sequence to their cognate amino acid in an ATP-dependent fashion by the appropriate amino-acid-tRNA synthetase, as shown in the [FRAME: TRNA-CHARGING-PWY]. Subsequently, these charged tRNAs interact with the ribosome and template mRNA to generate polypeptides. The discovery of the role of tRNA in protein synthesis is reviewed in detail in [CITS: [7033244]]."

COMMENT-INTERNAL

COMMON-NAME — tRNA^{lysT}

COMPONENT-OF

COMPONENTS

CREATION-DATE — 25-Mar-1997 12:16:58

CREDITS — SRI International — annotation LAST-CURATED — 3355079216

DBLINKS — shearer — annotation LAST-CURATED — 3355079216

GENE — lysT

HISTORY

INHIBITORS-ALLOSTERIC-OF
INHIBITORS-COMPETITIVE-OF
INHIBITORS-IRREVERSIBLE-OF
INHIBITORS-NONCOMPETITIVE-OF
INHIBITORS-OTHER-OF
INHIBITORS-UNCOMPETITIVE-OF

INHIBITORS-UNKMECH-OF — 3'-nucleotidase inherited from RNA
2',3'-cyclic nucleotide 2'-phosphodiesterase inherited from RNA

MODIFIED-FORM — L-lysyl-tRNA^{lysT}

MOLECULAR-WEIGHT

II-1-NAME

II-1-NAME

II-NAME

PROSTHETIC-GROUPS-OF

SCHEMA? — T inherited from LYS-tRNAs

SPLICE-FORM-ISOTHOHS

SYNONYMS

The RNA ontology is simple (right now)

