

Demos  
November 14, 2006



# Unified Medical Language System Semantic Navigator



*Olivier Bodenreider*

Lister Hill National Center  
for Biomedical Communications  
Bethesda, Maryland - USA



# Issues

## ◆ Size

- Large number of concepts (>1 million)

## ◆ Complexity

- Polyhierarchical structures
- Multiple information sources
- Multiple properties

## ◆ Lack of formality

- Redundant relations
- Hierarchies vs. hierarchical relations





# Challenges

- ◆ Restrict information space to selected information sources
- ◆ Reduce complexity
  - Group concepts by semantic groups
  - Transitive reduction on hierarchical relations
  - Select co-occurring concepts
- ◆ Reduce the cognitive burden on the user
  - Use graph-based rather than tree-based representations



# UMLS Semantic Navigator

## *SemNav*

<http://umlsks.nlm.nih.gov>\*

► SN Resources ► Semantic Navigator  
(\* free UMLS registration required)



# Unified Medical Language System®

- ◆ Developed at NLM since 1990
- ◆ 139 source vocabularies
  - 17 languages
- ◆ Broad coverage of biomedicine
  - 5.1M names
  - 1.3M concepts
  - 16M relations
- ◆ Integration
  - Synonymous terms are clustered in a concept
  - Hierarchies (trees) are combined in a graph structure





# Terminology integration Terms

Duchenne muscular dystrophy

MeSH, SNOMED  
CTV3, Jablonski,  
CRISP, DxPlain,  
MedDRA, LOINC

Duchenne's muscular dystrophy

COSTAR

Duchenne de Boulogne muscular dystrophy

Jablonski

Duchenne type progressive muscular dystrophy

SNOMED

pseudohypertrophic muscular dystrophy

MeSH, CTV3  
SNOMED

X-linked recessive muscular dystrophy

Jablonski

severe generalized familial muscular dystrophy

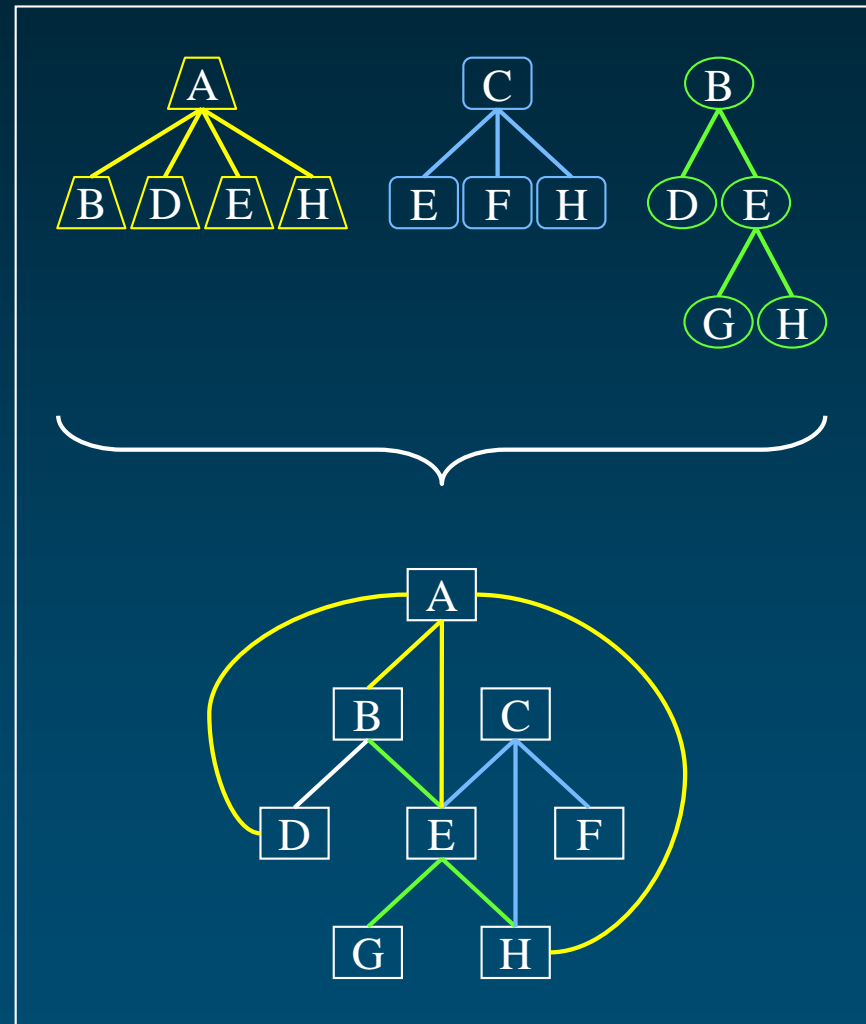
SNOMED





# Terminology integration Relationships

- ◆ Inter-concept relationships: hierarchies from the source vocabularies
- ◆ Redundancy: multiple paths
- ◆ One **graph** instead of multiple **trees** (multiple inheritance)

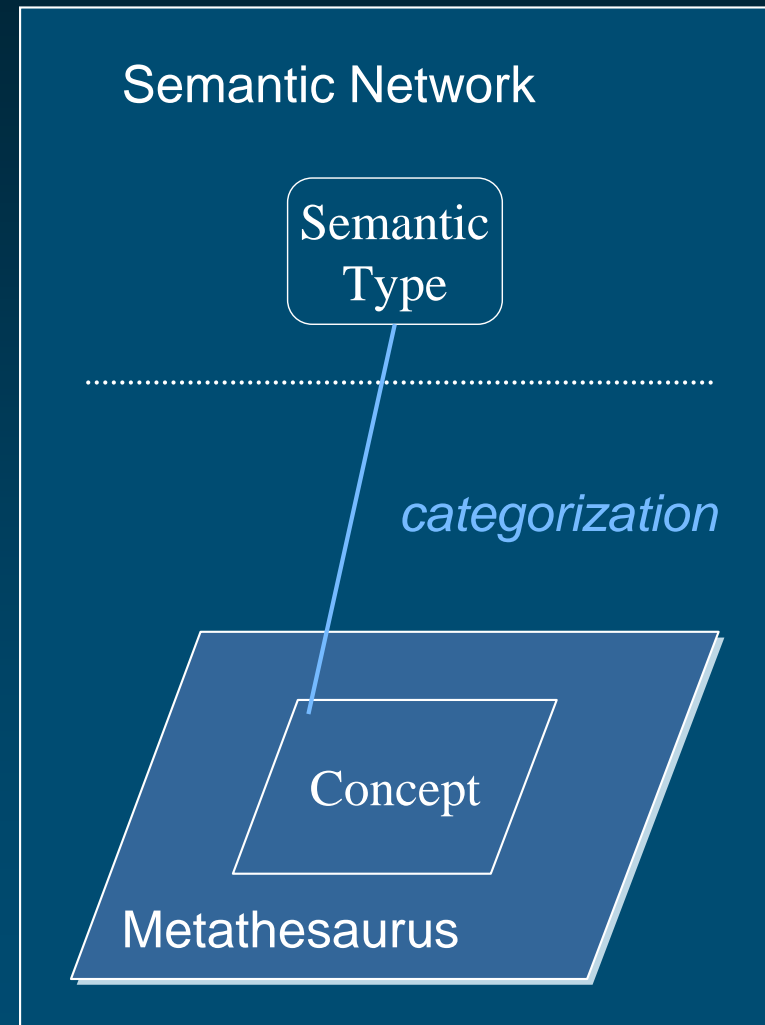




# UMLS A two-level structure

## ◆ Two-level structure

- Semantic Network
  - 135 Semantic Types (STs)
  - 54 types of relationships among STs
- Metathesaurus
  - >1M concepts
  - ~12 M inter-concept relationships
- Link = categorization





## Semantic Types

Anatomical  
Structure

Fully Formed  
Anatomical  
Structure

Embryonic  
Structure

Disease or  
Syndrome

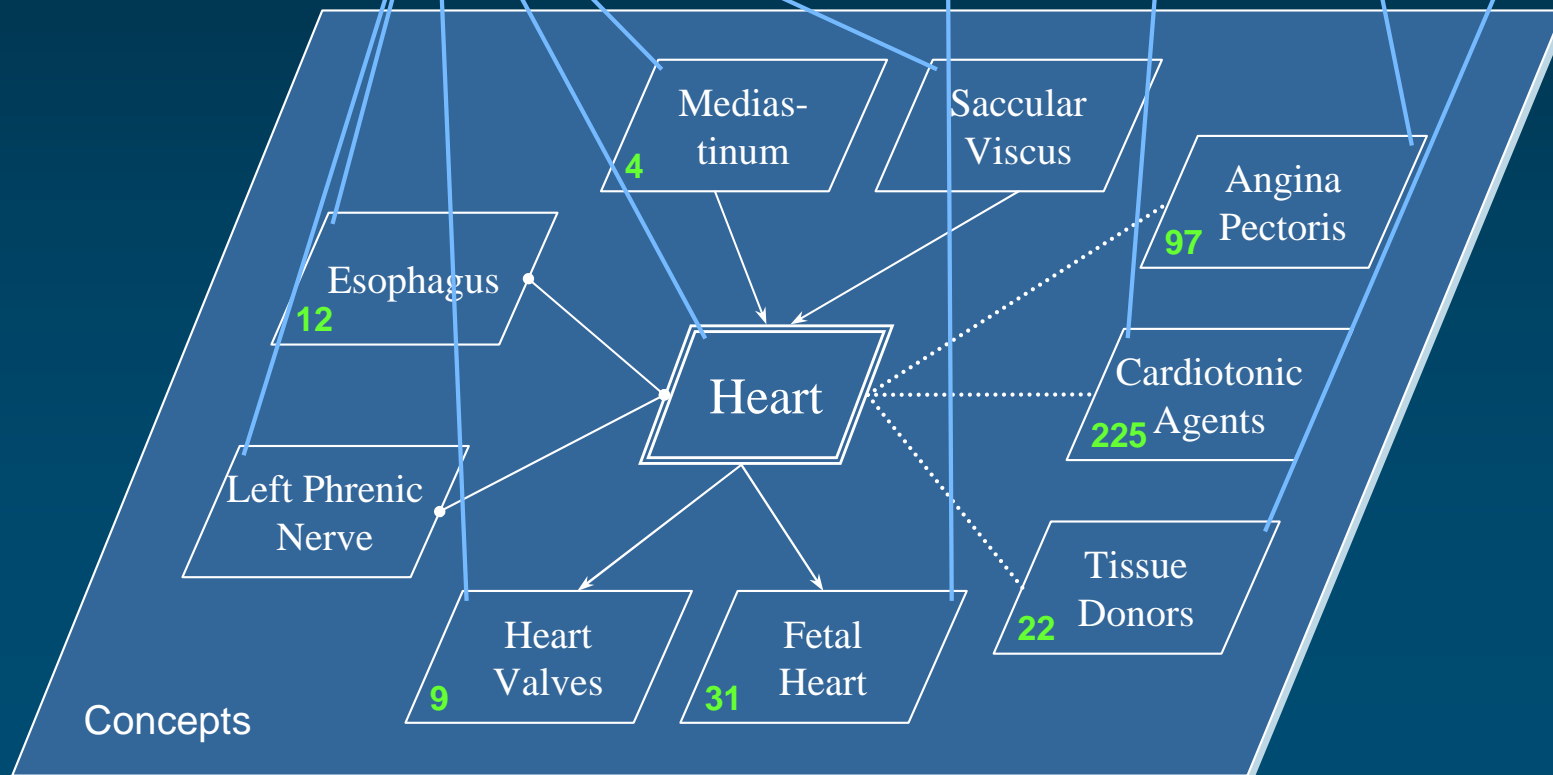
Body Part, Organ or  
Organ Component

Pharmacologic  
Substance

Population  
Group

*Semantic  
Network*

*Metathesaurus*






UMLS Knowledge Source Server (UMLSKS) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

← → ↺ × 🏠 📄

http://umlsks.nlm.nih.gov/kss/servlet/Turbine/action/DoNextSearch/template/meta%2Cconcept%2CResultScreen.vm;j:

🖨️ ➕ 🔍



# UMLS Knowledge Source Server (UMLSKS)

UMLSKS Version 5.0    UMLS Releases: 2002 2002AB 2002AC 2002AD 2003AA 2003AB 2003AC 2004AA 2004AB 2004AC 2005AA 2005AB 2005AC 2006AA 2006AB

[Metathesaurus](#)    [Semantic Network](#)    [SPECIALIST Lexicon](#)

[Home](#)    [Advanced Search](#)    [Logout](#)

Metathesaurus Search for: **dystrophin** in UMLS Release **2006AB**

[Display](#)   [Display All](#)

**Concept**

- ☒ Definition
- ☒ Synonyms
- ☐ Other Languages
- ☐ Suppressible Synonyms
- ☐ Sources

**Context**

- ☒ Ancestors
- ☐ Parents
- ☐ Siblings
- ☐ Children

**Relations**

- ☐ Narrower
- ☐ Broader
- ☐ Similar
- ☐ Other
- ☐ Related and possibly synonymous
  - ☐ Source asserted synonymy
  - ☐ Allowable Subheadings
  - ☐ Associated Expressions

**Co-occurring Concepts**

- ☐ Co-occurring MeSH
- ☐ Co-occurring AI/RHEUM

**Concept: Dystrophin**

CUI: [C0079259](#)

**Semantic Type:** [Amino Acid, Peptide, or Protein](#)  
[Biologically Active Substance](#)

**Definition:**  
A muscle protein localized in surface membranes which is the product of the Duchenne/Becker muscular dystrophy gene. Individuals with Duchenne muscular dystrophy usually lack dystrophin completely while those with Becker muscular dystrophy have dystrophin of an altered size. It shares features with other cytoskeletal proteins such as SPECTRIN and alpha-actinin but the precise function of dystrophin is not clear. One possible role might be to preserve the integrity and alignment of the plasma membrane to the myofibrils during muscle contraction and relaxation. MW 400 kDa. [\(MeSH\)](#)

large, structural, spectrin-like protein expressed in skeletal muscle; genetic defect is linked to Duchenne and Becker muscular dystrophy. [\(CRISP Thesaurus\)](#)

**Synonyms:**  
[Dystrophin](#)

**Ancestors:**

**MeSH**

- [MeSH Descriptors](#) []
- [Index Medicus Descriptor](#) []
- [Chemicals and Drugs \(MeSH Category\)](#) [D]
- [Amino Acids, Peptides, and Proteins](#) []

Done



UMLS Knowledge Source Server (UMLSKS) - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

← → ↺ × 🏠

http://umlsks.nlm.nih.gov/kss/servlet/Turbine/action/DoNextSearch/template/meta%2Cconcept%2CResultScreen.vm;j=

🖨️ ➕ 🔍

☐ Co-occurring MeSH

☐ Co-occurring AI/RHEUM

[MeSH Descriptors \[\]](#)

[Index Medicus Descriptor \[\]](#)

[Chemicals and Drugs \(MeSH Category\) \[D\]](#)

[Amino Acids, Peptides, and Proteins \[\]](#)

[Proteins \[\]](#)

[Contractile Proteins \[\]](#)

[Muscle Proteins \[\]](#)

[Dystrophin \[\]](#)

[...]

**CRISP Thesaurus**

[chemical \[\]](#)

[organic chemical \[\]](#)

[amide \[\]](#)

[peptide \[\]](#)

[protein \[\]](#)

[cytoskeletal protein \[\]](#)

[actin binding protein \[\]](#)

[dystrophin \[\]](#)

[...]

**Digital Anatomist**

[Anatomical entity \[\]](#)

[Physical anatomical entity \[\]](#)

[Material physical anatomical entity \[\]](#)

[Anatomical structure \[\]](#)

[Biological macromolecule \[\]](#)

[Protein \[\]](#)

[Actin-binding protein \[\]](#)

[Dystrophin \[\]](#)



[Amino Acids, Peptides, and Proteins \[D12\]](#)

[Proteins \[D12.776\]](#)

[Contractile Proteins \[D12.776.210\]](#)

[Muscle Proteins \[D12.776.210.500\]](#)

[Actinin \[D12.776.210.500.095\]](#)

[Actins \[D12.776.210.500.100\]](#)

[Actomyosin \[D12.776.210.500.154\]](#)

[Calsequestrin \[D12.776.210.500.220\]](#)

[CapZ Actin Capping Protein \[D12.776.210.500.227\]](#)

[Caveolin 3 \[D12.776.210.500.235\]](#)

[Cofilin 2 \[D12.776.210.500.242\]](#)

► [Dystrophin \[D12.776.210.500.250\]](#)

[Dystrophin-Associated Proteins \[D12.776.210.500.410\] +](#)

[Myogenic Regulatory Factors \[D12.776.210.500.570\] +](#)

[Myoglobin \[D12.776.210.500.588\]](#)

[Myosins \[D12.776.210.500.600\] +](#)

[Parvalbumins \[D12.776.210.500.750\]](#)

[Profilins \[D12.776.210.500.775\]](#)

[Ryanodine Receptor Calcium Release Channel \[D12.776.210.500.800\]](#)

[Tropomodulin \[D12.776.210.500.847\]](#)

[Tropomyosin \[D12.776.210.500.895\]](#)

[Troponin \[D12.776.210.500.910\] +](#)

[Amino Acids, Peptides, and Proteins \[D12\]](#)

[Proteins \[D12.776\]](#)

[Membrane Proteins \[D12.776.543\]](#)

[Ankyrins \[D12.776.543.080\]](#)

[Arrestins \[D12.776.543.090\] +](#)

[Bacterial Outer Membrane Proteins \[D12.776.543.100\] +](#)

[Calnexin \[D12.776.543.162\]](#)

[Connexins \[D12.776.543.225\] +](#)

► [Dystrophin \[D12.776.543.250\]](#)

[Dystrophin-Associated Proteins \[D12.776.543.268\] +](#)

[Ephrins \[D12.776.543.287\] +](#)

# MeSH Browser



[Amino Acids, Peptides, and Proteins \[D12\]](#)

[Proteins \[D12.776\]](#)

[Cytoskeletal Proteins \[D12.776.220\]](#)

[Adenomatous Polyposis Coli Protein \[D12.776.220.040\]](#)

[Catenins \[D12.776.220.145\] +](#)

► [Dystrophin \[D12.776.220.250\]](#)

[Dystrophin-Associated Proteins \[D12.776.220.362\] +](#)

[Intermediate Filament Proteins \[D12.776.220.475\] +](#)

[Microfilament Proteins \[D12.776.220.525\] +](#)

[Microtubule Proteins \[D12.776.220.600\] +](#)

[Plakins \[D12.776.220.790\] +](#)

[Plakophilins \[D12.776.220.885\]](#)

[Spectrin \[D12.776.220.980\]](#)

[Talin \[D12.776.220.985\]](#)

[Utrophin \[D12.776.220.987\]](#)

[Vinculin \[D12.776.220.990\]](#)





## Profilament Proteins

### Binding Protein

### Contractile Proteins

Protein, Organized by Location

proteins

Actin-Binding Protein

## Muscle Proteins

## Membrane Proteins

Dystrophin

Dp260 protein, rat

140-kDa dystrophin

DMD protein, rat

apo-dystrophin 1



- 1

Chemicals &amp; Drugs

- 120-kDa hemocyte-specific membrane protein, flesh fly ☐
- 15a protein, *Aedes aegypti* ☐
- 22.6-kDa antigen, *Schistosoma japonicum* ☐
- 22kDa polypeptide, *Nicotiana tabacum* ☐
- 3D3-lyric protein, human ☐
- 3D3-lyric protein, mouse ☐
- 4.1B protein, mouse ☐
- 5-lipoxygenase-act protein ☐
- 7 kDa protein, carlavirus ☐
- A14.5L protein,

Done



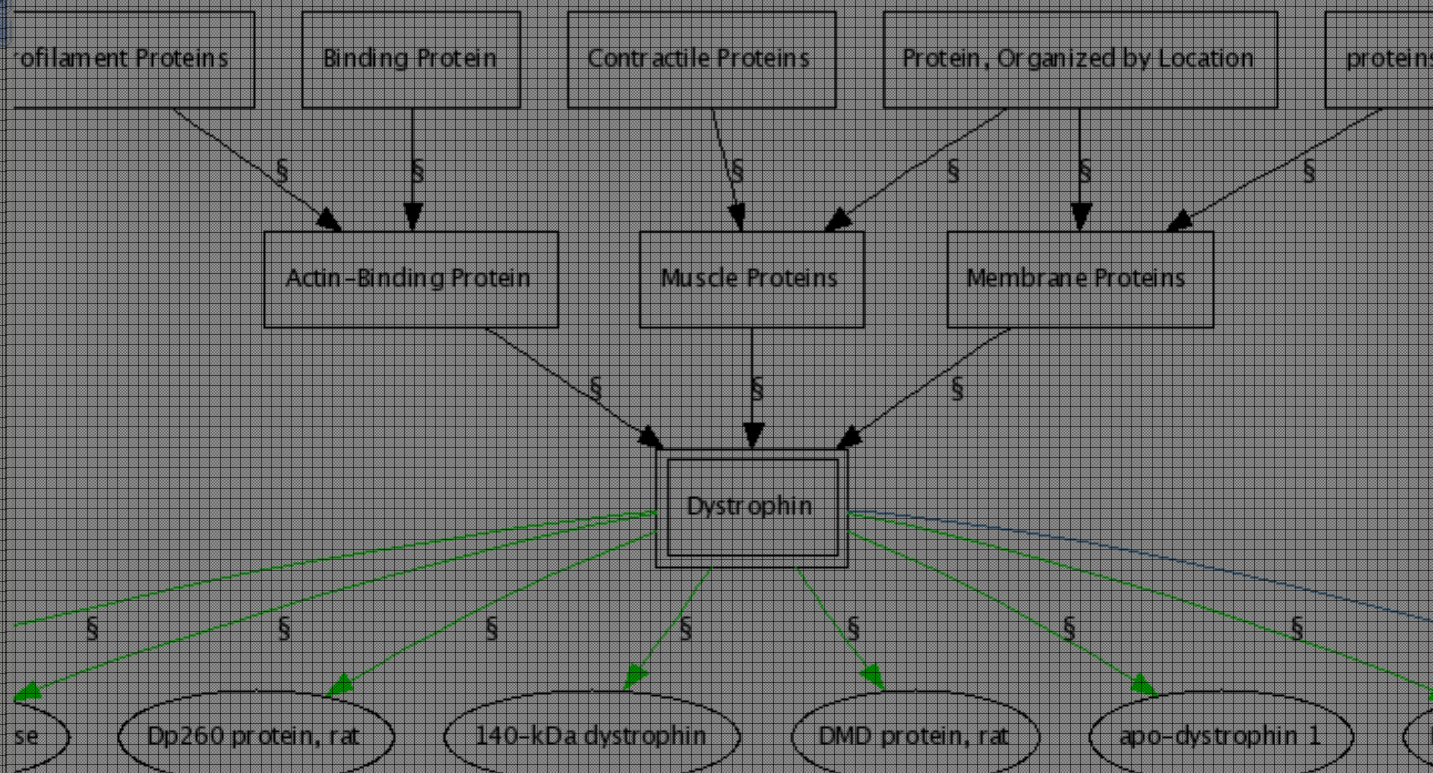




## Siblings

## Chemicals & Drugs

- 120-kDa hemocyte-specific membrane protein, flesh fly
- 15a protein, Aedes aegypti
- 22.6-kDa antigen, Schistosoma japonicum
- 22kDa polypeptide, Nicotiana tabacum
- 3D3-lync protein, human
- 3D3-lync protein, mouse
- 4.1B protein, mouse
- 5-lipoxygenase-act protein
- 7 kDa protein, carlavirus
- A14.5L protein,



## Other Related Concepts

### Disorders

- Muscular Dystrophies
- Muscular Dystrophy, Duchenne

### Living Beings

- Mice, Inbred mdx

## Co-occurring Concepts

### Anatomy

- Astrocytes [4]
- Brain [24]
- Cell membrane [7]
- Cytoskeleton [5]
- Diaphragm (Anatomy) [6]
- Heart [9]
- Hippocampus (Brain) [9]

BCI

Dystrophin

LEGEND

Start again

Apply new parameters

Restrict to vocabulary:

Show all

Highlight vocabulary:

Nothing

UMLS data:

UMLS\_2006AA

## Similar Concepts

(none)

## Allegedly

## Closest MeSH Terms

(none)



- 120-kDa hemocyte-specific membrane protein, flesh fly ☐
- 15a protein, *Aedes aegypti* ☐
- 22.6-kDa antigen, *Schistosoma japonicum* ☐
- 22kDa polypeptide, *Nicotiana tabacum* ☐
- 3D3-lyric protein, human ☐
- 3D3-lyric protein, mouse ☐
- 4.1B protein, mouse ☐
- 5-lipoxygenase-act protein ☐
- 7 kDa protein, calvarius ☐
- A14.5L protein,



- ◆ Astrocytes [4] ☐
- ◆ Brain [24] ☐
- ◆ Cell membrane [7] ☐
- ◆ Cytoskeleton [5]
- ◆ Diaphragm (Anatomy) [6] ☐
- ◆ Heart [9] ☐
- ◆ Hippocampus (Brain) [9] ☐

Done





10



- 



- 



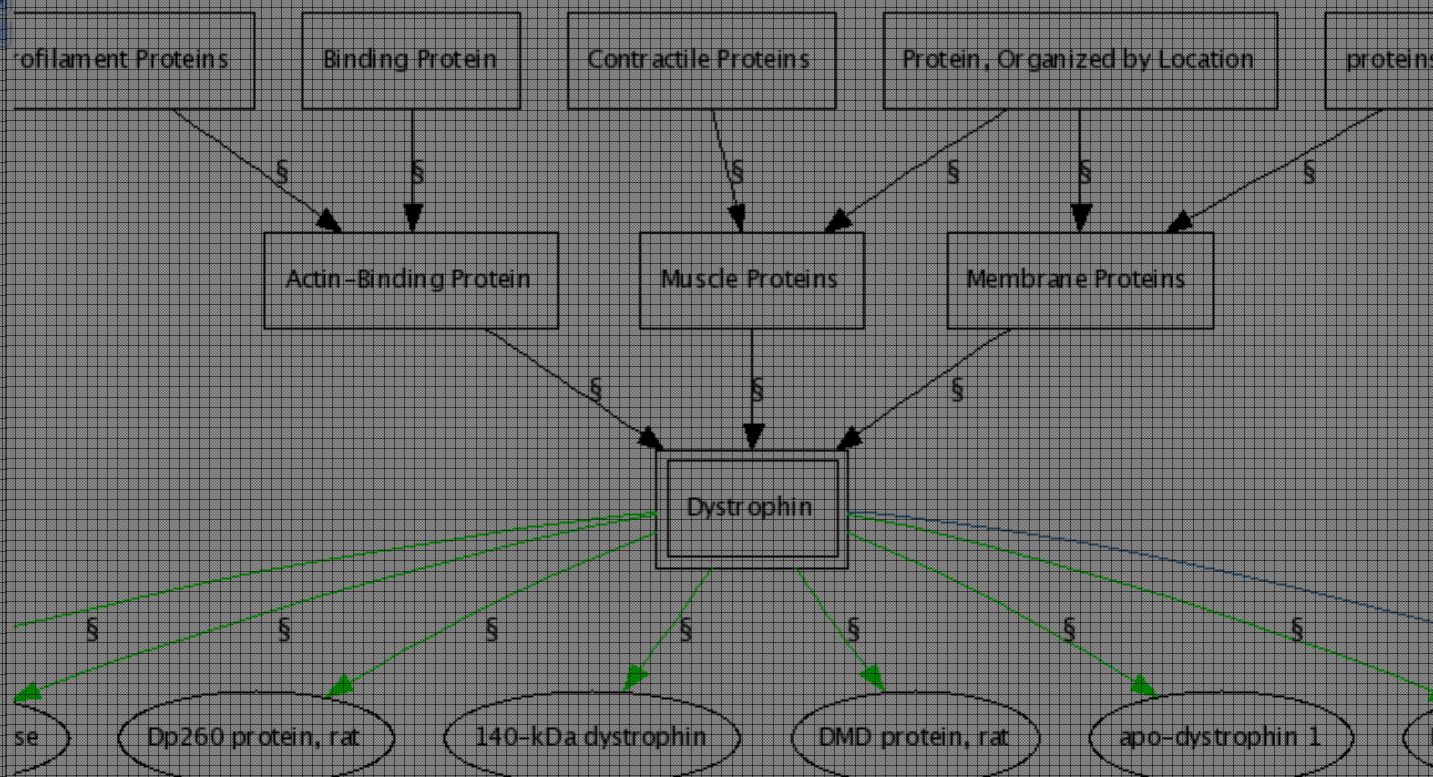
- 



## Siblings

## Chemicals & Drugs

- 120-kDa hemocyte-specific membrane protein, flesh fly ☐
- 15a protein, Aedes aegypti ☐
- 22.6-kDa antigen, Schistosoma japonicum ☐
- 22kDa polypeptide, Nicotiana tabacum ☐
- 3D3-lync protein, human ☐
- 3D3-lync protein, mouse ☐
- 4.1B protein, mouse ☐
- 5-lipoxygenase-act protein ☐
- 7 kDa protein, carlavirus ☐
- A14.5L protein, ☐



## Other Related Concepts

### Disorders

- Muscular Dystrophies ☐
- Muscular Dystrophy, Duchenne ☐

### Living Beings

- Mice, Inbred mdx ☐

## Co-occurring Concepts

### Anatomy

- Astrocytes [4] ☐
- Brain [24] ☐
- Cell membrane [7] ☐
- Cytoskeleton [5] ☐
- Diaphragm (Anatomy) [6] ☐
- Heart [9] ☐
- Hippocampus (Brain) [1] ☐

BCI

Dystrophin

LEGEND

Start again

Apply new parameters

Restrict to vocabulary:

Show all

Highlight vocabulary:

Nothing

UMLS data:

UMLS 2006AA

## Similar Concepts

(none)

## Allegedly

## Closest MeSH Terms

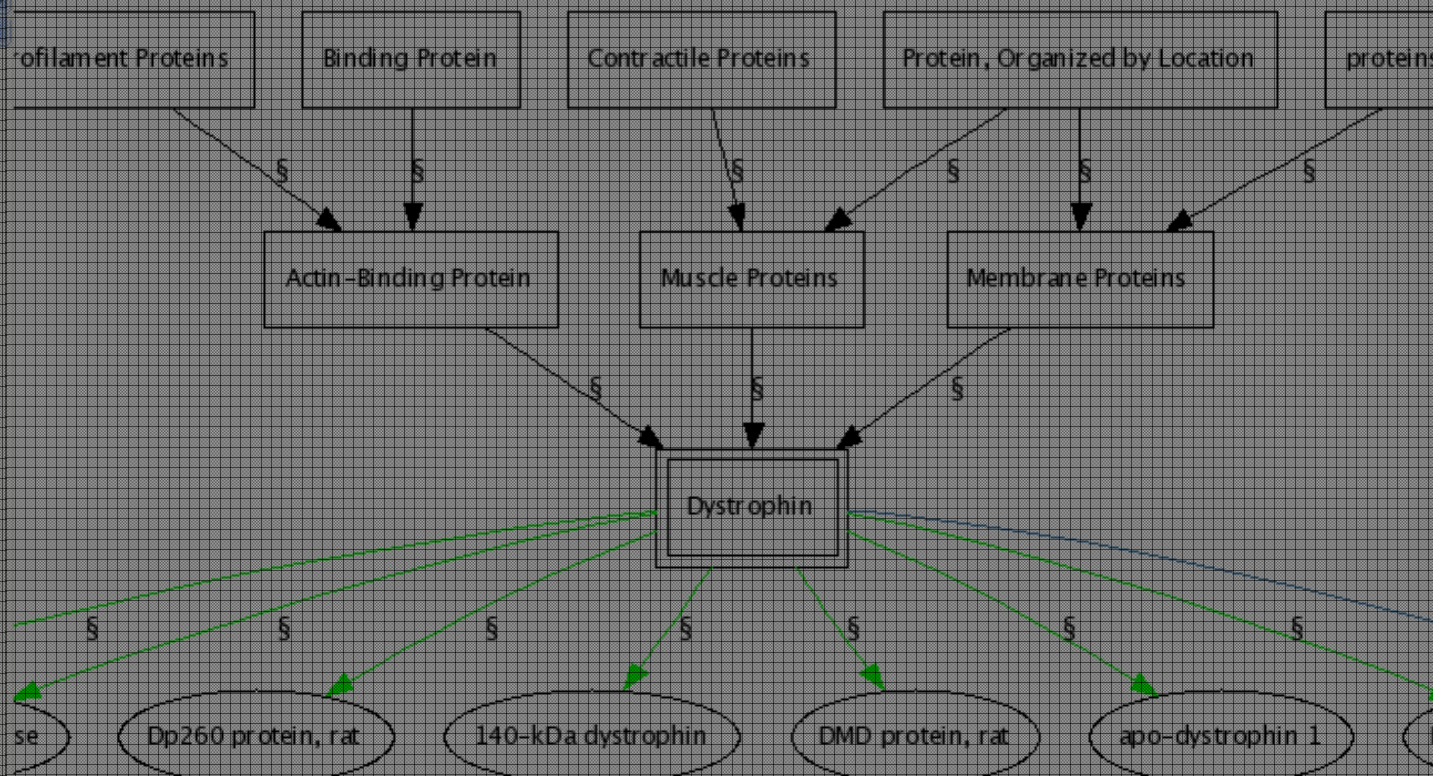
(none)



## Siblings

## Chemicals & Drugs

- 120-kDa hemocyte-specific membrane protein, flesh fly
- 15a protein, Aedes aegypti
- 22.6-kDa antigen, Schistosoma japonicum
- 22kDa polypeptide, Nicotiana tabacum
- 3D3-lync protein, human
- 3D3-lync protein, mouse
- 4.1B protein, mouse
- 5-lipoxygenase-act protein
- 7 kDa protein, carlavirus
- A14.5L protein,



## Other Related Concepts

### Disorders

- Muscular Dystrophies
- Muscular Dystrophy, Duchenne

### Living Beings

- Mice, Inbred mdx

## Co-occurring Concepts

### Anatomy

- Astrocytes [4]
- Brain [24]
- Cell membrane [7]
- Cytoskeleton [5]
- Diaphragm (Anatomy) [6]
- Heart [9]
- Hippocampus (Brain) [9]

BCI

Dystrophin

LEGEND

Start again

Apply new parameters

Restrict to vocabulary:

Show all

Highlight vocabulary:

Nothing

UMLS data:

UMLS\_2006AA

## Similar Concepts

(none)

## Allegedly

## Closest MeSH Terms

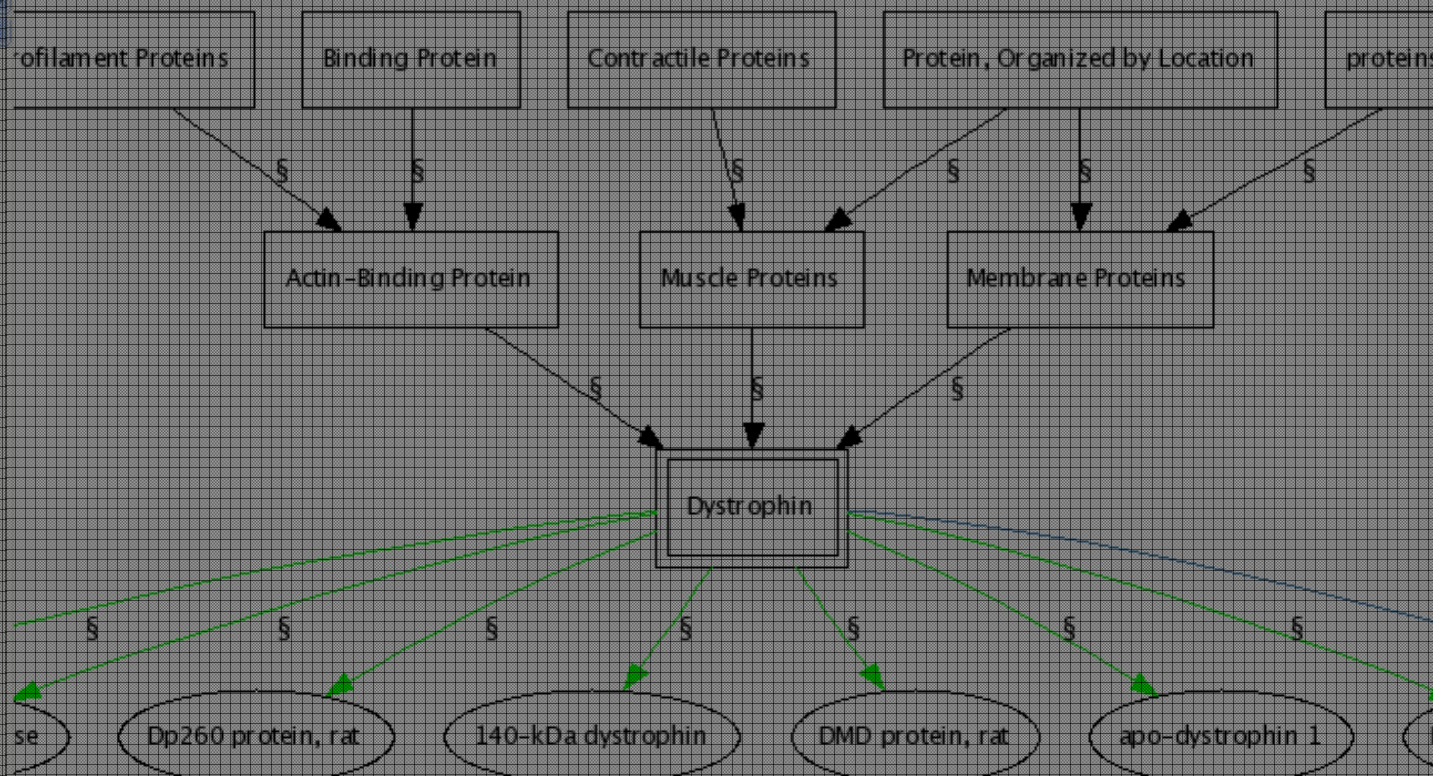
(none)



## Siblings

## Chemicals & Drugs

- 120-kDa hemocyte-specific membrane protein, flesh fly
- 15a protein, Aedes aegypti
- 22.6-kDa antigen, Schistosoma japonicum
- 22kDa polypeptide, Nicotiana tabacum
- 3D3-lync protein, human
- 3D3-lync protein, mouse
- 4.1B protein, mouse
- 5-lipoxygenase-act protein
- 7 kDa protein, carlavirus
- A14.5L protein,



## Other Related Concepts

### Disorders

- Muscular Dystrophies
- Muscular Dystrophy, Duchenne

### Living Beings

- Mice, Inbred mdx

## Co-occurring Concepts

### Anatomy

- Astrocytes [4]
- Brain [24]
- Cell membrane [7]
- Cytoskeleton [5]
- Diaphragm (Anatomy) [6]
- Heart [9]
- Hippocampus (Brain) [9]

BCI

Dystrophin

LEGEND

Start again

Apply new parameters

Restrict to vocabulary:

Show all

Highlight vocabulary:

Nothing

UMLS data:

UMLS 2006AA

## Similar Concepts

(none)

## Allegedly

## Closest MeSH Terms

(none)



# SemNav Visualization options

Start again    Apply new parameters

**Restrict to vocabulary:** Show all ▼

**Highlight vocabulary:** Nothing ▼

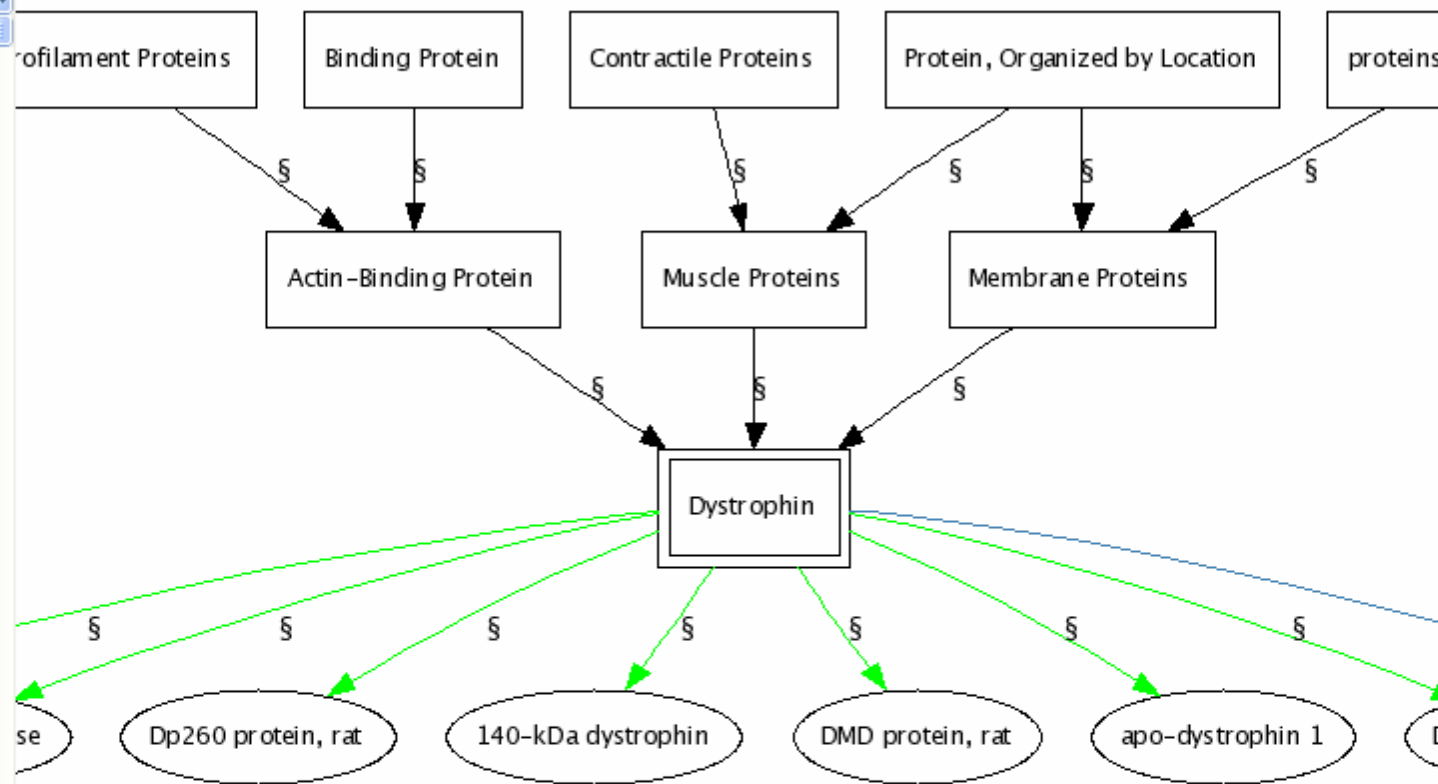
**UMLS data:** UMLS\_2002 ▼

**Type of hierarchical rel:** ☒ All ☐ Parent/Child only ☐ Broader/Narrower only

**Transitive reduction:** ☒ yes ☐ no

Start again    Apply new parameters





- ◆ 120-kDa hemocyte-specific membrane protein, flesh fly □
- ◆ 15a protein, *Aedes aegypti* □
- ◆ 22.6-kDa antigen, *Schistosoma japonicum* □
- ◆ 22kDa polypeptide, *Nicotiana tabacum* □
- ◆ 3D3-lyric protein, human □
- ◆ 3D3-lyric protein, mouse □
- ◆ 4.1B protein, mouse □
- ◆ 5-lipoxygenase-act protein □
- ◆ 7 kDa protein, carlavirus □
- ◆ A14.5L protein,

- Astrocytes [4] ☐
- Brain [24] ☐
- Cell membrane [7] ☐
- Cytoskeleton [5]
- Diaphragm (Anatomy) [6] ☐
- Heart [9] ☐
- Hippocampus (Brain) [9] ☐

Done



UMLS\_2006AA

UMLS® Semantic Navigator

[ 2.12 ] - Mozilla Firefox

File

Edit

View

Go

Bookmarks

Tools

Help

←

→

↺

✕

🏠

🔍

http://mor.nlm.nih.gov/perl/semnav.pl

🖨

📄

🔗

Siblings

Chemicals & Drugs

◆ Actinin ☐

◆ Actins ☐

◆ Actomyosin ☐

◆ Adenomatous Polyposis Coli Protein ☐

◆ Ankyrins ☐

◆ Arrestins ☐

◆ Bacterial Outer Membrane Proteins ☐

◆ Calnexin ☐

◆ Calsequestrin ☐

◆ CapZ Actin Capping Protein ☐

◆ Catenins ☐

◆ Caveolin 3 ☐

◆ Cofilin 2 ☐

◆ Connexins ☐

◆ Dystrophin-Associated Proteins ☐

◆ Ephrins ☐

◆ Heterotrimeric GTP-Binding Proteins ☐

Proteins

§

Contractile Proteins

§

Cytoskeletal Proteins

§

Membrane Proteins

§

Muscle Proteins

§

Dystrophin

Other Related Concepts

Disorders

◆ Muscular Dystrophy, Duchenne ☐

Living Beings

◆ Mice, Inbred mdx ☐

(2 other related concepts)

BCI

Dystrophin

LEGEND

Start again

Apply new parameters

Restrict to vocabulary:

MeSH

Highlight vocabulary:

Nothing

UMLS data:

UMLS 2006AA

Similar Concepts

(none)

Allegedly

Closest MeSH Terms

(none)

Co-occurring Concepts

(not displayed: Restrict to

Done



UMLS\_2006AA

UMLS® Semantic Navigator

[2.12] - Mozilla Firefox

File

Edit

View

Go

Bookmarks

Tools

Help

←

→

↺

✕

🏠

🌐

http://mor.nlm.nih.gov/perl/semnav.pl

🖨

📄

🔍

Siblings

Chemicals & Drugs

- 120-kDa hemocyte-specific membrane protein, flesh fly
- 15a protein, Aedes aegypti
- 22.6-kDa antigen, Schistosoma japonicum
- 22kDa polypeptide, Nicotiana tabacum
- 3D3-lyric protein, human
- 3D3-lyric protein, mouse
- 4.1B protein, mouse
- 5-lipoxygenase-act protein
- 7 kDa protein, carlavirus
- A14.5L protein,

Actin-binding protein

Membrane Proteins

Muscle Proteins

Dystrophin

DMD protein, human

dys-1 protein, C elegans

apo-dystrophin 1

dystrophin-related protein

Other Related Concepts

Disorders

- Muscular Dystrophies
- Muscular Dystrophy, Duchenne

Living Beings

- Mice, Inbred mdx

Co-occurring Concepts

Anatomy

- Astrocytes [4]
- Brain [24]
- Cell membrane [7]
- Cytoskeleton [5]
- Diaphragm (Anatomy) [6]
- Heart [9]
- Hippocampus (Brain) [9]

BCI

Dystrophin

LEGEND

Start again

Apply new parameters

Restrict to vocabulary:

Show all

Highlight vocabulary:

Digital Anatomist

UMLS data:

UMLS\_2006AA

Similar Concepts

(none)

Allegedly

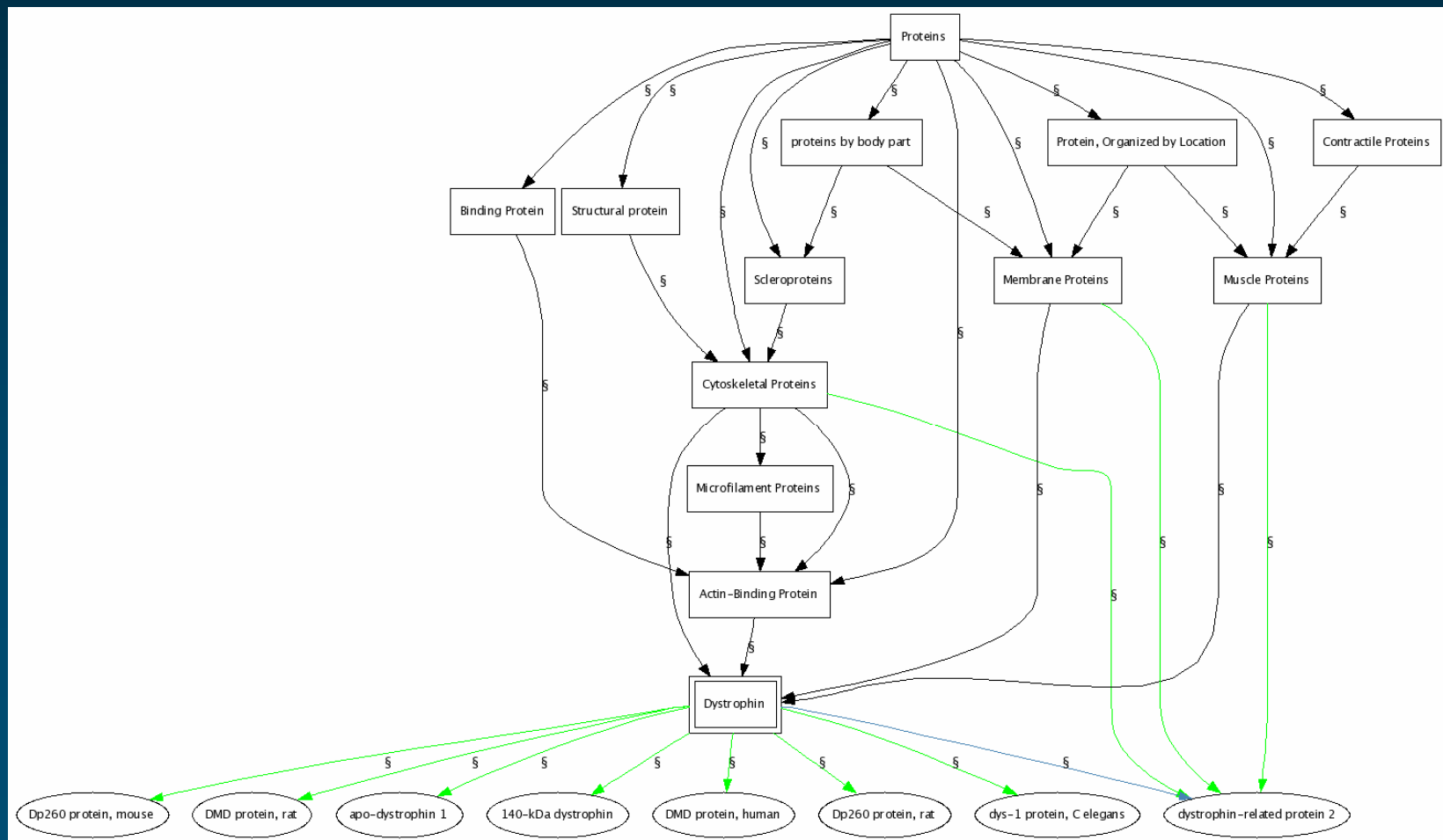
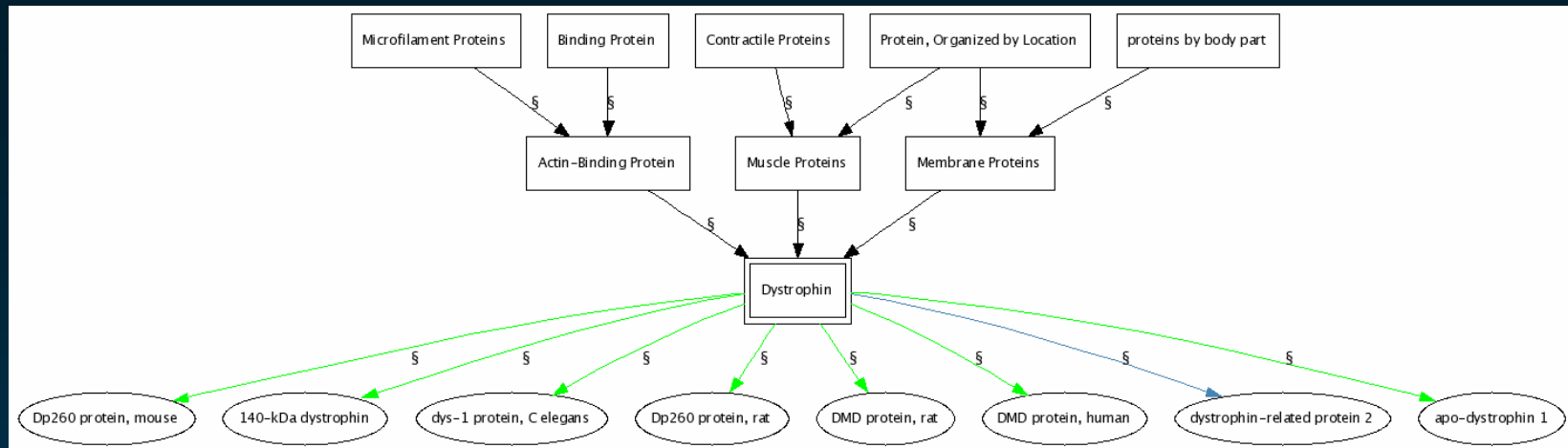
(none)

Closest MeSH Terms

(none)

javascript:rels\_on('C0079259', 'C0013264', 'UMLS\_2006AA')







http://mor.nlm.nih.gov - Relationship Viewer

Relationships  
of **Dystrophin** (C1)  
*Amino Acid, Peptide, or Protein*  
*Biologically Active Substance*  
to **Muscular Dystrophy, Duchenne** (C2)  
*Disease or Syndrome*

---

**Metathesaurus Relationships**

C1 otherwise related to C2

not defined	♦ MeSH
-------------	--------

C1 co-occurs with C2

Frequency = 190	♦ MEDLINE
-----------------	-----------

**Semantic Network Relationships**

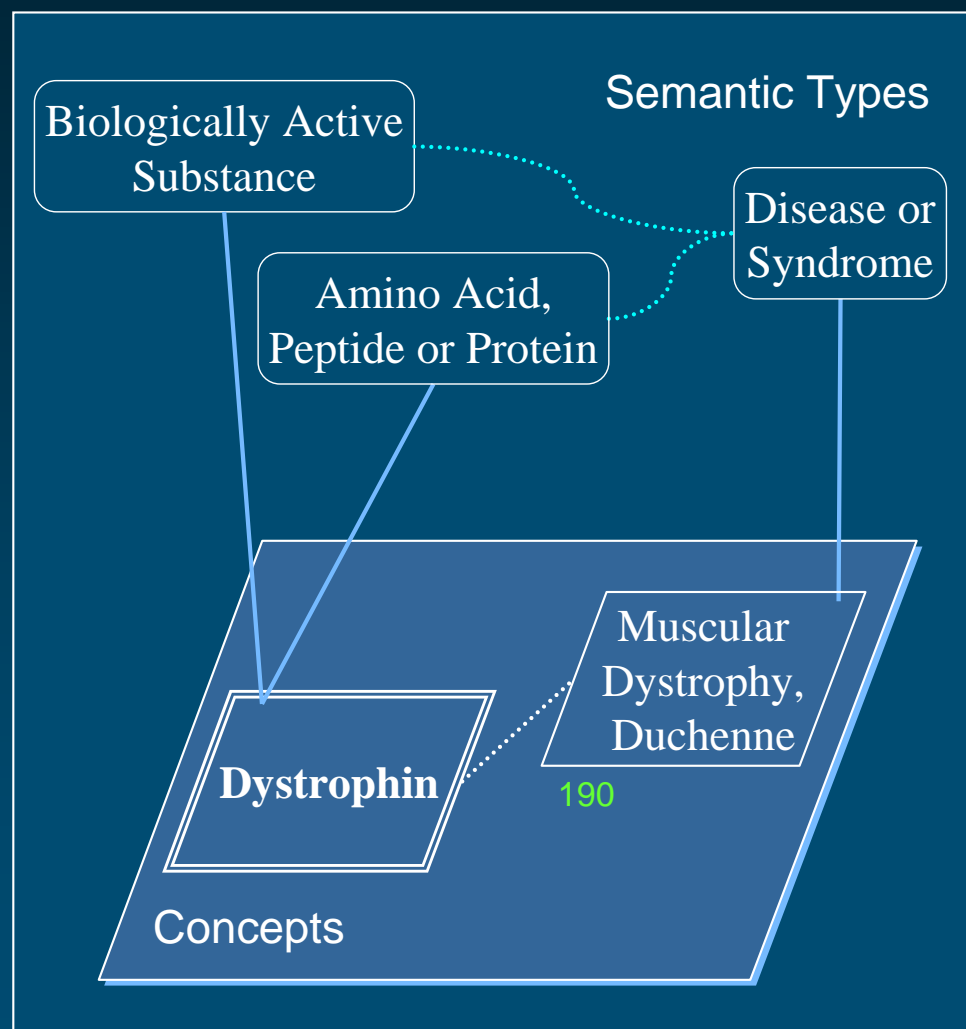
<i>Amino Acid, Peptide, or Protein</i>	♦ affects ♦ causes	<i>Disease or Syndrome</i>
<i>Biologically Active Substance</i>	♦ affects ♦ causes ♦ complicates ♦ produced_by	<i>Disease or Syndrome</i>

[Close this window](#)

Interface version: 2.01 UMLS data: UMLS\_2006AA

Done

# SemNav Relationships



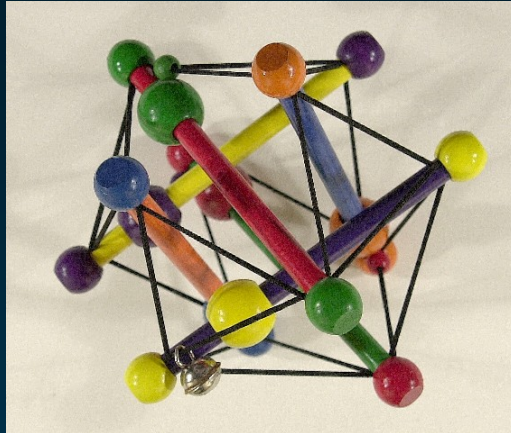


# Technical details

- ◆ Simple web/cgi technology (apache, Perl)
- ◆ dot (GraphViz)
  - PNG file (-Tpng)
  - Client-side map (-Tcmap)
- ◆ Precompute the transitive closure on hierarchical relations to perform the transitive closure fast
- ◆ Remove cycles (UMLS)







# Medical Ontology Research

Contact: [olivier@nlm.nih.gov](mailto:olivier@nlm.nih.gov)

Web: [mor.nlm.nih.gov](http://mor.nlm.nih.gov)



*Olivier Bodenreider*

Lister Hill National Center  
for Biomedical Communications  
Bethesda, Maryland - USA