

We have a table regarding various protein properties. We plan to use the data for discovering novel drug combinations for multi target drugs.

Please describe shortly each of the properties.

Please add a hashtag regarding: expression, localization, interaction, pathology.

Please add a hashtag regarding the protein property confidence: low, medium, high.

Please add an hashtag regarding the importance of a property for achieving our goal, discovering novel drug combinations for multi target drugs.

Table columns:

COG_ONTOLOGY,GOTERM_BP_DIRECT,GOTERM_CC_DIRECT,GOTERM_MF_DIRECT,INTERPRO,KEGG_PATHWAY,OMIM_DISEASE,PIR_SUPERFAMILY,SMART,UP_KW_BIOLOGICAL_PROCESS,UP_KW_CELLULAR_COMPONENT,UP_KW_DISEASE,UP_KW_DOMAIN,UP_KW_LIGAND,UP_KW_MOLECULAR_FUNCTION,UP_KW_PTM,UP_SEQ_FEATURE

Here's a concise description of each protein property with requested hashtags:

1. COG_ONTOLOGY

Evolutionary classification of proteins into functional groups.

#interaction (functional associations) | #high confidence | #high importance

Identifies conserved protein families for cross-species target validation^[1]

2. GOTERM_BP_DIRECT

Direct Gene Ontology annotations for biological processes.

#expression (pathway activity) | #high confidence | #high importance

Reveals protein roles in disease-relevant pathways^{[2] [3]}

3. GOTERM_CC_DIRECT

Direct Gene Ontology annotations for cellular components.

#localization | #high confidence | #medium importance

Identifies subcellular targets for drug delivery^{[3] [4]}

4. GOTERM_MF_DIRECT

Direct Gene Ontology annotations for molecular functions.

#interaction (binding/activity) | #high confidence | #high importance

Pinpoints druggable functional domains^{[2] [3]}

5. INTERPRO

Protein domain/family identification.

#interaction (structural motifs) | #high confidence | #high importance

Highlights multi-target binding sites^{[5] [6]}

6. KEGG_PATHWAY

Mapped biological pathway memberships.

#expression (pathway context) | #high confidence | #critical importance

Key for polypharmacology strategies^{[7] [8]}

7. OMIM_DISEASE

Genetic disease associations.

#pathology | #medium confidence | #high importance

Prioritizes clinically relevant targets^{[7] [9]}

8. PIR_SUPERFAMILY

Protein family classification.

#interaction (evolutionary relationships) | #high confidence | #medium importance

Supports target family-based screening^{[1] [5]}

9. SMART

Domain architecture analysis.

#interaction (functional modules) | #high confidence | #high importance

Identifies multi-domain targeting opportunities^{[5] [6]}

10. UP_KW_BIOLOGICAL_PROCESS

UniProt-curated process annotations.

#expression | #medium confidence | #medium importance

Contextualizes target roles^{[5] [9]}

11. UP_KW_CELLULAR_COMPONENT

UniProt-curated localization data.

#localization | #medium confidence | #medium importance

Informs drug delivery strategies^{[5] [9]}

12. UP_KW_DISEASE

UniProt-curated disease links.

#pathology | #medium confidence | #high importance

Identifies comorbid targets^{[7] [9]}

13. UP_KW_DOMAIN

Functional domain annotations.

#interaction | #medium confidence | #high importance

Reveals multi-target binding potential^{[5] [6]}

14. UP_KW_LIGAND

Binding partner information.

#interaction | #medium confidence | #critical importance

Critical for combination therapy design^{[8] [6]}

15. UP_KW_MOLECULAR_FUNCTION

Functional activity annotations.

#interaction | #medium confidence | #high importance

Identifies synergistic target activities^{[5] [3]}

16. UP_KW_PTM

Post-translational modification sites.

#interaction (regulation) | #medium confidence | #high importance

Highlights druggable regulatory nodes^[9] ^[6]

17. UP_SEQ_FEATURE

Sequence characteristics/domains.

#interaction (structural features) | #medium confidence | #high importance

Guides multi-epitope targeting^[5] ^[9]

Key patterns:

- #interaction properties (#high importance) dominate drug target discovery
- Pathway (#KEGG) and domain (#INTERPRO/SMART) data are most critical for multi-target strategies
- Disease annotations (#OMIM/UP_KW_DISEASE) provide clinical relevance filters^[7] ^[9]
- High-confidence annotations (#COG/GOTERM) enable reliable target prioritization^[1] ^[2] ^[3]

✱

Top 10 Protein Properties by Importance for Multi-Target Drug Discovery

Based on your goal-discovering novel drug combinations for multi-target drugs-the following properties are prioritized for their relevance to target selection, druggability, and polypharmacology:

1. KEGG_PATHWAY

Membership in biological pathways; reveals how proteins function together and identifies key nodes for multi-target intervention.

#expression #high #critical

2. INTERPRO

Protein domain/family identification; highlights structural features and binding sites for multi-target design.

#interaction #high #high

3. GOTERM_MF_DIRECT

Molecular function annotations; pinpoints enzymatic or binding activities relevant for drug targeting.

#interaction #high #high

4. COG_ONTOLOGY

Evolutionary classification; identifies conserved protein families and functional groups, supporting cross-species and multi-target strategies.

#interaction #high #high

5. SMART

Domain architecture analysis; identifies modular domains that can be targeted by multi-

functional drugs.

#interaction #high #high

6. UP_KW_LIGAND

Binding partner information; critical for understanding potential drug interactions and synergy.

#interaction #medium #critical

7. OMIM_DISEASE

Genetic disease associations; prioritizes clinically relevant targets and links target modulation to therapeutic outcomes.

#pathology #medium #high

8. UP_KW_DOMAIN

Functional domain annotations; provides additional details on structure-function relationships for multi-target approaches.

#interaction #medium #high

9. UP_KW_PTM

Post-translational modification sites; highlights regulatory nodes that may be druggable for combinatorial effects.

#interaction #medium #high

10. GOTERM_BP_DIRECT

Biological process annotations; contextualizes target roles within disease-relevant processes and pathways.

#expression #high #high

Rationale

- **Pathway and domain information** (KEGG_PATHWAY, INTERPRO, SMART, UP_KW_DOMAIN) are essential for identifying proteins that act together or share druggable features, enabling rational multi-target drug design^{[10] [11]}.
- **Molecular function and ligand binding** (GOTERM_MF_DIRECT, UP_KW_LIGAND) are critical for identifying actionable targets and understanding how drugs can modulate multiple proteins^{[12] [13]}.
- **Disease associations** (OMIM_DISEASE) ensure clinical relevance and therapeutic impact.
- **Post-translational modifications** (UP_KW_PTM) and **biological process context** (GOTERM_BP_DIRECT) provide insight into regulatory mechanisms and functional networks, supporting the design of effective drug combinations^{[12] [11]}.

These properties collectively maximize the likelihood of identifying effective, specific, and clinically relevant multi-target drug combinations.

✱✱

1. <https://pmc.ncbi.nlm.nih.gov/articles/PMC102395/>

2. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11536429/>

3. <http://geneontology.org/docs/go-annotations/>

4. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6668538/>
5. <https://www.sciencedirect.com/science/article/pii/S2667074724000260>
6. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8185533/>
7. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4379170/>
8. <https://www.frontiersin.org/journals/chemistry/articles/10.3389/fchem.2021.634663/full>
9. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11305224/>
10. <https://www.nature.com/articles/s41467-024-47120-y>
11. https://www.mdpi.com/journal/molecules/special_issues/Multitarget_Ligand
12. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4379170/>
13. <https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2022.1003480/full>