

Using XlinkCyNET to visualize protein cross-links in Cytoscape software

Diogo Borges Lima (✉ diogobor@gmail.com)

Department of Chemical Biology, Leibniz - Forschungsinstitut für Molekulare Pharmakologie (FMP), 13125, Berlin, Germany <https://orcid.org/0000-0001-6056-0825>

Ying Zhu

Department of Chemical Biology, Leibniz - Forschungsinstitut für Molekulare Pharmakologie (FMP), 13125, Berlin, Germany

Fan Liu (✉ fliu@fmp-berlin.de)

Department of Chemical Biology, Leibniz - Forschungsinstitut für Molekulare Pharmakologie (FMP), 13125, Berlin, Germany

Method Article

Keywords: cytoscape, protein-protein interaction, cross-linking mass spectrometry

Posted Date: November 30th, 2020

DOI: <https://doi.org/10.21203/rs.3.pex-1172/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

We present a high-performance app for Cytoscape to visualize cross-linking mass-spectrometry (XL-MS) data. **XlinkCyNET** is an open-source Java plugin that generates residue-to-residue connections provided by XL-MS in protein interaction networks. Importantly, it provides an interactive interface for the exploration of cross-links and offers various options to display protein domains. XlinkCyNET works well in complex networks containing thousands of proteins.

Introduction

In systems biology, it is essential to understand what biomolecules are involved in different biological processes and how they coordinate to precisely execute an immensely broad range of cellular functions through interactions[1]. An indispensable approach to visualize and interpret these interactions is to use network diagrams. So far, several software tools have been developed for this purpose, such as Biogrid[2] and Cytoscape[3]. Among various options, Cytoscape is arguably one of the most popular network visualization tools. In particular, it offers a great selection of plugins capable of customizing layouts to facilitate the interpretation of the data (*e.g.*, Cerebral [4] and stringApp[5]). However, a plugin that allows the detailed visualization of residue-to-residue connections from cross-linking mass spectrometry (XL-MS) data and annotation of protein domains is still lacking. In light of the suitability and versatility provided by the Cytoscape platform, here we introduce **XlinkCyNET**, a Cytoscape app that supports the visualization of XL-MS data. We present a step-by-step protocol to illustrate the usage of XlinkCyNET and describe the functionality of important parameters.

Reagents

Equipment

Software

- Cytoscape 3.0 or later
- Java 8 or later
- XlinkCyNET is freely available from Cytoscape App Store: <http://apps.cytoscape.org/apps/xlinkcynet>, and <https://www.theliulab.com/software/XlinkCyNET>.

Procedure

1. App installation:

1.1. Install [Cytoscape](#) on your computer.

1.2. Download XlinkCyNET app

1.2.1. Launch Cytoscape. Go to Apps menu -> App Manager... -> search for XlinkCyNET -> click on *Install*.

1.2.2. Click on the *Download* button at <https://www.theliulab.com/software/XlinkCyNET>.

1.2.2.1. Launch Cytoscape. Go to Apps menu -> App Manager... -> click on *Install from file* and select the jar plugin downloaded in step 1.2.2.

2. Workflow

The following workflow demonstrates how to use **XlinkCyNET** in Cytoscape.

2.1. In Cytoscape, load a protein interaction network by clicking on File menu -> Import -> Network from file.

2.1.1. The input file must have the following columns, which should be assigned to the correct column type in Cytoscape. (**Figure 1**)

Required columns:

gene_a -> represents the node name: *source node (Local)*,

gene_b -> represents the node name: *target node (Local)*,

length_protein_a -> represents the number of residues of protein *a*: *Source Node Attribute (Shared)*,

length_protein_b -> represents the number of residues of protein *b*: *Target Node Attribute (Shared)*,

protein_a -> represents the Uniprot protein accession number of protein *a* in the following format:
db|Unique_Identifier|Entry_Name: *Source Node Attribute (Shared)*,

protein_b -> represents the Uniprot protein accession number of protein *b* in the following format:
db|Unique_Identifier|Entry_Name: *Target Node Attribute (Shared)*,

pro_site_ab -> represents all links (inter and intra) between protein *a* and *b*. The links must be separated by '#': *Source Node Attribute (Shared)*,

pro_site_ba -> represents all links (inter and intra) between protein *b* and *a*. The links must be separated by '#': *Target Node Attribute (Shared)*.

Figure 1: Example of input file of an interaction network in Cytoscape.

2.2. Once an interaction network is loaded, XlinkCyNET can be applied automatically to display the residue interaction layout. Two options are currently available: applying the layout to a single node or a set of nodes.

2.2.1. *Apply the layout to a single node.* After selecting a node, the layout can be applied by clicking on Apps menu -> XlinkCyNET -> Apply layout (or by pressing CTRL + E). The node will be expanded based on the length of the protein (*i.e.*, number of residues). All cross-links will be displayed* as well as annotated protein domains** (**Figure 2**). Note that at least one node needs to be selected, otherwise a warning message will be displayed. To restore the original layout, press CTRL + E again.

*If 'Display Intralink' and/or 'Display Interlink' is checked. See step 2.4.1.1 for more details.

**If domain annotation is loaded for the protein. See step 2.3 for more details.

Figure 2: Example of applying layout of protein-protein interactions to single / multiple nodes

2.2.1.1. Another way to apply cross-link layouts on single nodes is by right-clicking on the node -> Apps -> XlinkCyNET (or by pressing CTRL + U). A new window will be opened (**Figure 3**).

2.2.1.1. *Troubleshooting.* If the *Node table* does not contain *length_protein_a* or *length_protein_b* columns, no window will be opened.

Figure 3: Applying protein-protein interaction layouts to single nodes.

2.2.1.1.2. The protein node can be expanded horizontally or vertically, according to the checked *Expansion* option.

2.2.1.1.3. The length of the protein expansion is determined by the value of scaling factor, which ranges between 0.1 (small) and 1 (original protein length).

2.2.1.1.4. Protein domain annotations can be downloaded from Supfam [6][7] or Pfam [8] databases by clicking on the respective button. All domain annotations will be automatically incorporated in the table under column “*Domain*”.

2.2.1.1.5. Protein domains can also be added by inserting domain information directly in the table[§]. The following information can be included: *Domain*, *Start Residue*, *End Residue*, and *Color*. Columns marked as (*) are required. Colors for protein domains can be chosen manually, otherwise random colors will be displayed^{***}. New rows can be inserted by CTRL + I and existing rows can be deleted by CTRL + D.

^{***}Specific colors for protein domains can also be set globally from Apps menu -> XlinkCyNET -> Protein Domains -> Set domains color (or by pressing CTRL + T). For more details, see step 2.3.3.

[§]Supfam/Pfam computes the domain reliability for the protein, which is represented by *e-value*.

2.2.1.1.6. To apply the layout, click on *OK* button. An example is shown in **Figure 2**.

2.2.1.1.6.1. Once protein domain annotations have been loaded into a node, they will be inserted into the *Cytoscape Node Table* in the column *domain_annotation*.

2.2.1.1.7. The original node layout can be restored by clicking on *Restore style* button.

2.2.2. *Apply the layout to a set of nodes:* If more than one node has been selected, the layout can be automatically applied by pressing CTRL + E (**Figure 2**). The original layout can be restored to all nodes by pressing CTRL + E again.

2.2.3. The process can be aborted by clicking on *Cancel* button.

2.3. In XlinkCyNET, protein domains can be loaded, exported and the color of a specific domain can be selected under domain global settings, Apps -> XlinkCyNET -> Protein Domains.

2.3.1. Protein domains can be loaded by clicking on Apps -> XlinkCyNET -> Protein Domains -> Load (or by pressing CTRL + R). A new window will be opened (**Figure 4**).

Figure 4: Load domains for a set of proteins.

Protein domains can be added manually, by inserting the corresponding information in the table or importing from a text file.

2.3.1.1. *Node Name*, which represents exactly the name set up to the node (case-sensitive).

2.3.1.2. *Domain(s)*, which represents all domains that are displayed for the protein. The following pattern is required: *name_of_domain[start_range-end_range]*, *other_name_of_domain[start_range-end_range]*. Different domains are comma-separated.

2.3.1.3. To insert a new row in the table, press CTRL + I; to delete an existing row, press CTRL + D.

2.3.1.4. Protein domain annotations can be downloaded from Supfam [6][7] or Pfam [8] databases by clicking on the respective button. All domain annotations will be automatically incorporated in the table under column "*Domain*".

2.3.1.5. An input file with node names and domains can be imported through File menu -> Import file.

2.3.1.5.1. Two formats of the input file are currently supported: comma-separated (*CSV*) and tab-delimited (*tab*) formats. The *CSV* files must align with the format in the table (Node Name, Domains^{****}). XlinkCyNET can also recognize *tab* files that are directly downloaded from Uniprot. However, they must contain the following columns: *Gene names (in Names & Taxonomy category)*; and *Topological domain (in Subcellular location category)*.

2.3.1.5.2. Domain information can also be copied and pasted from an external table following the same format (Node Name, Domains^{****}).

^{****} Protein domains must follow the same pattern of columns described in 2.3.1.2.

2.3.1.6. To load protein domains, click on *OK* button. A message will be displayed confirming the process.

2.3.1.6.1. Once the protein domain annotations have been loaded, they will be inserted into the *Cytoscape Node Table* in the column *domain_annotation*.

2.3.1.7. The process can be aborted by clicking on *Cancel* button.

2.3.2. To export the protein domains, click on Apps menu -> XlinkCyNET -> Protein Domains -> Export (or by pressing CTRL + D). A message will be displayed showing the selected network and a new window will be opened to set the file name. All protein domains will be exported to a CSV file.

2.3.3. To customize the colors of protein domains, click on Apps menu -> XlinkCyNET -> Protein Domains -> Set domains color, or press CTRL + T. A new window will be opened, and all loaded domains will be displayed in the table (**Figure 5**).

2.3.3.1. Click on the corresponding cell and select a color from the palette (**Figure 6**).

Figure 5: Customize protein domain color

Figure 6: Color palette in Cytoscape.

2.4. XlinkCyNET includes a set of features for the display of legends and nodes of cross-links. The *Settings* panel can be opened by clicking on Apps menu -> XlinkCyNET -> Settings, or by clicking on *XlinkCyNET Settings* tab in the Control panel on the right side of the Cytoscape window (**Figure 7**).

Figure 7: XlinkCyNET Settings

2.4.1. The settings are grouped into two categories: *Link* and *Node*.

2.4.1.1. Intralinks and interlinks can be displayed or hidden separately based on the respective check options. Once checked, their colors can be modified by clicking on the color palette (**Figure 6**).

2.4.1.1.1. The opacity of the link can be modified within the range of 0 (transparent) to 255 (opaque). The width of the link can also be modified within the range of 1 (thin) and 10 (thick).

2.4.1.1.2. The legend can be displayed or hidden by checking the respective option. Once checked, it's the font size and opacity can be modified. Opacity ranges between 0 (transparent) to 255 (opaque). Nodes need to be moved once to update the new settings in the network.

2.4.1.2. Some features of the nodes can also be modified, including the font size of the node name as well as the color, opacity and width of the node border (**Figure 7**).

2.5. For any further questions, please send an email to diogobor@gmail.com or visit our website at Apps -> XlinkCyNET -> About.

Troubleshooting

2.2.1.a. Troubleshooting: If no network is loaded or selected, no action will be done.

2.2.1.b. Troubleshooting: If the purpose of the analysis is to investigate intermolecular interactions, we recommend unchecking the '*Display Intralink*' option before applying the layout, in order to speed up the application of the layout.

2.2.1.c. Troubleshooting: Once the goal is to analyze as many proteins as possible at the same time, we recommend enabling the '*Always Show Graphics Details*' option, as shown in **Figure 8**.

2.2.1.1. Troubleshooting: If the *Node table* does not contain *length_protein_a* or *length_protein_b* columns, no window will be opened.

2.2.1.1.4. Troubleshooting: If there is no internet connection or Supfam/Pfam website is offline, an alert will be shown in *Cytoscape Task History*.

2.2.1.1.6.2. Troubleshooting: Once the goal is to analyze as many proteins as possible at the same time, we recommend enabling the '*Always Show Graphics Details*' option, as shown in **Figure 7**.

2.2.2. Troubleshooting: Once the goal is to analyze as many proteins as possible at the same time, we recommend enabling the '*Always Show Graphics Details*' option, as shown in **Figure 7**.

2.3.1.4. Troubleshooting: If there is no internet connection or Supfam/Pfam website is offline, an alert will be shown in *Cytoscape Task History*.

2.3.1.5.2. Troubleshooting: If there is a typo in the *Domain* column, information will not be loaded, and a warning message will be shown in *Cytoscape Task History*.

2.3.1.6. Troubleshooting: If the domains are being searched in the Supfam/Pfam database and the process has not been completed, a warning message will be displayed asking for confirmation.

2.3.1.7. Troubleshooting: If the domains are being searched in the Supfam/Pfam database and the process has not been completed, a warning message will be displayed asking for confirmation.

2.4. Troubleshooting: If control panel is hidden, go to View menu -> Show Control Panel.

Time Taken

Anticipated Results

References

1. A. Ma'ayan, A. D. Rouillard, N. R. Clark, Z. Wang, Q. Duan, and Y. Kou, "Lean Big Data integration in systems biology and systems pharmacology," *Trends Pharmacol. Sci.*, Aug. 2014, doi: 10.1016/j.tips.2014.07.001.
2. R. Oughtred *et al.*, "The BioGRID interaction database: 2019 update," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D529–D541, Jan. 2019, doi: 10.1093/nar/gky1079.
3. P. Shannon *et al.*, "Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks," *Genome Res.*, vol. 13, no. 11, pp. 2498–2504, Nov. 2003, doi: 10.1101/gr.1239303.
4. A. Barsky, J. L. Gardy, R. E. W. Hancock, and T. Munzner, "Cerebral: a Cytoscape plugin for layout of and interaction with biological networks using subcellular localization annotation," *Bioinformatics*, vol. 23, no. 8, pp. 1040–1042, Apr. 2007, doi: 10.1093/bioinformatics/btm057.
5. N. T. Doncheva, J. H. Morris, J. Gorodkin, and L. J. Jensen, "Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data," *J. Proteome Res.*, vol. 18, no. 2, pp. 623–632, Feb. 2019, doi: 10.1021/acs.jproteome.8b00702.
6. J. Gough, K. Karplus, R. Hughey, and C. Chothia, "Assignment of homology to genome sequences using a library of hidden Markov models that represent all proteins of known structure¹¹Edited by G. Von Heijne," *J. Mol. Biol.*, vol. 313, no. 4, pp. 903–919, Nov. 2001, doi: 10.1006/jmbi.2001.5080.
7. A. P. Pandurangan, J. Stahlhacke, M. E. Oates, B. Smithers, and J. Gough, "The SUPERFAMILY 2.0 database: a significant proteome update and a new webserver," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D490–D494, Jan. 2019, doi: 10.1093/nar/gky1130.
8. M. Punta *et al.*, "The Pfam protein families database," *Nucleic Acids Res.*, vol. 40, no. Database issue, pp. D290-301, Jan. 2012, doi: 10.1093/nar/gkr1065.

Figures

Import Network From Table

Preview

Click on a column to edit it.

Select All Select None

gene_a	gene_b	length_protein_a	length_protein_b	protein_a	protein_b	pro_site_ab	pro_site_ba
MRPS15	NOA1	257	698	sp P82914 MRPS15	sp Q8NC60 NOA1	0#MRPS15-235-NOA1-89#M...	0#MRPS15-235-NOA1-89#M...
MRPS15	MRPS27	257	414	sp P82914 MRPS15	sp Q8NC60 NOA1	0#MRPS15-235-NOA1-89#M...	0#MRPS15-235-MRPS27
MDH2	MRPS15	338	257	sp P40926 MDH2	sp P82914 MRPS15	0#MDH2-157-MRPS15-235#...	0#MDH2-157-MRPS15-235#...
LRPPRC	PNPT1	1394	783	sp P42704 LRPPRC	sp Q8TCS8 PNPT1	0#LRPPRC-763-PNPT1-106#...	0#LRPPRC-763-PNPT1-106#...
NDUF88	PDE12	186	609	sp O95169 NDUF88	sp Q6L8Q7 PDE12	0#NDUF88-53-PDE12-84#ND...	0#NDUF88-53-PDE12-84#ND...
MTHFS	MTHFS	203	203	sp P49914 MTHFS	sp P49914 MTHFS	0#MTHFS-10-MTHFS-18#MT...	0#MTHFS-10-MTHFS-18#MT...
TOMM20	TOMM70	145	608	sp Q15388 TOMM20	sp O94826 TOMM70	0#TOMM20-61-TOMM70-16...	0#TOMM20-61-TOMM70-16...

Advanced Options...

Cancel OK

Figure 1

Example of input file of an interaction network in Cytoscape.

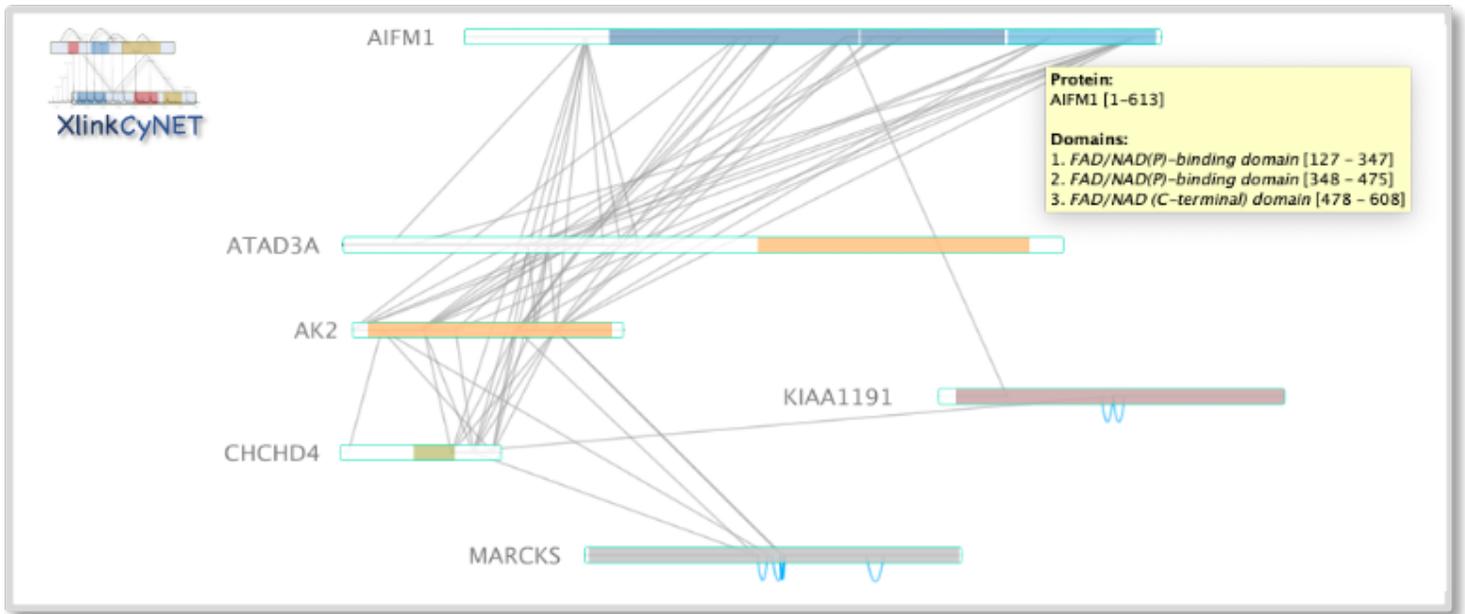


Figure 2

Example of applying layout of protein-protein interactions to single / multiple nodes.

XlinkCyNET - Single Node

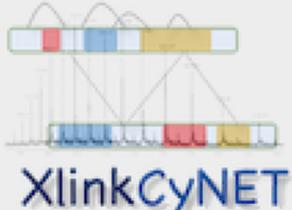
Protein

Name: **SLC25A13**

Size: **675 residues**

Expansion: Horizontal Vertical

Scaling factor: ⬆️ ⬇️ ⬆️



Search for domains in Pfam database: 

Status: **Done!**  Restore style

	Domain(*)	Start Residue(*)	End Residue(*)	e-value	Color
1	Mito_carr	325	423	4.5e-21	
2	Mito_carr	424	515	5e-12	
3	Mito_carr	516	611	2e-23	

(*) Required fields

 OK
 Cancel

Figure 3

Applying protein-protein interaction layouts to single nodes.

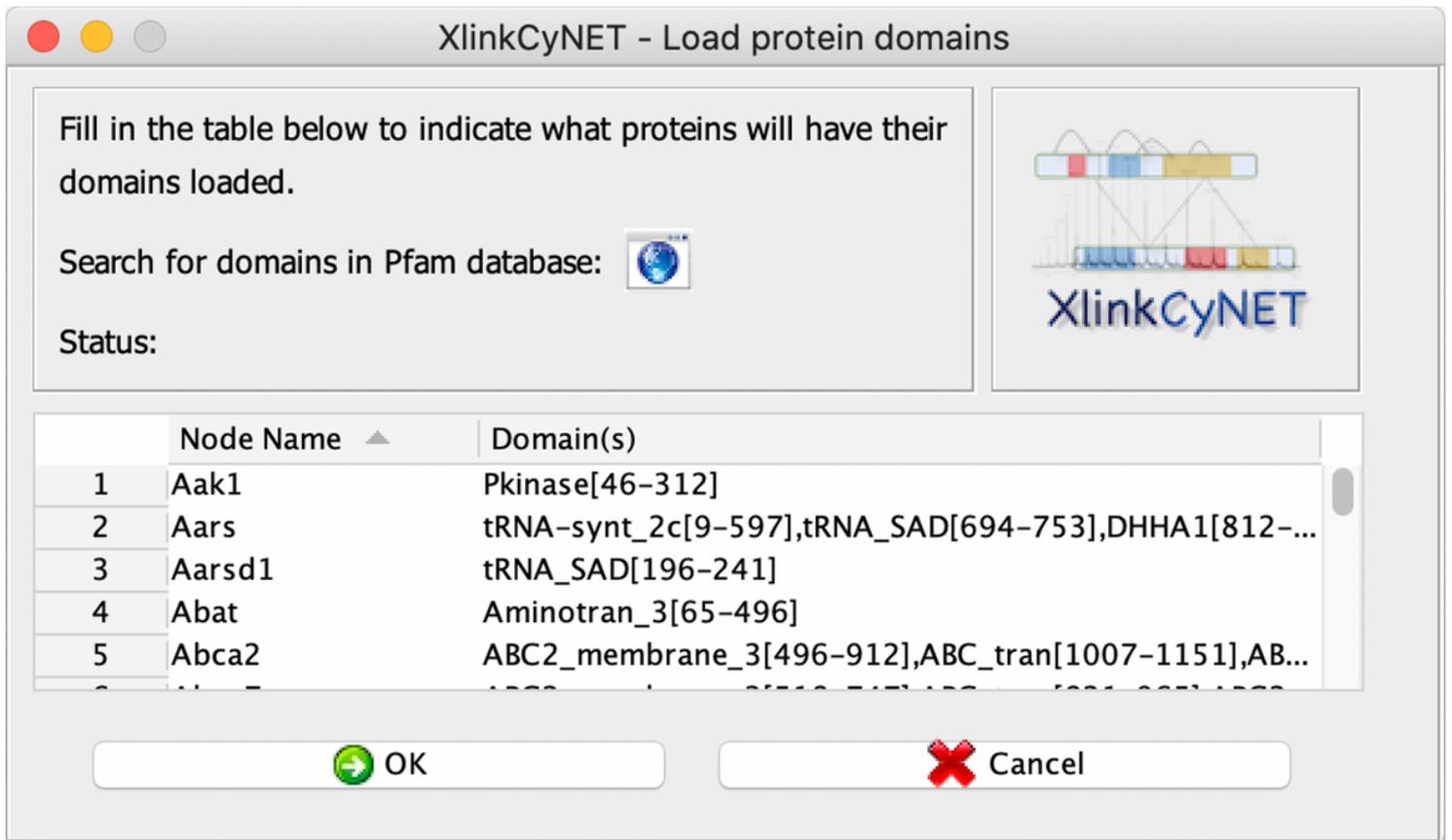


Figure 4

Load domains for a set of proteins.

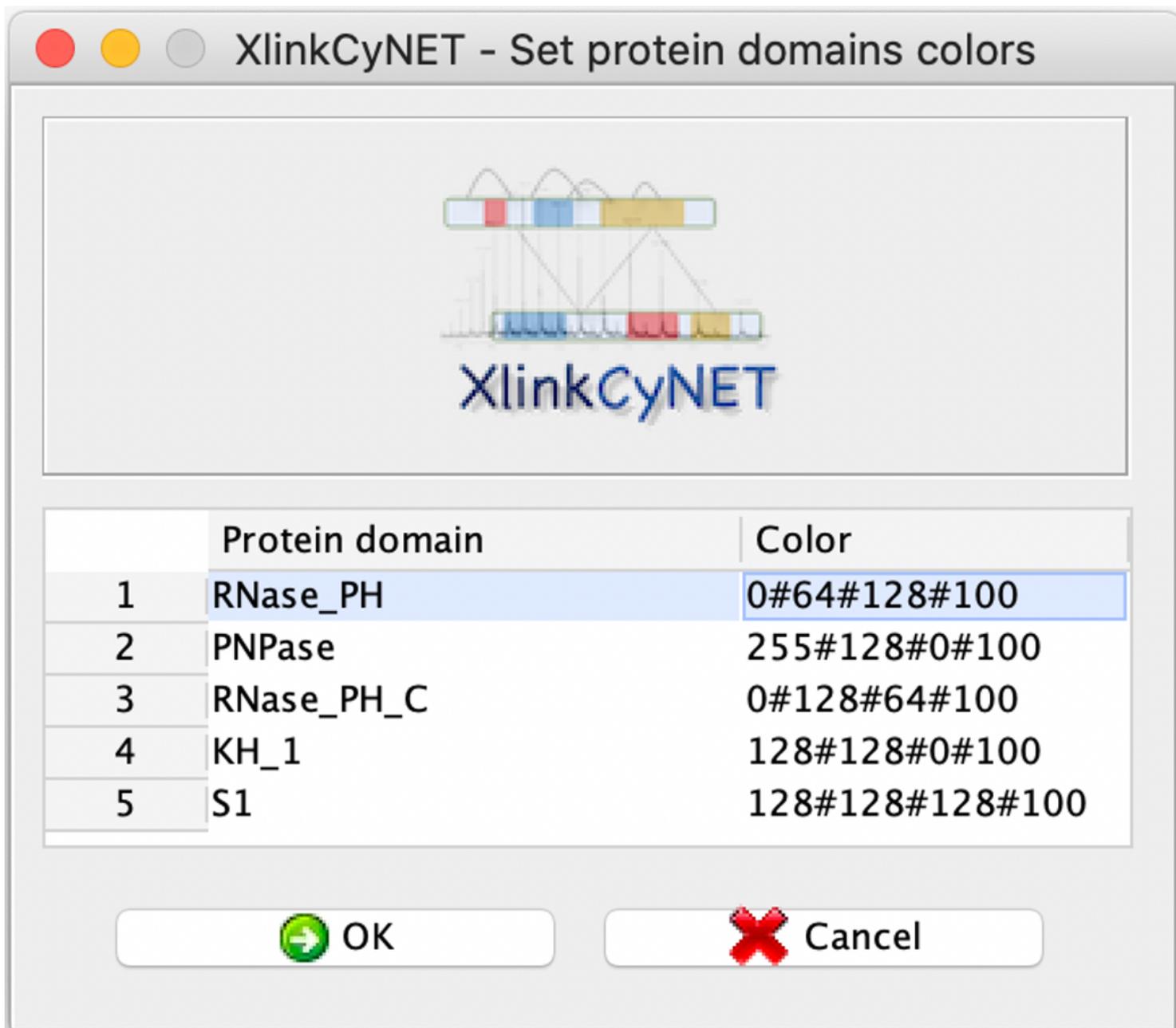


Figure 5

Customize protein domain color.

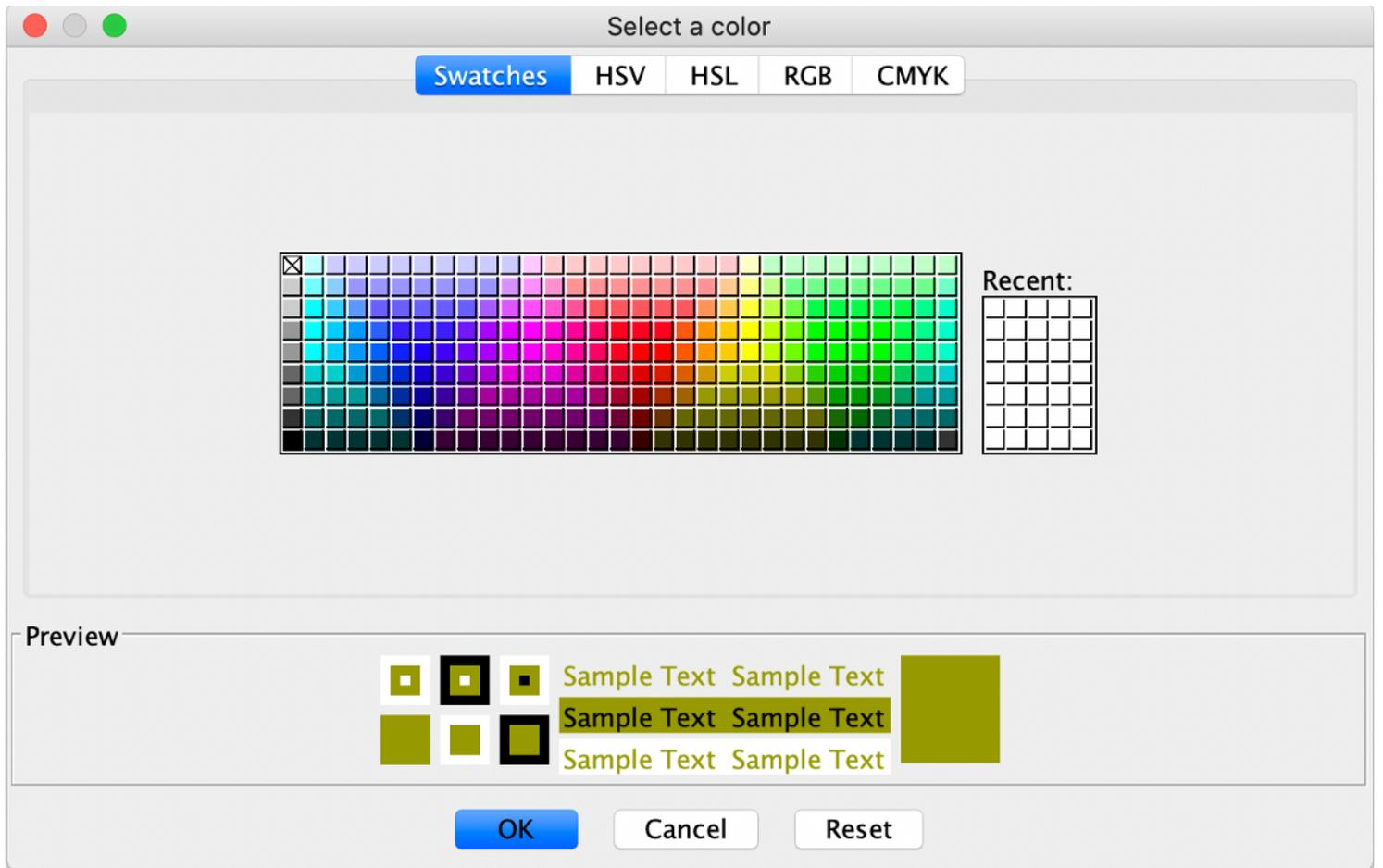


Figure 6

Color palette in Cytoscape.

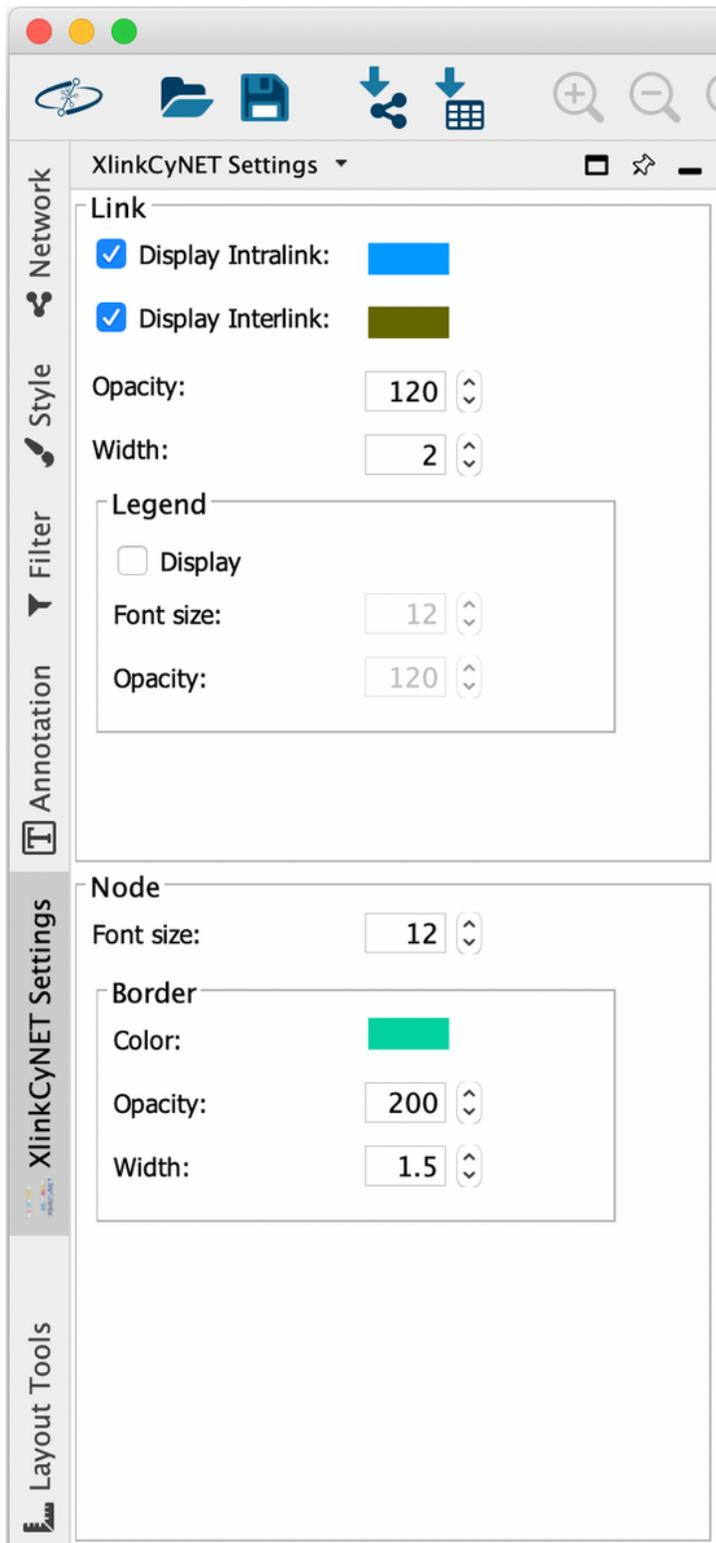


Figure 7

XlinkCyNET Settings.

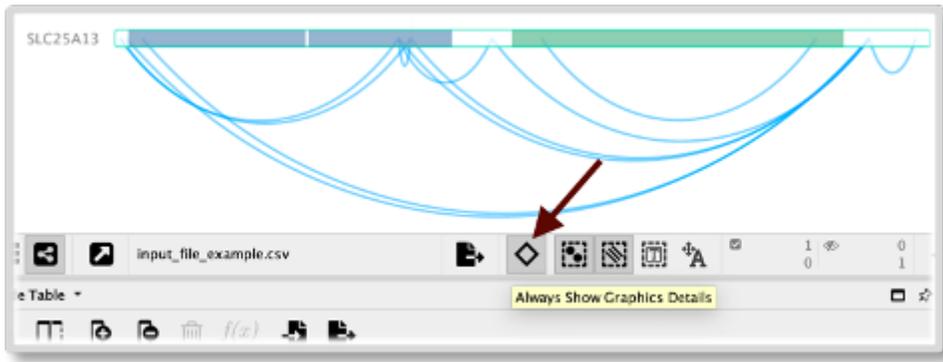


Figure 8

Enabling option to always render all graphics details in Cytoscape.