



**HAL**  
open science

## Protocol to identify human subcellular alternative protein interactions using cross-linking mass spectrometry.

Diego-Fernando Garcia del Rio, Isabelle Fournier, Tristan Cardon, Michel Salzet

### ► To cite this version:

Diego-Fernando Garcia del Rio, Isabelle Fournier, Tristan Cardon, Michel Salzet. Protocol to identify human subcellular alternative protein interactions using cross-linking mass spectrometry.. STAR Protocols, 2023, STAR Protocols, 4 (3), pp.102380. 10.1016/j.xpro.2023.102380 . hal-04400811

**HAL Id: hal-04400811**

**<https://hal.univ-lille.fr/hal-04400811>**

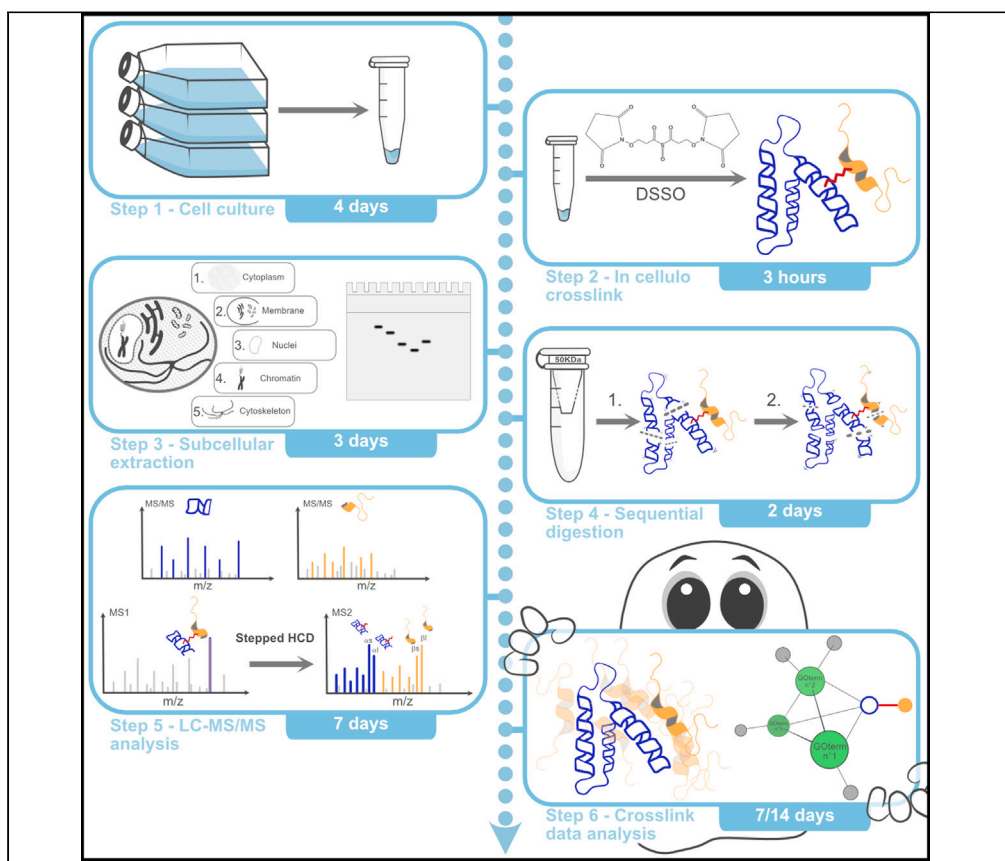
Submitted on 17 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Protocol

# Protocol to identify human subcellular alternative protein interactions using cross-linking mass spectrometry



Diego Fernando Garcia-del Rio, Isabelle Fournier, Tristan Cardon, Michel Salzet

isabelle.fournier@univ-lille.fr (I.F.)  
tristan.cardon@univ-lille.fr (T.C.)

### Highlights

Identify details related to possible function of unknown proteins

Steps described for using the cross-link to identify protein interactions

Determine subcellular localization information for unreferenced proteins

Highlight alternative proteins in cells by cross-link and subcellular fractionation

Since the start of mass-spectrometry-based proteomics, proteins from non-referenced open reading frames or alternative proteins (AltProts) have been overlooked. Here, we present a protocol to identify human subcellular AltProt and deciphering some interactions using cross-linking mass spectrometry. We describe steps for cell culture, *in cellulo* cross-link, subcellular extraction, and sequential digestion. We then detail both liquid chromatography-tandem mass spectrometry and cross-link data analyses. The implementation of a single workflow allows the non-targeted identification of signaling pathways involving AltProts.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Garcia-del Rio et al., STAR Protocols 4, 102380  
September 15, 2023 © 2023 The Author(s).  
<https://doi.org/10.1016/j.xpro.2023.102380>



## Protocol

## Protocol to identify human subcellular alternative protein interactions using cross-linking mass spectrometry

Diego Fernando Garcia-del Rio,<sup>1,2,3,5</sup> Isabelle Fournier,<sup>1,\*</sup> Tristan Cardon,<sup>1,4,5,\*</sup> and Michel Salzet<sup>1,4,6</sup><sup>1</sup>Université de Lille, Univ. Lille, CHU Lille, Inserm U1192 - Protéomique Réponse Inflammatoire Spectrométrie de Masse - PRISM, F-59000 Lille, France<sup>2</sup>VIB Center for Medical Biotechnology, VIB, Ghent 9052, Belgium<sup>3</sup>Department of Biomolecular Medicine, Ghent University, Ghent 9052, Belgium<sup>4</sup>These authors contributed equally<sup>5</sup>Technical contact: [diego.garciadelrio@univ-lille.fr](mailto:diego.garciadelrio@univ-lille.fr); [tristan.cardon@univ-lille.fr](mailto:tristan.cardon@univ-lille.fr)<sup>6</sup>Lead contact\*Correspondence: [isabelle.fournier@univ-lille.fr](mailto:isabelle.fournier@univ-lille.fr) (I.F.), [tristan.cardon@univ-lille.fr](mailto:tristan.cardon@univ-lille.fr) (T.C.)  
<https://doi.org/10.1016/j.xpro.2023.102380>

## SUMMARY

Since the start of mass-spectrometry-based proteomics, proteins from non-referenced open reading frames or alternative proteins (AltProts) have been overlooked. Here, we present a protocol to identify human subcellular AltProt and decipher some interactions using cross-linking mass spectrometry. We describe steps for cell culture, *in cellulo* cross-link, subcellular extraction, and sequential digestion. We then detail both liquid chromatography-tandem mass spectrometry and cross-link data analyses. The implementation of a single workflow allows the non-targeted identification of signaling pathways involving AltProts.

For complete details on the use and execution of this protocol, please refer to Garcia-del Rio et al.<sup>1</sup>

## BEFORE YOU BEGIN

The protocol described below outlines the detailed steps and resources required for a high throughput interactomic study of alternative proteins (AltProts). The study of this kind of protein has been disregarded because mRNA was considered monocistronic. AltProts, also known as ghost proteins<sup>2</sup> or short open reading frames (sORF)-encoded proteins (SEPs),<sup>3</sup> are translated from alternative open reading frames (AltORFs), such as 3' and 5' UTRs, reading frame shifts or long non-coding RNAs (LncRNAs). Despite their physiological presence in the cell, studying these proteins has been difficult due to the absence antibodies and databases that cover this type of proteins. The number of potential Human AltProts has been estimated to be around 450,000 sequences,<sup>4,5</sup> five times larger than the actual reference proteome available in Uniprot. However, these proteins represent a vast source of potential physiopathological biomarkers.<sup>2,6–10</sup> To overcome this challenge, we propose a methodology based on cross-linking mass spectrometry (XL-MS), subcellular fractionation, and bioinformatic tools, which enables the retrieval of functional information through network and gene ontology (GO) analysis.

In this protocol we propose the exploration of AltProt in non-pathological cells, however it can be adapted to any cell, adaptations in terms of quantity of cells will be expected, especially due to their size and their cellular content.



### Immortalized human ovarian epithelial cells (SV40) culture

⌚ Timing: 4 days

1. Seed immortalized human ovarian epithelial cells (SV40) into 25 cm<sup>2</sup> flasks.
  - a. Thaw a cryogenic vial with immortalized human ovarian epithelial cells (SV40) at 37°C.
  - b. Transfer the contents of the cryogenic vial to 9 mL of complete medium (Prigrow I medium with 10% of Hi-FBS and 100 U/mL of Penicillin-Streptomycin).
  - c. Centrifuge at 100 × g for 5 min at 20°C.
  - d. Remove supernatant by pipetting. Wash the cells gently pipetting up and down 5 mL of Dulbecco's phosphate-buffered saline (DPBS).
  - e. Pellet the cells by centrifuging at 100 × g for 5 min at 20°C and remove the supernatant.
  - f. Suspend cells with 5 mL of complete medium and thoroughly seed them in a 25 cm<sup>2</sup> flask.
  - g. Incubate cells at 37°C with 5% CO<sub>2</sub>.
  - h. Observe the cells in a microscope every day until the cells reach 80%–90% confluency.

**Note:** A cryogenic vial of immortalized human ovarian cells typically contains 1 million viable cells in 1 mL FBS /DMSO 10% and is made from cells that are approximately 80%–90% confluent. Pre-warm all the reagents in a water bath at 37°C, for 20 min.

2. Passage of immortalized human ovarian cells into 75 cm<sup>2</sup> flasks.
  - a. Once the cells reach ~80%–90% of confluency, remove the medium and wash the cell with 2.5 mL of DPBS.
  - b. Detach the cell from the flask using 0.5 mL of 0.05% Trypsin-EDTA (1×), phenol red.
  - c. Incubate for 5 min at 37°C, 5% CO<sub>2</sub>.
  - d. Add 1.5 mL of complete medium to inactivate Trypsin.
  - e. Transfer the cells into a conical centrifuge tube and spin at 100 × g for 5 min at 20°C.
  - f. Remove supernatant and wash the cells with 5 mL DPBS.
  - g. Pellet the cells by centrifuging at 100 × g for 5 min at 20°C and remove the supernatant.
  - h. Suspend cells with 10 mL of complete medium and thoroughly seed them in a 75 cm<sup>2</sup> flask.

**Note:** Repeat this step until the desired number of cells are reached with a confluency of ~80%–90%.

Pre-warm all the reagents in a water bath at 37°C, for 20 min.

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
<b>Antibodies</b>		
Goat Anti-Chicken IgG IgY (IgG) (H + L) (HRP) (1/5000)	Jackson Immuno Research	Cat# 103-035-155; RRID: AB_2337381
Monoclonal Mouse anti-Cytokeratin 18 (1/1000)	Dako	Cat# M7010; RRID: AB_2133299
Monoclonal Mouse anti-Histone H3 (1/1000)	Santa Cruz Biotechnology	Cat# sc-517576; RRID: AB_2848194
Monoclonal Mouse anti-Hsp70 (1/1000)	Abcam	Cat# ab2787; RRID: AB_303300
Monoclonal Mouse anti-SP1 (1/200)	Santa Cruz Biotechnology	Cat# sc-420; RRID: AB_628271
Peroxidase AffiniPure Goat Anti-Mouse IgG (H + L) (1/5000)	Jackson Immuno Research	Cat# 115-035-146; RRID: AB_2307392
Polyclonal Chicken anti-Calreticulin (1/200)	Abcam	Cat# ab2908; RRID: AB_303403
<b>Chemicals, peptides, and recombinant proteins</b>		
Acetonitrile	Carlo Erba Reagents	Cat# 412341
Acrylamide: Bis acrylamide 29:1 (40% solution / electrophoresis)	Euromedex	Cat# EU0063-B

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Amersham Protran Western blotting membranes, nitrocellulose	Merck	Cat# GE10600002
Ammonium persulfate (APS), BioUltra, for molecular biology, ≥ 98.0%	Sigma-Aldrich	Cat# 09913
Ammonium bicarbonate, BioUltra ≥ 99.5%	Sigma-Aldrich	Cat# 09830
Bovine serum albumin	Merck	Cat# A3059
Bromophenol blue sodium salt	Sigma-Aldrich	Cat# B5525
Chymotrypsin, sequencing grade	Promega	Cat# V1062
Dimethyl sulfoxide (DMSO)	Sigma-Aldrich	Cat# D5879
Disuccinimidyl sulfoxide (DSSO)	Thermo Fisher Scientific	Cat# A33545
DL-Dithiothreitol (DTT)	VWR Life Science	Cat# 97063-760
DPBS, no calcium, no magnesium	Thermo Fisher Scientific	Cat# 14190-094
Fetal bovine serum, qualified, heat inactivated, E.U.-approved, South America origin	Gibco	Cat# 10500064
Formic acid (for LC-MS)	TCI America	Cat#F0654
Glycerol, 99+%, extra pure	Thermo Fisher Scientific	Cat# 10562524
Glycine, 99+%, for analysis	Acros Organics	Cat# 10358210
Hydrochloric acid 37%, Iodoacetamide (IAA)	VWR Chemicals	Cat# 20252.420
2-Mercaptoethanol	Sigma-Aldrich	Cat# I1149
Methanol ≥ 98.5%, technical	Sigma-Aldrich	Cat# M6250
PageBlue™ Protein Staining Solution	VWR Chemicals	Cat# 20903.368
Penicillin-Streptomycin (10,000 U/mL)	Thermo Fisher Scientific	Cat# 24620
Prigrow I Medium	Gibco	Cat# 15140122
Sodium chloride	Applied Biological Materials	Cat# TM001
Sodium dodecyl sulfate (SDS), UltraPure™	Fisher Chemical	Cat# S/3161/60
SuperSignal™ West Dura Extended Duration Substrate	Invitrogen	Cat# 15525017
Tetramethylethylenediamine (TEMED)	Thermo Fisher Scientific	Cat#34075
TG-SDS 10x	Bio-Rad	Cat# 1610801
Trifluoroacetic acid	Euromedex	Cat# EU0510
TRIS biotech grade	Sigma-Aldrich	Cat# 302031
Trypsin-EDTA (0.05%), phenol red	Interchim	Cat# UP031657
Trypsin/Lys-C Mix, Mass Spec Grade	Gibco	Cat# 5300054
Tween 20	Promega	Cat# V5073
Urea Ultra-Pure	Sigma-Aldrich	Cat# P2287
Water, UHPLC-MS	Euromedex	Cat# EU0014B
	Thermo Fisher Scientific	Cat# 15339865
<b>Critical commercial assays</b>		
Detergent Removal Spin Columns HiPPR™	Thermo Fisher Scientific	Cat# 88306
Subcellular protein fractionation for cultured cells	Thermo Fisher Scientific	Cat# 78840
<b>Experimental models: Cell lines</b>		
Human Immortalized Ovarian Epithelial Cell line (SV40)	Applied Biological Materials	Cat# T1074
<b>Software and algorithms</b>		
Biological General Repository for Interaction Datasets (BioGRID)	Oughtred et al. <sup>34</sup>	RRID:SCR_007393; <a href="http://www.thebiogrid.org/">http://www.thebiogrid.org/</a>
ClueGO	Bindea et al. <sup>32</sup>	RRID:SCR_005748; <a href="https://apps.cytoscape.org/apps/cluego">https://apps.cytoscape.org/apps/cluego</a>
CluePedia	Bindea <sup>33</sup>	RRID:SCR_015784; <a href="https://apps.cytoscape.org/apps/cluepedia">https://apps.cytoscape.org/apps/cluepedia</a>
ClusPro 2.0	Kozakov et al. <sup>24</sup>	RRID:SCR_018248; <a href="https://cluspro.bu.edu/login.php">https://cluspro.bu.edu/login.php</a>
Cytoscape 3.9.1	Shannon et al. <sup>30</sup>	RRID:SCR_003032; <a href="https://cytoscape.org">https://cytoscape.org</a>
IntAct	Orchard et al. <sup>35</sup>	RRID:SCR_006944; <a href="http://www.ebi.ac.uk/intact">http://www.ebi.ac.uk/intact</a>

(Continued on next page)

**Continued**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
I-TASSER (Iterative Threading ASSEmblY Refinement)	Yang et al. <sup>21</sup>	RRID:SCR_014627; <a href="https://zhanggroup.org/I-TASSER/">https://zhanggroup.org/I-TASSER/</a>
NetMHC - 4.0	Andreatta and Nielsen 2016	RRID:SCR_021651; <a href="https://services.healthtech.dtu.dk/service.php?NetMHC-4.0">https://services.healthtech.dtu.dk/service.php?NetMHC-4.0</a>
OpenProt Protein Database 1.6	Brunet et al. <sup>5</sup>	<a href="https://www.openprot.org/p/ng/Home">https://www.openprot.org/p/ng/Home</a>
OriginPro, Version 2022b	OriginLab Corporation	RRID:SCR_014212; <a href="https://www.originlab.com/">https://www.originlab.com/</a>
Proteome Discoverer 2.5	Thermo Fisher Scientific	RRID:SCR_014477; <a href="https://www.thermofisher.com/order/catalog/product/OPTON-31040">https://www.thermofisher.com/order/catalog/product/OPTON-31040</a>
STRING app	Doncheva et al. <sup>31</sup>	<a href="http://apps.cytoscape.org/apps/stringapp">http://apps.cytoscape.org/apps/stringapp</a>
UniProtKB	The UniProt Consortium	RRID:SCR_004426 <a href="https://www.uniprot.org/uniprotkb?facets=model_organism%3A9606&amp;query=%2A">https://www.uniprot.org/uniprotkb?facets=model_organism%3A9606&amp;query=%2A</a>
XlinkX 2.5 nodes for Proteome Discoverer 2.5	Thermo Fisher Scientific	<a href="https://www.thermofisher.com/order/catalog/product/OPTON-31047">https://www.thermofisher.com/order/catalog/product/OPTON-31047</a>
YASARA view	YASARA Biosciences	RRID:SCR_017591; <a href="http://www.yasara.org/">http://www.yasara.org/</a>
yFiles Layout Algorithms	yWorks	<a href="https://apps.cytoscape.org/apps/yfileslayoutalgorithms">https://apps.cytoscape.org/apps/yfileslayoutalgorithms</a>

**Other**

Amicon Ultra-0.5 Centrifugal Filter Unit 50 kDa	Merck	Cat# UFC505024
BB15 CO <sub>2</sub> Incubator	Thermo Fisher Scientific	Cat# 51023121
ECLIPSE Ts2 inverted microscope	Nikon	Cat# Ts2-FL
iBright CL750 Imaging System	Thermo Fisher Scientific	Cat# A44116
LD79 Digital Test-Tube Rotator	Labinco	Cat# 79000
Mini orbital shaker	ClearLine	Cat# 060956CL
Mini-PROTEAN® Tetra Handcast System	Bio-Rad	Cat# 1658003FC
Mini tube rotator	Thermo Fisher Scientific	Cat# 15534080
Oven 100-800	Memmert	Cat# 200718
Refrigerated centrifuge 5804 R	Eppendorf	Cat# 5805000010
PowerPac 1000	Bio-Rad	Cat# 4006038
Trans-Blot Cell	Bio-Rad	Cat# 20179
TW8 Water Bath	Julabo	Cat# 9550108
Vacuuge Concentrator System 5301	Eppendorf	Cat# 000210
ZipTip with 0.6 µL C18 resin	Merck	Cat# ZTC18S096

**MATERIALS AND EQUIPMENT**

**Complete cell culture media**

Reagent	Final concentration	Amount
Hi-FBS (Heat-Inactivated Fetal Bovine Serum)	10%	5 mL
Penicillin/streptomycin	100 U/mL	500 µL
Prigrow I medium	N/A	Up to 50 mL
<b>Total</b>	<b>N/A</b>	<b>50 mL</b>

Store at 4°C up to 1 month.

**Tris/HCl 1.5 M**

Reagent	Final concentration	Amount
Tris base	1.5 M	908 mg
HCl	Up to pH 8.1	N/A
ddH <sub>2</sub> O	N/A	Up to 5 mL
<b>Total</b>	<b>N/A</b>	<b>5 mL</b>

Store at 20°C up to 6 months.

### 2× Laemmli buffer

Reagent	Final concentration	Amount
Tris base	125 mM	747 mg
SDS	4%	2 g
Glycerol	20%	10 mL
2-mercapto-ethanol	10%	5 mL
Bromophenol blue	N/A	100 mg
HCl	N/A	Up to pH   6.8
ddH <sub>2</sub> O	N/A	Up to 50 mL
<b>Total</b>	<b>N/A</b>	<b>50 mL</b>

Store 500 μL aliquots at –20°C up to 1 year.

2-mercapto-ethanol is seriously irritating and toxic if swallowed or inhaled, so it is advised to handle it in an active fume hood.

### 4% concentration SDS-PAGE gel

Reagent	Final concentration	Amount
Acrylamide: Bis Acrylamide 29:1 (40% Solution / Electrophoresis)	4%	1 mL
Tris-HCl (0.5 M) pH 6.8	125 mM	2.5 mL
SDS (10%)	0.1%	100 μL
APS (10%)	0.05%	50 μL
TEMED	0.1%	10 μL
ddH <sub>2</sub> O	N/A	Up to 10 mL
<b>Total</b>	<b>N/A</b>	<b>10 mL</b>

Once the gel is solid, it can be stored for one week at 4°C keeping it humid.

Invitrogen™ Novex™ 4%–12% Tris-Glycine Plus, 1.0 mm, Midi Protein Gels (Cat# WXP41220BOX) can be also use.

### 12% migration SDS-PAGE gel

Reagent	Final concentration	Amount
Acrylamide: Bis Acrylamide 29:1 (40% Solution / Electrophoresis)	12%	3 mL
Tris-HCl (1.5 M) pH 8.8	375 mM	2.5 mL
SDS (10%)	0.1%	100 μL
APS (10%)	0.062%	62.2 μL
TEMED	0.062%	6.2 μL
ddH <sub>2</sub> O	N/A	Up to 10 mL
<b>Total</b>	<b>N/A</b>	<b>10 mL</b>

Once the gel is solid, it can be stored for one week at 4°C keeping it humid.

### Towbin transfer buffer

Reagent	Final concentration	Amount
Tris base	25 mM	3.03 g
Glycine	192 mM	14.4 g
Methanol	20%	200 mL
ddH <sub>2</sub> O	N/A	Up to 1 L
<b>Total</b>	<b>N/A</b>	<b>1 L</b>

Store at 4°C up to 3 months.

### 10× TBS-T buffer

Reagent	Final concentration	Amount
Tris base	200 mM	24.23 g
NaCl	1.5 M	87.66 g
Tween 20	1%	10 mL
ddH <sub>2</sub> O	N/A	Up to 1 L
<b>Total</b>	<b>N/A</b>	<b>1 L</b>

Store at 20°C up to 6 months.

Denaturing buffer		
Reagent	Final concentration	Amount
Urea	8 M	24 g
Tris base	0.1 M	600 mg
HCl	N/A	Up to pH 8.5
ddH <sub>2</sub> O	N/A	Up to 50 mL
<b>Total</b>	<b>N/A</b>	<b>50 mL</b>

Store at 20°C up to 6 months.

- 50 mM DSSO: 1 mg of DSSO (pre-weighed tube) in 51.5  $\mu$ L of DMSO.
- 10  $\mu$ M BSA for cross-linking control: 6.6 mg of BSA in 10 mL of DPBS. Store at 4°C for up to 6 months.
- 1 $\times$  Tris-Glycine-SDS running buffer: Dilute 100 mL of 10 $\times$  TG-SDS in 900 mL of ddH<sub>2</sub>O.
- 1 $\times$  TBS-T buffer: Dilute 100 mL of 10 $\times$  TBS-T buffer in 900 mL of ddH<sub>2</sub>O.
- 5% milk blocking buffer: Dissolve 5 g of powder milk in 100 mL of 1 $\times$  TBS-T.
- 100 mM reduction buffer: Dissolve 15.4 mg of dithiothreitol (DTT) in 1 mL of denaturing buffer.
- 50 mM alkylation buffer: Dissolve 9.3 mg of iodoacetamide (IAA) in 1 mL of denaturing buffer.
- 50 mM ammonium bicarbonate buffer: Dissolve 197.6 mg of ammonium bicarbonate in 50 mL of ddH<sub>2</sub>O.
- 1% trifluoroacetic acid (TFA): Dilute 100  $\mu$ L of TFA in 10 mL of UHPLC grade water.
- 0.1% TFA: Dilute 1 mL of 1% TFA in 9 mL of UHPLC-MS grade water.
- Mobile phase A: 0.1% formic acid in UHPLC-MS grade water.
- Mobile phase B: 0.1% formic acid in HPLC grade acetonitrile.

## STEP-BY-STEP METHOD DETAILS

### *In cellulo* cross-linking

⌚ Timing: 3 h

This step consists of detachment, collection of the cells and *in-cellulo* cross-linking.

#### 1. Cell harvesting

- Once the cells reach ~80%–90% of confluency, remove the medium and wash the cells with 5 mL of DPBS.
- Detach the cell from the flask using 1 mL of 0.05% Trypsin-EDTA (1 $\times$ ), phenol red.
- Incubate for 5 min at 37°C, 5% CO<sub>2</sub>.
- Add 2 mL of complete cell culture media to inactivate Trypsin-EDTA (0.05%), phenol red.
- Transfer the cells into a centrifuge tube and centrifugate at 100  $\times$  g for 5 min at 20°C.
- Remove supernatant and wash the cells with 5 mL of DPBS.
- Pellet the cells by centrifuging at 100  $\times$  g for 5 min at 20°C and remove the supernatant by aspiration.
- Repeat the DPBS wash two more times.
- Count the cells and aliquot to 3 million cells.
- Pellet the cells by centrifuging at 100  $\times$  g for 5 min at 20°C and aspirate the supernatant.
- Keep the dry pellet on ice.

**Note:** In our experience the optimal number of cells is 3 million, this number has to be adapted and tested for other kinds of cells.

#### 2. Cross-linking reaction.

- Resuspend 3 million cells in 196  $\mu$ L of DPBS.
- Prepare two 10  $\mu$ M BSA solution as positive and negative control.



- c. Prepare a 50 mM stock solution of disuccinimidyl sulfoxide (DSSO) in DMSO.
- d. Add 4  $\mu\text{L}$  of 50 mM DSSO to the suspended cells and positive BSA control (Final DSSO concentration of 2 mM).
- e. Incubate at 37°C with continuous shaking (10 RPM) for 1 h.
- f. To quench the reaction, add 10  $\mu\text{L}$  of Tris-HCl 1.5 M.
- g. Incubate for 30 min at 20°C with continuous shaking (10 RPM).
- h. Centrifuge at 2,000  $\times g$  for 10 min at 4°C. Remove the supernatant.
- i. Wash the cells with 200  $\mu\text{L}$  of ice-cold DPBS.
- j. Pellet the cells by centrifuging at 100  $\times g$  for 5 min at 4°C and remove as much supernatant as possible.

**Note:** It is recommended to use a BSA cross-linking positive control (96  $\mu\text{L}$  of 10  $\mu\text{M}$  BSA). Negative control is performed using 4  $\mu\text{L}$  of DMSO instead of DSSO. The optimal working final concentration of DSSO is between 1–5 mM.

**△ CRITICAL:** Cross-linkers are moisture sensitive. Prepare these cross-linkers immediately before use. Use amine-free buffers (PBS, 20 mM HEPES, 100 mM carbonate/bicarbonate, or 50 mM borate). Cross-linking reactions (acylation) are favored near neutral pH (pH 6–9) and with concentrated protein solutions.

### Subcellular protein fractionation of cross-linked and non-cross-linked cells.

⌚ **Timing:** 3 days

The following methodology describes the steps of subcellular protein fractionation after the cross-linking reaction. This methodology is based on the instructions from the Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Scientific, Cat# 78840). For more details and troubleshooting, please refer to the manual on [Thermo website](#).

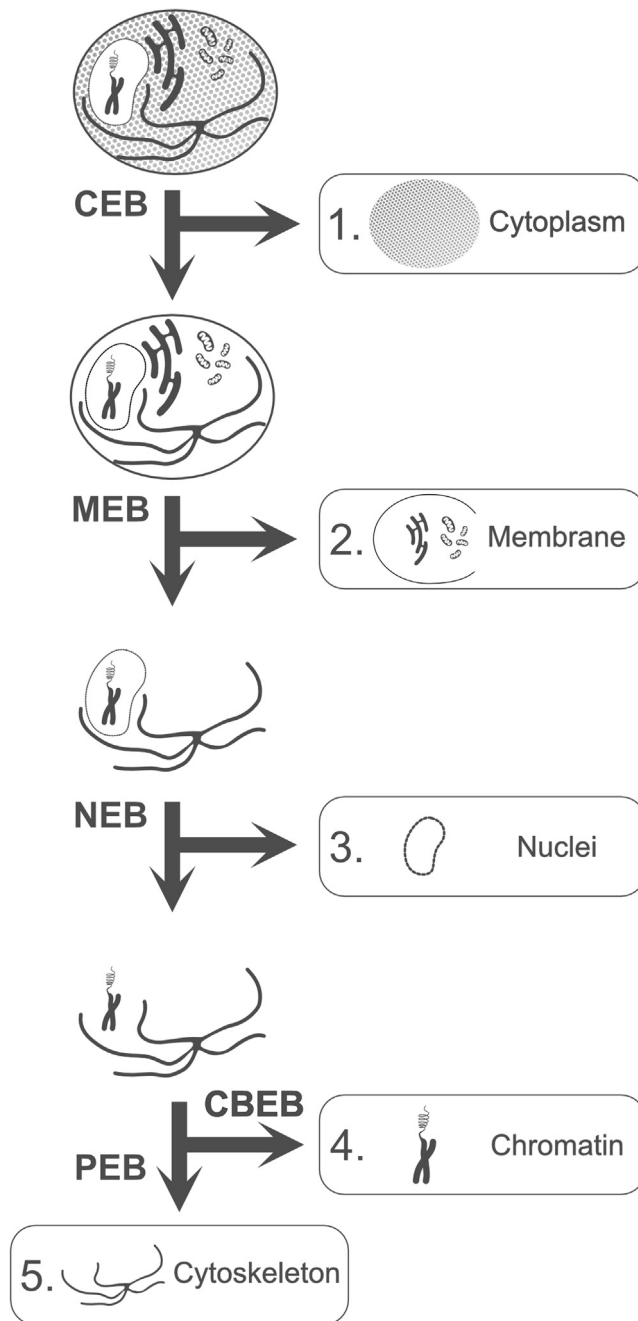
DMSO control of cells are treated according to the same protocol. These controls are key to determining the experimental subcellular location of the AltProts.

Other subcellular fractionation kits in the market are Abcam's Cell Fractionation Kit - Standard (Cat# ab109719) and Cell Signaling's Cell Fractionation Kit (Cat# 9038). These kits are designed to fractionate the cells in three subcellular fractions which will not decrease the complexity of the samples as much.

**Note:** Thaw all buffers using a 20°C water bath. Keep CEB, MEB, and NEB buffers on ice until use. Use a rotary shaker to avoid clumping of insoluble material during incubations.

**△ CRITICAL:** Immediately before use, add Thermo Scientific Halt Protease Inhibitor Cocktail at a 1:100 dilution into each volume of buffer required. Keep all protein extracts on ice.

3. Subcellular Protein Fractionation of 3 million cells ([Figure 1](#)).
  - a. Lyse the cells by adding 300  $\mu\text{L}$  of Cytoplasmic Extraction Buffer (CEB). Incubate at 4°C for 10 min with gentle shaking (10 RPM).
  - b. Centrifuge at 2,000  $\times g$  for 5 min. Aspirate by pipetting and immediately transfer the supernatant (cytoplasmic extract) to a clean, pre-chilled (4°C in ice) 1.5 mL microcentrifuge tube.
  - c. Resuspend the pellet in 300  $\mu\text{L}$  of ice-cold Membrane Extraction Buffer (MEB). Vortex for 5 s and incubate at 4°C for 10 min with gentle shaking (10 RPM).
  - d. Centrifuge at 5,000  $\times g$  for 5 min. Aspirate and immediately transfer the supernatant (membrane extract) to a clean, pre-chilled 1.5 mL microcentrifuge tube.



**Figure 1. Subcellular protein fractionation workflow**

Pelleted cells, both cross-linked and non-cross-linked, are resuspended in CEB buffer. After incubation and centrifugation, the resulting cytoplasmic extract is removed and stored. The remaining pellet is then retaken in MEB buffer, incubated, and centrifuged to obtain the membrane extract, which is also removed and stored. Next, the pellet is resuspended in NEB buffer, incubated, and centrifuged to obtain the nuclear extract, which is similarly removed and stored. The pellet from the previous step is then retaken in CBEB buffer, incubated, and centrifuged to obtain the chromatin-bound extract, which is also removed and stored. Finally, the remaining pellet is taken back in PEB buffer, incubated, and centrifuged to obtain the cytoskeletal extract, which is also removed and stored.

- e. Add 150  $\mu\text{L}$  of ice-cold Nuclear Extraction Buffer (NEB). Roughly vortex (highest vortex setting) for 15 s and incubate at 4°C for 30 min with gentle shaking (10 RPM).

**Note:** During the 30-min incubation time, prepare the Chromatin-Bound Extraction Buffer (CBEB) by adding 15  $\mu\text{L}$  of 100 mM  $\text{CaCl}_2$  and 9  $\mu\text{L}$  of Micrococcal Nuclease (300 units) in 150  $\mu\text{L}$  of 20°C NEB.

- f. Centrifuge at 7,000  $\times$  g for 5 min. Aspirate and immediately transfer the supernatant (nuclear extract) to a clean, pre-chilled 1.5 mL microcentrifuge tube.
- g. Resuspend the pellet in 150  $\mu\text{L}$  of 20°C CBEB. Roughly vortex for 15 s and incubate at 20°C for 15 min with gentle shaking (10 RPM).
- h. After incubation, roughly vortex 15 s and centrifuge at 16,000  $\times$  g for 5 min. Aspirate and immediately transfer the supernatant (chromatin-bound extract) to a clean, pre-chilled 1.5 mL microcentrifuge tube.
- i. Add 150  $\mu\text{L}$  of Pellet Extraction Buffer (PEB) to the remaining pellet. Roughly vortex for 15 s and incubate at 20°C for 10 min with gentle shaking (10 RPM).
- j. After incubation, roughly vortex for 15 s and centrifuge at 16,000  $\times$  g for 5 min. Aspirate and immediately transfer the supernatant (cytoskeletal extract) to a clean pre-chilled 1.5 mL microcentrifuge tube.
- k. Aliquot 10  $\mu\text{L}$  of each extract for SDS-PAGE and western blotting.

**Note:** Performing protein quantification employing Thermo Scientific Pierce BCA Protein Assay (Cat# 23225) is recommended to calculate the correct protein: protease ratio, for the sequential enzymatic digestion.

Keep the extracts in ice or for long-term storage keep them at  $-80^\circ\text{C}$ .

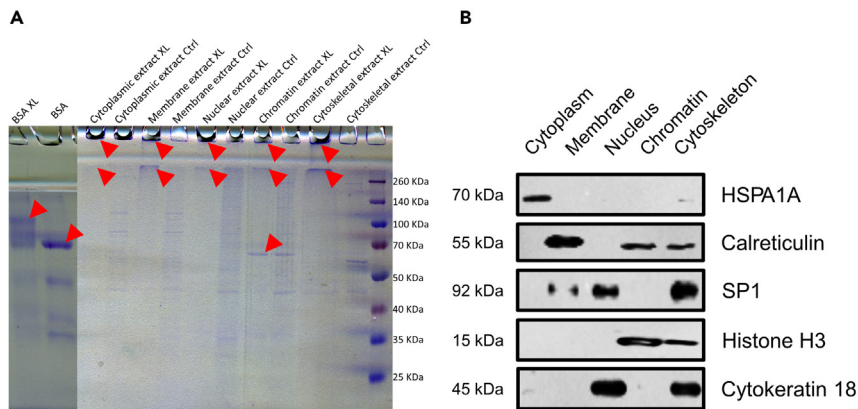
4. Cross-linking and subcellular fractionation confirmation.
  - a. Mix 10  $\mu\text{L}$  of 2 $\times$  Laemmli buffer with the 10  $\mu\text{L}$  protein aliquot.
  - b. Load each sample of subcellular protein fraction and BSA on to a 4%–12% SDS-PAGE gel.
  - c. Migrate the gels for 15 min at 70 V and for 90 min at 120 V in Tris-glycine-SDS buffer.
  - d. After migration, stain the gels with PageBlue™ Protein Staining Solution (Coomassie blue) for 1 h.
  - e. Destain the gels by discarding the excess staining solution. Rinse the gels two times with water.
  - f. Wash the gel for 16 h in an orbital shaker at 60 RPM. Placing a folded Kimwipes Tissue in the container to absorb excess dye will accelerate the destaining process.
  - g. Visualize the destained gels using your preferred system (Figure 2A).

**Note:** Only non-cross-linked samples will continue to the western blot analysis.

**△ CRITICAL:** Use one membrane with the five subcellular protein fractions for one compartment specific primary antibody.

- h. Transfer the gels onto a 0.45  $\mu\text{m}$  nitrocellulose membrane. Employ a tank transfer system for 2 h at 290 mA in Towbin buffer.
- i. Wash the membranes for 5 min in an orbital shaker three times with 20 mL of 1 $\times$  TBS-T buffer.
- j. Block the membranes for 1 h in an orbital shaker with 20 mL of milk blocking solution.
- k. Meanwhile, prepare the primary antibody dilution in milk blocking solution. The concentration of antibody has been adjusted as following: Cytokeratin 18 (1/1000), SP1 (1/200), Histone H3 (1/1000), Hsp70 (1/1000) and Calreticulin (1/200).

**△ CRITICAL:** Depending on the antibody's supplier, the dilution of it must be adjusted.



**Figure 2. Cross-linking reaction and subcellular fractionation confirmation**

(A) The Coomassie blue stained SDS-PAGE displays each cross-linked subcellular fraction, compared to a non-cross-linked fraction. BSA cross-linked and not cross-linked are used as controls. Red arrows display cross-linked signals demonstrating that the reaction takes place.

(B) The subcellular fractionation was confirmed by Western blot with compartment specific markers. The cytoplasm fraction showed the presence of HSPA1A signal. Calreticulin signal was detected at chromatin, cytoskeleton, and exhibited a stronger signal at the membrane-bounded fraction. SP1 was observed in the nucleus and cytoskeleton, while Histone H3 was found in chromatin and cytoskeleton. Similarly, Cytokeratin 18 was detected in the nucleus and cytoskeleton. These findings are consistent with the results reported in UniProtKB, COMPARTMENTS, and the literature.

- l. Incubate each membrane for each antibody for 16 h at 4°C in an orbital shaker.
- m. After incubation, wash the membrane for 5 min in an orbital shaker three times with 20 mL of 1 × TBS-T buffer.
- n. Incubate for 1 h with the matched HRP anti-Chicken (1/5000) or anti-Mouse (1/5000) secondary antibody.
- o. After incubation, wash the membrane for 5 min in an orbital shaker three times with 1 × TBS-T buffer.
- p. Perform the horseradish-peroxidase reaction, by preparing the SuperSignal™ West Dura Extended Duration Substrate (1 mL of Luminol/Enhancer Solution and 1 mL of Stable Peroxide Solution).
- q. Incubate the membrane with the substrate working solution for 5 min in the dark.
- r. Remove the membrane from the substrate working solution and place it in a plastic sheet protector.
- s. Remove the excess liquid with an absorbent tissue pressing out bubbles.
- t. Scan the membranes using the Invitrogen iBright Imaging System or other compatible imaging system (Figure 2B).
  - i. Mode: Chemi Blots.
  - ii. Exposure mode: Normal.
  - iii. After the autoexposure, the exposure time adjusted for each membrane: Cytokeratin 18 (5085 ms), SP1 (30 s), Histone H3 (4106 ms), Hsp70 (8213 ms) and Calreticulin (10 s).
  - iv. Resolution: 4 × 4.
  - v. Optical zoom: 1 ×.
  - vi. Digital zoom: 2 ×.
  - vii. Focus Level: 220.
  - viii. Sensitivity: Frame 1:100.

### Sequential enzymatic digestion

⌚ Timing: 2 days

The subsequent steps described the filter aided sample preparation (FASP)<sup>11</sup> sequential enzymatic digestion using LysC/Trypsin followed by Chymotrypsin for cross-linked samples. A 50 kDa cut-off Amicon filter is suggested to eliminate as many as possible non-cross-linked proteins.

### 5. FASP and sequential digestion

- a. Concentrate the five subcellular fractions in the 50 kDa Amicon filter by centrifugation at 4°C for 15 min at 14,000 × g. As a result, a 20 μL protein concentrate will remain as dead volume in the filter.
- b. Add 80 μL of denaturing buffer to the filter and pipette up and down gently inside the filter.
- c. Add 100 μL of reduction buffer.
- d. Incubate at 56°C for 40 min.

**Note:** Don't incubate at 95°C. At above 60°C urea can produce protein carbamylation. Additionally, the filter could melt.

- e. Centrifuge 15 min at 14,000 × g.
- f. Add 200 μL of denaturing buffer and centrifugate 15 min at 14,000 × g.
- g. Repeat step 5e at least two times.
- h. Add 100 μL of alkylation buffer.
- i. Incubate for 20 min at 20°C in the dark.
- j. Centrifugate 15 min at 14,000 × g.
- k. Add 200 μL of ammonium bicarbonate buffer and centrifugate 15 min at 14,000 × g.
- l. Repeat the previous step at least two times.
- m. Add Trypsin/Lys-C Mix Mass Spec Grade to the vendor recommended 25:1 protein: protease ratio (w/w). Incubate for 16 h at 37°C.
- n. After incubation, add Chymotrypsin, Sequencing Grade at a 100:1 protein: protease ratio (w/w). Incubate for 4 h at 20°C.
- o. Place the Amicon filter into a new clean tube.
- p. Add 50 μL of ammonium bicarbonate buffer and centrifugate 15 min at 14,000 × g.
- q. Repeat the previous step.
- r. Discard the Amicon filter.
- s. Acidify the filtered peptides with TFA 1% until pH < 7.
- t. Vacuum dry the samples in a SpeedVac concentrator and store them at –20°C if needed.
- u. The membrane fraction contains a large amount of polymer. It requires the use of a HiPPR™ Detergent Removal Resin column (Thermo Scientific, Cat# 88305) following the [vendor's protocol](#) to be compatible for MS analysis.

**Note:** –80°C is recommended for long term sample storage.

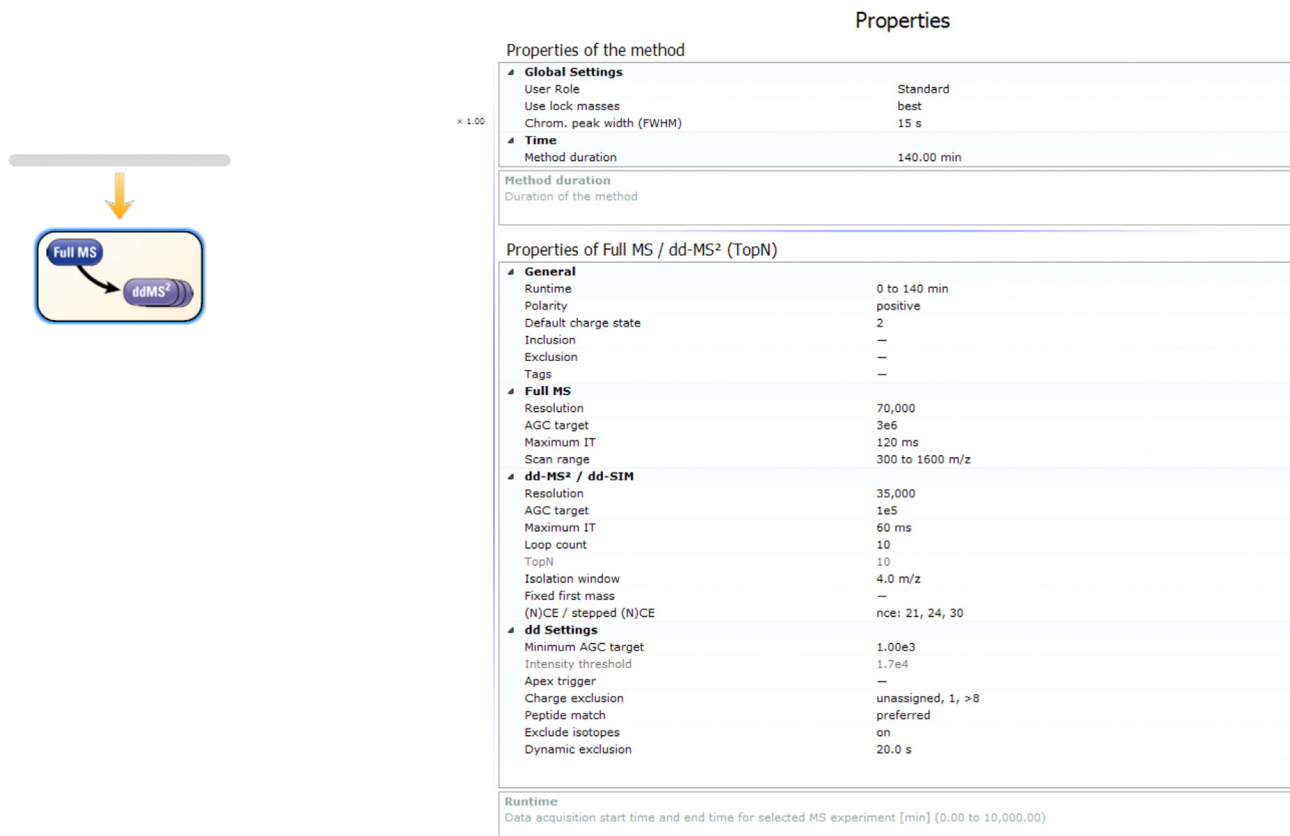
### NanoLC-MS/MS analysis

⌚ Timing: 1 week

The following section describes the parameters used in the nanoLC-MS/MS sample analysis and shotgun protein interrogation of non-cross-linked samples. We recommend the use of Sequest HT<sup>12</sup> search algorithm at Thermo Fisher's Proteome Discoverer.

### 6. NanoLC-MS/MS

- a. Resuspend the dried samples in 0.1% TFA.
- b. Desalt the peptides using C18 resin ZipTips.



**Figure 3. Settings for MS and MS<sup>2</sup> acquisition method**  
Parameters used for data dependent acquisition method.

**Note:** we recommend following the protocol described in the product insert (Merck Cat# ZTC18S096). Another alternative is to use Affinisep AttractSPE@Tips - C18, 200 $\mu$ L (Cat #Tips-C18.T1.200.96); Thermo Fisher Pierce™ C18 Tips (Cat # 87782), or home-made stage tips based on C18 membrane.

- c. Vacuum dry the desalted peptides.
- d. Resuspend in 20  $\mu$ L of ACN/0.1% FA (2:98, v/v) and then transfer to a clean autosampler vial.
- e. Inject 5  $\mu$ L of sample onto the nanoLC-MS/MS system.
- f. Analyze samples on a nanoAcquity (Waters) coupled to a Q Exactive mass spectrometer (Thermo Fisher Scientific).
  - i. The injected sample is trapped on a ACQUITY UPLC M-Class Symmetry C18 Trap Column (100  $\text{\AA}$ , 5  $\mu$ m, 180  $\mu$ m  $\times$  20 mm, 2G, V/M, Waters Part No: 186007496).
  - ii. The peptides are separated on a ACQUITY UPLC M-Class Peptide BEH C18 Column (75  $\mu$ m  $\times$  250 mm, 1.7  $\mu$ m, 130  $\text{\AA}$ , Waters Part No: 186007484) with a flow of 300 nL/min.
  - iii. Mobile phase A is ultrapure water, 0.1% formic acid, while mobile phase B is acetonitrile, 0.1% formic acid.
  - iv. The eluent gradient is set to go from 5% to 20% of mobile phase A in 100 min, then from 20% to 30% in 20 min, and finally to 90% in 20 min.
  - v. Settings for MS and MS/MS acquisitions (Figure 3): range is m/z 300–1,600, resolution 70,000 at FWHM (m/z 400), positive mode, AGC target of  $3 \times 10^6$  and stepped NCE of 21, 24 and 30.

MS/MS spectra are acquired using a Top-10 DDA (data-dependent acquisition) method, with a resolution of 35,000 FWHM. Dynamic exclusion is enabled, and only MS/MS spectra from peptide ions with charge states between +2 and +8 are selected.

### 7. Shotgun data analysis of non-cross-linked samples

**Note:** as previously described, the databased [OpenProt<sup>5</sup>](#) can be oversized for some identification nodes, we recommend the use of the limited [AltProt database](#) with at least 1 identification in other MS data or Riboseq analysis.<sup>5</sup> According to size limitation we recommend the use of SequestHT (Thermo ProteomeDiscoverer V2.5) which is not size limited.

Two different databases and consensus steps are employed to analyze AltProts (in a FASTA file combined AltProt, new isoforms and RefProt) and RefProts alone. The common consensus and processing parameters are enumerated at 7a. Specific parameters for RefProts (7b) and AltProts (7c) are displayed below.

- a. Analyze the RAW LC-MS/MS data using Proteome Discoverer V2.5 (Thermo Fisher Scientific) with the Sequest HT search engine.
  - i. Select LysC-trypsin and chymotrypsin as cleaving enzymes and 2 possible missed cleavages.
  - ii. Variable modifications: methionine oxidation and protein N-terminus acetylation.
  - iii. Static modifications: carbamidomethylation of cysteines.
  - iv. Minimum peptide length: six amino acids.
  - v. Minimum precursor tolerance: 10 ppm.
  - vi. PSM and peptide validator: between 0.01 and 0.05 FDR.
  - vii. Fragment mass tolerance: 0.02 Da.
  - viii. Validation is done with Percolator using strict FDR = 0.01 and relaxed FDR = 0.05
- b. For RefProts identification.
  - i. Protein database: [UniProtKB v.2022\\_02](#) reviewed and unreviewed. (77,895 sequences, downloaded from Uniprot website, 25 feb 2022)
  - ii. At least two peptides per sequence.
- c. For AltProts identification.
  - i. Protein database: Homo sapiens OpenProt v1.6 (184,706 sequences), containing RefProts and predicted AltProts detected in mass spectrometry experiments with at least one unique peptide.
  - ii. At least one peptide per sequence.

⚠ **CRITICAL:** To eliminate false positives, use protein [BLASTP<sup>13</sup>](#)

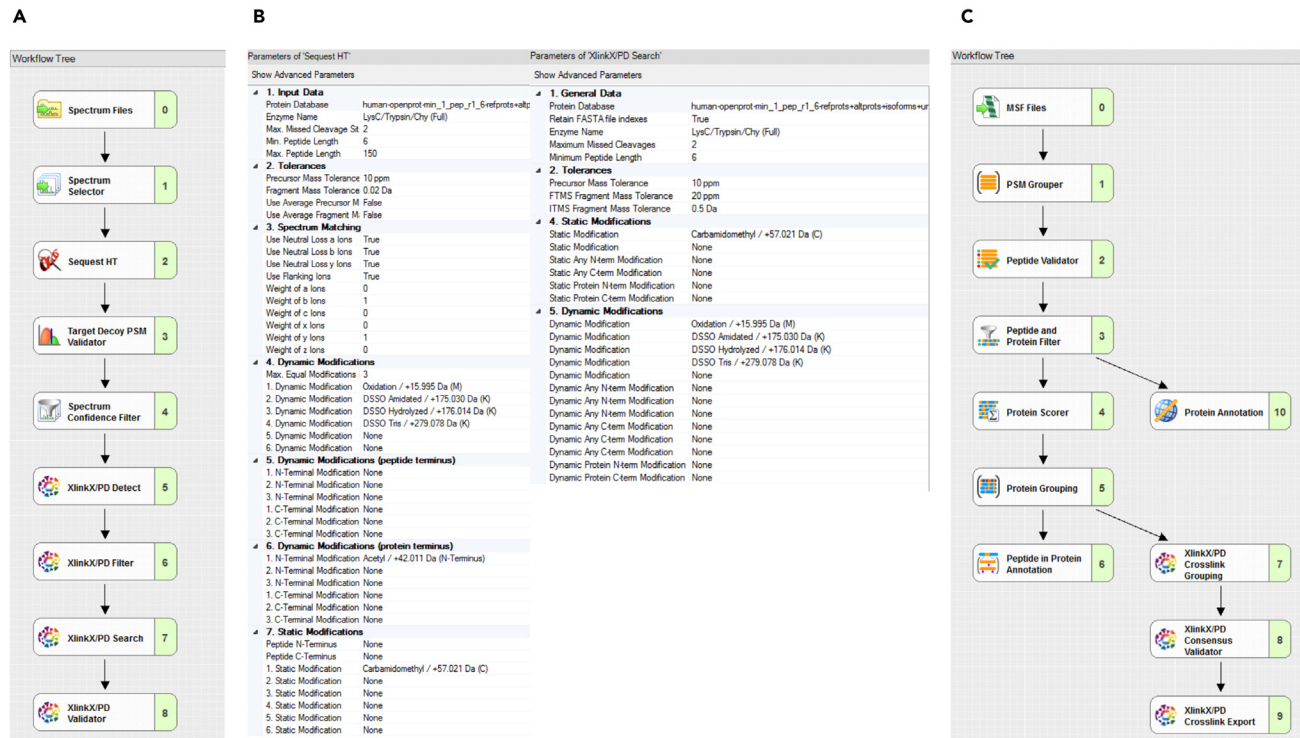
The basic parameters of BlastP are preserved (or the automated adaptation for small protein) and the database used for the comparison is "non-redundant protein sequences". An AltProt is considered "false positive" if it presents a sequence homology (identity+coverage) > 80%, ideally no identification should be detected, if an alignment <80% homology is identified, it should be checked that the peptide/PSM identified in MS is specific to the AltProt.

Additionally, check the identified PSMs of the AltProt with [the NextProt Peptide uniqueness checker<sup>14</sup>](#) tool. Parameters from the NextProt uniqueness tool have been kept unchanged and the expected result is no sequence homology for the specific peptide of the AltProt previously identified and tested.

### Cross-link data analysis and interaction modeling

⌚ Timing: 1 week





**Figure 4. Proteome Discoverer cross-link identification workflow**

(A and B) (A) Displays the processing step used to identify cross-links. The parameters used in the Sequest HT are displayed in the (B) panel. (B) The parameters used at Sequest HT and XlinkX/DP search are displayed. (C) Shows the consensus step for cross-link identification.

To identify the cross-links, we employed the XlinkX<sup>15–17</sup> node in Proteome Discoverer. Additionally, the validation of the cross-links can be performed by docking the protein-protein interactions (PPIs) and measuring the distances between the residues involved in each cross-link. Here, we describe the modeling of the 3D structure of AltProts and docking them to the RefProts to which they were cross-linked.

## 8. Cross-linking identification

- Analyze the RAW LC-MS/MS data using Proteome Discoverer V2.5 (Thermo Fisher Scientific) with the Sequest HT search engine at the processing step (Figure 4A).
  - Protein database: Homo sapiens OpenProt v1.6, which contains RefProts and predicted AltProts detected in mass spectrometry experiments with at least one unique peptide.
  - Select LysC-trypsin and chymotrypsin as cleaving enzymes and allow for 2 possible missed cleavages.
  - Set minimum peptide length to six amino acids and at least one peptide per sequence.
  - Set minimum precursor tolerance to 10 ppm.
  - Set fragment mass tolerance to 0.02 Da.
  - Maximum equal modifications per peptide: 3
  - Maximum dynamic modifications per peptide: 4
  - Variable modifications: methionine oxidation and N-terminus acetylation, DSSO amidated, hydrolyzed, and Tris form.
  - Static modification: carbamidomethylation of cysteines.
- Set the Target Decoy PSM Validator:
  - Target/decoy selection: concatenated
  - FDR set between 0.01 and 0.05.



- c. Set a spectrum confidence filter: worse than high.
- d. Detect the cross-links using the XlinkX/PD Detect node in Proteome Discoverer V2.5 with the following parameters (Figure 4B):
  - i. Set acquisition strategy as MS2.
  - ii. Set DSSO (158.0037 Da) as cross-linker.
- e. Set the following parameters at the XlinkX/PD Search node:
  - i. Set same parameters as the Sequest HT node.
  - ii. Precursor mass tolerance of 10 ppm.
  - iii. FTMS fragment of 20 ppm.
  - iv. ITMS fragment of 0.5 Da.
- f. Set XlinkX/PD Validator to FDR: 0.05.
- g. At the consensus step set a peptide validator node with a target FDR for PSMs and peptides between 0.01 to 0.05 (Figure 4C).
- h. Add a Peptide and Protein Filter with the next parameters:
  - i. Peptide Confidence At Least: High
  - ii. Minimum Number of Peptide Sequences: 1
- i. At the XlinkX/PD Consensus Validator set the cross-link spectrum match (CSM) and cross-link FDR threshold as 0.05.
- j. Perform manual curation of the identified cross-links:
  - i. Verify the quality of the CSMs. The cross-linking, b and y ions should be visible and describe the amino acid sequence of the two peptides identified in the CSM. A clear example can be observed at Garcia-del Rio et al.<sup>1</sup>

**Note:** As cross-linking technology has been evolving in the last 20 years, a community-wide effort has been done to development of methodological standards which are available for the reader<sup>18–20</sup>

- ii. Eliminate the cross-link spectrum matches that involved N-terminal residues.

**△ CRITICAL:** Verify that the peptides identified at the CSMs correspond to the attributed proteins using the NextProt peptide uniqueness checker tool.

9. Modeling and prediction of interactions between AltProts and RefProts (Figure 5A).
  - a. Retrieve the AltProts sequences from OpenProt database.
  - b. Generate the 3D models at I-TASSER (Iterative Threading ASSEmbly Refinement).<sup>21</sup>

**Note:** I-TASSER generates five models with the lowest free energy and highest confidence, and the first model usually has the highest score and better quality. However, lower-ranked models might have better quality. For more information, visit the I-TASSER server website.

- c. Download the Alphafold<sup>22</sup> or PDB<sup>23</sup> 3D structures of the RefProts involved in the cross-links.
- d. For RefProt-AltProt docking ClusPro<sup>24</sup> tool is used, submit the RefProt as a receptor and the AltProt as a ligand at the ClusPro protein-protein docking server. Do not use any restraints in the docking.
- e. After the docking is finished, display all the balanced models, and download them. Additionally, download the coefficients for these models.

**Note:** If you do not have any prior knowledge of what forces dominate in your complex, use the balanced coefficients models.

- f. Open the complex in YASARA<sup>25</sup> view and identify the number of the atoms involved in the cross-linked complex.
- g. To verify the cross-linking distance, use the command:

```
>DistanceAtomA,AtomB,bound=No
```

- h. If the distance corresponds to constrains of DSSO, use the following command to label and join both cross-linked atoms (Figure 5B):

```
>LabelDisAtomA,AtomB,Format=DIS,Height=0.7,Color=Black,X=0.0,Y=0.0,Z=0.0,bound=Yes
```

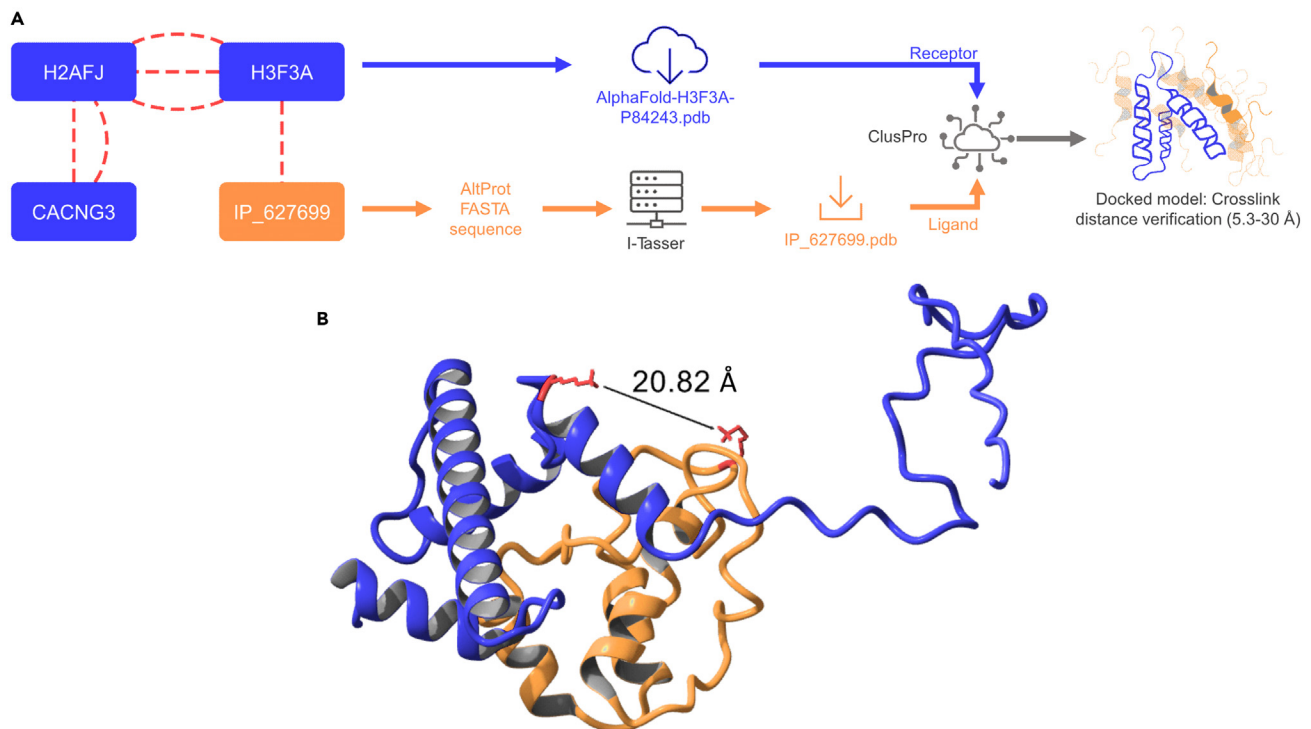
**Note:** For DSSO, the distances described in the literature are from 5.3 Å<sup>26</sup> to 30 Å.<sup>27</sup> Other molecular viewers such as PyMol<sup>28</sup> or ChimeraX<sup>29</sup> can also be used.

### Cross-linking network analysis

⌚ Timing: 1 week

For the visualization and gene ontology (GO) enrichment of the network obtained by cross-link identification, we recommend the use of Cytoscape.<sup>30</sup> Additional apps have to be downloaded at Cytoscape App Store: STRING,<sup>31</sup> ClueGO,<sup>32</sup> CluePedia<sup>33</sup> and yFiles Layout Algorithms.

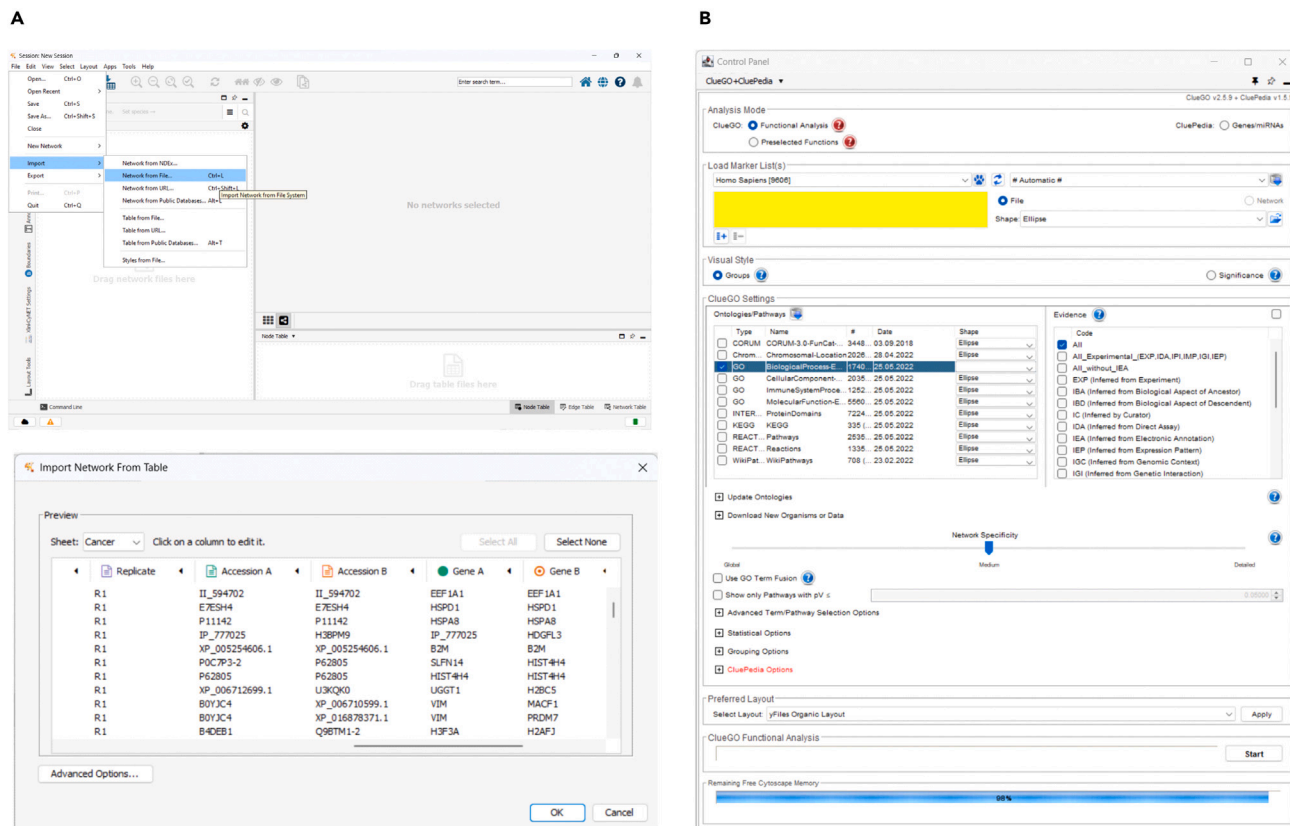
⚠ **CRITICAL:** Before starting your network analysis, we recommend performing the Cytoscape and ClueGo tutorials found on their websites. This will provide you with a general



**Figure 5. Cross-link interaction modeling workflow and result of the interaction found between H3F3A and the AltProt IP\_627699**

(A) Workflow used to generate an interaction model. H3F3A is identified cross-linked to IP\_627699 in MS analysis. H3F3A 3D model is obtained from AlphaFold databased. IP\_627699 FASTA sequence is obtained from OpenProt and is used to generate a 3D model at I-Tasser. Then the two models are docked in ClusPro. Finally, the best interaction model processed, the distances between the cross-linked residues are measured.

(B) 3D ribbon model showing the interaction between H3F3A (blue) and IP\_627699 (orange). The cross-linked residues are displayed in red and the distance of the interaction [20.82 Å], confirmed the possible cross-link identification as its fits in the restricted distance of DSSO [5 to 30 Å].



**Figure 6. Cytoscape and ClueGo interface windows**

(A) Displays how to import a cross-linking network from a file of identification by Proteome Discoverer 2.5. Additionally, it shows how to label the columns as target and source nodes during this process.

(B) Presents the control panel of the ClueGo app at Cytoscape. The loading marker square, where to load the query protein list, is highlighted in yellow.

panorama of the commands, formatting options, and different analyses that can be performed in the software.

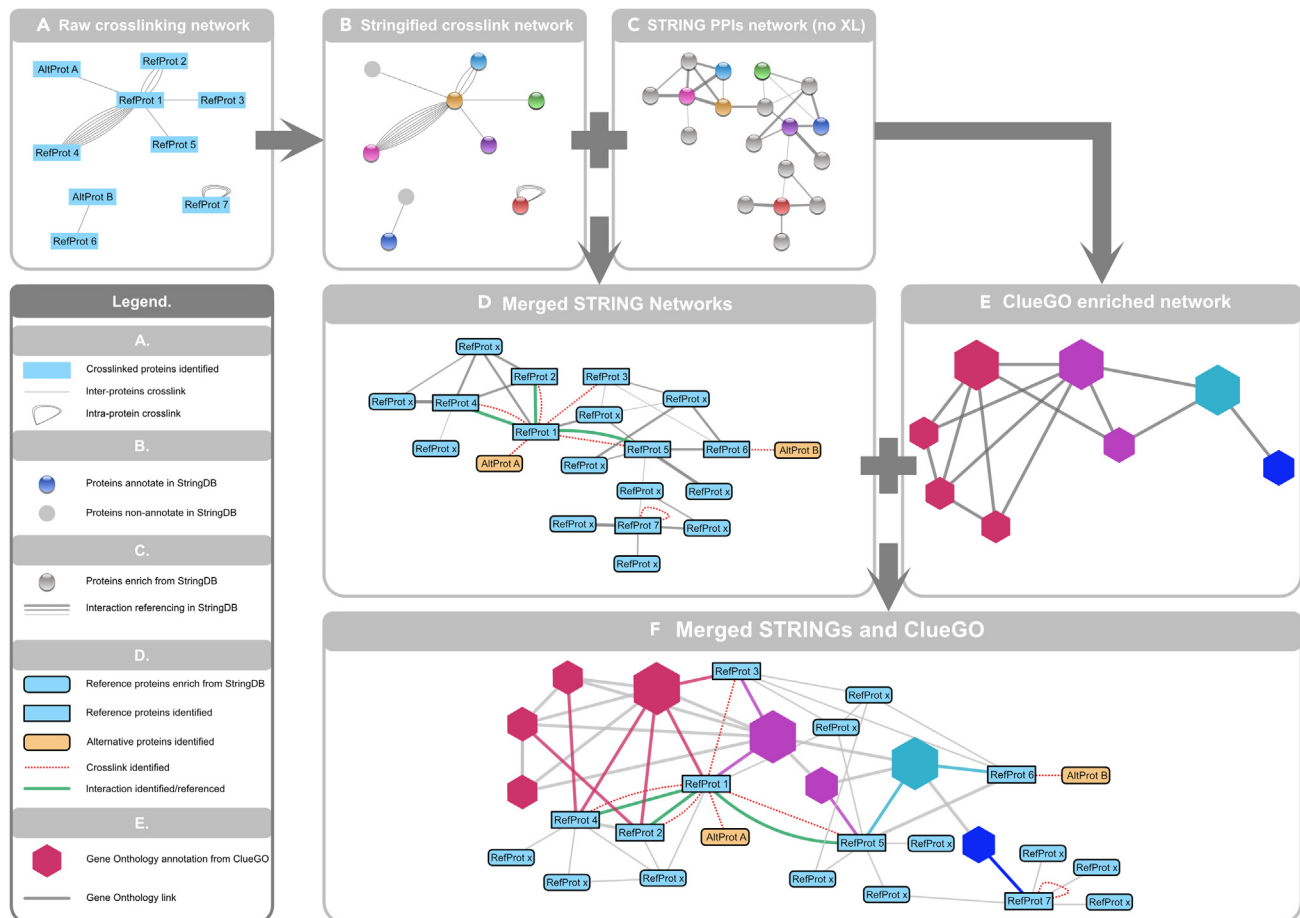
10. PPIs network treatment.

- Export the cross-links identified from Proteome Discoverer 2.5 as an Excel file (Figure 6A).
- Open the file and split the column description to obtain the gene symbols of the proteins identified in cross-link.
- Write in two new columns (Gene A and Gene B) the gene annotations.
- Import this network from the file to Cytoscape.
  - Assign the source node to the Gene A column and the target node to Gene B column.
- Select the network (Figure 7A) and STRINGify it (STRING App).

**Note:** Verify that all RefProt nodes are now STRING nodes. If not (plain gray nodes at Figure 7B), verify that there are no spaces in the accession numbers or genes. For Cytoscape a space after the gene/accession creates a duplicate non-referenced node.

- Once all the RefProts have a STRING node, copy the column of accession numbers, and input them in the STRING protein query. This step will identify the already described interactions between the RefProts (Figure 7C).

**Note:** If there are RefProt nodes that do not have any interactors, select them one by one and add known interactors in the STRING menu (Figure 7C, bright gray STRING nodes).



**Figure 7. Cross-linking network evolution during the processing steps**

(A) Shows a raw cross-linking network after import.

(B) Stringified cross-linking network. Note that the AltProts are not in the STRING format, meaning they are not indexed at string database.

(C) Displays only the interaction between the RefProts without cross-links. The enriched nodes are shown in gray STRING nodes.

(D) After merging B and C, all the interactions are displayed. Formatting is done for the nodes and edges.

(E) Presents a ClueGo enriched network. Only GO term nodes are displayed, the protein nodes are kept hidden.

(F) Resulting network after merging D and E. Combined the information of enriched proteins, query RefProts and AltProt, connected to the GO term. As well as the cross-link identified interactions, StringDB and other databased enriched interaction existing between the proteins of the network.

g. Merge the networks.

**Note:** The resulting merged network (Figure 7D) will simplify the redundant cross-links. We recommend formatting this network to visualize the interactions of interest.

h. Cross-validate the RefProt interactions in other databases. We recommend [BioGrid<sup>34</sup>](#) and [IntAct<sup>35</sup>](#)

i. GO term enrichment (Figure 6B).

- i. Open ClueGo App.
- ii. Select the functional analysis as analysis mode.
- iii. Input the whole list of RefProt genes at the load marker list box.
- iv. Select the ontologies/pathways to use for the enrichment.
- v. Select the specificity of the network.
- vi. Enable the GO term fusion.
- vii. Run the app.

- j. Merge the ClueGO network (Figure 7E) and the merged STRING network (Figure 7D).
- k. At the resulting network (Figure 7F) select the preferred network layout and edit the node's properties. We recommend formatting the network in a way that you can obtain the information you need.

### EXPECTED OUTCOMES

This methodology enables the identification of AltProts, their subcellular localization, and protein-protein interactions (PPIs) in cell lines. An easy way to verify that a cross-linking reaction has occurred is by using SDS-PAGE. After a cross-linking reaction, larger protein complexes will be present in the sample. These larger complexes cannot enter the concentration gel and will be observable at the top of the wells and some at the interface between concentration and migration gels. Additionally, the disappearance of some protein bands in the cross-linked sample, compared to a negative cross-linking control, can indicate that a cross-linking reaction has taken place (Figure 2A). To validate the subcellular protein fractionation, we recommend using western blotting and compartment-specific antibodies. The optimal outcome should be that the signal of the antibody is present in just one fraction, but traces of specific markers may appear in other fractions (Figure 2B).

For AltProts identification, an AltProt/RefProt ratio of 5%–10% can be expected. It is always recommended to validate identifications by performing a BLAST search and using the NextProt peptide uniqueness checker. Additionally, different properties and characteristics of the AltProts identified can be retrieved from the OpenProt database. Although peptide fractionation using SEC<sup>36</sup> and SCX<sup>37</sup> columns can increase the number of cross-links identified, our fast and non-fractionated method can only be expected to identify a couple of hundred interactions at most. Even a small number of AltProt-RefProt cross-links is sufficient to infer their possible function or pathway involvement. This methodology was exemplified in a study of PPIs in immortalized human ovarian epithelial cells (SV40), and the characterization of the AltProts identified in the cell line.<sup>1</sup> In this study, the subcellular localization of 112 AltProts was observed, and subcellular protein fractionation decreased the complexity of the cross-linked sample, allowing us to identify a network of 220 cross-links without peptide enrichment, 16 of which were AltProt-RefProt interactions. Furthermore, the possible involvement of these AltProts in some cellular processes, such as antigen processing and presentation of peptide antigen via MHC class I, mRNA transcription by RNA polymerase II, and regulation of mitochondrial outer membrane permeabilization involved in apoptotic signaling pathway, was investigated.

### LIMITATIONS

The methodology described above has some limitations. The first limitation (I) is related to the lack of information on AltProts. Still, the concept of alternative proteins has limited spread, and few tools and databases have been developed for the analysis of this ghost proteome. Therefore, the methodology used here provides us with a snapshot of some possible functions of a limited number of AltProts. Targeted studies need to be conducted to confirm or expand the information about these AltProt "hits".

The second limitation (II) is linked to the protein fractionation technique. The kit employed in this protocol is based on pelleting the non-extracted fraction and removing the supernatant. If the removal of the supernatant is not optimal, traces of proteins that don't correspond to the next fraction could remain. Additionally, there is limited information about the buffer composition and detergents employed. To avoid the use of the kit, other subcellular fractionation techniques can be used (e.g., sucrose gradient), since they are compatible with MS.

The third limitation (III) is the need to decrease the complexity of the cross-linking sample before injection into the nLC-MS/MS system, increasing the detection of cross-linked peptides. For this,

fractionation techniques like SEC and SCX chromatography or the use of enrichable cross-linkers like NNP9,<sup>38</sup> tBu-PhoX,<sup>39</sup> and alkyne-A-DSBSO<sup>40</sup> can help identify more cross-linked peptides.

## TROUBLESHOOTING

### Problem 1

No cross-linking patterns are observed in the SDS-PAGE gels (related to Step 1).

#### Potential solution

- Verify the BSA positive control. If there is no cross-link in BSA positive control, it means DSSO was hydrolyzed.
- Avoid buffers that contain primary amines.
- Repeat the cross-linking reaction with a new vial of DSSO. Follow the vendor's storage and use recommendations.

### Problem 2

No extraction is observed in the SDS-PAGE gels after subcellular fractionation (related to Step 2).

#### Potential solution

- Verify that the volumes used for the extraction are appropriate (read the [vendor's manual](#)).
- Remove DPBS completely before starting and keep the pellet as dry as possible (according to the kit manufacturer).
- Increase the incubation times.
- Vortex at the highest setting.
- Add the appropriate volume of Halt Protease Inhibitor Cocktail.

### Problem 3

The extracted proteins are not compartmentalized (related to Step 2).

#### Potential solution

- Vortex longer to disperse completely the cell pellets.
- Increase the incubation times.
- Carefully remove all extracts before proceeding to the next step. Remove the remaining buffer with a smaller pipette.
- Re-centrifuge sample and remove excess extract.
- Primary antibodies are not specific.
- Verify if the protein selected as compartment-specific is reported in literature to be in other compartments.

### Problem 4

Urea is not dissolved at the Denaturing solution (related to Step 3).

#### Potential solution

- Place the solution in an ultrasonic bath and sonicate it for 5–10 min.
- Freeze and thaw the buffer to solubilize the urea.

### Problem 5

Presence of polymer traces in the membrane fractions (related to Step 4).

### Potential solution

- Verify by MALDI-MS that all the solutions prepared are not contaminated by polymers.
- Repeat the HiPPR™ Detergent Removal Resin protocol and dilute the sample.

### Problem 6

No cross-links identified by XlinkX at Proteome Discoverer (related to Step 5).

### Potential solution

- If the confirmation of the cross-linking reaction by SDS-PAGE was skipped, there is no certainty that the cross-linking reaction happened.
- If only cross-link dead ends are found in the sample, perform peptide fractionation (SEC or SCX) of the samples after the digestion.

### Problem 7

Unable to find the lysine residue involved in the cross-link after modeling the interaction in ClusPro (related to Step 5).

### Potential solution

- The cross-linking description found in Proteome Discoverer is based on the sequence found in the FASTA file. However, some PDB accessions do not present all the amino acids described in the protein databases. Therefore, we recommend finding the 3D structures in which the residues involved are present. In UniProt, under the structure menu, we can observe the different 3D models and coverage for each protein.
- Please note that after docking in ClusPro, the model numbering starts with the receptor, followed by the ligand protein. To find the position of the ligand residue involved in the cross-link, simply add the residues of the receptor to the position of the cross-link.

## RESOURCE AVAILABILITY

### Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Michel Salzet ([michel.salzet@univ-lille.fr](mailto:michel.salzet@univ-lille.fr)).

### Materials availability

This study did not generate new unique reagents.

### Data and code availability

- The mass spectrometry proteomics data from Garcia-del Rio et al.<sup>1</sup> have been deposited to the ProteomeXchange Consortium via the PRIDE<sup>41</sup> partner repository with the dataset identifier PXD035764, study following this protocol.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## ACKNOWLEDGMENTS

This research was supported by funding from I-SITE of “Université Lille Nord-Europe” (ULNE), the European metropole of Lille (MEL), the Regional Council of “Hauts de France”, the “Institut National de la Santé et de la Recherche Médicale” (Inserm), and Université de Lille.



## AUTHOR CONTRIBUTIONS

Conceptualization, M.S., T.C.; Methodology, D.F.G.-d.R., T.C.; Software, D.F.G.-d.R., T.C.; Formal Analysis, D.F.G.-d.R.; Resources, I.F., M.S.; Data curation, D.F.G.-d.R., T.C.; Writing - Original Draft, D.F.G.-d.R., T.C.; Writing - Review & Editing, I.F., M.S.; Supervision, T.C., M.S., I.F.; Project Administration, M.S., I.F.; Funding Acquisition, M.S., I.F.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## REFERENCES

- Garcia-del Rio, D.F., Cardon, T., Eyckerman, S., Fournier, I., Bonnefond, A., Gevaert, K., and Salzter, M. (2023). Employing non-targeted interactomics approach and subcellular fractionation to increase our understanding of the ghost proteome. *iScience* 26, 105943. <https://doi.org/10.1016/j.isci.2023.105943>.
- Cardon, T., Salzter, M., Franck, J., and Fournier, I. (2019). Nuclei of HeLa cells interactomes unravel a network of ghost proteins involved in proteins translation. *Biochim. Biophys. Acta. Gen. Subj.* 1863, 1458–1470. <https://doi.org/10.1016/j.bbagen.2019.05.009>.
- Martinez, T.F., Chu, Q., Donaldson, C., Tan, D., Shokhirev, M.N., and Saghatelian, A. (2020). Accurate annotation of human protein-coding small open reading frames. *Nat. Chem. Biol.* 16, 458–468. <https://doi.org/10.1038/s41589-019-0425-0>.
- Brunet, M.A., Brunelle, M., Lucier, J.-F., Delcourt, V., Levesque, M., Grenier, F., Samandi, S., Leblanc, S., Aguilar, J.-D., Dufour, P., et al. (2019). OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteomes. *Nucleic Acids Res.* 47, D403–D410. <https://doi.org/10.1093/nar/gky936>.
- Brunet, M.A., Lucier, J.-F., Levesque, M., Leblanc, S., Jacques, J.-F., Al-Saedi, H.R.H., Guillo, N., Grenier, F., Avino, M., Fournier, I., et al. (2021). OpenProt 2021: deeper functional annotation of the coding potential of eukaryotic genomes. *Nucleic Acids Res.* 49, D380–D388. <https://doi.org/10.1093/nar/gkaa1036>.
- Aboulouard, S., Wisztorski, M., Duhamel, M., Saudemont, P., Cardon, T., Narducci, F., Lemaire, A.-S., Kobeissy, F., Leblanc, E., Fournier, I., and Salzter, M. (2021). In-depth proteomics analysis of sentinel lymph nodes from individuals with endometrial cancer. *Cell Rep. Med.* 2, 100318. <https://doi.org/10.1016/j.xcrm.2021.100318>.
- Hajjaji, N., Aboulouard, S., Cardon, T., Bertin, D., Robin, Y.-M., Fournier, I., and Salzter, M. (2021). Path to clonal theranostics in luminal breast cancers. *Front. Oncol.* 11, 802177. <https://doi.org/10.3389/fonc.2021.802177>.
- Le Rhun, E., Duhamel, M., Wisztorski, M., Gimeno, J.-P., Zairi, F., Escande, F., Reyns, N., Kobeissy, F., Muraige, C.-A., Salzter, M., and Fournier, I. (2017). Evaluation of non-supervised MALDI mass spectrometry imaging combined with microproteomics for glioma grade III classification. *Biochim. Biophys. Acta. Proteins Proteom.* 1865, 875–890. <https://doi.org/10.1016/j.bbapap.2016.11.012>.
- Cardon, T., Fournier, I., and Salzter, M. (2021). Shedding light on the ghost proteome. *Trends Biochem. Sci.* 46, 239–250. <https://doi.org/10.1016/j.tibs.2020.10.003>.
- Leblanc, S., and Brunet, M.A. (2020). Modelling of pathogen-host systems using deeper ORF annotations and transcriptomics to inform proteomics analyses. *Comput. Struct. Biotechnol. J.* 18, 2836–2850. <https://doi.org/10.1016/j.csbj.2020.10.010>.
- Wiśniewski, J.R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat. Methods* 6, 359–362. <https://doi.org/10.1038/nmeth.1322>.
- Tabb, D.L., Eng, J.K., and Yates, J.R. (2001). Protein identification by SEQUEST. In *Proteome Research: Mass Spectrometry Principles and Practice*, P. James, ed. (Springer), pp. 125–142. [https://doi.org/10.1007/978-3-642-56895-4\\_7](https://doi.org/10.1007/978-3-642-56895-4_7).
- McGinnis, S., and Madden, T.L. (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res.* 32, W20–W25. <https://doi.org/10.1093/nar/gkh435>.
- Schaeffer, M., Gateau, A., Teixeira, D., Michel, P.-A., Zahn-Zabal, M., and Lane, L. (2017). The neXtProt peptide uniqueness checker: a tool for the proteomics community. *Bioinforma. Oxf. Engl.* 33, 3471–3472. <https://doi.org/10.1093/bioinformatics/btx318>.
- Liu, F., van Breukelen, B., and Heck, A.J.R. (2014). Facilitating protein disulfide mapping by a combination of pepsin digestion, electron transfer higher energy dissociation (ETHCD), and a dedicated search algorithm SlinkS. *Mol. Cell. Proteomics* 13, 2776–2786. <https://doi.org/10.1074/mcp.O114.039057>.
- Liu, F., Rijkers, D.T.S., Post, H., and Heck, A.J.R. (2015). Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. *Nat. Methods* 12, 1179–1184. <https://doi.org/10.1038/nmeth.3603>.
- Liu, F., Lössl, P., Scheltema, R., Viner, R., and Heck, A.J.R. (2017). Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. *Nat. Commun.* 8, 15473. <https://doi.org/10.1038/ncomms15473>.
- Leitner, A., Bonvin, A.M.J.J., Borchers, C.H., Chalkley, R.J., Chamot-Rooke, J., Combe, C.W., Cox, J., Dong, M.-Q., Fischer, L., Götze, M., et al. (2020). Toward increased reliability, transparency, and accessibility in cross-linking mass spectrometry. *Structure* 28, 1259–1268. <https://doi.org/10.1016/j.str.2020.09.011>.
- Iacobucci, C., Piotrowski, C., Aebersold, R., Amaral, B.C., Andrews, P., Bernfur, K., Borchers, C., Brodie, N.I., Bruce, J.E., Cao, Y., et al. (2019). First community-wide, comparative cross-linking mass spectrometry study. *Anal. Chem.* 91, 6953–6961. <https://doi.org/10.1021/acs.analchem.9b00658>.
- Piersimoni, L., Kastritis, P.L., Arlt, C., and Sinz, A. (2022). Cross-linking mass spectrometry for investigating protein conformations and protein–protein interactions—a method for all seasons. *Chem. Rev.* 122, 7500–7531. <https://doi.org/10.1021/acs.chemrev.1c00786>.
- Yang, J., Yan, R., Roy, A., Xu, D., Poisson, J., and Zhang, Y. (2015). The I-TASSER Suite: protein structure and function prediction. *Nat. Methods* 12, 7–8. <https://doi.org/10.1038/nmeth.3213>.
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. <https://doi.org/10.1038/s41586-021-03819-2>.
- Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The protein data bank. *Nucleic Acids Res.* 28, 235–242. <https://doi.org/10.1093/nar/28.1.235>.
- Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhorny, D., Yueh, C., Beglov, D., and Vajda, S. (2017). The ClusPro web server for protein–protein docking. *Nat. Protoc.* 12, 255–278. <https://doi.org/10.1038/nprot.2016.169>.
- Land, H., and Humble, M.S. (2018). YASARA: a tool to obtain structural guidance in biocatalytic investigations. *Methods Mol. Biol.* 1685, 43–67. [https://doi.org/10.1007/978-1-4939-7366-8\\_4](https://doi.org/10.1007/978-1-4939-7366-8_4).
- Kao, A., Chiu, C.I., Vellucci, D., Yang, Y., Patel, V.R., Guan, S., Randall, A., Baldi, P., Rychnovsky, S.D., and Huang, L. (2011). Development of a novel cross-linking strategy for fast and accurate identification of cross-linked peptides of protein complexes. *Mol. Cell. Proteomics* 10. <https://doi.org/10.1074/mcp.M110.002212>.
- Hevler, J.F., Lukassen, M.V., Cabrera-Orefice, A., Arnold, S., Pronker, M.F., Franc, V., and Heck, A.J.R. (2021). Selective cross-linking of coinciding protein assemblies by in-gel cross-linking mass spectrometry. *EMBO J.* 40,



- e106174. <https://doi.org/10.15252/embj.2020106174>.
28. Schrödinger, LLC (2015). The PyMOL Molecular Graphics System, Version 1.8.
  29. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. <https://doi.org/10.1002/pro.3943>.
  30. Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. <https://doi.org/10.1101/gr.1239303>.
  31. Doncheva, N.T., Morris, J.H., Gorodkin, J., and Jensen, L.J. (2019). Cytoscape StringApp: network analysis and visualization of proteomics data. *J. Proteome Res.* 18, 623–632. <https://doi.org/10.1021/acs.jproteome.8b00702>.
  32. Bindea, G., Mlecnik, B., Hackl, H., Charoentong, P., Tosolini, M., Kirilovsky, A., Fridman, W.-H., Pagès, F., Trajanoski, Z., and Galon, J. (2009). ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 25, 1091–1093. <https://doi.org/10.1093/bioinformatics/btp101>.
  33. Bindea, G., Galon, J., and Mlecnik, B. (2013). CluePedia Cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinforma. Oxf. Engl.* 29, 661–663. <https://doi.org/10.1093/bioinformatics/btt019>.
  34. Oughtred, R., Rust, J., Chang, C., Breitkreutz, B.-J., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., Zhang, F., et al. (2021). The BioGRID database: a comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Sci.* 30, 187–200. <https://doi.org/10.1002/pro.3978>.
  35. Orchard, S., Ammari, M., Aranda, B., Breuza, L., Briganti, L., Broackes-Carter, F., Campbell, N.H., Chavali, G., Chen, C., del-Toro, N., et al. (2014). The MIntAct project—IntAct as a common curation platform for 11 molecular interaction databases. *Nucleic Acids Res.* 42, D358–D363. <https://doi.org/10.1093/nar/gkt1115>.
  36. Mendes, M.L., Fischer, L., Chen, Z.A., Barbon, M., O'Reilly, F.J., Giese, S.H., Bohlke-Schneider, M., Belsom, A., Dau, T., Combe, C.W., et al. (2019). An integrated workflow for crosslinking mass spectrometry. *Mol. Syst. Biol.* 15, e8994. <https://doi.org/10.15252/msb.20198994>.
  37. Fritzsche, R., Ihling, C.H., Götze, M., and Sinz, A. (2012). Optimizing the enrichment of cross-linked products for mass spectrometric protein analysis. *Rapid Commun. Mass Spectrom.* 26, 653–658. <https://doi.org/10.1002/rcm.6150>.
  38. Nury, C., Redeker, V., Dautrey, S., Romieu, A., van der Rest, G., Renard, P.-Y., Melki, R., and Chamot-Rooke, J. (2015). A novel bio-orthogonal cross-linker for improved protein/protein interaction analysis. *Anal. Chem.* 87, 1853–1860. <https://doi.org/10.1021/ac503892c>.
  39. Jiang, P.-L., Wang, C., Diehl, A., Viner, R., Etienne, C., Nandhikonda, P., Foster, L., Bomgardner, R.D., and Liu, F. (2022). A membrane-permeable and immobilized metal affinity chromatography (IMAC) enrichable cross-linking reagent to advance in vivo cross-linking mass spectrometry. *Angew. Chem. Int. Ed.* 61, e202113937. <https://doi.org/10.1002/anie.202113937>.
  40. Burke, A.M., Kandur, W., Novitsky, E.J., Kaake, R.M., Yu, C., Kao, A., Vellucci, D., Huang, L., and Rychnovsky, S.D. (2015). Synthesis of two new enrichable and MS-cleavable cross-linkers to define protein-protein interactions by mass spectrometry. *Org. Biomol. Chem.* 13, 5030–5037. <https://doi.org/10.1039/C5OB00488H>.
  41. Perez-Riverol, Y., Bai, J., Bandla, C., García-Seisdedos, D., Hewapathirana, S., Kamatchinathan, S., Kundu, D.J., Prakash, A., Frericks-Zipper, A., Eisenacher, M., et al. (2022). The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res.* 50, D543–D552. <https://doi.org/10.1093/nar/gkab1038>.