CASP 13 internal ID: T0975/X0975

Cross-linking mass spectrometry data

**Protein information (as provided)**

CASP13 internal ID: regular target T0987, XL-MS assisted X0987.

Protein Name: Exo5

Organism Name: Homo sapiens

Amino acid sequence:

10 20 30 40 50 60   
LEDAQESKAL VNMPGPSSES LGKDDKPISL QNWKRGLDIL SPMERFHLKY LYVTDLATQN   
  
 70 80 90 100 110 120   
WCELQTAYGK ELPGFLAPEK AAVLDTGASI HLARELELHD LVTVPVTTKE DAWAIKFLNI   
  
 130 140 150 160 170 180   
LLLIPTLQSE GHIREFPVFG EGEGVLLVGV IDELHYTAKG ELELAELKTR RRPMLPLEAQ   
  
 190 200 210 220 230 240   
KKKDCFQVSL YKYIFDAMVQ GKVTPASLIH HTKLCLEKPL GPSVLRHAQQ GGFSVKSLGD   
  
 250 260 270 280 290 300   
LMELVFLSLT LSDLPVIDIL KIEYIHQETA TVLGTEIVAF KEKEVRAKVQ HYMAYWMGHR   
  
 310 320 330 340   
EPQGVDVEEA WKCRTCTYAD ICEWRKGSGV LSSTLAPQVK KAK

**Methods**

The target protein or protein complex was cross-linked and analyzed by mass spectrometry as described here:

*Lysine-specific chemical cross-linking of protein complexes and identification of cross-linking sites using LC-MS/MS and the xQuest/xProphet software pipeline*. Leitner, Walzthoeni and Aebersold. *Nature Protocols*, 2014. DOI: 10.1038/nprot.2013.168

*Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes*. Leitner, Joachimiak, Unverdorben, Walzthoeni, Frydman, Förster and Aebersold. *Proceedings of the National Academy of Sciences of the United States of America,* 2014. DOI: 10.1073/pnas.1320298111

The concentration the protein or protein complex was adjusted to avoid over-cross-linking, e.g. introduction of non-native oligomerization states. All cross-linking reactions were followed by SDS-PAGE.

**Cross-links identified by mass spectrometry**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Intra protein cross-links** | | | | | |
| **Protein1** | **Protein2** | **AbsPos1** | **AbsPos2** | **ld-Score\*** | **Chemistry** |
| Exo5 | Exo5 | 288 | 183 | 45.78 | DSS |
| Exo5 | Exo5 | 288 | 181 | 44.92 | DSS |
| Exo5 | Exo5 | 26 | 71 | 40.96 | ZL |
| Exo5 | Exo5 | 34 | 71 | 39.36 | ZL |
| Exo5 | Exo5 | 168 | 183 | 39.27 | DSS |
| Exo5 | Exo5 | 80 | 168 | 38.38 | DSS |
| Exo5 | Exo5 | 183 | 192 | 37.79 | DSS |
| Exo5 | Exo5 | 326 | 25 | 36.71 | ZL |
| Exo5 | Exo5 | 34 | 38 | 36.21 | ZL |
| Exo5 | Exo5 | 288 | 283 | 36.1 | DSS |
| Exo5 | Exo5 | 288 | 182 | 35.27 | DSS |
| Exo5 | Exo5 | 116 | 100 | 35.01 | ZL |
| Exo5 | Exo5 | 326 | 163/166 | 34.76 | ZL |
| Exo5 | Exo5 | 340 | 71 | 34.61 | ZL |
| Exo5 | Exo5 | 38 | 71 | 34.25 | PDH |
| Exo5 | Exo5 | 116 | 85 | 34.16 | ZL |
| Exo5 | Exo5 | 281 | 288 | 33.57 | DSS |
| Exo5 | Exo5 | 109 | 116 | 33.54 | DSS |
| Exo5 | Exo5 | 26 | 34 | 33.53 | DSS |
| Exo5 | Exo5 | 71 | 283 | 33.45 | ZL |
| Exo5 | Exo5 | 276 | 181 | 33.12 | ZL |
| Exo5 | Exo5 | 80 | 183 | 32.8 | DSS |
| Exo5 | Exo5 | 116 | 97 | 32.5 | ZL |
| Exo5 | Exo5 | 312 | 168 | 29.72 | DSS |
| Exo5 | Exo5 | 168 | 181 | 29.43 | DSS |
| Exo5 | Exo5 | 326 | 26 | 28.35 | DSS |
| Exo5 | Exo5 | 340 | 288 | 27.63 | DSS |
| Exo5 | Exo5 | 301 | 38 | 27.27 | PDH |
| Exo5 | Exo5 | 301 | 71 | 26.62 | PDH |

\* The score is a measure of confidence for the identification of the two connected peptides (i.e. computational assignment) that are identified by MS (the higher, the better). It is generally NOT correlated with the distance between the cross-linked residues. In addition, physicochemical properties of the peptides may affect the identification, so that some cross-linked peptides intrinsically have lower scores.

For the reported data, we expect a **false positive rate of identification of approximately 5%.**

\*\* Cross-linking chemistries:

DSS: disuccinimidyl suberate – a lysine specific cross-linker.  
ZL: Zero-length cross-links formed between lysine and an aspartate/glutamate residue by the coupling reagent 4-(4,6-dimethoxy-1,3,5- triazin-2-yl)-4-methylmorpholinium chloride (DMTMM).  
PDH: pimelic acid dihydrazide – a carboxylic acid specific cross-linker (aspartate and glutamate).

For experimentally observed distance restraints, see the following plot (ADH is not used here):



(taken from Leitner et al., PNAS, 2014)

**Sub-optimal sequence regions for conventional cross-linking mass spectrometry**

Red residues: Lysine residues. Can be cross-linked by DSS and the zero-length cross-linking reagent DMTMM. Cleavage sites for trypsin (protease used in the experimental process).

Black residues: Arginine residues. Cleavage sites for trypsin (protease used in the experimental process).

Green residues: Aspartate and glutamate residues. Can be cross-linked by PDH and the zero-length cross-linking reagent DMTMM.

Residues highlighted in yellow are sub-optimal regions for mass spectrometric analysis.

10 20 30 40 50 60

L**ED**AQ**E**S**K**AL VNMPGPSS**E**S LG**KDDK**PISL QNW**KR**GL**D**IL SPM**ER**FHL**K**Y LYVT**D**LATQN

70 80 90 100 110 120

WC**E**LQTAYG**K E**LPGFLAP**EK** AAVL**D**TGASI HLA**RE**L**E**LH**D** LVTVPVTT**KE D**AWAI**K**FLNI

130 140 150 160 170 180

LLLIPTLQS**E** GHI**RE**FPVFG **E**G**E**GVLLVGV I**DE**LHYTA**K**G **E**L**E**LA**E**L**K**T**R RR**PMLPL**E**AQ

190 200 210 220 230 240

**KKKD**CFQVSL Y**K**YIF**D**AMVQ G**K**VTPASLIH HT**K**LCL**EK**PL GPSVL**R**HAQQ GGFSV**K**SLG**D**

250 260 270 280 290 300

LM**E**LVFLSLT LS**D**LPVI**D**IL **K**I**E**YIHQ**E**TA TVLGT**E**IVAF **KEKE**V**R**A**K**VQ HYMAYWMGH**R**

310 320 330 340

**E**PQGV**D**V**EE**A W**K**C**R**TCTYA**D** IC**E**W**RK**GSGV LSSTLAPQV**K K**A**K**

**Residues labelled by cross-linking reagents**

Red residues: Residues labeled by either by DSS (Lysine reactive) or PDH (carboxylic acid reactive), but not cross-linked. These residues are expected to be solvent exposed.

Green residues: Reactive unlabeled residues.

Notes:

Absence of a modification may also mean that the corresponding modified peptide is present, but not identified by MS.

10 20 30 40 50 60

L**ED**AQ**E**S**K**AL VNMPGPSS**E**S LG**KDDK**PISL QNW**K**RGL**D**IL SPM**E**RFHL**K**Y LYVT**D**LATQN

70 80 90 100 110 120

WC**E**LQTAYG**K E**LPGFLAP**EK** AAVL**D**TGASI HLAR**E**L**E**LH**D** LVTVPVTT**KE D**AWAI**K**FLNI

130 140 150 160 170 180

LLLIPTLQS**E** GHIR**E**FPVFG **E**G**E**GVLLVGV I**DE**LHYTA**K**G **E**L**E**LA**E**L**K**TR RRPMLPL**E**AQ

190 200 210 220 230 240

**KKKD**CFQVSL Y**K**YIF**D**AMVQ G**K**VTPASLIH HT**K**LCL**EK**PL GPSVLRHAQQ GGFSV**K**SLG**D**

250 260 270 280 290 300

LM**E**LVFLSLT LS**D**LPVI**D**IL **K**I**E**YIHQ**E**TA TVLGT**E**IVAF **KEKE**VRA**K**VQ HYMAYWMGHR

310 320 330 340

**E**PQGV**D**V**EE**A W**K**CRTCTYA**D** IC**E**WR**K**GSGV LSSTLAPQV**K K**A**K**