CASP 13 internal ID: T0985/X0985

Cross-linking mass spectrometry data

**Protein information (as provided)**

CASP13 internal ID: regular target T0985, XL-MS assisted X0985.

Protein Name: ACL\_1061

Organism Name: Acholeplasma laidlawii PG-8A

Amino acid sequence:

10 20 30 40 50 60   
MAHHHHHHVG TGSNDDDDKS PMKLKQDVIS IYQKISLFES GQLNITKLAS GAYYLDDELT   
  
 70 80 90 100 110 120   
LITDPVNSGA RFPYAVNGMT IWAYASGYIS INHSSYYILP PNLEGKEPFL DFFGIEQDGN   
  
 130 140 150 160 170 180   
NTYPVSLLGV SERNDEIENK RYTVFSKNIA YYITVTKNFL YAVTVYISKD FKIYFNTVAH   
  
 190 200 210 220 230 240   
NLTGETKQIT LSSFFNMLFK YDSGESIETK WFKKVSYENN MFIYDAPEDI DRHTRIENYG   
  
 250 260 270 280 290 300   
VVKRHLHTKP KNIQNTTSRI DYVGKRYRSV RNALSIRSLK FEKAPLVTNF TDTAINADLI   
  
 310 320 330 340 350 360   
NYEVKAYDTI ISSYRIETCH DKDTLNKMMA SDLTDKEIKK VYEGLSNTQS YDFDNFGISF   
  
 370 380 390 400 410 420   
KGVNDNRVDD KVLNQFLKLV NYQIHFSSLS SNSGTVFLGV RDVMQQLESS LIWDRKNVRS   
  
 430 440 450 460 470 480   
KILEVLSFID PSGLPPRQYA LPPKEGNPRM DLRPFIDQGL WIISTLHTYL AYTEDYDILN   
  
 490 500 510 520 530 540   
EVCGYYERIE PNSAKKSKVE NSVLEHLIRV TNYLVSNIDP STYGLKALYG DWNDALDGLG   
  
 550 560 570 580 590 600   
LIEGSSGYGN GVSVMATLQL YENLERMIEI LKLVDPQNEH INTYEVVRHN LSLGINKYAV   
  
 610 620 630 640 650 660   
VIKQDEKRVL HGWGHDRSYF VGSFNDPDGH SRNSLTSNAF YIISDMIKNT PEMKPHLLHA   
  
 670 680 690 700 710 720   
FHNLDSKYGL KTFDPAMQDF HGFGRIINLP PGTAENAATY VHATLFGVLA LYMLGEGDFA   
  
 730 740 750 760 770 780   
NEQVLKVLPI TKKEMSTSPF IMPNSYVHNE ELNMDGESMS DWYTGSANTL LKTLIRGLFG   
  
 790 800 810 820 830 840   
LEVKFDHLRL RPSKAFFSKE ATLMVSIGNK LTRIVYKNNN NGNRTFKLNG KVIEAKLDTL   
  
 850 860   
SGLLYIDINK SILEHQNVIH IQD

**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 of 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**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Inter-chain cross-links\*** | | | | | |
| **Protein1** | **Protein2** | **AbsPos1** | **AbsPos2** | **ld-Score\*\*** | **Chemistry\*\*\*** |
| ACL\_1061 | ACL\_1061 | 225 | 214 | 34.76 | ZL |
| ACL\_1061 | ACL\_1061 | 19 | 25 | 28.97 | DSS |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Intra-chain cross-links** | | | | | |
| **Protein1** | **Protein2** | **AbsPos1** | **AbsPos2** | **ld-Score\*\*** | **Chemistry\*\*\*** |
| ACL\_1061 | ACL\_1061 | 39 | 140 | 45.65 | ZL |
| ACL\_1061 | ACL\_1061 | 799 | 817 | 43.4 | DSS |
| ACL\_1061 | ACL\_1061 | 25 | 140 | 43.24 | DSS |
| ACL\_1061 | ACL\_1061 | 327 | 444 | 42.75 | DSS |
| ACL\_1061 | ACL\_1061 | 140 | 265 | 42.67 | DSS |
| ACL\_1061 | ACL\_1061 | 322 | 444 | 42.12 | DSS |
| ACL\_1061 | ACL\_1061 | 817 | 827 | 41.77 | DSS |
| ACL\_1061 | ACL\_1061 | 800 | 817 | 41.63 | ZL |
| ACL\_1061 | ACL\_1061 | 25 | 249 | 41.32 | DSS |
| ACL\_1061 | ACL\_1061 | 800 | 794 | 41.17 | ZL |
| ACL\_1061 | ACL\_1061 | 251 | 280 | 41.07 | DSS |
| ACL\_1061 | ACL\_1061 | 243 | 280 | 39.95 | DSS |
| ACL\_1061 | ACL\_1061 | 322 | 25 | 39.73 | DSS |
| ACL\_1061 | ACL\_1061 | 265 | 280 | 39.64 | DSS |
| ACL\_1061 | ACL\_1061 | 243 | 249 | 38.84 | DSS |
| ACL\_1061 | ACL\_1061 | 243 | 280 | 38.77 | DSS |
| ACL\_1061 | ACL\_1061 | 322 | 336 | 38.62 | DSS |
| ACL\_1061 | ACL\_1061 | 800 | 794 | 38.06 | ZL |
| ACL\_1061 | ACL\_1061 | 25 | 243 | 37.61 | DSS |
| ACL\_1061 | ACL\_1061 | 336 | 25 | 37.52 | DSS |
| ACL\_1061 | ACL\_1061 | 147 | 138 | 37.52 | ZL |
| ACL\_1061 | ACL\_1061 | 305 | 25 | 36.95 | DSS |
| ACL\_1061 | ACL\_1061 | 838 | 831 | 35.94 | ZL |
| ACL\_1061 | ACL\_1061 | 654 | 794 | 35.28 | DSS |
| ACL\_1061 | ACL\_1061 | 336 | 243 | 35.13 | DSS |
| ACL\_1061 | ACL\_1061 | 243 | 265 | 34.15 | DSS |
| ACL\_1061 | ACL\_1061 | 361 | 371 | 34.08 | DSS |
| ACL\_1061 | ACL\_1061 | 322 | 265 | 33.4 | DSS |
| ACL\_1061 | ACL\_1061 | 231 | 208 | 33.32 | PDH |
| ACL\_1061 | ACL\_1061 | 498 | 495 | 33.28 | DSS |
| ACL\_1061 | ACL\_1061 | 597 | 605 | 33.15 | ZL |
| ACL\_1061 | ACL\_1061 | 249 | 280 | 32.98 | DSS |
| ACL\_1061 | ACL\_1061 | 214 | 25 | 32.49 | DSS |
| ACL\_1061 | ACL\_1061 | 498 | 444 | 32.31 | DSS |
| ACL\_1061 | ACL\_1061 | 336 | 322 | 29.7 | DSS |
| ACL\_1061 | ACL\_1061 | 444 | 265 | 29.48 | DSS |
| ACL\_1061 | ACL\_1061 | 444 | 243 | 29.41 | DSS |
| ACL\_1061 | ACL\_1061 | 339 | 243 | 27.47 | DSS |
| ACL\_1061 | ACL\_1061 | 444 | 140 | 27.31 | DSS |
| ACL\_1061 | ACL\_1061 | 421 | 444 | 26.63 | DSS |

\* These cross-links connect two identical residues or peptides with overlapping sequence, pointing to a homo-dimeric contact (native or non-native).

\*\* 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 process).

Black residues: Arginine residues. Cleavage sites for trypsin.

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

MAHHHHHHVG TGSN**DDDDK**S PM**K**L**K**Q**D**VIS IYQ**K**ISLF**E**S GQLNIT**K**LAS GAYYL**DDE**LT

70 80 90 100 110 120

LIT**D**PVNSGA **R**FPYAVNGMT IWAYASGYIS INHSSYYILP PNL**E**G**KE**PFL **D**FFGI**E**Q**D**GN

130 140 150 160 170 180

NTYPVSLLGV S**ER**N**DE**I**E**N**K R**YTVFS**K**NIA YYITVT**K**NFL YAVTVYIS**KD** F**K**IYFNTVAH

190 200 210 220 230 240

NLTG**E**T**K**QIT LSSFFNMLF**K** Y**D**SG**E**SI**E**T**K** WF**KK**VSY**E**NN MFIY**D**AP**ED**I **DR**HT**R**I**E**NYG

250 260 270 280 290 300

VV**KR**HLHT**K**P **K**NIQNTTS**R**I **D**YVG**KR**Y**R**SV **R**NALSI**R**SL**K** F**EK**APLVTNF T**D**TAINA**D**LI

310 320 330 340 350 360

NY**E**V**K**AY**D**TI ISSY**R**I**E**TCH **DKD**TLN**K**MMA S**D**LT**DKE**I**KK** VY**E**GLSNTQS Y**D**F**D**NFGISF

370 380 390 400 410 420

**K**GVN**D**N**R**V**DD K**VLNQFL**K**LV NYQIHFSSLS SNSGTVFLGV **RD**VMQQL**E**SS LIW**DRK**NV**R**S

430 440 450 460 470 480

**K**IL**E**VLSFI**D** PSGLPP**R**QYA LPP**KE**GNP**R**M **D**L**R**PFI**D**QGL WIISTLHTYL AYT**ED**Y**D**ILN

490 500 510 520 530 540

**E**VCGYY**ER**I**E** PNSA**KK**S**K**V**E** NSVL**E**HLI**R**V TNYLVSNI**D**P STYGL**K**ALYG **D**WN**D**AL**D**GLG

550 560 570 580 590 600

LI**E**GSSGYGN GVSVMATLQL Y**E**NL**ER**MI**E**I L**K**LV**D**PQN**E**H INTY**E**VV**R**HN LSLGIN**K**YAV

610 620 630 640 650 660

VI**K**Q**DEKR**VL HGWGH**DR**SYF VGSFN**D**P**D**GH S**R**NSLTSNAF YIIS**D**MI**K**NT P**E**M**K**PHLLHA

670 680 690 700 710 720

FHNL**D**S**K**YGL **K**TF**D**PAMQ**D**F HGFG**R**IINLP PGTA**E**NAATY VHATLFGVLA LYMLG**E**G**D**FA

730 740 750 760 770 780

N**E**QVL**K**VLPI T**KKE**MSTSPF IMPNSYVHN**E E**LNM**D**G**E**SMS **D**WYTGSANTL L**K**TLI**R**GLFG

790 800 810 820 830 840

L**E**V**K**F**D**HL**R**L **R**PS**K**AFFS**KE** ATLMVSIGN**K** LT**R**IVY**K**NNN NGN**R**TF**K**LNG **K**VI**E**A**K**L**D**TL

850 860

SGLLYI**D**IN**K** SIL**E**HQNVIH IQ**D**

**Residues labelled by cross-linking reagents**

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

Green residues: reactive unlabeled residues.

Notes:

Complex formation in the cross-linked sample may not be quantitative, so that exposed regions could only be accessible in the free binding partners.

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

MAHHHHHHVG TGSN**DDDDK**S PM**K**L**K**Q**D**VIS IYQ**K**ISLF**E**S GQLNIT**K**LAS GAYYL**DDE**LT

70 80 90 100 110 120

LIT**D**PVNSGA RFPYAVNGMT IWAYASGYIS INHSSYYILP PNL**E**G**KE**PFL **D**FFGI**E**Q**D**GN

130 140 150 160 170 180

NTYPVSLLGV S**E**RN**DE**I**E**N**K** RYTVFS**K**NIA YYITVT**K**NFL YAVTVYIS**KD** F**K**IYFNTVAH

190 200 210 220 230 240

NLTG**E**T**K**QIT LSSFFNMLF**K** Y**D**SG**E**SI**E**T**K** WF**KK**VSY**E**NN MFIY**D**AP**ED**I **D**RHTRI**E**NYG

250 260 270 280 290 300

VV**K**RHLHT**K**P **K**NIQNTTSRI **D**YVG**K**RYRSV RNALSIRSL**K** F**EK**APLVTNF T**D**TAINA**D**LI

310 320 330 340 350 360

NY**E**V**K**AY**D**TI ISSYRI**E**TCH **DKD**TLN**K**MMA S**D**LT**DKE**I**KK** VY**E**GLSNTQS Y**D**F**D**NFGISF

370 380 390 400 410 420

**K**GVN**D**NRV**DD K**VLNQFL**K**LV NYQIHFSSLS SNSGTVFLGV R**D**VMQQL**E**SS LIW**D**R**K**NVRS

430 440 450 460 470 480

**K**IL**E**VLSFI**D** PSGLPPRQYA LPP**KE**GNPRM **D**LRPFI**D**QGL WIISTLHTYL AYT**ED**Y**D**ILN

490 500 510 520 530 540

**E**VCGYY**E**RI**E** PNSA**KK**S**K**V**E** NSVL**E**HLIRV TNYLVSNI**D**P STYGL**K**ALYG **D**WN**D**AL**D**GLG

550 560 570 580 590 600

LI**E**GSSGYGN GVSVMATLQL Y**E**NL**E**RMI**E**I L**K**LV**D**PQN**E**H INTY**E**VVRHN LSLGIN**K**YAV

610 620 630 640 650 660

VI**K**Q**DEK**RVL HGWGH**D**RSYF VGSFN**D**P**D**GH SRNSLTSNAF YIIS**D**MI**K**NT P**E**M**K**PHLLHA

670 680 690 700 710 720

FHNL**D**S**K**YGL **K**TF**D**PAMQ**D**F HGFGRIINLP PGTA**E**NAATY VHATLFGVLA LYMLG**E**G**D**FA

730 740 750 760 770 780

N**E**QVL**K**VLPI T**KKE**MSTSPF IMPNSYVHN**E E**LNM**D**G**E**SMS **D**WYTGSANTL L**K**TLIRGLFG

790 800 810 820 830 840

L**E**V**K**F**D**HLRL RPS**K**AFFS**KE** ATLMVSIGN**K** LTRIVY**K**NNN NGNRTF**K**LNG **K**VI**E**A**K**L**D**TL

850 860

SGLLYI**D**IN**K** SIL**E**HQNVIH IQ**D**