CASP 13 internal ID: T0999/X0999

Cross-linking mass spectrometry data

**Protein information (as provided)**

CASP13 internal ID: regular target T0999, XL-MS assisted X0999.

Protein Name: Pentafunctional AROM polypeptide

Organism Name: Chaetomium thermophilum

**Amino acid sequence:**

 10 20 30 40 50 60
MATANVAGAG GSGSEPTRIA ILGKEDIIVD HGIWLNFVAH DLLQTLPSST YVLITDTNLY

 70 80 90 100 110 120
TTYVPPFQAV FEAAAPRDVR LLTYAIPPGE YSKSRETKAE IEDWMLSHAC TRDTVIIALG

 130 140 150 160 170 180
GGVIGDMIGY VAATFMRGVR FVQVPTTLLA MVDSSIGGKT AIDTPMGKNL IGAFWQPRRI

 190 200 210 220 230 240
YIDLAFLETL PVREFINGMA EVIKTAAIWN ETEFTALEEN AAAILEAVRS KASSPAARLA

 250 260 270 280 290 300
PIRHILKRIV LGSARVKAEV VSADEREGGL RNLLNFGHSI GHAYEAILAP QVLHGECVAI

 310 320 330 340 350 360
GMVKEAELAR YLGVLRPSAV ARLTKLIASY DLPTSVHDKR IAKLSAGKEC PVDVLLQKMA

 370 380 390 400 410 420
VDKKNEGRKK KIVLLSAIGK TYEKKATVVD DRAIRLVLSP SVRVTPGVPK GLSVTVTPPG

 430 440 450 460 470 480
SKSISNRALV LAALGEGTTR IHGLLHSDDV QYMLAAIEQL HGADFSWEDA GEILVVTGKG

 490 500 510 520 530 540
GKLQASKEPL YLGNAGTASR FLTSVVALCA PSAVSSTVLT GNARMKVRPI GALVDALRAN

 550 560 570 580 590 600
GVGVKYLEKE KSLPVEVDAA GGFAGGVIEL AATVSSQYVS SILMAAPYAH QPVTLRLVGG

 610 620 630 640 650 660
KPISQPYIDM TIAMMASFGI KVERSAEDPN TYLIPKGVYK NPPEYVVESD ASSATYPLAV

 670 680 690 700 710 720
AAITGTTCTI PNIGSESLQG DARFAVEVLR PMGCAVEQTA TSTTVTGPPI GTLKAIPHVD

 730 740 750 760 770 780
MEPMTDAFLT AAVLAAVADG TTQITGIANQ RVKECNRIAA MKDQLAKFGV QCNELEDGIE

 790 800 810 820 830 840
VIGKPYQELR NPVEGIYCYD DHRVAMSHSV LSTISPHPVL ILERECTAKT WPGWWDILSQ

 850 860 870 880 890 900
FFKVQLDGEE DPTKRTTQST QQVRKGTDRS IFIVGMRGAG KSTAGRWMSE LLKRPLVDLD

 910 920 930 940 950 960
AELERREGMT IPEIIRGERG WEGFRQAELE LLQDVIKNQS KGYIFSCGGG IVETEAARKL

 970 980 990 1000 1010 1020
LIDYHKNGGP VLLVHRDTDQ VVEYLMRDKT RPAYSENIRE VYERRKPWFY ECSNLQYHSP

 1030 1040 1050 1060 1070 1080
HEDGSEALLQ PPADFARFVK LIAGQSTHLE DVRAKKHSFF VSLTVPNVAD ALDIIPRVVV

 1090 1100 1110 1120 1130 1140
GSDAVELRVD LLESYEPEFV ARQVALLRAA AQVPIVYTVR TQSQGGKFPD EDYDLALRLY

 1150 1160 1170 1180 1190 1200
QTGLRSGVEY LDLEMTMPDH ILQAVTDAKG FTSIIASHHD PQCKLSWKSG SWIPFYNKAL

 1210 1220 1230 1240 1250 1260
QYGDVIKLVG VAREMADNFA LTNFKAKMLA AHDNKPMIAL NMGTAGKLSR VLNGFLTPVS

 1270 1280 1290 1300 1310 1320
HPALPSKAAP GQLSATEIRQ ALSLIGEIEP KSFYLFGKPI SASRSPALHN TLFYKTGLPH

 1330 1340 1350 1360 1370 1380
HYSRFETDEA SKALESLIRS PDFGGASVTI PLKLDIMPLL DSATDAARTI GAVNTIIPQT

 1390 1400 1410 1420 1430 1440
RDGSTTTLVG DNTDWRGMVH ALLHSSGSGS VVQRTAAPRG AAMVVGSGGT ARAAIYALHD

 1450 1460 1470 1480 1490 1500
LGFAPIWIVA RSEERVAELV RGFDGYDLRR MTSPHQGKDN MPSVVISTIP ATQPIDPSMR

 1510 1520 1530 1540 1550 1560
EVIVEVLKHG HPSAEGKVLL EMAYQPPRTP LMTLAEDQGW RTVGGLEVLA AQGWYQFQLW

 1570 1580
TGITPLYEEA RAAVMGEDSV ELEHHHHHH

**Methods**

**IMPORTANT NOTE: In contrast to all other cross-linking data provided by this laboratory, the AROM data was not acquired explicitly for CASP 13. Therefore, only one cross-linking chemistry (DSS) was used for the experiments. In addition, no information about modified lysine residues (other than those found in cross-links) is provided.**

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

**Cross-links identified by mass spectrometry**

|  |
| --- |
| **Intra-chain cross-links** |
| **Protein1** | **Protein2** | **AbsPos1** | **AbsPos2** | **ld-Score\*** | **Chemistry\*\*** |
| AROM | AROM | 339 | 343 | 35.26 | DSS |
| AROM | AROM | 371 | 385 | 34.51 | DSS |
| AROM | AROM | 526 | 549 | 34.19 | DSS |
| AROM | AROM | 371 | 384 | 34.06 | DSS |
| AROM | AROM | 989 | 989 | 33.93 | DSS |
| AROM | AROM | 526 | 343 | 33.78 | DSS |
| AROM | AROM | 371 | 363 | 32.86 | DSS |
| AROM | AROM | 966 | 343 | 32.51 | DSS |
| AROM | AROM | 371 | 385 | 32.4 | DSS |
| AROM | AROM | 762 | 343 | 32.2 | DSS |
| AROM | AROM | 371 | 384 | 31.79 | DSS |
| AROM | AROM | 959 | 549 | 31.71 | DSS |
| AROM | AROM | 339 | 854 | 31.69 | DSS |
| AROM | AROM | 1040 | 881 | 31.58 | DSS |
| AROM | AROM | 526 | 753 | 31.45 | DSS |
| AROM | AROM | 854 | 343 | 31.23 | DSS |
| AROM | AROM | 526 | 482 | 31.06 | DSS |
| AROM | AROM | 487 | 231 | 30.94 | DSS |
| AROM | AROM | 1127 | 231 | 30.86 | DSS |
| AROM | AROM | 358 | 364 | 30.62 | DSS |
| AROM | AROM | 989 | 343 | 30.56 | DSS |
| AROM | AROM | 1235 | 343 | 30.48 | DSS |
| AROM | AROM | 343 | 959 | 30.31 | DSS |
| AROM | AROM | 989 | 549 | 30.25 | DSS |
| AROM | AROM | 1188 | 231 | 30.11 | DSS |
| AROM | AROM | 966 | 549 | 29.9 | DSS |
| AROM | AROM | 989 | 363 | 29.88 | DSS |
| AROM | AROM | 370 | 384 | 29.79 | DSS |
| AROM | AROM | 339 | 231 | 29.74 | DSS |
| AROM | AROM | 636 | 482 | 29.71 | DSS |
| AROM | AROM | 358 | 364 | 29.68 | DSS |
| AROM | AROM | 941 | 549 | 29.62 | DSS |
| AROM | AROM | 989 | 854 | 29.6 | DSS |
| AROM | AROM | 343 | 959 | 29.39 | DSS |
| AROM | AROM | 545 | 231 | 29.28 | DSS |
| AROM | AROM | 526 | 959 | 29.26 | DSS |
| AROM | AROM | 937 | 762 | 29.14 | DSS |
| AROM | AROM | 989 | 762 | 29.08 | DSS |
| AROM | AROM | 526 | 231 | 29.03 | DSS |
| AROM | AROM | 1040 | 865 | 28.71 | DSS |
| AROM | AROM | 371 | 364 | 28.68 | DSS |
| AROM | AROM | 937 | 343 | 28.62 | DSS |
| AROM | AROM | 966 | 231 | 28.46 | DSS |
| AROM | AROM | 339 | 526 | 28.31 | DSS |
| AROM | AROM | 937 | 549 | 28.31 | DSS |
| AROM | AROM | 937 | 959 | 27.99 | DSS |
| AROM | AROM | 966 | 231 | 27.7 | DSS |
| AROM | AROM | 410 | 343 | 27.62 | DSS |
| AROM | AROM | 545 | 959 | 27.61 | DSS |
| AROM | AROM | 231 | 959 | 27.48 | DSS |
| AROM | AROM | 1040 | 231 | 27.41 | DSS |
| AROM | AROM | 410 | 231 | 27.35 | DSS |
| AROM | AROM | 98 | 168 | 27.15 | DSS |
| AROM | AROM | 1127 | 1127 | 27.12 | DSS |
| AROM | AROM | 487 | 487 | 27.05 | DSS |
| AROM | AROM | 1235 | 1188 | 27 | DSS |
| AROM | AROM | 339 | 959 | 26.94 | DSS |
| AROM | AROM | 989 | 343 | 26.94 | DSS |
| AROM | AROM | 487 | 989 | 26.53 | DSS |
| AROM | AROM | 371 | 363 | 26.43 | DSS |
| AROM | AROM | 937 | 231 | 26.42 | DSS |
| AROM | AROM | 937 | 753 | 26.35 | DSS |
| AROM | AROM | 1040 | 343 | 26.28 | DSS |
| AROM | AROM | 1127 | 549 | 26.27 | DSS |
| AROM | AROM | 1235 | 231 | 26.14 | DSS |
| AROM | AROM | 168 | 98 | 26.09 | DSS |
| AROM | AROM | 410 | 385 | 26.08 | DSS |
| AROM | AROM | 1040 | 989 | 26.05 | DSS |
| AROM | AROM | 941 | 231 | 25.97 | DSS |
| AROM | AROM | 941 | 339 | 25.86 | DSS |
| AROM | AROM | 941 | 854 | 25.7 | DSS |
| AROM | AROM | 1127 | 989 | 25.7 | DSS |
| AROM | AROM | 1188 | 989 | 25.65 | DSS |
| AROM | AROM | 1184 | 881 | 25.39 | DSS |
| AROM | AROM | 339 | 989 | 25.39 | DSS |
| AROM | AROM | 1235 | 231 | 25.37 | DSS |
| AROM | AROM | 989 | 385 | 25.26 | DSS |
| AROM | AROM | 358 | 989 | 24.89 | DSS |
| AROM | AROM | 989 | 959 | 24.45 | DSS |
| AROM | AROM | 893 | 1040 | 24.11 | DSS |
| AROM | AROM | 1227 | 348 | 24 | DSS |
| AROM | AROM | 1235 | 358 | 23.93 | DSS |
| AROM | AROM | 893 | 865 | 23.78 | DSS |
| AROM | AROM | 1227 | 168 | 23.77 | DSS |
| AROM | AROM | 1227 | 1227 | 23.76 | DSS |
| AROM | AROM | 989 | 881 | 23.73 | DSS |
| AROM | AROM | 526 | 343 | 23.65 | DSS |
| AROM | AROM | 1235 | 343 | 23.6 | DSS |
| AROM | AROM | 1184 | 989 | 23.56 | DSS |
| AROM | AROM | 762 | 959 | 23.55 | DSS |
| AROM | AROM | 1198 | 231 | 23.52 | DSS |
| AROM | AROM | 966 | 339 | 23.41 | DSS |
| AROM | AROM | 1235 | 364 | 23.4 | DSS |
| AROM | AROM | 854 | 959 | 23.37 | DSS |
| AROM | AROM | 941 | 231 | 23.26 | DSS |
| AROM | AROM | 1184 | 231 | 22.81 | DSS |
| AROM | AROM | 1235 | 168 | 22.76 | DSS |
| AROM | AROM | 487 | 636 | 22.63 | DSS |
| AROM | AROM | 1184 | 343 | 22.45 | DSS |
| AROM | AROM | 1227 | 1188 | 22.26 | DSS |
| AROM | AROM | 487 | 636 | 22.21 | DSS |
| AROM | AROM | 168 | 371 | 22.2 | DSS |
| AROM | AROM | 784 | 410 | 22.09 | DSS |
| AROM | AROM | 1040 | 854 | 21.95 | DSS |
| AROM | AROM | 1225 | 1188 | 21.88 | DSS |
| AROM | AROM | 1267 | 1127 | 21.51 | DSS |
| AROM | AROM | 168 | 358 | 21.4 | DSS |
| AROM | AROM | 358 | 854 | 21.18 | DSS |
| AROM | AROM | 487 | 526 | 21.13 | DSS |
| AROM | AROM | 941 | 343 | 21.04 | DSS |
| AROM | AROM | 865 | 881 | 21.03 | DSS |
| AROM | AROM | 1184 | 1184 | 20.66 | DSS |
| AROM | AROM | 1188 | 1188 | 20.64 | DSS |
| AROM | AROM | 1235 | 1235 | 20.57 | DSS |
| AROM | AROM | 941 | 343 | 20.51 | DSS |
| AROM | AROM | 487 | 545 | 20.37 | DSS |
| AROM | AROM | 410 | 487 | 20.36 | DSS |
| AROM | AROM | 410 | 881 | 20.31 | DSS |
| AROM | AROM | 784 | 989 | 20.2 | DSS |
| AROM | AROM | 98 | 1188 | 19.32 | DSS |
| AROM | AROM | 1227 | 989 | 18.7 | DSS |
| AROM | AROM | 168 | 1188 | 18.54 | DSS |
| AROM | AROM | 1227 | 231 | 18.53 | DSS |
| AROM | AROM | 941 | 854 | 18.09 | DSS |
| AROM | AROM | 545 | 545 | 17.8 | DSS |
| AROM | AROM | 1184 | 1127 | 17.1 | DSS |
| AROM | AROM | 482 | 636 | 16.94 | DSS |
| AROM | AROM | 937 | 989 | 16.77 | DSS |
| AROM | AROM | 941 | 989 | 16.65 | DSS |

\* 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-links reflecting different conformations may exist.

\*\*Cross-linking chemistries:

DSS: disuccinimidyl suberate – a lysine specific cross-linker.

For experimentally observed distance restraints, see the following plot (only DSS was used for this protein):



(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 are cleavage sites for trypsin (protease used in the experimental process).

Black residues: Arginine residues. Cleavage sites for trypsin.

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

 10 20 30 40 50 60
MATANVAGAG GSGSEPT**R**IA ILG**K**EDIIVD HGIWLNFVAH DLLQTLPSST YVLITDTNLY

 70 80 90 100 110 120
TTYVPPFQAV FEAAAP**R**DV**R** LLTYAIPPGE YS**K**S**R**ET**K**AE IEDWMLSHAC T**R**DTVIIALG

 130 140 150 160 170 180
GGVIGDMIGY VAATFM**R**GV**R** FVQVPTTLLA MVDSSIGG**K**T AIDTPMG**K**NL IGAFWQP**RR**I

 190 200 210 220 230 240
YIDLAFLETL PV**R**EFINGMA EVI**K**TAAIWN ETEFTALEEN AAAILEAV**R**S **K**ASSPAA**R**LA

 250 260 270 280 290 300
PI**R**HIL**KR**IV LGSA**R**V**K**AEV VSADE**R**EGGL **R**NLLNFGHSI GHAYEAILAP QVLHGECVAI

 310 320 330 340 350 360
GMV**K**EAELA**R** YLGVL**R**PSAV A**R**LT**K**LIASY DLPTSVHD**KR** IA**K**LSAG**K**EC PVDVLLQ**K**MA

 370 380 390 400 410 420
VD**KK**NEG**RKK** **K**IVLLSAIG**K** TYE**KK**ATVVD D**R**AI**R**LVLSP SV**R**VTPGVP**K** GLSVTVTPPG

 430 440 450 460 470 480
S**K**SISN**R**ALV LAALGEGTT**R** IHGLLHSDDV QYMLAAIEQL HGADFSWEDA GEILVVTG**K**G

 490 500 510 520 530 540
G**K**LQAS**K**EPL YLGNAGTAS**R** FLTSVVALCA PSAVSSTVLT GNA**R**M**K**V**R**PI GALVDAL**R**AN

 550 560 570 580 590 600
GVGV**K**YLE**K**E **K**SLPVEVDAA GGFAGGVIEL AATVSSQYVS SILMAAPYAH QPVTL**R**LVGG

 610 620 630 640 650 660
**K**PISQPYIDM TIAMMASFGI **K**VE**R**SAEDPN TYLIP**K**GVY**K** NPPEYVVESD ASSATYPLAV

 670 680 690 700 710 720
AAITGTTCTI PNIGSESLQG DA**R**FAVEVL**R** PMGCAVEQTA TSTTVTGPPI GTL**K**AIPHVD

 730 740 750 760 770 780
MEPMTDAFLT AAVLAAVADG TTQITGIANQ **R**V**K**ECN**R**IAA M**K**DQLA**K**FGV QCNELEDGIE

 790 800 810 820 830 840
VIG**K**PYQEL**R** NPVEGIYCYD DH**R**VAMSHSV LSTISPHPVL ILE**R**ECTA**K**T WPGWWDILSQ

 850 860 870 880 890 900
FF**K**VQLDGEE DPT**KR**TTQST QQV**RK**GTD**R**S IFIVGM**R**GAG **K**STAG**R**WMSE LL**KR**PLVDLD

 910 920 930 940 950 960
AELE**RR**EGMT IPEII**R**GE**R**G WEGF**R**QAELE LLQDVI**K**NQS **K**GYIFSCGGG IVETEAA**RK**L

 970 980 990 1000 1010 1020
LIDYH**K**NGGP VLLVH**R**DTDQ VVEYLM**R**D**K**T **R**PAYSENI**R**E VYE**RRK**PWFY ECSNLQYHSP

 1030 1040 1050 1060 1070 1080
HEDGSEALLQ PPADFA**R**FV**K** LIAGQSTHLE DV**R**A**KK**HSFF VSLTVPNVAD ALDIIP**R**VVV

 1090 1100 1110 1120 1130 1140
GSDAVEL**R**VD LLESYEPEFV A**R**QVALL**R**AA AQVPIVYTV**R** TQSQGG**K**FPD EDYDLAL**R**LY

 1150 1160 1170 1180 1190 1200
QTGL**R**SGVEY LDLEMTMPDH ILQAVTDA**K**G FTSIIASHHD PQC**K**LSW**K**SG SWIPFYN**K**AL

 1210 1220 1230 1240 1250 1260
QYGDVI**K**LVG VA**R**EMADNFA LTNF**K**A**K**MLA AHDN**K**PMIAL NMGTAG**K**LS**R** VLNGFLTPVS

 1270 1280 1290 1300 1310 1320
HPALPS**K**AAP GQLSATEI**R**Q ALSLIGEIEP **K**SFYLFG**K**PI SAS**R**SPALHN TLFY**K**TGLPH

 1330 1340 1350 1360 1370 1380
HYS**R**FETDEA S**K**ALESLI**R**S PDFGGASVTI PL**K**LDIMPLL DSATDAA**R**TI GAVNTIIPQT

 1390 1400 1410 1420 1430 1440
**R**DGSTTTLVG DNTDW**R**GMVH ALLHSSGSGS VVQ**R**TAAP**R**G AAMVVGSGGT A**R**AAIYALHD

 1450 1460 1470 1480 1490 1500
LGFAPIWIVA **R**SEE**R**VAELV **R**GFDGYDL**RR** MTSPHQG**K**DN MPSVVISTIP ATQPIDPSM**R**

 1510 1520 1530 1540 1550 1560
EVIVEVL**K**HG HPSAEG**K**VLL EMAYQPP**R**TP LMTLAEDQGW **R**TVGGLEVLA AQGWYQFQLW

 1570 1580
TGITPLYEEA **R**AAVMGEDSV ELEHHHHHH