
STRUCTURE FUNCTION ANALYSIS OF THE DEUBIQUITYLATING ENZYME FAM

Poon-Yu Khut, B.Sc

**School of Molecular & Biomedical Science (Biochemistry)
University of Adelaide
Adelaide, South Australia 5005**

October 2006

APPENDIX A

Tables listing the results of searches using the monoisotopic masses $[M^+H]^+$ of the observed spectra from bands analysed by mass spectrometry. Bands isolated from the partial proteolysis of FAM-Cys-GST with the proteases trypsin, V8 protease (Endoproteinase Glu-C) and chymotrypsin were cut out of the gel, treated with Iodoacetamide which forms carbamidomethyl-cysteine (Cys_CAM), and fully digested with trypsin. The resultant peptides were then analysed by mass spectrometry and the processed peptide mass data was analysed by Findmod (http://ca.expasy.org/cgi-bin/findmod_form.pl) and matched against the amino acid sequence of FAM-GST. The search took into account the likely carbamidomethyl modification to cysteine and oxidation of methionine (methionine sulfoxide, MSO). Matching peptides were found with an error of ± 50 ppm (parts per million) and allowed for up to 2 missed cleavages (#MC). Amino acid position is given in relation to FAM-GST (2804 amino acids). The position of the tightest clustering of peptides are given next to the name of the sample (*corresponds to bands in figure 4.13*). In some instances, the peptides cluster around several regions of FAM, all of which were plotted against the GST-FAM sequence in *Appendix B*.

Trypsin 1 (FAM: 566-869 aa)

Matching peptides:

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|---------------------|-----|--|-----------|-----------------------|
| 5558.656 | 5558.7333 | 13.906 | 1 | LAQDWFPLELLAMALNPCKFHIYNGTRPCESVSSSV QLPEDELFAR | 399-446 | (1xCys_CAM, 1xMSO) |
| 791.448 | 791.4410 | -8.844 | 0 | NLEIFR | 566-571 | |
| 1311.650 | 1311.6691 | 14.562 | 0 | VISSVSYTHR | 597-607 | |
| 1645.718 | 1645.7088 | -5.589 | 0 | HGSEDEEWLTAER | 608-621 | |
| 1827.938 | 1827.9599 | 11.981 | 0 | ALTQLDLNIWAAQAGK | 661-677 | |
| 992.490 | 992.5159 | 26.096 | 1 | ASWTNASKK | 710-718 | |
| 1444.650 | 1444.6736 | 16.336 | 1 | LAEDDKDGVMAHK | 730-742 | (MSO: 739) |
| 1284.700 | 1284.7310 | 24.13 | 1 | TNDKWVIPALK | 797-807 | |
| 826.523 | 826.5185 | -5.443 | 0 | WVIPALK | 801-807 | |
| 3393.603 | 3393.6678 | 19.095 | 0 | HDLINQLQHNHALVTLVAENLATYMESMR | 841-869 | (2xMSO) |
| 981.576 | 981.5839 | 8.048 | 1 | IKANVAHTK | 1164-1172 | |
| 1343.754 | 1343.7715 | 13.023 | 0 | SGGLPLVLSMLTR | 1361-1373 | |
| 1700.834 | 1700.8602 | 15.404 | 0 | NHQNLDSLEQYVK | 1987-2000 | |
| 1699.836 | 1699.7995 | -21.472 | 0 | FSEYLLECPSAEVR | 2341-2354 | (Cys_CAM: 2348) |
| 881.452 | 881.4662 | 16.11 | 1 | RAYQCIK | 2624-2630 | |

Trypsin 2 (FAM: 464-827 aa)

Matching peptides:

| User mass | DB mass | Δmass (ppm) | #MC | peptide | position | known modifications |
|------------------|----------------|--------------------------------------|------------|-------------------|-----------------|----------------------------|
| 1048.532 | 1048.5819 | 47.59 | 0 | LTQSMAIR | 65-73 | (MSO: 69) |
| 1048.532 | 1048.5092 | -21.744 | 0 | MNALNEVNK | 588-596 | (MSO: 588) |
| 981.547 | 981.5251 | -22.311 | 0 | DGLTISFTK | 358-366 | |
| 981.547 | 981.5839 | 37.594 | 1 | IKANVAHTK | 1164-1172 | |
| 1517.774 | 1517.7859 | 7.84 | 0 | FGTLNGFQLHDR | 464-476 | |
| 986.536 | 986.5153 | -20.982 | 0 | VPQEETVK | 557-565 | |
| 1311.640 | 1311.6691 | 22.186 | 0 | VISSVSYTHR | 597-607 | |
| 1645.708 | 1645.7088 | 0.486 | 0 | HGSSEDEEWLTAER | 608-621 | |
| 1343.638 | 1343.6590 | 15.629 | 0 | DSLHQPQYVEK | 638-648 | |
| 696.398 | 696.4039 | 8.472 | 0 | HEAIVK | 678-683 | |
| 864.444 | 864.4210 | -26.606 | 0 | ASWTNASK | 710-717 | |
| 992.482 | 992.5159 | 34.157 | 1 | ASWTNASKK | 710-718 | |
| 1428.664 | 1428.6787 | 10.289 | 1 | LAEDDKDGVMAHK | 730-742 | |
| 1444.662 | 1444.6736 | 8.03 | 1 | LAEDDKDGVMAHK | 730-742 | (MSO: 739) |
| 830.437 | 830.4519 | 17.942 | 0 | IQWIDR | 785-790 | |
| 806.441 | 806.4406 | -0.495 | 0 | FIEELR | 791-796 | |
| 1284.690 | 1284.7310 | 31.914 | 1 | TNDKWWIPALK | 797-807 | |
| 1894.916 | 1894.8963 | -10.395 | 0 | EICSLFGEAPQNLSSSR | 811-827 | (Cys_CAM: 813) |

Trypsin 3 (NA)

Matching peptides:

| User mass | DB mass | Δmass (ppm) | #MC | peptide | position | known modifications |
|------------------|----------------|--------------------------------------|------------|------------------------------|-----------------|----------------------------|
| 1235.610 | 1235.5659 | -35.69 | 1 | MFEDRLCHK | 132-140 | (1xCys_CAM) |
| 1709.934 | 1709.8652 | -40.234 | 1 | FEIHRCIINNTHR | 378-390 | (Cys_CAM: 383) |
| 1446.728 | 1446.7621 | 23.57 | 1 | NEAKNDALSMIIK | 536-548 | |
| 3354.735 | 3354.6173 | -35.084 | 2 | DFFESNVLQLDPSLLTENGMKCFERFFK | 959-986 | |
| 3354.735 | 3354.6092 | -37.498 | 2 | VLTVLREYINCESDSYHEERTILPMSR | 1094-1120 | (1xCys_CAM, 1xMSO) |
| 976.453 | 976.4709 | 18.332 | 1 | CFERFFK | 980-986 | |
| 1359.732 | 1359.7664 | 25.299 | 0 | SGGLPLVLSMLTR | 1361-1373 | (MSO: 1370) |
| 1312.728 | 1312.7735 | 34.661 | 0 | LLLTAIGYGHVR | 1398-1409 | |
| 1423.660 | 1423.6713 | 7.937 | 2 | NGGDGEKNRWYK | 2129-2140 | |
| 1682.884 | 1682.8206 | -37.672 | 0 | GSASDWYDALCILLR | 2306-2320 | |

Trypsin 4 (GST: 19-218, FAM: 328-390 aa)

Matching peptides:

| User mass | DB mass | Δmass (ppm) | #MC | peptide | position | known modifications |
|------------------|----------------|--------------------------------------|------------|-----------------------|-----------------|----------------------------|
| 2269.113 | 2269.1386 | 11.282 | 1 | LLLEYLEEKYEEHLYER | 19-35 | |
| 2269.128 | 2269.1386 | 4.671 | 1 | LLLEYLEEKYEEHLYER | 19-35 | |
| 2357.190 | 2357.2063 | 6.915 | 1 | KFELGLEFPNLPYYIDGDVK | 45-64 | |
| 2229.098 | 2229.1113 | 5.967 | 0 | FELGLEFPNLPYYIDGDVK | 46-64 | |
| 1314.654 | 1314.6041 | -37.956 | 1 | HNMLGGCPKER | 79-89 | (1xCys_CAM, 1xMSO) |
| 1314.654 | 1314.6940 | 30.426 | 1 | IAYSKDFETLK | 109-119 | |
| 1817.930 | 1817.9425 | 6.876 | 1 | ERAEISMLEGAVLDIR | 88-103 | (MSO: 94) |
| 1532.768 | 1532.7988 | 20.094 | 0 | AEISMLEGAVLDIR | 90-103 | (MSO: 94) |
| 1182.672 | 1182.6841 | 10.231 | 1 | RIEAIQIDK | 182-191 | |
| 2326.138 | 2326.1403 | 0.989 | 0 | YIAWPLQGQWQATFGGGDHPK | 198-218 | |
| 1599.830 | 1599.8839 | 33.691 | 0 | GELEVLEAAIDLSK | 328-342 | |
| 1652.822 | 1652.8438 | 13.19 | 1 | FEIHRCIINNTHR | 378-390 | |
| 1517.792 | 1517.7859 | -4.018 | 0 | FGTLNGFQLHDR | 464-476 | |

| | | | | | |
|----------|-----------|---------|---|-----------------------|---------------------------|
| 1583.842 | 1583.9115 | 43.881 | 2 | LLELIRRLAEDDK | 723-735 |
| 1399.708 | 1399.6409 | -47.938 | 0 | ASYDTLCVLDGDK | 1066-1078 |
| 1207.610 | 1207.6616 | 42.729 | 1 | TILPMSRAFR | 1114-1123 (MSO: 1118) |
| 2790.322 | 2790.2292 | -33.257 | 2 | FDYDWERECAIKFNDYFEFPR | 2041-2061 |
| 1637.784 | 1637.7686 | -9.402 | 1 | KIIEDCSNSDETVK | 2529-2542 (Cys_CAM: 2534) |

Trypsin 5 (FAM: 811-991 aa)

Matching peptides:

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|---------------------|-----|------------------------|-----------|---------------------|
| 1831.968 | 1831.8941 | -40.338 | 2 | YIADKHNMLGGCPKER | 74-89 | |
| 2571.322 | 2571.2256 | -37.489 | 1 | EICSLFGAPQNLSSRFQSQR | 811-833 | |
| 2571.322 | 2571.2404 | -31.734 | 0 | CMVALFSSCPVAYQILQNGDLK | 2631-2653 | (2xCys_CAM) |
| 2443.244 | 2443.3318 | 35.936 | 2 | FLLKDGQLWLCAPQAKQIWK | 904-923 | (Cys_CAM: 914) |
| 1924.910 | 1924.9408 | 16.001 | 1 | QIWKCLAENAVYLCDR | 920-935 | |
| 1576.770 | 1576.7511 | -11.986 | 2 | CFERFFKAVNCR | 980-991 | (1xCys_CAM) |
| 1576.770 | 1576.7060 | -40.588 | 0 | GDLLEGANAYHCEK | 2001-2014 | (Cys_CAM: 2012) |
| 1576.765 | 1576.7511 | -8.815 | 2 | CFERFFKAVNCR | 980-991 | (1xCys_CAM) |
| 1576.765 | 1576.7060 | -37.417 | 0 | GDLLEGANAYHCEK | 2001-2014 | (Cys_CAM: 2012) |
| 1186.567 | 1186.5276 | -33.204 | 1 | RFDYDWER | 2040-2047 | |
| 1630.684 | 1630.7280 | 26.983 | 0 | MQYSLEYFQFMK | 2244-2255 | (1xMSO) |

V8 Protease 1 (NA)

No peptides were observed in the mass spectrum

V8 Protease 2 (GST: 12-232, FAM: 1856-2075 aa)

Matching peptides:

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|---------------------|-----|----------------------|----------|---------------------|
| 770.446 | 770.4519 | 7.658 | 0 | GLVQPTR | 12-18 | |
| 2269.047 | 2269.1386 | 40.369 | 1 | LLLEYLEEKYEEHLYER | 19-35 | |
| 2269.204 | 2269.1386 | -28.82 | 1 | LLLEYLEEKYEEHLYER | 19-35 | |
| 2357.242 | 2357.2063 | -15.144 | 1 | KFELGLEFPNLPYYIDGDVK | 45-64 | |
| 2229.138 | 2229.1113 | -11.977 | 0 | FELGLEFPNLPYYIDGDVK | 46-64 | |
| 1032.609 | 1032.5870 | -21.304 | 0 | LTQSMAIIR | 65-73 | |
| 1032.609 | 1032.6200 | 10.653 | 1 | NLEIFRLK | 566-573 | |
| 1048.611 | 1048.5819 | -27.75 | 0 | LTQSMAIIR | 65-73 | (MSO: 69) |
| 1048.598 | 1048.5819 | -15.353 | 0 | LTQSMAIIR | 65-73 | (MSO: 69) |
| 1619.750 | 1619.7668 | 10.372 | 1 | YIADKHNMLGGCPK | 74-87 | (1xCys_CAM, 1xMSO) |
| 1801.956 | 1801.9476 | -4.661 | 1 | ERAEISMLEGAVLDIR | 88-103 | |
| 1817.926 | 1817.9425 | 9.076 | 1 | ERAEISMLEGAVLDIR | 88-103 | (MSO: 94) |
| 1532.764 | 1532.7988 | 22.704 | 0 | AEISMLEGAVLDIR | 90-103 | (MSO: 94) |
| 581.279 | 581.3042 | 43.353 | 0 | YGVSR | 104-108 | |
| 752.372 | 752.3825 | 13.956 | 0 | DFETLK | 114-119 | |
| 752.372 | 752.3686 | -4.518 | 0 | FSQSQR | 828-833 | |
| 708.369 | 708.3926 | 33.316 | 0 | VDFLSK | 120-125 | |
| 708.382 | 708.3926 | 14.964 | 0 | VDFLSK | 120-125 | |
| 730.397 | 730.4167 | 26.972 | 0 | LPEMLK | 126-131 | |
| 746.402 | 746.4117 | 12.996 | 0 | LPEMLK | 126-131 | (MSO: 129) |
| 713.283 | 713.2923 | 13.038 | 0 | MFEDR | 132-136 | (MSO: 132) |
| 666.335 | 666.3643 | 43.972 | 0 | LVCFK | 176-180 | (Cys_CAM: 178) |
| 1026.593 | 1026.5829 | -9.837 | 0 | IEAIPQIDK | 183-191 | |
| 2326.078 | 2326.1403 | 26.783 | 0 | YIAWPLQGWQATFGGGDHPK | 198-218 | |

| | | | | | | |
|----------|-----------|---------|---|--------------------|-----------|--------------------|
| 686.358 | 686.3831 | 36.57 | 0 | SDLVPR | 219-224 | |
| 990.517 | 990.5003 | -16.859 | 0 | HNQTSLYK | 225-232 | |
| 937.620 | 937.5869 | -35.301 | 0 | WVVPVLPK | 320-327 | |
| 1311.620 | 1311.6691 | 37.435 | 0 | VISSVSYTHR | 597-607 | |
| 1645.660 | 1645.7088 | 29.654 | 0 | HGSSEDEEWLTAER | 608-621 | |
| 1343.640 | 1343.6590 | 14.141 | 0 | DSLHQPQYVEK | 638-648 | |
| 1182.616 | 1182.5646 | -43.462 | 0 | YMPDICVIR | 1473-1481 | (1xCys_CAM, 1xMSO) |
| 2110.050 | 2110.0127 | -17.676 | 0 | DDVFGYPQQFEDKPPLSK | 1856-1873 | |
| 1448.774 | 1448.8120 | 26.229 | 0 | HLQVIFGHLAASR | 1887-1899 | |
| 1448.770 | 1448.8120 | 28.99 | 0 | HLQVIFGHLAASR | 1887-1899 | |
| 938.531 | 938.5094 | -23.014 | 0 | LQYYVPR | 1900-1906 | |
| 938.531 | 938.4876 | -46.241 | 1 | RAYQCIK | 2624-2630 | (Cys_CAM: 2628) |
| 1700.828 | 1700.8602 | 18.932 | 0 | NHQNLDSLEQYVK | 1987-2000 | |
| 1576.666 | 1576.7060 | 25.37 | 0 | GDLLEGANAYHCEK | 2001-2014 | (Cys_CAM: 2012) |
| 1538.730 | 1538.7406 | 6.889 | 0 | ELDMEPYTVAGVAK | 2062-2075 | (MSO: 2065) |

V8 Protease 3 (NA)

Matching peptides:

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|----------|---------------------|-----|----------|-----------|---------------------|
| 581.286 | 581.3042 | 31.31 | 0 | YGVSR | 104-108 | |
| 988.542 | 988.5356 | -6.473 | 1 | QEAVRMVR | 1086-1093 | |
| 901.492 | 901.5213 | 32.502 | 2 | RIRDDVK | 1650-1656 | |
| 901.492 | 901.5213 | 32.502 | 2 | IRDDVKR | 1651-1657 | |
| 728.355 | 728.3508 | -5.765 | 1 | RMSYR | 2177-2181 | (MSO: 2178) |
| 728.355 | 728.3508 | -5.765 | 1 | MSYRR | 2178-2182 | (MSO: 2178) |

V8 Protease 4 (GST: 12-232 aa)

Matching peptides:

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|---------------------|-----|----------------------|-----------|---------------------|
| 770.424 | 770.4519 | 36.214 | 0 | GLVQPTR | 12-18 | |
| 2269.041 | 2269.1386 | 43.014 | 1 | LLLEYLEEKYEELHYER | 19-35 | |
| 905.411 | 905.4111 | 0.11 | 1 | DEGDKWR | 36-42 | |
| 2357.196 | 2357.2063 | 4.37 | 1 | KFELGLEFPNLPYYIDGDVK | 45-64 | |
| 1032.579 | 1032.5870 | 7.748 | 0 | LTQSMAIIR | 65-73 | |
| 1032.579 | 1032.6200 | 39.706 | 1 | NLEIFRLK | 566-573 | |
| 1048.568 | 1048.5819 | 13.256 | 0 | LTQSMAIIR | 65-73 | (MSO: 69) |
| 1801.900 | 1801.9476 | 26.417 | 1 | ERAEISMLEGAVLDIR | 88-103 | |
| 1817.894 | 1817.9425 | 26.679 | 1 | ERAEISMLEGAVLDIR | 88-103 | (MSO: 94) |
| 1516.734 | 1516.8039 | 46.086 | 0 | AEISMLEGAVLDIR | 90-103 | |
| 1516.734 | 1516.8039 | 46.086 | 1 | LMPPDSTTIEKLR | 1287-1299 | (MSO: 1288) |
| 752.350 | 752.3825 | 43.198 | 0 | DFETLK | 114-119 | |
| 752.350 | 752.3686 | 24.723 | 0 | FSQSQR | 828-833 | |
| 708.360 | 708.3926 | 46.022 | 0 | VDFLSK | 120-125 | |
| 1026.589 | 1026.5829 | -5.941 | 0 | IEAIPQIDK | 183-191 | |
| 686.349 | 686.3831 | 49.683 | 0 | SDLVPR | 219-224 | |
| 990.502 | 990.5003 | -1.715 | 0 | HNQTSLYK | 225-232 | |
| 988.540 | 988.5356 | -4.45 | 1 | QEAVRMVR | 1086-1093 | |
| 1182.602 | 1182.5646 | -31.624 | 0 | YMPDICVIR | 1473-1481 | (1xCys_CAM, 1xMSO) |
| 901.490 | 901.5213 | 34.72 | 2 | RIRDDVK | 1650-1656 | |
| 901.490 | 901.5213 | 34.72 | 2 | IRDDVKR | 1651-1657 | |

Chymotrypsin 1 (GST: 12-224, FAM: 311-382, 540-1205, 1374-1686, 1796-2075 aa)**Matching peptides:**

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|---------------------|-----|----------------|-----------|---------------------|
| 770.444 | 770.4519 | 10.254 | 0 | GLVQPTR | 12-18 | |
| 1048.544 | 1048.5819 | 36.145 | 0 | LTQSMAIR | 65-73 | (MSO: 69) |
| 1048.544 | 1048.5092 | -33.188 | 0 | MNALNEVNK | 588-596 | (MSO: 588) |
| 752.388 | 752.3825 | -7.309 | 0 | DFETLK | 114-119 | |
| 752.388 | 752.3686 | -25.784 | 0 | FSQSQR | 828-833 | |
| 708.373 | 708.3926 | 27.669 | 0 | VDFLSK | 120-125 | |
| 697.327 | 697.2974 | -42.447 | 0 | MFEDR | 132-136 | |
| 557.311 | 557.2864 | -44.14 | 0 | LCHK | 137-140 | (Cys_CAM: 138) |
| 1182.660 | 1182.6841 | 20.378 | 1 | RIEAIQIDK | 182-191 | |
| 1026.568 | 1026.5829 | 14.514 | 0 | IEAIQIDK | 183-191 | |
| 686.384 | 686.3831 | -1.31 | 0 | SDLVPR | 219-224 | |
| 1129.555 | 1129.5782 | 20.539 | 0 | LDDMINRPR | 311-319 | |
| 1145.554 | 1145.5731 | 16.673 | 0 | LDDMINRPR | 311-319 | (MSO: 314) |
| 937.565 | 937.5869 | 23.358 | 0 | WVVPVLPK | 320-327 | |
| 1262.587 | 1262.6157 | 22.731 | 1 | GLDVKSEACQR | 344-354 | (Cys_CAM: 352) |
| 981.508 | 981.5251 | 17.422 | 0 | DGLTISFTK | 358-366 | |
| 1218.598 | 1218.6364 | 31.512 | 0 | ILTDEAVSGWK | 367-377 | |
| 1218.598 | 1218.5612 | -30.198 | 1 | EACFKWYSK | 936-944 | (Cys_CAM: 938) |
| 701.365 | 701.3729 | 11.264 | 0 | FEIHR | 378-382 | |
| 1004.536 | 1004.5445 | 8.462 | 0 | NDALSMIIK | 540-548 | |
| 1004.536 | 1004.5305 | -5.474 | 1 | QEAVRMVR | 1086-1093 | (MSO: 1091) |
| 1020.528 | 1020.5394 | 11.171 | 0 | NDALSMIIK | 540-548 | (MSO: 545) |
| 560.318 | 560.3151 | -5.175 | 0 | NLASR | 552-556 | |
| 560.318 | 560.3151 | -5.175 | 1 | QREK | 719-722 | |
| 986.534 | 986.5153 | -18.954 | 0 | VPGQEETVK | 557-565 | |
| 791.434 | 791.4410 | 8.845 | 0 | NLEIFR | 566-571 | |
| 532.320 | 532.3275 | 14.089 | 0 | MILR | 574-577 | |
| 1106.624 | 1106.6204 | -3.252 | 0 | LLQISSFNGK | 578-587 | |
| 1032.510 | 1032.5142 | 4.068 | 0 | MNALNEVNK | 588-596 | |
| 1311.640 | 1311.6691 | 22.186 | 0 | VISSVSYTHR | 597-607 | |
| 1645.695 | 1645.7088 | 8.386 | 0 | HGSSEDEEWLTAER | 608-621 | |
| 1343.631 | 1343.6590 | 20.839 | 0 | DSLHQPQYVEK | 638-648 | |
| 506.329 | 506.3337 | 9.283 | 0 | FVIK | 655-658 | |
| 696.407 | 696.4039 | -4.45 | 0 | HEAIVK | 678-683 | |
| 909.493 | 909.5152 | 24.409 | 0 | NVHDLAK | 684-691 | |
| 756.477 | 756.4978 | 27.496 | 0 | LLELIR | 723-728 | |
| 1256.536 | 1256.5575 | 17.111 | 0 | ILDYSCSQDR | 771-780 | (Cys_CAM: 776) |
| 830.449 | 830.4519 | 3.492 | 0 | IQWIDR | 785-790 | |
| 806.423 | 806.4406 | 21.825 | 0 | FIEELR | 791-796 | |
| 1284.708 | 1284.7310 | 17.903 | 1 | TNDKWWIPALK | 797-807 | |
| 905.471 | 905.4628 | -9.055 | 0 | SPHVFYR | 834-840 | |
| 905.471 | 905.4873 | 18.002 | 1 | DGARVLMK | 1279-1286 | (MSO: 1285) |
| 1351.547 | 1351.5760 | 21.457 | 0 | DNEDYDPQTVR | 874-884 | |
| 662.386 | 662.3984 | 18.72 | 0 | LNFLR | 899-903 | |
| 520.337 | 520.3493 | 23.639 | 0 | FLLK | 904-907 | |
| 1587.677 | 1587.7206 | 27.462 | 0 | LMGDEPDLPDINK | 945-958 | (MSO: 946) |
| 1344.717 | 1344.7416 | 18.294 | 2 | AVNCREGKLVAK | 987-998 | (Cys_CAM: 990) |
| 1062.551 | 1062.5578 | 6.4 | 0 | EIYTNLGR | 1036-1044 | |
| 1399.615 | 1399.6409 | 18.505 | 0 | ASYDTLCVLDGDK | 1066-1078 | |
| 700.449 | 700.4715 | 32.122 | 0 | VLTVLR | 1094-1099 | |
| 833.449 | 833.4549 | 7.079 | 0 | TILPMSR | 1114-1120 | (MSO: 1118) |

| | | | | | |
|----------|-----------|---------|---|---------------------|---------------------------|
| 871.509 | 871.5148 | 6.655 | 0 | HLSFIVR | 1126-1132 |
| 618.323 | 618.3392 | 26.2 | 0 | CILNR | 1159-1163 |
| 632.373 | 632.3977 | 39.059 | 0 | SLITAK | 1200-1205 |
| 1421.609 | 1421.6478 | 27.293 | 0 | NNFLPNADMETR | 1374-1385 |
| 1437.630 | 1437.6427 | 8.834 | 0 | NNFLPNADMETR | 1374-1385 (MSO: 1382) |
| 849.483 | 849.4828 | -0.234 | 0 | GAYLNALK | 1387-1394 |
| 1312.740 | 1312.7735 | 25.519 | 0 | LLLTAIGYGHVR | 1398-1409 |
| 1217.637 | 1217.6120 | -20.531 | 0 | LAQQISDEASR | 1462-1472 |
| 503.304 | 503.2936 | -20.662 | 1 | ERAK | 1608-1611 |
| 1407.671 | 1407.7379 | 47.525 | 1 | AKHSGDYFTLLR | 1610-1621 |
| 1970.038 | 1970.0076 | -15.43 | 0 | TGETGVEETILEGHLGVTK | 1658-1676 |
| 1303.691 | 1303.7256 | 26.54 | 1 | ELLAFQTPEKK | 1677-1687 |
| 620.372 | 620.3766 | 7.415 | 0 | GFVGLK | 1796-1801 |
| 963.521 | 963.5258 | 4.982 | 0 | EYNIGVLR | 1879-1886 |
| 1576.669 | 1576.7060 | 23.467 | 0 | GDLLEGANAYHCEK | 2001-2014 (Cys_CAM: 2012) |
| 1234.559 | 1234.5527 | -5.102 | 0 | FNDYFEFPR | 2053-2061 |
| 1538.715 | 1538.7406 | 16.637 | 0 | ELDMEPYTVAGVAK | 2062-2075 (MSO: 2065) |
| 1699.793 | 1699.7995 | 3.824 | 0 | FSEYLLECPSEAVR | 2341-2354 (Cys_CAM: 2348) |

Chymotrypsin 2 (FAM: 311-377, 540-907 aa)

Matching peptides:

| User mass | DB mass | Amass (ppm) | #MC | peptide | position | known modifications |
|------------------|----------------|--------------------|------------|------------------|-----------------|----------------------------|
| 905.441 | 905.4111 | -33.022 | 1 | DEGDKWR | 36-42 | |
| 905.441 | 905.4628 | 24.077 | 0 | SPHVFYR | 834-840 | |
| 500.277 | 500.2649 | -24.186 | 0 | LCHK | 137-140 | |
| 1145.544 | 1145.5731 | 25.403 | 0 | LDDMINRPR | 311-319 | (MSO: 314) |
| 1431.759 | 1431.7630 | 2.794 | 1 | FFRDGLTISFTK | 355-366 | |
| 981.515 | 981.5251 | 10.29 | 0 | DGLTISFTK | 358-366 | |
| 1218.626 | 1218.6364 | 8.534 | 0 | ILTDEAVSGWK | 367-377 | |
| 1020.517 | 1020.5394 | 21.95 | 0 | NDALSMIIK | 540-548 | (MSO: 545) |
| 560.319 | 560.3151 | -6.959 | 0 | NLASR | 552-556 | |
| 560.319 | 560.3151 | -6.959 | 1 | QREK | 719-722 | |
| 791.434 | 791.4410 | 8.845 | 0 | NLEIFR | 566-571 | |
| 1106.610 | 1106.6204 | 9.398 | 0 | LLQISSFNGK | 578-587 | |
| 1311.680 | 1311.6691 | -8.309 | 0 | VISSVSYTHR | 597-607 | |
| 1343.657 | 1343.6590 | 1.488 | 0 | DSLHQPPYVEK | 638-648 | |
| 506.327 | 506.3337 | 13.233 | 0 | FVIK | 655-658 | |
| 696.413 | 696.4039 | -13.066 | 0 | HEAIVK | 678-683 | |
| 909.504 | 909.5152 | 12.314 | 0 | NVHDLAK | 684-691 | |
| 756.497 | 756.4978 | 1.058 | 0 | LLELIR | 723-728 | |
| 1256.605 | 1256.5575 | -37.799 | 0 | ILDYSCSQDR | 771-780 | (Cys_CAM: 776) |
| 806.431 | 806.4406 | 11.904 | 0 | FIEELR | 791-796 | |
| 1351.565 | 1351.5760 | 8.139 | 0 | DNEDYDPQTVR | 874-884 | |
| 662.390 | 662.3984 | 12.681 | 0 | LNFLR | 899-903 | |
| 1887.977 | 1888.0149 | 20.074 | 1 | FLLKDGQLWLCAPQAK | 904-919 | (Cys_CAM: 914) |
| 1887.977 | 1888.0016 | 13.03 | 2 | LQYYVPRGFWKQFR | 1900-1913 | |
| 700.465 | 700.4715 | 9.28 | 0 | VLTVLR | 1094-1099 | |
| 833.451 | 833.4549 | 4.679 | 0 | TILPMSR | 1114-1120 | (MSO: 1118) |
| 1207.617 | 1207.6616 | 36.932 | 1 | TILPMSRAFR | 1114-1123 | (MSO: 1118) |
| 578.354 | 578.3409 | -22.649 | 1 | AFRGK | 1121-1125 | |
| 831.459 | 831.4617 | 3.247 | 1 | RCILNR | 1158-1163 | (Cys_CAM: 1159) |
| 1247.608 | 1247.6188 | 8.657 | 0 | LMPPDSTTIEK | 1287-1297 | (MSO: 1288) |
| 1312.711 | 1312.7735 | 47.611 | 0 | LLLTAIGYGHVR | 1398-1409 | |
| 1217.596 | 1217.6120 | 13.141 | 0 | LAQQISDEASR | 1462-1472 | |

| | | | | | |
|----------|-----------|---------|---|---------------|-----------------------|
| 503.299 | 503.2936 | -10.728 | 1 | ERAK | 1608-1611 |
| 537.261 | 537.2820 | 39.087 | 0 | GFWK | 1907-1910 |
| 1659.890 | 1659.8496 | -24.338 | 2 | SVRKQNVQFMHNR | 2231-2243 (MSO: 2240) |
| 1177.667 | 1177.6939 | 22.842 | 0 | LYSVVSQLIR | 2469-2478 |

Chymotrypsin 3 (FAM: 311-382, 540-1163 aa)

Matching peptides:

| User mass | DB mass | Δmass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|-------------|-----|-----------------|-----------|---------------------|
| 2024.922 | 2024.9096 | -6.123 | 2 | YEEHLYERDEGDKWR | 28-42 | |
| 697.305 | 697.2974 | -10.898 | 0 | MFEDR | 132-136 | |
| 500.285 | 500.2649 | -40.176 | 0 | LCHK | 137-140 | |
| 557.299 | 557.2864 | -22.608 | 0 | LCHK | 137-140 | (Cys_CAM: 138) |
| 557.299 | 557.2790 | -35.886 | 0 | SHSAR | 2701-2705 | |
| 1129.556 | 1129.5782 | 19.654 | 0 | LDDMINRPR | 311-319 | |
| 1145.556 | 1145.5731 | 14.927 | 0 | LDDMINRPR | 311-319 | (MSO: 314) |
| 1431.731 | 1431.7630 | 22.351 | 1 | FFRDGLTISFTK | 355-366 | |
| 981.509 | 981.5251 | 16.403 | 0 | DGLTISFTK | 358-366 | |
| 1218.605 | 1218.6364 | 25.767 | 0 | ILTDEAVSWGK | 367-377 | |
| 1218.605 | 1218.5612 | -35.942 | 1 | EACFKWYSK | 936-944 | (Cys_CAM: 938) |
| 701.352 | 701.3729 | 29.8 | 0 | FEIHR | 378-382 | |
| 917.533 | 917.5050 | -30.516 | 2 | KEAKNEAK | 532-539 | |
| 1020.546 | 1020.5394 | -6.466 | 0 | NDALSMIIK | 540-548 | (MSO: 545) |
| 560.322 | 560.3151 | -12.313 | 0 | NLASR | 552-556 | |
| 560.322 | 560.3151 | -12.313 | 1 | QREK | 719-722 | |
| 986.512 | 986.5153 | 3.345 | 0 | VPGQEETVK | 557-565 | |
| 791.438 | 791.4410 | 3.791 | 0 | NLEIFR | 566-571 | |
| 1106.618 | 1106.6204 | 2.169 | 0 | LLQISSFNGK | 578-587 | |
| 1048.504 | 1048.5092 | 4.959 | 0 | MNALNEVNK | 588-596 | (MSO: 588) |
| 1311.639 | 1311.6691 | 22.948 | 0 | VISSVSYTHR | 597-607 | |
| 1645.684 | 1645.7088 | 15.07 | 0 | HGSSEDEEWLTAER | 608-621 | |
| 1343.659 | 1343.6590 | 0 | 0 | DSLHQPQYVEK | 638-648 | |
| 506.331 | 506.3337 | 5.332 | 0 | FVIK | 655-658 | |
| 696.408 | 696.4039 | -5.886 | 0 | HEAIVK | 678-683 | |
| 909.496 | 909.5152 | 21.111 | 0 | NVHDLAK | 684-691 | |
| 756.502 | 756.4978 | -5.551 | 0 | LLELIR | 723-728 | |
| 846.395 | 846.4315 | 43.124 | 1 | RLAEDDK | 729-735 | |
| 1256.604 | 1256.5575 | -37.003 | 0 | ILDYSCSQDR | 771-780 | (Cys_CAM: 776) |
| 806.438 | 806.4406 | 3.224 | 0 | FIEELR | 791-796 | |
| 1284.717 | 1284.7310 | 10.897 | 1 | TNDKWWIPALK | 797-807 | |
| 542.260 | 542.2391 | -38.541 | 0 | MYGR | 870-873 | (MSO: 870) |
| 1351.553 | 1351.5760 | 17.017 | 0 | DNEDYDPQTVR | 874-884 | |
| 1274.609 | 1274.6124 | 2.667 | 0 | YSHVQEVQER | 889-898 | |
| 662.393 | 662.3984 | 8.152 | 0 | LNFLR | 899-903 | |
| 520.324 | 520.3493 | 48.624 | 0 | FLLK | 904-907 | |
| 1062.549 | 1062.5578 | 8.282 | 0 | EIYTNLGPR | 1036-1044 | |
| 1399.680 | 1399.6409 | -27.934 | 0 | ASYDTLCVLDGDK | 1066-1078 | |
| 774.445 | 774.4403 | -6.068 | 1 | RCILNR | 1158-1163 | |
| 831.499 | 831.4617 | -44.858 | 1 | RCILNR | 1158-1163 | (Cys_CAM: 1159) |
| 618.321 | 618.3392 | 29.435 | 0 | CILNR | 1159-1163 | |
| 675.335 | 675.3606 | 37.907 | 0 | CILNR | 1159-1163 | (Cys_CAM: 1159) |
| 1247.611 | 1247.6188 | 6.252 | 0 | LMPPDSTTIEK | 1287-1297 | (MSO: 1288) |
| 1437.588 | 1437.6427 | 38.05 | 0 | NNFLPNADMETR | 1374-1385 | (MSO: 1382) |
| 503.305 | 503.2936 | -22.649 | 1 | ERAK | 1608-1611 | |
| 537.308 | 537.2820 | -48.388 | 0 | GFWK | 1907-1910 | |

| | | | | | |
|----------|-----------|---------|---|----------------|---------------------------|
| 1538.715 | 1538.7406 | 16.637 | 0 | ELDMEPYTVAGVAK | 2062-2075 (MSO: 2065) |
| 1659.905 | 1659.8496 | -33.374 | 2 | SVRKQNVQFMHNR | 2231-2243 (MSO: 2240) |
| 1630.800 | 1630.7280 | -44.149 | 0 | MQYSLEYFQFMK | 2244-2255 (1xMSO) |
| 1699.793 | 1699.7995 | 3.824 | 0 | FSEYLLECPSAEVR | 2341-2354 (Cys_CAM: 2348) |
| 845.430 | 845.4376 | 8.99 | 1 | NHYQKR | 2619-2624 |

Chymotrypsin 4 (FAM: 311-382, 464-1205, 1287-1686 aa)

Matching peptides:

| User mass | DB mass | Amass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|-------------|-----|----------------|-----------|---------------------|
| 905.435 | 905.4111 | -26.395 | 1 | DEGDKWR | 36-42 | |
| 905.435 | 905.4628 | 30.703 | 0 | SPHVFYR | 834-840 | |
| 1032.580 | 1032.5870 | 6.779 | 0 | LTQSMAIR | 65-73 | |
| 1032.580 | 1032.6200 | 38.738 | 1 | NLEIFRLK | 566-573 | |
| 500.284 | 500.2649 | -38.177 | 0 | LCHK | 137-140 | |
| 557.312 | 557.2864 | -45.934 | 0 | LCHK | 137-140 | (Cys_CAM: 138) |
| 1129.548 | 1129.5782 | 26.736 | 0 | LDDMINRPR | 311-319 | |
| 1145.548 | 1145.5731 | 21.911 | 0 | LDDMINRPR | 311-319 | (MSO: 314) |
| 937.578 | 937.5869 | 9.493 | 0 | WVVPVLPK | 320-327 | |
| 1333.654 | 1333.6892 | 26.394 | 2 | KGLDVKSEACQR | 343-354 | |
| 1143.537 | 1143.5364 | -0.524 | 1 | SEACQRFFR | 349-357 | |
| 1431.750 | 1431.7630 | 9.08 | 1 | FFRDGLTISFTK | 355-366 | |
| 981.514 | 981.5251 | 11.309 | 0 | DGLTISFTK | 358-366 | |
| 1218.633 | 1218.6364 | 2.79 | 0 | ILTDEAVSGWK | 367-377 | |
| 701.368 | 701.3729 | 6.986 | 0 | FEIHR | 378-382 | |
| 1517.750 | 1517.7859 | 23.653 | 0 | FGTLNGFQILHDR | 464-476 | |
| 1020.535 | 1020.5394 | 4.311 | 0 | NDALSMIIK | 540-548 | (MSO: 545) |
| 560.323 | 560.3151 | -14.098 | 0 | NLASR | 552-556 | |
| 560.323 | 560.3151 | -14.098 | 1 | QREK | 719-722 | |
| 986.508 | 986.5153 | 7.4 | 0 | VPQEETVK | 557-565 | |
| 791.413 | 791.4410 | 35.38 | 0 | NLEIFR | 566-571 | |
| 1048.517 | 1048.5092 | -7.438 | 0 | MNALNEVNK | 588-596 | (MSO: 588) |
| 1311.665 | 1311.6691 | 3.126 | 0 | VISSVSYTHR | 597-607 | |
| 1645.708 | 1645.7088 | 0.486 | 0 | HGSSEDEEWLTAER | 608-621 | |
| 1343.665 | 1343.6590 | -4.464 | 0 | DSLHQPQYVEK | 638-648 | |
| 506.337 | 506.3337 | -6.516 | 0 | FVIK | 655-658 | |
| 696.413 | 696.4039 | -13.066 | 0 | HEAIVK | 678-683 | |
| 909.497 | 909.5152 | 20.011 | 0 | NVHDLAK | 684-691 | |
| 756.490 | 756.4978 | 10.311 | 0 | LLELIR | 723-728 | |
| 1728.716 | 1728.7857 | 40.319 | 1 | ILDYSCSQDRDTQK | 771-784 | (Cys_CAM: 776) |
| 662.394 | 662.3984 | 6.643 | 0 | LNFLR | 899-903 | |
| 520.325 | 520.3493 | 46.702 | 0 | FLLK | 904-907 | |
| 1332.649 | 1332.6576 | 6.453 | 0 | VVIQSNDIACR | 1018-1029 | |
| 1389.637 | 1389.6790 | 30.224 | 0 | VVIQSNDIACR | 1018-1029 | (Cys_CAM: 1028) |
| 1062.553 | 1062.5578 | 4.517 | 0 | EIYTNLGPR | 1036-1044 | |
| 1399.626 | 1399.6409 | 10.646 | 0 | ASYDTLCVLDGDK | 1066-1078 | |
| 700.460 | 700.4715 | 16.418 | 0 | VLTVLR | 1094-1099 | |
| 833.450 | 833.4549 | 5.879 | 0 | TILPMSR | 1114-1120 | (MSO: 1118) |
| 1207.622 | 1207.6616 | 32.792 | 1 | TILPMSRAFR | 1114-1123 | (MSO: 1118) |
| 578.357 | 578.3409 | -27.836 | 1 | AFRGK | 1121-1125 | |
| 871.503 | 871.5148 | 13.54 | 0 | HLSFIVR | 1126-1132 | |
| 831.490 | 831.4617 | -34.034 | 1 | RCILNR | 1158-1163 | (Cys_CAM: 1159) |
| 632.379 | 632.3977 | 29.571 | 0 | SLITAK | 1200-1205 | |
| 1231.601 | 1231.6238 | 18.512 | 0 | LMPPDSTTIEK | 1287-1297 | |
| 1196.696 | 1196.6568 | -32.756 | 1 | LRAICLDHAK | 1298-1307 | (Cys_CAM: 1302) |

| | | | | | |
|----------|-----------|--------|---|---------------------|-----------------------|
| 1421.629 | 1421.6478 | 13.224 | 0 | NNFLPNADMETR | 1374-1385 |
| 1437.613 | 1437.6427 | 20.659 | 0 | NNFLPNADMETR | 1374-1385 (MSO: 1382) |
| 1217.598 | 1217.6120 | 11.498 | 0 | LAQQISDEASR | 1462-1472 |
| 503.298 | 503.2936 | -8.741 | 1 | ERAK | 1608-1611 |
| 1969.971 | 1970.0076 | 18.579 | 0 | TGETGVEETILEGHLGVTK | 1658-1676 |
| 1303.688 | 1303.7256 | 28.841 | 1 | ELLAQTPEKK | 1677-1687 |
| 1538.744 | 1538.7406 | -2.209 | 0 | ELDMEPYTVAGVAK | 2062-2075 (MSO: 2065) |
| 1177.665 | 1177.6939 | 24.54 | 0 | LYSVVSQLIR | 2469-2478 |
| 845.420 | 845.4376 | 20.818 | 1 | NHYQKR | 2619-2624 |

Chymotrypsin 5 (FAM: 311-382, 464-1163, 1287-1686, 1879-2208, 2400-2793 aa)

Matching peptides:

| User mass | DB mass | Δmass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|-------------|-----|-------------------------|-----------|---------------------|
| 2617.277 | 2617.3411 | 24.491 | 2 | LTQSMAIIRYIADKHNMLGGCPK | 65-87 | (1xCys_CAM) |
| 2617.277 | 2617.2359 | -15.702 | 2 | QIWKCLAENAVYLCDREACFK | 920-940 | (2xCys_CAM) |
| 609.296 | 609.3242 | 46.283 | 0 | YIADK | 74-78 | |
| 1129.567 | 1129.5782 | 9.915 | 0 | LDDMINRPR | 311-319 | |
| 1145.562 | 1145.5731 | 9.69 | 0 | LDDMINRPR | 311-319 | (MSO: 314) |
| 937.567 | 937.5869 | 21.225 | 0 | WVVPVLPK | 320-327 | |
| 981.512 | 981.5251 | 13.347 | 0 | DGLTISFTK | 358-366 | |
| 1218.609 | 1218.6364 | 22.485 | 0 | ILTDEAVSGWK | 367-377 | |
| 1218.609 | 1218.5612 | -39.224 | 1 | EACFKWYSK | 936-944 | (Cys_CAM: 938) |
| 701.368 | 701.3729 | 6.986 | 0 | FEIHR | 378-382 | |
| 1517.744 | 1517.7859 | 27.607 | 0 | FGTLNGFQILHDR | 464-476 | |
| 1020.540 | 1020.5394 | -0.587 | 0 | NDALSMIIK | 540-548 | (MSO: 545) |
| 560.322 | 560.3151 | -12.313 | 0 | NLASR | 552-556 | |
| 560.322 | 560.3151 | -12.313 | 1 | QREK | 719-722 | |
| 986.516 | 986.5153 | -0.709 | 0 | VPGQEETVK | 557-565 | |
| 791.433 | 791.4410 | 10.108 | 0 | NLEIFR | 566-571 | |
| 532.309 | 532.3275 | 34.754 | 0 | MILR | 574-577 | |
| 1106.625 | 1106.6204 | -4.156 | 0 | LLQISSFNGK | 578-587 | |
| 1032.511 | 1032.5142 | 3.099 | 0 | MNALNEVNK | 588-596 | |
| 1048.507 | 1048.5092 | 2.098 | 0 | MNALNEVNK | 588-596 | (MSO: 588) |
| 1311.649 | 1311.6691 | 15.324 | 0 | VISSVSYTHR | 597-607 | |
| 1645.677 | 1645.7088 | 19.323 | 0 | HGSSEDEEWLTAER | 608-621 | |
| 1343.667 | 1343.6590 | -5.953 | 0 | DSLHQPYVEK | 638-648 | |
| 506.324 | 506.3337 | 19.158 | 0 | FVIK | 655-658 | |
| 506.324 | 506.3007 | -46.017 | 0 | VLMK | 1283-1286 | (MSO: 1285) |
| 696.408 | 696.4039 | -5.886 | 0 | HEAIVK | 678-683 | |
| 909.495 | 909.5152 | 22.21 | 0 | NVHDLAK | 684-691 | |
| 756.490 | 756.4978 | 10.311 | 0 | LLELIR | 723-728 | |
| 1199.512 | 1199.5361 | 20.092 | 0 | ILDYSCSQDR | 771-780 | |
| 1256.536 | 1256.5575 | 17.111 | 0 | ILDYSCSQDR | 771-780 | (Cys_CAM: 776) |
| 1671.746 | 1671.7642 | 10.887 | 1 | ILDYSCSQDRDTQK | 771-784 | |
| 806.431 | 806.4406 | 11.904 | 0 | FIEELR | 791-796 | |
| 1576.690 | 1576.7511 | 38.752 | 2 | CFERFFKAVNCR | 980-991 | (1xCys_CAM) |
| 1576.690 | 1576.7060 | 10.148 | 0 | GDLLEGANAYHCEK | 2001-2014 | (Cys_CAM: 2012) |
| 1633.829 | 1633.7726 | -34.519 | 2 | CFERFFKAVNCR | 980-991 | (2xCys_CAM) |
| 1062.526 | 1062.5578 | 29.929 | 0 | EIYTNLGPR | 1036-1044 | |
| 1399.635 | 1399.6409 | 4.215 | 0 | ASYDTLCVLDGDK | 1066-1078 | |
| 831.455 | 831.4617 | 8.058 | 1 | RCILNR | 1158-1163 | (Cys_CAM: 1159) |
| 1231.609 | 1231.6238 | 12.017 | 0 | LMPPDSTTIEK | 1287-1297 | |
| 1577.678 | 1577.7489 | 44.939 | 1 | NNFLPNADMETR | 1374-1386 | |
| 1217.568 | 1217.6120 | 36.138 | 0 | LAQQISDEASR | 1462-1472 | |

| | | | | | |
|----------|-----------|---------|---|------------------------|---------------------------|
| 503.297 | 503.2936 | -6.754 | 1 | ERAK | 1608-1611 |
| 1970.013 | 1970.0076 | -2.74 | 0 | TGETGVEETILEGHLGVTK | 1658-1676 |
| 1303.683 | 1303.7256 | 32.677 | 1 | ELLAFTQPEKK | 1677-1687 |
| 963.510 | 963.5258 | 16.398 | 0 | EYNIGVLR | 1879-1886 |
| 1448.804 | 1448.8120 | 5.522 | 0 | HLQVIFGHLAASR | 1887-1899 |
| 938.492 | 938.5094 | 18.54 | 0 | LQYYVPR | 1900-1906 |
| 938.492 | 938.4876 | -4.687 | 1 | RAYQCIK | 2624-2630 (Cys_CAM: 2628) |
| 537.272 | 537.2820 | 18.613 | 0 | GFWK | 1907-1910 |
| 1083.580 | 1083.5945 | 13.382 | 0 | LWGEPVNLK | 1914-1922 |
| 1024.535 | 1024.5608 | 25.182 | 0 | ALGHPAMLSK | 1944-1953 |
| 1040.534 | 1040.5557 | 20.855 | 0 | ALGHPAMLSK | 1944-1953 (MSO: 1950) |
| 1021.509 | 1021.5312 | 21.733 | 0 | VLGGSFADQK | 1954-1963 |
| 1700.872 | 1700.8602 | -6.937 | 0 | NHQNLDSLEQYVK | 1987-2000 |
| 1219.785 | 1219.8136 | 23.447 | 1 | KLPPVLAIQLK | 2029-2039 |
| 1186.521 | 1186.5276 | 5.562 | 1 | RFDYDWER | 2040-2047 |
| 1030.422 | 1030.4264 | 4.27 | 0 | FDYDWER | 2041-2047 |
| 1234.537 | 1234.5527 | 12.717 | 0 | FNDYFEFPR | 2053-2061 |
| 1522.718 | 1522.7457 | 18.191 | 0 | ELDMEPYTVAGVAK | 2062-2075 |
| 1538.708 | 1538.7406 | 21.187 | 0 | ELDMEPYTVAGVAK | 2062-2075 (MSO: 2065) |
| 2615.214 | 2615.2067 | -2.79 | 0 | LEGDNVNPESLIQQNEQSESEK | 2076-2098 |
| 1185.462 | 1185.4728 | 9.11 | 0 | FDDGDVTECK | 2141-2150 (Cys_CAM: 2149) |
| 1560.691 | 1560.7634 | 46.39 | 0 | WVNAYILFYER | 2186-2196 |
| 1400.627 | 1400.6474 | 14.565 | 0 | MDTIGHDDEVIR | 2197-2208 |
| 1416.630 | 1416.6423 | 8.683 | 0 | MDTIGHDDEVIR | 2197-2208 (MSO: 2197) |
| 798.513 | 798.5196 | 8.265 | 0 | AVLNLLR | 2400-2406 |
| 971.506 | 971.4832 | -23.468 | 0 | YQYAEK | 2461-2468 |
| 1177.669 | 1177.6939 | 21.143 | 0 | LYSVVSQLIR | 2469-2478 |
| 1717.822 | 1717.8503 | 16.474 | 1 | GIPDDRDLGFDTIQR | 2602-2616 |
| 1862.812 | 1862.8402 | 15.138 | 0 | AQENYEGGEEVSPQTK | 2777-2793 |

Chymotrypsin 6 (FAM: 311-382, 464-1205, 1287-1686, 1796-2075 aa)

Matching peptides:

| User mass | DB mass | Amass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|-------------|-----|----------------|-----------|---------------------|
| 1039.544 | 1039.5023 | -40.113 | 1 | KAGCMTATTR | 233-242 | |
| 1129.575 | 1129.5782 | 2.833 | 0 | LDDMINRPR | 311-319 | |
| 1145.525 | 1145.5731 | 41.989 | 0 | LDDMINRPR | 311-319 | (MSO: 314) |
| 981.507 | 981.5251 | 18.441 | 0 | DGLTISFTK | 358-366 | |
| 1218.625 | 1218.6364 | 9.355 | 0 | ILTDEAVSGWK | 367-377 | |
| 701.363 | 701.3729 | 14.115 | 0 | FEIHR | 378-382 | |
| 1517.713 | 1517.7859 | 48.033 | 0 | FGTLNGFQILHDR | 464-476 | |
| 1004.539 | 1004.5445 | 5.475 | 0 | NDALSMIIK | 540-548 | |
| 1004.539 | 1004.5305 | -8.461 | 1 | QEAVRMVR | 1086-1093 | (MSO: 1091) |
| 1020.532 | 1020.5394 | 7.251 | 0 | NDALSMIIK | 540-548 | (MSO: 545) |
| 560.321 | 560.3151 | -10.529 | 0 | NLASR | 552-556 | |
| 560.321 | 560.3151 | -10.529 | 1 | QREK | 719-722 | |
| 986.521 | 986.5153 | -5.777 | 0 | VPQEETVK | 557-565 | |
| 791.435 | 791.4410 | 7.581 | 0 | NLEIFR | 566-571 | |
| 532.323 | 532.3275 | 8.454 | 0 | MILR | 574-577 | |
| 1106.618 | 1106.6204 | 2.169 | 0 | LLQISSFNGK | 578-587 | |
| 1048.500 | 1048.5092 | 8.774 | 0 | MNALNEVNK | 588-596 | (MSO: 588) |
| 1311.634 | 1311.6691 | 26.761 | 0 | VISSVSYTHR | 597-607 | |
| 1645.705 | 1645.7088 | 2.309 | 0 | HGSSEDEEWLTAER | 608-621 | |
| 1343.626 | 1343.6590 | 24.56 | 0 | DSLHQPQYVEK | 638-648 | |
| 506.328 | 506.3337 | 11.258 | 0 | FVIK | 655-658 | |

| | | | | | | |
|----------|-----------|---------|---|---------------------|-----------|-----------------|
| 696.424 | 696.4039 | -28.861 | 0 | HEAIVK | 678-683 | |
| 909.495 | 909.5152 | 22.21 | 0 | NVHDLLAK | 684-691 | |
| 756.480 | 756.4978 | 23.53 | 0 | LLELIR | 723-728 | |
| 1256.534 | 1256.5575 | 18.702 | 0 | ILDYSCSQDR | 771-780 | (Cys_CAM: 776) |
| 830.448 | 830.4519 | 4.696 | 0 | IQWIDR | 785-790 | |
| 806.421 | 806.4406 | 24.305 | 0 | FIEELR | 791-796 | |
| 1284.719 | 1284.7310 | 9.341 | 1 | TNDKWVIPALK | 797-807 | |
| 1351.544 | 1351.5760 | 23.677 | 0 | DNEDYDPQTVR | 874-884 | |
| 662.368 | 662.3984 | 45.896 | 0 | LNFLR | 899-903 | |
| 520.335 | 520.3493 | 27.482 | 0 | FLLK | 904-907 | |
| 1587.686 | 1587.7206 | 21.793 | 0 | LMGDEPDLDPDINK | 945-958 | (MSO: 946) |
| 1633.800 | 1633.7726 | -16.77 | 2 | CFERFFKAVNCR | 980-991 | (2xCys_CAM) |
| 1344.719 | 1344.7416 | 16.806 | 2 | AVNCREGKLVAK | 987-998 | (Cys_CAM: 990) |
| 1062.550 | 1062.5578 | 7.341 | 0 | EIYTNLGPR | 1036-1044 | |
| 700.448 | 700.4715 | 33.55 | 0 | VLTVLR | 1094-1099 | |
| 833.450 | 833.4549 | 5.879 | 0 | TILPMSR | 1114-1120 | (MSO: 1118) |
| 1207.637 | 1207.6616 | 20.37 | 1 | TILPMSRAFR | 1114-1123 | (MSO: 1118) |
| 871.508 | 871.5148 | 7.803 | 0 | HLSFIVR | 1126-1132 | |
| 618.325 | 618.3392 | 22.965 | 0 | CILNR | 1159-1163 | |
| 632.375 | 632.3977 | 35.896 | 0 | SLITAK | 1200-1205 | |
| 1247.595 | 1247.6188 | 19.077 | 0 | LMPPDSTTIEK | 1287-1297 | (MSO: 1288) |
| 1312.721 | 1312.7735 | 39.993 | 0 | LLTAIGYGHVR | 1398-1409 | |
| 1217.589 | 1217.6120 | 18.89 | 0 | LAQQISDEASR | 1462-1472 | |
| 503.302 | 503.2936 | -16.689 | 1 | ERAK | 1608-1611 | |
| 1970.005 | 1970.0076 | 1.32 | 0 | TGETGVEETILEGHLGVTK | 1658-1676 | |
| 1303.679 | 1303.7256 | 35.745 | 1 | ELLAFTPEKK | 1677-1687 | |
| 620.367 | 620.3766 | 15.475 | 0 | GFVGLK | 1796-1801 | |
| 963.503 | 963.5258 | 23.664 | 0 | EYNIGVLR | 1879-1886 | |
| 1448.803 | 1448.8120 | 6.212 | 0 | HLQVIFGHLAASR | 1887-1899 | |
| 1021.483 | 1021.5312 | 47.186 | 0 | VLGGSFADQK | 1954-1963 | |
| 1700.888 | 1700.8602 | -16.343 | 0 | NHQNLDSLEQYVK | 1987-2000 | |
| 1576.662 | 1576.7060 | 27.907 | 0 | GDLLEGANAYHCEK | 2001-2014 | (Cys_CAM: 2012) |
| 1091.681 | 1091.7186 | 34.442 | 0 | LPPVLAQLK | 2030-2039 | |
| 1234.529 | 1234.5527 | 19.198 | 0 | FNDYFEFPR | 2053-2061 | |
| 1522.707 | 1522.7457 | 25.415 | 0 | ELDMEPYTVAGVAK | 2062-2075 | |
| 1538.715 | 1538.7406 | 16.637 | 0 | ELDMEPYTVAGVAK | 2062-2075 | (MSO: 2065) |
| 1699.783 | 1699.7995 | 9.707 | 0 | FSEYLLECPSAEVR | 2341-2354 | (Cys_CAM: 2348) |

Chymotrypsin 7 (GST: 79-191, FAM: 1287-1694 aa)

Matching peptides:

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|---------------------|-----|--------------|----------|---------------------|
| 956.443 | 956.4440 | 1.046 | 0 | HNMLGGCPK | 79-87 | |
| 752.373 | 752.3825 | 12.627 | 0 | DFETLK | 114-119 | |
| 752.373 | 752.3686 | -5.847 | 0 | FSQSQR | 828-833 | |
| 708.376 | 708.3926 | 23.434 | 0 | VDFLSK | 120-125 | |
| 730.387 | 730.4167 | 40.663 | 0 | LPEMLK | 126-131 | |
| 1408.647 | 1408.6963 | 34.998 | 1 | LPEMLKMFEDR | 126-136 | |
| 1026.533 | 1026.5829 | 48.61 | 0 | IEAIPQIDK | 183-191 | |
| 1431.743 | 1431.7630 | 13.969 | 1 | FFRDGLTISFTK | 355-366 | |
| 1218.627 | 1218.6364 | 7.714 | 0 | ILTDEAVSGWK | 367-377 | |
| 560.324 | 560.3151 | -15.883 | 0 | NLASR | 552-556 | |
| 560.324 | 560.3151 | -15.883 | 1 | QREK | 719-722 | |
| 1106.631 | 1106.6204 | -9.578 | 0 | LLQISSFNGK | 578-587 | |
| 696.410 | 696.4039 | -8.758 | 0 | HEAIVK | 678-683 | |

| | | | | | | |
|----------|-----------|---------|---|---------------|-----------|-----------------|
| 909.486 | 909.5152 | 32.106 | 0 | NVHDLLAK | 684-691 | |
| 846.418 | 846.4315 | 15.95 | 1 | RLAEDDK | 729-735 | |
| 773.385 | 773.3610 | -31.031 | 0 | DGVMAHK | 736-742 | (MSO: 739) |
| 520.325 | 520.3493 | 46.702 | 0 | FLLK | 904-907 | |
| 700.462 | 700.4715 | 13.562 | 0 | VLTVLR | 1094-1099 | |
| 833.437 | 833.4549 | 21.477 | 0 | TILPMSR | 1114-1120 | (MSO: 1118) |
| 833.437 | 833.3974 | -47.513 | 0 | FHIGCEK | 1688-1694 | |
| 831.440 | 831.4617 | 26.099 | 1 | RCILNR | 1158-1163 | (Cys_CAM: 1159) |
| 618.310 | 618.3392 | 47.226 | 0 | CILNR | 1159-1163 | |
| 1231.680 | 1231.6238 | -45.628 | 0 | LMPPDSTTIEK | 1287-1297 | |
| 1437.613 | 1437.6427 | 20.659 | 0 | NNFLPNADMETR | 1374-1385 | (MSO: 1382) |
| 1217.556 | 1217.6120 | 45.994 | 0 | LAQQISDEASR | 1462-1472 | |
| 503.298 | 503.2936 | -8.741 | 1 | ERAK | 1608-1611 | |
| 632.356 | 632.3362 | -31.31 | 1 | DDVKR | 1653-1657 | |
| 963.503 | 963.5258 | 23.664 | 0 | EYNIGVLR | 1879-1886 | |
| 1173.583 | 1173.5582 | -21.131 | 0 | QNVQFMHNR | 2235-2243 | |
| 1699.792 | 1699.7995 | 4.412 | 0 | FSEYLLECPSEVR | 2341-2354 | (Cys_CAM: 2348) |

Chymotrypsin 8 (GST: 79-180 aa)

Matching peptides:

| User mass | DB mass | Amass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|-------------|-----|---------------|-----------|---------------------|
| 1257.580 | 1257.5827 | 2.147 | 1 | HNMLGGCPKER | 79-89 | (1xMSO) |
| 752.353 | 752.3825 | 39.21 | 0 | DFETLK | 114-119 | |
| 752.353 | 752.3686 | 20.735 | 0 | FSQSQR | 828-833 | |
| 708.376 | 708.3926 | 23.434 | 0 | VDFLSK | 120-125 | |
| 666.350 | 666.3643 | 21.46 | 0 | LVCFK | 176-180 | (Cys_CAM: 178) |
| 560.320 | 560.3151 | -8.744 | 0 | NLASR | 552-556 | |
| 560.320 | 560.3151 | -8.744 | 1 | QREK | 719-722 | |
| 1106.615 | 1106.6204 | 4.88 | 0 | LLQISSFNGK | 578-587 | |
| 696.411 | 696.4039 | -10.194 | 0 | HEAIVK | 678-683 | |
| 846.444 | 846.4315 | -14.767 | 1 | RLAEDDK | 729-735 | |
| 1351.565 | 1351.5760 | 8.139 | 0 | DNEDYDPQTVR | 874-884 | |
| 662.398 | 662.3984 | 0.604 | 0 | LNFLR | 899-903 | |
| 520.328 | 520.3493 | 40.936 | 0 | FLLK | 904-907 | |
| 1062.566 | 1062.5578 | -7.716 | 0 | EIYTNLGPR | 1036-1044 | |
| 700.463 | 700.4715 | 12.135 | 0 | VLTVLR | 1094-1099 | |
| 833.456 | 833.4549 | -1.319 | 0 | TILPMSR | 1114-1120 | (MSO: 1118) |
| 1207.616 | 1207.6616 | 37.76 | 1 | TILPMSRAFR | 1114-1123 | (MSO: 1118) |
| 831.442 | 831.4617 | 23.694 | 1 | RCILNR | 1158-1163 | (Cys_CAM: 1159) |
| 1312.742 | 1312.7735 | 23.996 | 0 | LLLTAIGYGHVR | 1398-1409 | |
| 1217.598 | 1217.6120 | 11.498 | 0 | LAQQISDEASR | 1462-1472 | |
| 503.301 | 503.2936 | -14.702 | 1 | ERAK | 1608-1611 | |
| 632.359 | 632.3362 | -36.054 | 1 | DDVKR | 1653-1657 | |
| 1423.642 | 1423.6713 | 20.581 | 2 | NGGDGEKNRWYK | 2129-2140 | |
| 868.434 | 868.4570 | 26.484 | 2 | RMSYRR | 2177-2182 | |
| 1699.782 | 1699.7995 | 10.295 | 0 | FSEYLLECPSEVR | 2341-2354 | (Cys_CAM: 2348) |
| 798.490 | 798.5196 | 37.07 | 0 | AVLNLLR | 2400-2406 | |
| 1177.659 | 1177.6939 | 29.635 | 0 | LYSVVSQLIR | 2469-2478 | |

Chymotrypsin 9 (GST: 79-224, FAM: 311-382 aa)

Matching peptides:

| User mass | DB mass | Δ mass (ppm) | #MC | peptide | position | known modifications |
|-----------|-----------|---------------------|-----|--------------|-----------|---------------------|
| 1257.638 | 1257.5827 | -43.97 | 1 | HNMLGGCPKER | 79-89 | (1xMSO) |
| 1314.677 | 1314.6940 | 12.931 | 1 | IAYSKDFETLK | 109-119 | |
| 686.379 | 686.3831 | 5.973 | 0 | SDLVPR | 219-224 | |
| 1145.594 | 1145.5731 | -18.243 | 0 | LDDMINRPR | 311-319 | (MSO: 314) |
| 937.567 | 937.5869 | 21.225 | 0 | WVVPVLPK | 320-327 | |
| 1333.656 | 1333.6892 | 24.894 | 2 | KGLDVKSEACQR | 343-354 | |
| 1390.693 | 1390.7107 | 12.727 | 2 | KGLDVKSEACQR | 343-354 | (Cys_CAM: 352) |
| 1218.637 | 1218.6364 | -0.491 | 0 | ILTDEAVSGWK | 367-377 | |
| 701.375 | 701.3729 | -2.993 | 0 | FEIHR | 378-382 | |
| 560.316 | 560.3151 | -1.605 | 0 | NLASR | 552-556 | |
| 560.316 | 560.3151 | -1.605 | 1 | QREK | 719-722 | |
| 909.486 | 909.5152 | 32.106 | 0 | NVHDLAK | 684-691 | |
| 864.445 | 864.4210 | -27.762 | 0 | ASWTNASK | 710-717 | |
| 757.396 | 757.3661 | -39.476 | 0 | DGVMAHK | 736-742 | |
| 773.386 | 773.3610 | -32.324 | 0 | DGVMAHK | 736-742 | (MSO: 739) |
| 662.400 | 662.3984 | -2.414 | 0 | LNFLR | 899-903 | |
| 520.338 | 520.3493 | 21.717 | 0 | FLLK | 904-907 | |
| 700.482 | 700.4715 | -14.989 | 0 | VLTVLR | 1094-1099 | |
| 833.460 | 833.4549 | -6.118 | 0 | TILPMSR | 1114-1120 | (MSO: 1118) |
| 578.333 | 578.3409 | 13.66 | 1 | AFRGK | 1121-1125 | |
| 831.469 | 831.4617 | -8.779 | 1 | RCILNR | 1158-1163 | (Cys_CAM: 1159) |
| 859.476 | 859.5182 | 49.1 | 1 | CILNRIK | 1159-1165 | |
| 632.378 | 632.3977 | 31.152 | 0 | SLITAK | 1200-1205 | |
| 1247.637 | 1247.6188 | -14.587 | 0 | LMPPDSTTIEK | 1287-1297 | (MSO: 1288) |
| 1196.597 | 1196.6568 | 49.975 | 1 | LRAICLDHAK | 1298-1307 | (Cys_CAM: 1302) |
| 1217.618 | 1217.6120 | -4.927 | 0 | LAQQISDEASR | 1462-1472 | |
| 503.311 | 503.2936 | -34.57 | 1 | ERAK | 1608-1611 | |
| 1423.688 | 1423.6713 | -11.729 | 2 | NGGDGEKNRWYK | 2129-2140 | |

APPENDIX B

The occurrence of the matched peptides from the partial proteolysis products (Trypsin 1-5, V8 protease 2-4, Chymotrypsin 1-9) were manually plotted against a trypsin digested FAM-Cys-GST sequence (<http://ca.expasy.org/cgi-bin/peptide-mass.pl>). This table includes potential carbamidomethyl modification to cysteine (Cys_CAM) and oxidation of methionine (MSO) modifications but does not include peptides < 500 Da. Secondary structural elements are highlighted as predicted by the JUFO sever (www.jens-meiler.de). Residues that are likely to contribute to a *helix* are marked in blue while those that contribute to a *strand* are green. Note, secondary elements in peptides 234-242 and 2794-2802 result only from the extra amino acids inserted from the vector.

| Position | Artif. Mods | Peptide Sequence | Trypsin | | | | | V8 | | | | Chymotrypsin | | | | | | | | |
|----------|------------------------------------|--|---------|---|---|---|---|----|---|---|---|--------------|---|---|---|---|---|---|---|--|
| | | | 1 | 2 | 3 | 4 | 5 | 2 | 3 | 4 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | |
| 1-9 | MSO: 1 | MSPILGYWK | | | | | | | | | | | | | | | | | | |
| 12-18 | | GLVQPTR | | | | | | | X | | X | X | | | | | | | | |
| 19-27 | | ILLEYLEEK | | | | X | | X | | X | | | | | | | | | | |
| 28-35 | | YEEHLYER | | | X | | X | | X | | | | | X | | | | | | |
| 36-40 | | DEGDK | | | | | | | | | X | | X | | | | | | | |
| 46-64 | | FELGLEFPNLPYYIDGDVK | | | X | | X | | X | | | | | | | | | | | |
| 65-73 | MSO: 69 | LTQSMAIR | X | | | | X | | X | | X | | X | X | | | | | | |
| 74-78 | | YIADK | | | | | X | X | | | | | | | X | | | | | |
| 79-87 | Cys_CAM: 85 MSO: 81 | HNMLGGCPK | | X | X | X | | | | | | | | | X | X | X | X | | |
| 90-103 | MSO: 94 | AEISMLEGAVLDIR | | X | | X | | X | | | | | | | | | | | | |
| 104-108 | | YGVSR | | | | | | X | X | | | | | | | | | | | |
| 109-113 | | IAYSK | | X | | | | | | | | | | | | | | X | | |
| 114-119 | | DFETLK | | X | | X | | X | X | | | | | X | X | X | | | | |
| 120-125 | | VDFLSK | | | | | X | | X | X | | | | X | X | | | | | |
| 126-131 | MSO: 129 | EPPEMLK | | | | | X | | | | | | | X | | | | | | |
| 132-136 | MSO: 132 | MFEDR | | X | | | X | | | X | | | X | | | | X | | | |
| 137-140 | Cys_CAM: 138 | ICHK | | X | | | | | | X | X | X | | | | | | | | |
| 141-175 | Cys_CAM: 169 MSO: 154, 165, 168 | TYLNGDHVT HPDFMLYDALDVVLYMDP MCLDAF PK | | | | | | | | | | | | | | | | | | |
| 176-180 | Cys_CAM: 178 | LVCFK | | | | | | X | | | | | | | | | | X | | |
| 183-191 | | IEAIPQIDK | | | | | | X | | X | | | | | | X | | | | |
| 198-218 | | YIAWPLQGQWQATFGGGDHPPK | | X | | | | X | | | | | | | | | | | | |
| 219-224 | Cys_CAM: 236 | SDLVPR | | | | | | X | | X | | | | | | | | X | | |
| 225-232 | MSO: 237 | HNQTSLYK | | | | | | X | | X | | | | | | | | | | |
| 234-242 | Cys_CAM: 236 MSO: 237 | AGCMTATTR | | | | | | | | | | | | | | | | X | | |
| 243-310 | | GSPVGGNDNQQA PDGQSQPPLQQNTSSPDSSNENS PAT PPDEQQGDAPPQIEDEEPAPHTDLAK | | | | | | | | | | | | | | | | | | |
| 311-319 | MSO: 314 | LDDMNRPR | | | | | | | | X | X | X | X | X | X | | | X | | |
| 320-327 | | WVVPVLPK | | | | | | | X | | X | | X | X | | | | X | | |
| 328-342 | | GELEVLLEAAIDLSK | | X | | | | | | | | | | | | | | | | |
| 344-348 | | GLDVK | | | | | | | | X | | | X | | | | | X | | |

| | | | | | | | | | | | | | | | | | | | | | | |
|---------|--------------------------|--|--|---|---|--|--|--|--|--|--|--|---|---|---|---|---|---|---|---|---|---|
| 349-354 | Cys_CAM: 352 | <i>SEACQR</i> | | | | | | | | | | | X | | | | | | | | X | |
| 358-366 | | DGLTISFTK | | X | | | | | | | | | | | X | X | X | X | X | X | X | X |
| 367-377 | | <i>ILTDEAVSGWK</i> | | | | | | | | | | | | | | X | | | | | | X |
| 378-382 | | <i>FEIHR</i> | | X | X | | | | | | | | | | X | | X | X | X | X | X | X |
| 383-390 | Cys_CAM: 383 | <i>CIINNTHR</i> | | X | X | | | | | | | | | | | | | | | | | |
| 391-398 | Cys_CAM: 395 | <i>LVELCVAK</i> | | | | | | | | | | | | | | | | | | | | |
| 399-419 | Cys_CAM: 418 MSO: 412 | <i>LAQDWFPILLELAMALNPHCK</i> | | X | | | | | | | | | | | | | | | | | | |
| 420-446 | Cys_CAM: 429 | FHIYNGTRPCESVSSSVQLPEDELFAR | | X | | | | | | | | | | | | | | | | | | |
| 447-451 | | SPDPR | | | | | | | | | | | | | | | | | | | | |
| 455-463 | | <i>GWLVDLLNK</i> | | | | | | | | | | | | | | | | | | | | |
| 464-476 | | <i>FGTLNGFQILHDR</i> | | X | X | | | | | | | | | | | | | | | | | X |
| 477-508 | Cys_CAM: 498 | <i>FINGSALNVQIAALIKPFGQCYEFLTLHTVK</i> | | | | | | | | | | | | | | | | | | | | |
| 510-531 | MSO: 517 | <i>YFLPIEMVPQFLENLTDEELK</i> | | | | | | | | | | | | | | | | | | | | |
| 540-548 | MSO: 545 | <i>NDALSMIK</i> | | | X | | | | | | | | X | X | X | X | X | X | X | X | X | X |
| 552-556 | | NLASR | | | | | | | | | | | X | X | X | X | X | X | X | X | X | X |
| 557-565 | | <i>VPGQEETVK</i> | | X | | | | | | | | | X | | | | | | | | | |
| 566-571 | | <i>NLEIFR</i> | | X | | | | | | | | | X | | | | | | | | | |
| 574-577 | MSO: 574 | <i>MILR</i> | | | | | | | | | | | X | | | | | | | | | |
| 578-587 | | <i>LLQISSFNGK</i> | | | | | | | | | | | X | X | X | | | | | | | X |
| 588-596 | MSO: 588 | <i>MNALNEVNK</i> | | X | | | | | | | | | X | | | | | | | | | X |
| 597-607 | | VISSVSYJHR | | X | X | | | | | | | | X | | | | | | | | | X |
| 608-621 | | HGSSDEEWLTAER | | X | X | | | | | | | | X | | | | | | | | | X |
| 622-637 | MSO: 622 | <i>MAEWIQNNLSIVLR</i> | | | | | | | | | | | | | | | | | | | | |
| 638-648 | | <i>DSLHQPYVEK</i> | | X | | | | | | | | | X | X | X | X | X | X | X | X | X | X |
| 655-658 | | <i>FVIK</i> | | | | | | | | | | | X | X | X | X | X | X | X | X | X | X |
| 661-677 | | <i>ALTLDNIWAAQAGK</i> | | X | | | | | | | | | | | | | | | | | | |
| 678-683 | | <i>HEAIVK</i> | | X | | | | | | | | | X | X | X | X | X | X | X | X | X | X |
| 684-691 | | <i>NVHDLAK</i> | | | | | | | | | | | X | X | X | X | X | X | X | X | X | X |
| 692-709 | Cys_CAM: 707 | <i>LAWDFSPEQLDHLFDCFK</i> | | | | | | | | | | | | | | | | | | | | |
| 710-717 | | <i>ASWTNASK</i> | | X | X | | | | | | | | | | | | | | | | | X |
| 723-728 | | <i>LLELIR</i> | | | | | | | | | | | X | | | | | | | | | X |
| 730-735 | | <i>LAEDDK</i> | | X | X | | | | | | | | X | | | | | | | | | X |
| 736-742 | MSO: 739 | <i>DGVMAHK</i> | | X | X | | | | | | | | | | | | | | | | | X |
| 743-770 | MSO: 761 | <i>VLNLLWNLAHSDDVPVDMDLALSAHIK</i> | | | | | | | | | | | | | | | | | | | | |
| 771-780 | Cys_CAM: 776 | <i>ILDYSCSQDR</i> | | | | | | | | | | | | | | | | | | | | X |
| 785-790 | | <i>IQWIDR</i> | | X | | | | | | | | | X | | | | | | | | | X |
| 791-796 | | <i>FIEELR</i> | | X | | | | | | | | | X | X | X | | | | | | | X |
| 801-807 | | <i>WVIPALK</i> | | X | X | | | | | | | | X | | | | | | | | | X |
| 811-827 | Cys_CAM: 813 | <i>EICSLFGAPQNLSSSR</i> | | X | | | | | | | | | X | | | | | | | | | |
| 828-833 | | FSQSQR | | | | | | | | | | | X | X | X | | | | | | | X |
| 834-840 | | SPHVFYR | | | | | | | | | | | X | X | | | | | | | | X |
| 841-869 | MSO: 865, 868 | <i>HDLINQLQHNHALVTLVAENLATYMESMR</i> | | X | | | | | | | | | | | | | | | | | | |
| 870-873 | MSO: 870 | <i>MYGR</i> | | | | | | | | | | | | | | | | | | | | X |
| 874-884 | | <i>DNEDYDPQTVR</i> | | | | | | | | | | | X | X | X | | | | | | | X |
| 889-898 | | <i>YSHVQEVQER</i> | | | | | | | | | | | X | | | | | | | | | X |
| 899-903 | | <i>LNFLR</i> | | | | | | | | | | | X | X | X | X | | | | | | X |
| 904-907 | | <i>FLK</i> | | | | | | | | | | | X | X | X | X | | | | | | X |
| 908-919 | Cys_CAM: 914 | <i>DGQLWLCAPQAK</i> | | | | | | | | | | | X | | | | | | | | | X |
| 920-923 | | <i>QIWK</i> | | | | | | | | | | | X | | | | | | | | | X |
| 924-935 | Cys_CAM: 924, 933 | <i>CLAE NAVYLCDR</i> | | | | | | | | | | | X | | | | | | | | | X |
| 936-940 | Cys_CAM: 938 | <i>EACFK</i> | | | | | | | | | | | X | | X | | | | | | | X |
| 941-944 | | <i>WYSK</i> | | | | | | | | | | | X | | X | | | | | | | X |
| 945-958 | MSO: 946 | <i>LMGDEPLDPDINK</i> | | | | | | | | | | | X | | | | | | | | | X |
| 959-979 | MSO: 978 | <i>DFFESNVLQLDPSSLTENGMK</i> | | | | | | | | | | | X | | | | | | | | | X |
| 980-983 | Cys_CAM: 980 | <i>CFER</i> | | X | X | | | | | | | | | | | | | | | | | X |

| | | | | | | | | | | | | | |
|-----------|---|---|--|---|---|---|---|---|---|---|---|---|---|
| 987-991 | Cys_CAM: 990 | <i>AVNCR</i> | | X | | X | | | X | X | | | |
| 1001-1017 | MSO: 1003, 1004 | <i>AYMMDDLELIGLDYLWR</i> | | | | | | | | | | | |
| 1018-1029 | Cys_CAM: 1028 | <i>VVIQSNDDIACR</i> | | | | | | | | X | | | |
| 1030-1035 | | <i>AIDLLK</i> | | | | | | | | | | | |
| 1036-1044 | | <i>EIYTNLGPR</i> | | | | X | | X | X | X | X | | X |
| 1045-1063 | Cys_CAM: 1060 | <i>LQVNQVVIHEDFIQSCFDR</i> | | | | | | | | | | | |
| 1066-1078 | Cys_CAM: 1072 | <i>ASYDTLCVLDGDK</i> | | X | | X | | X | X | X | | | |
| 1079-1085 | Cys_CAM: 1083 | <i>DSINCAR</i> | | | | | | | | | | | |
| 1086-1090 | | <i>QEAVR</i> | | | X | X | | | | | | X | |
| 1094-1099 | | <i>VLTVLR</i> | | X | | X | X | X | X | X | X | X | X |
| 1100-1113 | Cys_CAM: 1105 | <i>EYINECDSDYHEER</i> | | X | | | | | | | | | |
| 1114-1120 | MSO: 1118 | <i>TILPMSR</i> | | X | X | | | X | | | X | X | X |
| 1126-1132 | | <i>HLSFIVR</i> | | | | X | | X | | X | | | |
| 1133-1138 | | <i>FPNQGR</i> | | | | | | | | | | | |
| 1139-1157 | | <i>QVDDLEVWSHTNDTIGSVR</i> | | | | | | | | | | | |
| 1159-1163 | Cys_CAM: 1159 | <i>CILNR</i> | | | | X | X | X | X | X | X | X | X |
| 1166-1172 | | <i>ANVAHTK</i> | | X | X | | | | | | | | |
| 1173-1188 | | <i>IELFVGGELIDPGDDR</i> | | | | | | | | | | | |
| 1190-1197 | | <i>LIGQLNLK</i> | | | | | | | | | | | |
| 1200-1205 | | <i>SLITAK</i> | | | | | | X | | X | | | X |
| 1206-1257 | Cys_CAM: 1246 MSO: 1213, 1252 | <i>LTQISSNPSSPDSSSDSSTGSPGNHGNHYSKGPNPEVE SCLPGVIMSLHPR</i> | | | | | | | | | | | |
| 1258-1278 | MSO: 1274 | <i>YISFLWQVADLGSSLNMPPLR</i> | | | | | | | | | | | |
| 1287-1297 | MSO: 1288 | <i>LMPPDSTIEK</i> | | | | | | X | X | X | X | X | X |
| 1300-1307 | Cys_CAM: 1302 | <i>AICLDHAK</i> | | | | | | | | X | | | X |
| 1308-1360 | MSO: 1341 | <i>LGESSLSPSLDSLFFGPSASQVLYLTVVYALLMPAGAPL TDDSDFQFHLK</i> | | | | | | | | | | | |
| 1361-1373 | MSO: 1370 | <i>SGGLPLVLSMLTR</i> | | X | X | | | | | | | | |
| 1374-1385 | MSO: 1382 | <i>NNFLPNADMETR</i> | | | | | | X | | X | X | X | X |
| 1387-1394 | | <i>GAYLNALK</i> | | | | | | X | | | | | |
| 1398-1409 | | <i>LLLTA/GYGHVR</i> | | X | | | | X | X | | | X | X |
| 1410-1456 | Cys_CAM: 1415, 1453 MSO: 1425, 1454 | <i>AVAEACQPGVEGVNPMTSVNQVTHDQAVVLQ^SSALQ^SSIP NPSS^EECMLR</i> | | | | | | | | | | | |
| 1457-1461 | | <i>NVSVR</i> | | | | | | | | | | | |
| 1462-1472 | | <i>LAQQ^SISDEASR</i> | | | | | | X | X | X | X | X | X |
| 1473-1481 | Cys_CAM: 1478 MSO: 1474 | <i>YMPDICVIR</i> | | | | | | X | X | | | | |
| 1486-1507 | Cys_CAM: 1492 | <i>IHWTSGCGGLQLVFS^NNE^EVTK</i> | | | | | | | | | | | |
| 1508-1511 | | <i>IYEK</i> | | | | | | | | | | | |
| 1512-1549 | Cys_CAM: 1526, 1527, 1536 MSO: 1533 | <i>TNAGNEPDLE^EQVCC^EALEVMTLCFAL^IPTALD^ALSK</i> | | | | | | | | | | | |
| 1552-1567 | Cys_CAM: 1564 | <i>AWQTFH^IDLLL^HCHSK</i> | | | | | | | | | | | |
| 1571-1583 | Cys_CAM: 1581 MSO: 1580 | <i>QVAQE^QF^LM^CTR</i> | | | | | | | | | | | |
| 1584-1607 | Cys_CAM: 1584, 1585 MSO: 1586 | <i>CCMGHRPL^LFFITLL^LFTVL^GSTAR</i> | | | | | | | | | | | |
| 1612-1621 | | <i>HSGDYFTLLR</i> | | | | | | X | | | | | |
| 1622-1649 | | <i>HLLNYAYNSNINVPNAEVLL^NNE^IDWLK</i> | | | | | | | | | | | |
| 1658-1676 | | <i>TGETGVE^ETH^EGHLGV^TK</i> | | | | | | X | | X | X | X | |
| 1677-1686 | | <i>ELLAFQ^TPEK</i> | | | | | | X | | X | X | X | |
| 1688-1694 | Cys_CAM: 1692 | <i>FHIGCEK</i> | | | | | | | | | | | X |
| 1695-1701 | | <i>GGANLIK</i> | | | | | | | | | | | |
| 1702-1720 | MSO: 1719 | <i>ELIDDFIFPASNVYL^QYMR</i> | | | | | | | | | | | |
| 1721-1755 | Cys_CAM: 1733, 1753 | <i>NGELPAEQAI^PVC^GS^PATINAG^FFELLVAL^AVGC^VR</i> | | | | | | | | | | | |
| 1759-1795 | Cys_CAM: 1777 MSO: 1767 | <i>QIVD^SL^TEM^YY^IGTAITTC^EALTEWEYL^PPPV^GPR^PPK</i> | | | | | | | | | | | |
| 1796-1801 | | <i>GFVGLK</i> | | | | | | X | | | | X | |
| 1802-1823 | MSO: 1809, 1818 | <i>NAGAT^SY^MN^SV^IQ^LL^SMIP^SIR</i> | | | | | | | | | | | |

| | | | | | | | | | |
|-----------|---|---|---|---|---|---|---|---|---|
| 1824-1845 | MSO: 1840 | <i>NGILAIEGTGSDVDDDDMSGDEK</i> | | | | | | | |
| 1846-1855 | | <i>QDNESNVDP</i> | | | | | | | |
| 1856-1873 | | <i>DDVFGYPQQFEDKPLSK</i> | | X | | | | | |
| 1874-1877 | | <i>TEDR</i> | | | | | | | |
| 1879-1886 | | <i>EYNIGVLR</i> | | | X | | X | X | X |
| 1887-1899 | | <i>HLQVIFGHLAASR</i> | | X | | | X | X | |
| 1900-1906 | | <i>LQYYVPR</i> | | X | | X | | | |
| 1907-1910 | | <i>GFWK</i> | | | | X | X | | |
| 1914-1922 | | <i>LWGEPVNL</i> | | | | | | X | |
| 1923-1943 | | <i>EQHDALEFFNSLVDSLDEALK</i> | | | | | | | |
| 1944-1953 | MSO: 1950 | <i>ALGHPAMLSK</i> | | | | | | X | |
| 1954-1963 | | <i>VLGGSFADQK</i> | | | | | | X | X |
| 1964-1971 | Cys_CAM: 1965, 1968 | <i>ICQGCPHR</i> | | | | | | | |
| 1972-1986 | Cys_CAM: 1974 | <i>YECEESFTTLNVDIR</i> | | | | | | | |
| 1987-2000 | | <i>NHQNLDSLEQYVK</i> | X | X | | | | X | X |
| 2001-2014 | Cys_CAM: 2012 | <i>GDLLEGANAYHCEK</i> | | X | X | X | | X | X |
| 2019-2023 | | <i>VDTVK</i> | | | | | | | |
| 2030-2039 | | <i>LPPVLAIQLK</i> | | | | | | X | X |
| 2041-2047 | | <i>FDYDWER</i> | | X | X | | | X | |
| 2048-2052 | Cys_CAM: 2049 | <i>ECAK</i> | | X | | | | | |
| 2053-2061 | | <i>FNDYFEFPR</i> | | X | | X | | X | X |
| 2062-2075 | MSO: 2065 | <i>ELDMEPYTVAGVAK</i> | | X | | X | X | X | X |
| 2076-2098 | | <i>LEGDNVNPESQLIQNEQSESEK</i> | | | | | | X | |
| 2106-2128 | | <i>LVGVLVHSGQASGGHYYSIIQR</i> | | | | | | | |
| 2129-2135 | | <i>NGGDGEK</i> | | X | | | | | X |
| 2141-2150 | Cys_CAM: 2149 | <i>FDDGDVTECK</i> | | | | | | X | |
| 2151-2158 | MSO: 2151, 2157 | <i>MDDEEMK</i> | | | | | | | |
| 2159-2176 | Cys_CAM: 2161 MSO: 2167, 2174, 2175 | <i>NQCFGGEYMGEVFDHMMK</i> | | | | | | | |
| 2178-2181 | MSO: 2178 | <i>MSYR</i> | | | X | | | | X |
| 2186-2196 | | <i>WWNAYILFYER</i> | | | | | | X | |
| 2197-2208 | MSO: 2197 | <i>MDTIGHDDEVIR</i> | | | | | | X | |
| 2209-2230 | MSO: 2224 | <i>YISEIAITTRPHQIVMPSAER</i> | | | | | | | |
| 2235-2243 | MSO: 2240 | <i>QNVQFMHNR</i> | | | | X | X | | X |
| 2244-2255 | MSO: 2244, 2254 | <i>MQYSLEYFQFMK</i> | | X | | X | | | |
| 2257-2291 | Cys_CAM: 2260 MSO: 2283 | <i>LLTCNGVYLNPPGQDHLSPAEAEITMISIQLAAR</i> | | | | | | | |
| 2292-2301 | | <i>FLFTIGFHTK</i> | | | | | | | |
| 2306-2320 | Cys_CAM: 2316 | <i>GSASDWYDALCILLR</i> | | X | | | | | |
| 2327-2340 | | <i>FWFAHNVLFNVSNR</i> | | | | | | | |
| 2341-2354 | Cys_CAM: 2348 | <i>FSEYLLECPSAEVR</i> | X | | | X | X | | X |
| 2360-2399 | Cys_CAM: 2374 | <i>LIVFIAHFSLQDGPCSPFASPGSSQAYDNLSLSDHLLR</i> | | | | | | | |
| 2400-2406 | | <i>AVLNLLR</i> | | | | | | X | X |
| 2408-2414 | | <i>EVSEHGR</i> | | | | | | | |
| 2415-2434 | MSO: 2425 | <i>HLQQYFNLVVMYANLGVAEK</i> | | | | | | | |
| 2435-2439 | | <i>TQLK</i> | | | | | | | |
| 2440-2460 | MSO: 2447 | <i>LSVPATFMLVSLDEGPGPIK</i> | | | | | | | |
| 2461-2468 | | <i>YQYAEIGK</i> | | | | | | X | |
| 2469-2478 | | <i>LYSVVSQLIR</i> | | | | X | X | X | X |
| 2479-2485 | Cys_CAM: 2479, 2480 | <i>CCNVSSR</i> | | | | | | | |
| 2486-2523 | MSO: 2486, 2510 | <i>MQSSINGNPSLPNPFDPNLSQPIMPIQQNVVDILFVR</i> | | | | | | | |
| 2524-2528 | | <i>TSYVK</i> | | | | | | | |
| 2530-2542 | Cys_CAM: 2534 | <i>IIDCSNSDETVK</i> | | X | | | | | |
| 2546-2595 | Cys_CAM: 2547, 2548 | <i>FCCWENPQFSSTVLSSELLWQVAYSYYELRPYLDLLQLL LIEDSWQTHR</i> | | | | | | | |
| 2596-2601 | | <i>IHNALK</i> | | | | | | | |

| | | | | | | |
|-----------|----------------------------------|---|---|---|-----|---|
| 2602-2607 | | GIPDDR | | | | X |
| 2608-2616 | | DGLFDTIQR | | | | X |
| 2619-2623 | | NHYQK | | | X X | |
| 2625-2630 | Cys_CAM: 2628 | AYQCIK | X | X | | X |
| 2631-2653 | Cys_CAM: 2631, 2639 MSO: 2632 | CMVALFSSCPVAYQILQGNGDLK | | X | | |
| 2656-2669 | | WTWAVEWLGDELER | | | | |
| 2670-2700 | | RPYTGPNQYTYNNWSPPVQSNETSNGYFLER | | | | |
| 2701-2705 | | SHSAR | | | | X |
| 2706-2710 | MSO: 2706 | MTLAK | | | | |
| 2711-2772 | Cys_CAM: 2712, 2715 MSO: 2767 | ACELCPREEEPDDQDAPDEHESPPPEDAPLYPHSPGSQY QQ NNHVHGQPYTGPAAHMNNPQR | | | | |
| 2777-2793 | | AQENYEGGEEVSPPTK | | | | X |
| 2794-2802 | | DQYPAFLYK | | | | |

REFERENCES

- Amerik AY, Hockstrasser M (2004) Mechanism and Function of Deubiquitinating Enzymes. *Biochim. Biophys. Acta.* 1695:189-207
- Amrein KE, Takacs B, Stieger M, Molnos J, Flint NA, Burn P (1995) Purification and characterization of recombinant human p50csk protein-tyrosine kinase from an *Escherichia coli* expression system overproducing the bacterial chaperones GroES and GroEL. *Proc. Natl. Acad. Sci. USA.* 92:1048-1052
- Baek KH (2003) Conjugation and deconjugation of ubiquitin regulating the destiny of proteins. *Exp. Mol. Med.* 35:1-7
- Baker RT, Tobias JT, Varshavsky A (1992) Ubiquitin-specific proteases of *Saccharomyces cerevisiae*. Cloning of UBP2 and UBP3, and functional analysis of the UBP gene family. *J. Biol. Chem* 267:23364-23375
- Baker RT, Wang X-W, Woollatt E, White JA, Sutherland GR (1999) Identification, functional characterization, and chromosomal localization of USP15, a novel human ubiquitin-specific protease related to the UNP oncoprotein, and a systematic nomenclature for human ubiquitin-specific proteases. *Genomics* 59:264-274
- Blanchette P, Gilchrist CA, Baker RT, Gray DA (2001) Association of UNP, a ubiquitin-specific protease, with the pocket proteins pRb, p107 and p130. *Oncogene* 20:5538-5542
- Brendel V, Bucher P, Nourbakhsh I, Blaisdell BE, Karlin S. (1992) Methods and algorithms for statistical analysis of protein sequences. *Proc. Natl. Acad. Sci. USA* 89:2002-2006

Brown GM, Furlong RA, Sargent CA, Erickson RP, Longepied G, Mitchell M, Jones MH, Hargreave TB, Cooke HJ, Affara NA (1998) Characterisation of the coding sequence and fine mapping of the human *DFFRY* gene and comparative expression analysis and mapping to the *Sxr^b* interval of the mouse Y chromosome of the *Dffry* gene. *Hum. Mol. Genet.* 7:97-107

Brown MH, Furlong RA, Burkin H, Chalmers IJ, Brown GM, Khwaja O, Affara NA (1996) The Drosophila developmental gene *fat facets* has a human homologue in Xp11.4 which escapes X-inactivation and has related sequences on Yq11.2. *Hum. Mol. Genet.* 5:1696-1701

Burridge K, Nuckolls G, Otey C, Pavalko FM, Simon K, Turner C (1990). Actin-membrane interaction in focal adhesions. *Cell Differ. Dev.* 32:337-342

Cadavid, AL, Ginzel A, Fischer JA (2000) The function of the Drosophila *Fat facets* deubiquitinating enzyme in limiting photoreceptor cell number is intimately associated with endocytosis. *Development* 127:1727-1736

Chen H, Polo S, Di Fiore PP, De Camilli PV (2003) Rapid Ca²⁺-dependent decrease of protein ubiquitination at synapses. *Proc. Natl. Acad. Sci. USA* 101:14908-14913

Chen X, Fischer JA (2000) In vivo structure/function analysis of the Drosophila *fat facets* deubiquitinating enzyme gene. *Genetics* 156:1829-1836

Chen X, Overstreet E, Wood SA, Fischer JA (2000) On the conservation of function of the Drosophila *Fat facets* deubiquitinating enzyme and *Fam*, its mouse homolog. *Development, Genes and Evolution* 210:603-610

Chen X, Zhang B, Fischer JA (2002) A specific protein substrate for a deubiquitinating enzyme: liquid facets is the substrate of fat facets. *Genes, Dev.* 16:289-294

Chesshyre JA, Hipkiss AR (1989) Low temperatures stabilize interferon α -2 against proteolysis in *Methylophilus methylotrophus* and *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 31:158-162

Ciechanover A, Orian A, Schwartz AL (2000) Ubiquitin-mediated proteolysis: Biological Regulation via Destruction. *BioEssays* 22:442-451

Cleland WW (1964) Dithiothreitol, a new protective reagent for SH groups, *Biochemistry* 3:480-2

Cole LK, Ross LS (2001) Apoptosis in the developing zebrafish embryo. *Dev. Biol.* 240:123-142

Collins BM, McCoy AJ, Kent HM, Evans PR, Owen DJ (2002) Molecular architecture and functional model of the endocytic AP2 complex. *Cell* 109:523-535

Coux O, Tanaka K, Goldberg AL (1996) Structure and functions of the 20S and 26S proteasomes. *Annu. Rev. Biochem.* 65:801-847

Daim S (1999) Molecular characterisation of the ubiquitin-specific protease. *Fam University of Adelaide Honours Thesis*

Desalle LM, Latres E, Lin D, Graner E, Montagnoli A, Baker RT, Pagano M, Loda M (2001) The de-ubiquitinating enzyme Unp interacts with the retinoblastoma protein. *Oncogene* 20:5533-5537

Deshaies RJ (1999) SCF and Cullin/Ring H2-based ubiquitin ligases. *Annu. Rev. Cell Dev. Biol.* 15:435-467

Deveraux Q, Ustrell V, Pickart C, Reichsteiner M (1994) A 26S protease subunit that binds polyubiquitin conjugates. *J. Biol. Chem.* 269:7059-7061

Evans PC, Smith TS, Lai MJ, Williams MG, Burke DF, Heyninck K, Kreike MM, Beyaert R, Blundell TL, Kilshaw PJ (2003) A Novel Type of Deubiquitinating Enzyme. *J. Biol. Chem.* 278:23180-23186

Fahrner RL, Cascio D, Lake JA, Slesarev A (2001) An Ancestral Nuclear Protein Assembly: Crystal Structure of the Methanopyrus Kandleri Histone. *Protein Sci.* 10: 2002-2007

Felsenstein J (1989) PHYLIP -- Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164-166

Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN, Timmis KN (2002) Chaperonins govern growth of *Escherichia coli* at low temperature. *Nat. Biotechnol.* 21:1266-1267

Fischer-Vize JA, Rubin G, Lehmann R (1992) The *fat facets* gene is required for *Drosophila* eye and embryo development. *Development* 116:985-1000

Fontana A, Polverino de Laureto P, De Filippis V, Scaramella E, Zambonin M (1997) Probing the partly folded states of proteins by limited proteolysis. *Folding Des.* 2:R17-26

Fontana A, Polverino de Laureto P, Spolaore B, Frare E, Picotti P, Zambonin M (2004) Probing protein structure by limited proteolysis *Acta. Biochim. Pol.* 51:299-321

Friocourt G, Chafey P, Billuart P, Koulakoff A, Vinet MC, Schaar B, McConnell SK, Francis F, Chelly J (2001) Doublecortin interacts with μ subunits of Clathrin adaptor complexes in the developing nervous system. *Mol. Cell. Neurosci.* 18:307-319

Friocourt G, Kappeler C, Saillour Y, Fauchereau F, Rodriguez MS, Bahi N, Vinet MC, Chafey P, Poirier K, Taya S, Wood SA, Dargemont C, Francis F, Chelly J (2005) Doublecortin interacts with the ubiquitin protease DFFRX, which associates with microtubules in neuronal processes. *Mol Cell Neurosci.* 28:153-64

Galan JM, Haguenaer-Tsapis R (1997) Ubiquitin Lys63 is involved in ubiquitination and endocytosis of a yeast plasma membrane protein. *EMBO J.* 16:5847-5854

GibcoBRL (1993) Guide to Baculovirus Expression Vector Systems (BEVS) and Insect Cell Culture Techniques. *Life Technologies Inc.*

GibcoBRL (1999) Bac-To-Bac Baculovirus Expression Systems. *Life Technologies Inc*

Ginalski K, Rychlewski L, Baker D, Grishin NV (2004) Protein structure prediction for the male-specific region of the human Y chromosome. *Proc. Natl. Acad. Sci. USA* 101:2305-2310

Goodman RH, Smolik S (2000) CBP/p300 in cell growth, transformation, and development. *Genes & Development* 14:1553-1577

Graham TA, Ferkey DM, Mao F, Kimelman D, Xu W (2001) Tcf4 can specifically recognize β -catenin using alternative conformations. *Nat. Struct. Biol.* 8:1048-1052

Groves MR, Hanlon N, Turowski P, Hemmings BA, Barford D (1999) The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell* 96:99-110

Gruenberg J (2001) The endocytic pathway: a mosaic of domains. *Nat. Rev. Mol. Cell Biol.* 2:721-730

Held D, Yaeger K, Novy R (2003) New coexpression vectors for expanded compatibilities in *E. coli*. *InNovations* 18:4-6

Heyninck K, Beyaert R (2005) A20 inhibits NF- κ B activation by dual ubiquitin-editing functions. *Trends Biochem. Sci* 30:1-4

Hicke L (1999) Gettin' down with ubiquitin: turning off cell surface receptors, transporters and channels. *Trends Cell Biol.* 9:107-112

Hicke L (2001a) Protein regulation by monoubiquitin. *Nat. Rev. Mol. Cell Biol.* 2:195-201

Hicke L (2001b) A new ticket for entry into budding vesicles – ubiquitin. *Cell* 106:527-530

Hicke L, Dunn R (2003) Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu. Rev. Cell Dev. Bio.* 19:141-172

Hochstrasser M (1995) Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. *Curr. Opin. Cell Biol.* 7:215-223

Hochstrasser M (1996) Ubiquitin-dependent protein degradation. *Ann. Rev. Genet.* 30:405-439

Hochstrasser M (2000) Evolution and function of ubiquitin-like protein-conjugation systems. *Nat. Cell Biol.* 2:E153-E157

Hopfner KP, Kopetzki E, Krebe GB, Bode S, Huber R, Engh RA (1998) New enzyme lineages by subdomain shuffling. *Proc. Natl. Acad. Sci. USA* 95:9813–9818

Hu M, Li P, Li M, Li W, Yao T, Wu JW, Gu W, Cohen RE, Shi Y (2002) Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. *Cell* 111:1041-1054

Huang Y, Baker RT, Fischer-Vize JA (1995) Control of cell fate by a deubiquitinating enzyme encoded by the *fat facets* gene. *Science* 270:1828-1831

Hubbard SJ, Eisenmenger F, Thornton JM (1994) Modelling studies of the change in conformation required for cleavage of limited proteolytic sites. *Protein Sci.* 3:757–68

Hubbard SJ (1998) The structural aspects of limited proteolysis of native proteins. *Biochim Biophys Acta.* 1382:191–206

Invitrogen Life Technologies (2003) Gateway Technology enables rapid gene expression and analysis. *Invitrogen Corporation*

Invitrogen Life Technologies (2004) E. coli expression system with Gateway technology. *Invitrogen Corporation*

Izaurrealde E, Lewis J, McGuigan C, Jankowska M, Darzynkiewicz E, Mattaj IW (1994) A nuclear cap binding protein complex involved in pre-mRNA splicing. *Cell* 78:657–668

Johnston SC, Larsen CN, Cook WJ, Wilkinson KD, Hill CP (1997) Crystal structure of a deubiquitinating enzyme (human UCH-L3) at 1.8 Å resolution. *EMBO J.* 16:3787-3796

Johnston SC, Riddle SM, Cohen RE, Hill CP (1999) Structural basis for the specificity of ubiquitin C-terminal hydrolases. *EMBO J.* 18:3877-3887

Jones DT (1999) Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292:195-202

Jowett T (1997) Tissue in situ hybridization. *Wiley, New York*

Kanai-Azuma M, Mattick JS, Kaibuchi K, Wood SA (2000) Co-localization of FAM and AF-6, the mammalian homologues of *Drosophila* *faf* and *canoe*, in mouse eye development. *Mech. Dev.* 91:383-386

Kapust RB, Waugh DS (1999) *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci.* 8:1668-1674

Kato M, Miyazawa K, Kitamura N (2000) A deubiquitinating enzyme UBPY interacts with the Src homology domain of Hrs-binding protein via a novel binding motif PX(V/I)(D/N)RXXKP. *J. Biol. Chem.* 48:37481-37487

Khut PY (2000) Investigation of novel FAM interacting proteins. *University of Adelaide Honours Thesis*

Kiefhaber T, Rudolph R, Kohler HH, Buchner J (1991) Protein aggregation *in vitro* and *in vivo*: a quantitative model of the kinetic competition between folding and aggregation. *Biotechnology (NY)* 9:825-829

Kim JH, Park KC, Chung SS, Bang O, Chung CH (2003) Deubiquitinating enzymes as cellular regulators. *J. Biochem.* 134:9-18

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.* 203:253-310

Krishna SS, Grishin NV (2004) The finger domain of the human deubiquitinating enzyme HAUSP is a zinc ribbon. *Cell Cycle* 3:1046-9

Krokoszynska I, Otlewski J (1996) Thermodynamic stability effects of single peptide bond hydrolysis in protein inhibitors of serine proteases. *J. Mol. Biol.* 256:793–802

Koeg M, Hoppe T, Schlenker S, Ulrich HD, Mayer TU, Jentsch A (1999) A Novel Ubiquitination factor, E4, is involved in multiubiquitin chain assembly. *Cell* 96:635-644

Kuriyama M, Harada N, Kuroda S, Yamamoto T, Nakafuku M, Iwamatsu A, Yamamoto D, Prasad R, Croce C, Canaani E, Kaibuchi K (1996) Identification of AF-6 and canoe as putative targets for Ras. *J. Biol. Chem.* 271:607-610

Layfield R, Alban A, Mayer RJ, Lowe J (2001) The ubiquitin protein catabolic disorders. *Neuropath. App. Neurobiol.* 27:171-179

Lee KH, Song GJ, Kang IS, Kim SW, Paick JS, Chung CH, Rhee K (2003) Ubiquitin-specific protease activity of *USP9Y*, a male infertility gene on the Y chromosome. *Reprod. Fert. Dev.* 15:129-133

Li M, Chen D, Shiloh A, Luo J, Nikolaev AY, Qin J, Gu W (2002) Deubiquitylation of p53 by HAUSP is an important pathway for p53 stabilization. *Nature* 416:648-653

Li Q, Hariharan IK, Chen F, Huang Y, Fischer JA (1997) Genetic interactions with Rap1 and Ras1 reveal a second function for the Fat facets deubiquitinating enzyme in *Drosophila* eye development. *Proc. Natl. Acad. Sci. USA* 94:12515-12520

Lima CD (2003) Regulating UBP-mediated ubiquitin deconjugation. *Structure* 11:3-4

Lin H, Keriell A, Morales CR, Bedard N, Zhao Q, Hingamp P, Lefrancois S, Combaret Y, Wing SS (2000) Divergent N-terminal sequences target an inducible testis deubiquitinating enzyme to distinct subcellular structures. *Mol. Cell. Biol.* 20:6568-6578

Lin H, Yin J, Reid J, Wilkinson KD, Wing SS (2001) Divergent N-terminal sequences of a deubiquitinating enzyme modulate substrate specificity. *J. Biol. Chem.* 276:20357-20363

Linding R, Russell RB, Neduva V, Gibson TJ (2003) GlobPlot: Exploring protein sequences for globularity and disorder. *Nucleic Acids Res.* 31:3701-8

Liu N, Dansereau DA, Lasko P (2003) Fat facets interacts with Vasa in the Drosophila pole plasm and protects it from degradation. *Curr. Biol.* 13:1905-1909

Liu X, Dai S, Zhu Y, Marrack P, Kappler JW (2003) The Structure of a Bcl-Xl/Bim Fragment Complex: Implications for Bim Function. *Immunity* 19:341-352

Llinas M, Marqusee M (1998) Subdomain interactions as a determinant in the folding and stability of T4 lysozyme. *Protein Sci.* 7:96-104

Lord JM, Davey J, Frigerio L, Roberts LM (2000) Endoplasmic reticulum-associated protein degradation. *Semin. Cell Dev. Bio.* 11:159-164

Lytle BL, Peterson FC, Qiu SH, Luo M, Zhao Q, Markley JL, Volkman BF (2004) Solution structure of a ubiquitin-like domain from tubulin-binding Cofactor B. *J.Biol.Chem.* 279:46787-46793

Mann M, Jensen ON (2003) Proteomic analysis of post-translational modifications. *Nat. Biotech.* 21:255-261

Mazza C, Ohno M, Segref A, Mattaj IW, Cusack S (2001) Crystal Structure of the Human Nuclear CAP Binding Complex. *Mol. Cell* 8:383-396

McGregor A, Blanchard AD, Rowe AJ, Critchley DR (1994) Identification of the vinculin binding site in the cytoskeletal protein alpha-actinin. *Biochem. J.* 301:225-233.

Meiler J, Mueller M, Zeidler A, Schmaeschke F (2002) JUFO: Secondary structure prediction for proteins. *www.jens-meiler.de*

Mogk A, Mayer MP, Deuerling E (2002) Mechanisms of protein folding: molecular chaperones and their application in biotechnology. *Chembiochem.* 3:807-814

Moretti PAB, Davidson AJ, Baker E, Lilley B, Zon LI, D'Andrea RJ (2001) Molecular cloning of a human *Vent*-like homeobox gene. *Genomics* 76:21-29

Morimoto M, Nishida T, Honda R, Yasuda H (2000) Modification of Cullin-1 by ubiquitin-like protein Nedd8 enhances the activity of SCF^{skp2} toward p27^{kip1}. *Biochem. Biophys. Res. Commun.* 270:1093–1096

Murray RZ, Jolly LA, Wood SA (2004) The FAM deubiquitylating enzyme localizes to multiple points of protein trafficking in epithelia where it associates with E-cadherin and β -catenin. *Mol. Biol. Cell* 15:1592-1599

Nicholas KB, Nicholas HB Jr., Deerfield DW II (1997) GeneDoc: Analysis and Visualization of Genetic Variation. *EMBNEW NEWS* 4

Nishihara K, Kanemori M, Kitagawa M, Yanagi H, Yura T (1998) Chaperone coexpression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl. Environ. Microbiol.* 64:1694-1699

Noma T, Kanai Y, Kanani-Azuma M, Ishii M, Fujisawa M, Kurohmaru M, Kawakami H, Wood SA, Hayashi Y (2002) Stage- and sex-dependent expressions of USP9x, an X-linked mouse ortholog of *Drosophila* Fat facets, during gonadal development and oogenesis in mice. *Mech. Dev.* 119 Suppl 1:S91-S95

Overstreet E, Chen X, Wendland B, Fischer JA (2003) Either part of a *Drosophila* epsin protein divided after the ENTH domain, functions in endocytosis of delta in the developing eye. *Curr. Biol.* 13:854-860

Overstreet E, Fitch E, Fischer JA (2004) Fat facets and liquid facets promote endocytosis and delta signalling in the signalling cells. *Development* 131:5355-5366

Pantaleon M, Kanai-Asuma M, Mattick JS, Kaibuchi K, Kaye PL, Wood SA (2001) FAM deubiquitylating enzyme is essential for preimplantation mouse embryo development. *Mech. Dev.* 109:151-60

Pickart CM (2000) Ubiquitin in chains. *Trends Biochem Sci.* 25:544-548

Pickart CM, Eddins MJ (2004) Ubiquitin: structures, functions, mechanisms. *Biochim. Biophys. Acta.* 1695:55-72

Pires-daSilva A, Sommer RJ (2003) The evolution of signalling pathways in animal development. *Nat. Rev. Genet.* 4:39-49

Plafker SM, Plafker KS, Weissman AM, Macara IG (2004) Ubiquitin charging of human class III ubiquitin-conjugating enzymes triggers their nuclear import. *J. Cell Biol.* 167:649-659

Pol A, Ortega D, Enrich C (1997) Identification of cytoskeleton-associated proteins in isolated rat liver endosomes. *Biochem. J.* 327:741-746

Posner M, Kantorow M, Horwitz J (1999) Cloning, sequencing and differential expression of alphaB-crystallin in the zebrafish, *Danio rerio*. *Biochim Biophys Acta.* 447:271-277

Raymond PA, Barthel LK (2004) A moving wave patterns the cone photoreceptor mosaic array in the zebrafish retina. *In. J. Dev. Biol.* 48:935-945

Read MA, Brownell JE, Gladysheva TB, Hottelet M, Parent LA, Coggins MB, Pierce JW, Podust VN, Luo RS, Chau V, Palombella VJ. (2000) Nedd8 modification of cul-1 activates SCF(beta(TrCP))-dependent ubiquitination of IkappaBalpha. *Mol. Cell Biol.* 20:2326-2333

Ready DF (1989) A multifaceted approach to neural development. *Trends Neurosci.* 12:102-110

Sambrook J, Russell D (2001) Molecular cloning: a laboratory manual. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor*

Sato T, Kanai Y, Noma T, Kanai-Azuma M, Taya S, Matsui T, Ishii M, Kawakami H, Kurohmaru M, Kaibuchi K, Wood SA, Hayashi Y (2004) A close correlation in the expression patterns of Af-6 and USP9x in Sertoli and granulosa cells of mouse testis and ovary. *Proc. Natl. Acad. Sci. USA* 101:2305-2310

Schubert U, Anton LC, Gibbs J, Norbury CC, Yewdell WJ, Bennink JR (2000) Rapid degradation of a large fraction of newly synthesised proteins by proteasomes. *Nature* 404:770-774

Semple CAM, (2003) The comparative proteomics of ubiquitination in mouse. *Genome Res.* 13:1389-1394

Sørensen HP, Mortensen KK (2005) Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microb. Cell Fact.* 4:1

Spector DL, Goldman RD, Leinwand LA (1998) Cells, a laboratory manual – volume 2: light microscopy and cell structure. *Cold Spring Harbor Laboratory Press, Cold Spring Harbor*

Spence J, Sadis S, Haas AL, Finley D (1995) A ubiquitin mutant with specific defects in DNA repair and multiubiquitination. *Mol. Cell. Biol.* 15:1265-1273

Tanaka K, Suzuki T, Hattori N, Mizuno Y (2004) Ubiquitin, proteasome and parkin. *Biochim Biophys Acta.* 1695:226-238

Tanaka TS, Jaradat SA, Lim MK, Kargul GJ, Wang X, Grahovac MJ, Pantano S, Sano Y, Piao Y, Nagaraja R, Doi H, Wood WH, Becker KG, Ko MSH (2000) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray. *Proc. Natl. Acad. Sci.* 97:9127-9132

Taya S, Yamamoto T, Kano K, Kawano Y, Iwamatsu A, Tsuchiya T, Tanaka K, Kanai-Azuma M, Wood SA, Mattick JS, Kaibuchi K (1998) The Ras target AF-6 is a substrate of the *Fam* deubiquitinating enzyme. *J. Cell Biol.* 142:1053-1062

Taya S, Yamamoto T, Kani-Azuma M, Wood SA, Kaibuchi K (1999) The deubiquitinating enzyme *Fam* interacts with and stabilizes β -catenin. *Genes to Cells* 4:757-767

Teodorescu O, Galor T, Pillardy J, Elber R (2004) Enriching the sequence substitution matrix by structural information. *Prot. Struct. Funct. Genet.* 54:41-48

Tintignac LA, Lagirand J, Batonnet S, Sirri V, Leibovitch MP, Leibovitch SA. (2005) Degradation of MyoD mediated by the SCF (MAFbx) ubiquitin ligase. *J Biol. Chem.* 280:2847-56

Varshavsky A (1997) The ubiquitin system. *Trends Biochem. Sci.* 22, 383-387.

Wang C, Deng L, Hong M, Akkaraju GR, Inoue J, Chen ZJ (2001) TAK1 is a ubiquitin-dependent kinase of MKK and IKK. *Nature* 412:346-351

Wang G, Yang J, Huibregtse JM (1999) Functional Domains of the Rsp5 Ubiquitin-Protein Ligase. *Mol. Cell Biol.* 19:342-352

Wang W, Struhl (2004) *Drosophila* epsin mediates a select endocytic pathway that DSL ligands must enter to activate notch. *Development* 131:5367-5380

Wilkinson KD (1997) Regulation of ubiquitin-dependent processes by deubiquitinating enzymes. *FASEB J.* 11:1245-1256

Wing SS (2003) Deubiquitinating enzymes-the importance of driving in reverse along the ubiquitin-proteasome pathway. *Int. J. Biochem Cell Biol.* 35:590-605

Wong BR, Parlanti F, Qu K, Demo S, Pray T, Huang J, Payan DG, Bennett MK (2003) Drug discovery in the ubiquitin regulatory pathway. *Drug Discov. Today* 8:746-754

Wood SA, Pascoe WS, Ru KL, Yamada T, Hirchenhain J, Kemier R, Mattick JS (1997) Cloning and expression analysis of a novel mouse gene with sequence similarity to the *Drosophila fat facets* gene. *Mech. Dev.* 63:29-38

Wu LC, Grandori R, Carey J (1994) Autonomous subdomains in protein folding. *Protein Sci.* 3:369-371

Wu Z, Li Q, Fortini M, Fischer JA (1999) Genetic analysis of the role of the *Drosophila fat facets* gene in the ubiquitin pathway. *Dev. Gen.* 25:312-320

Zhadanov AB, Provance DW, Speer CA, Coffin JD, Goss D, Blixt JA, Reichert CM, Mercer JA (1999) Absence of the tight junctional protein AF-6 disrupts epithelial cell-cell junctions and cell polarity during mouse development. *Curr. Biol.* 9:880-888

Zhang XD, Matunis MJ (2005) Ub in charge: regulating E2 enzyme nuclear import. *Nat. Cell Biol.* 7:12-14

Zheng N, Schulman BA, Song L, Miller JJ, Jeffrey PD, Wang P, Chu C, Koepp DM, Elledge SJ, Pagano M, Conaway RC, Conaway JW, Harper JW, Pavletich NP (2002) Structure of the Cul1-Rbx1-Skp1-F Box^{skp2} Scf Ubiquitin Ligase Complex. *Nature* 416:703-709

Zhu Z, Bulgakov OV, Scott SS, Dalton JT (2001) Recombinant expression and purification of human androgen receptor in a baculovirus system. *Biochem. Biophys. Res. Commun.* 284:828-835