

Mass Spectrometry Methods and Results

Mass spectrometry of affinity purified EpCAM: Each digest was analysed using a 60 minute LC-MS/MS run on a maXis impact UHR-TOF mass spectrometer (*Bruker Daltonics*). A 0-40% acetonitrile gradient in 0.1% formic was used to elute peptides from a 25 cm x 75 μ M C18 Pepmap column (*Dionex*) at 300 nl/min. MASCOT 2.3.0 was used to search the Swiss-Prot human database. Mass tolerances for parent and fragment ions were 10 ppm and 0.01 Da respectively, and only peptides with a MOWSE score >30 were accepted. Carbamidomethylation of cysteines was considered as a fixed modification and deamidation and methionine oxidation as variable modifications. Deglycosylation of tryptic peptides with PNGase was performed after in-gel digestion.

Table S1. Tryptic peptides identified in affinity purified urinary EpCAM. The sequence of each peptide is shown as well as position in the protein (residues = numbering assuming the mature polypeptide has had the signal peptide removed) and MS and MSMS parameters. The peptides in *italic* were only identified following PNGase treatment.

Sequence	Residues	m/z meas.	Mr calc.	z	Δ m/z [ppm]	RMS90 [ppm]	Score	IntCov. [%]	Modifications
L.AVNCFVNNNR.Q	12-20	604.2854	1206.5564	2	-0.12	2.50	43.7	35	Carbamidomethyl: 4
K.LAVNCFVNNNR.Q	11-21	660.8267	1319.6405	2	-1.22	3.67	59.2	52	Carbamidomethyl: 5
R.QCQCTSVGAQNTVICSK.L	22-38	970.9348	1939.8550	2	0.02	3.69	101.3	30	Carbamidomethyl: 2, 4, 15
R.AKPEGALQNNDGLYDPDCDESLFK.A	59-83	918.4168	2752.2286	3	-0.02	2.86	66.4	31	Carbamidomethyl: 18
<i>K.QCNGTSMCWCVNTAGVR</i>	<i>86-102</i>	<i>1001.4102</i>	<i>2000.7961</i>	<i>2</i>	<i>4.86</i>	<i>2.18</i>	<i>45.6</i>	<i>32</i>	<i>Carbamidomethyl: 2, 8, 10, Deamidated 3</i>
<i>K.QCNGTSMCWCVNTAGVR</i>	<i>86-102</i>	<i>1009.9049</i>	<i>2016.7910</i>	<i>2</i>	<i>497.18</i>	<i>4.14</i>	<i>82.8</i>	<i>38</i>	<i>Carbamidomethyl: 2, 8, 10, Deamidated 3, Oxidation;7</i>
R.TDKDEITCSEK.V	104-115	727.8241	1453.6355	2	-1.26	5.41	60.6	55	Carbamidomethyl: 9
K.TQNDVDIADVAYYFEK.D	180-195	946.4342	1890.8523	2	0.80	3.92	96.8	41	Deamidated: 3
K.TQNDVDIADVAYYFEK.D	180-195	945.9413	1889.8683	2	-0.15	2.99	112.2	48	
K.DVKGESLFHSK.K	196-206	623.8251	1245.6354	2	0.24	5.27	53.7	30	
K.KMDLTVNGEQLDLDPGQTLIYYVDEK.A	207-232	1005.1608	3012.4638	3	-1.07	5.15	47.8	31	Oxidation: 2
K.APEFSMQGLK.A	233-242	554.2776	1106.5430	2	-2.15	2.38	58.6	41	
K.APEFSMQGLK.A	233-243	589.7964	1177.5801	2	-1.61	2.09	34.8	35	

Table S2. Tryptic peptides identified in affinity purified cell line EpCAM. The peptide marked with an asterisk indicates that the N-terminal signal peptide is 2 residues shorter than shown in UniProtKB. .

The peptides in italic were only identified following PNGase treatment.

Sequence	Residues	m/z meas.	Mr calc.	z	Δ m/z [ppm]	RMS90 [ppm]	Scores	IntCov. [%]	Modifications
A.AAQEECVENYK.L*	-2-10	750.8072	1499.6021	2	-1.49	4.16	55.1	25	Carbamidomethyl: 6, 8
K.LAVNCFVNNR.Q	11-21	660.8272	1319.6405	2	-0.46	1.48	82.8	61	Carbamidomethyl: 5
K.LAVNCFVNNR.Q	11-21	661.3199	1320.6245	2	0.59	3.96	76.7	52	Carbamidomethyl: 5; Deamidated: 4
LAVNCFVNNR.Q	12-21	604.2854	1206.5564	2	-0.12	2.81	51.0	43	Carbamidomethyl: 4
R.QQCQTSVGAQNTVCSK.L	22-38	970.9344	1939.855	2	-0.4	1.36	116.1	48	Carbamidomethyl: 2, 4, 15
<i>K.CLVMKAEMNGSK.L</i>	43-54	684.8194	1367.6247	2	-0.35	2.57	40.2	18	Carbamidomethyl: 1; Deamidated: 9
R.AKPEGALQNNGLYDPDCDESLFK.A	59-83	918.4168	2752.2286	3	-0.02	1.18	95.3	29	Carbamidomethyl: 18
R.AKPEGALQNNGLYDPDCDESLFK.A	59-83	918.7455	2753.2127	3	0.73	1.19	65.1	28	Carbamidomethyl: 18; Deamidated: 9
D.PDCDESLFK.A	74-83	584.2531	1166.4914	2	0.22	4.61	33.5	23	Carbamidomethyl: 3
R.RTDKDEITCSEK.V	103-115	805.8761	1609.7366	2	0.65	3.56	51.3	8	Carbamidomethyl: 10
R.TDKDEITCSEK.V	104-115	727.8243	1453.6355	2	-0.99	2.63	79.8	34	Carbamidomethyl: 9
<i>K.QCNGTSMCWCVNTAGVR</i>	86-102	1001.4055	2000.7961	2	0.16	1.34	130.2	41	Carbamidomethyl: 2, 8, 10, Deamidated 3
<i>K.QCNGTSMCWCVNTAGVR</i>	86-102	1009.4029	2016.7910	2	0.10	1.61	109.3	34	Carbamidomethyl: 2, 8, 10, Deamidated 3, Oxidation: 7
K.DTEITCSEK.V	107-115	555.7398	1109.4659	2	-0.77	1.75	53.8	49	Carbamidomethyl: 6
K.TQNDVDIADVAYYFEK.D	180-195	945.9399	1889.8683	2	-1.63	3.05	102.8	47	
R.TALQKEITR.Y	141-150	580.8352	1159.6561	2	-0.21	1.42	67.6	39	
K.TQNDVDIADVAYYFEK.D	180-195	945.9429	1889.8683	2	1.55	1.31	121.8	59	
K.TQNDVDIADVAYYFEK.D	180-195	946.4352	1890.8523	2	1.85	1.70	115.9	43	Deamidated: 3
K.TQNDVDIADVAYYFEKDVK.G	180-198	745.026	2232.0586	3	-1.11	6.01	25.1	23	
K.TQNDVDIADVAYYFEKDVKG.E	180-199	764.033	2289.0801	3	-1.28	3.51	64.6	35	
K.DVKGESLFHSK.K	196-206	623.853	1245.6354	2	0.56	6.37	52.1	31	
K.GESLFHSK.K	199-206	452.7329	903.4450	2	6.86	4.48	33.7	40	
K.KMDLTVNGEQLDLPGQTLIYYVDEK.A	207-232	1005.1611	3012.4638	3	-0.77	3.23	81.2	29	Oxidation: 2
D.PGQTLIYYVDEK.A	221-232	713.3641	1424.7187	2	-3.57	2.39	62.8	35	
K.APEFSMQGLK.A	233-242	554.2778	1106.543	2	-1.79	6.21	64.4	34	
K.APEFSMQGLK.A	233-242	562.2767	1122.5379	2	0.8	2.84	46.2	25	Oxidation: 6
K.AEIKEMGEMHR.E	277-287	665.8154	1329.6169	2	-0.52	4.3	73.6	40	
K.AEIKEMGEMHR.E	277-287	673.8136	1345.6119	2	0.59	6.67	59.5	32	Oxidation: 6

Table S3. Asp-N peptides identified in affinity purified urinary EpCAM. The peptides which define the major EpCAM shedding cleavage are marked with #.

Sequence	Residues	m/z meas.	Mr calc.	z	Δ m/z [ppm]	RMS90 [ppm]	Score	IntCov. [%]	Modifications
Y.DSKSLRTALQKEITTRYQL.D	135-153	751.0835	2250.2332	3	-2.01	2.09	33.9	8	
A.DVAYYFEK.D	188-195	517.7437	1033.4757	2	-2.72	1.47	40.6	43	
K.DVKGESLFHSKKM.D	196-208	753.3915	1504.7708	2	-1.56	2.64	75.8	16	
K.DVKGESLFHSKKM.D	196-208	761.3895	1520.7657	2	-0.83	4.14	69.3	19	Oxidation: 13
V.DEKAPEFSMQGLK.A	230-242	740.3577	1478.7075	2	-4.51	4.21	36.2	29	
V.DEKAPEFSMQGLK.A#	230-243	783.8755	1565.7396	2	-1.98	2.72	65.7	33	Oxidation: 9
V.DEKAPEFSMQGLK.A#	230-243	775.8761	1549.7446	2	-4.51	2.91	78	45	

Table S4. Asp-N peptides identified in affinity purified cell line EpCAM.

Sequence	Residues	m/z meas.	Mr calc.	z	Δ m/z [ppm]	RMS90 [ppm]	Score	IntCov. [%]	Modifications
C.DESGLFKAKQ.C	77-86	561.7918	1121.5717	2	-2.35	4.33	61.5	28	
Y.DSKSLRTALQKEITTRYQL.D	135-153	751.0842	2250.2332	3	-1.08	0.58	48.2	6	
L.DPKFITSILY.E	154-163	598.8311	1195.6489	2	-1.03	6.95	59.5	28	
L.DPKFITSILY.E.N	154-164	663.3517	1324.6915	2	-1.98	2.01	67.1	30	
V.DIADVAYYFEK.D	185-195	667.3175	1332.6238	2	-2.5	4.16	31.2	19	
K.DVKGESLFHSHK.M.D	196-208	753.3923	1504.7708	2	-0.5	0.86	80.5	16	
K.DVKGESLFHSHK.M.D	196-208	761.3899	1520.7657	2	-0.31	2.16	85	22	Oxidation: 13
G.ESLFHSHK.M.D	200-208	553.7862	1105.559	2	-1.05	4.33	42.3	17	
L.DPGQTLIYV.D	220-229	584.7973	1167.5812	2	-0.98	1.66	45.1	41	
Y.EKAEIKEMG.E	275-283	517.7626	1033.5114	2	-0.72	1.24	42	25	
Y.EKAEIKEMGEMHR.E	275-287	529.9252	1586.7545	3	-0.46	2.44	45.1	8	
Y.EKAEIKEMGEMHRELNA.-	275-291	672.3264	2013.9612	3	-1.9	3.41	57.6	6	
A.EIKEMGEMHRELNA.-	278-291	851.897	1701.7814	2	-1.17	1.47	46.7	9	Oxidation: 8
A.EIKEMGEMHR	278-286	552.2454	1102.4787	2	-2.24	1.69	34	21	
A.EIKEMGEMHR.E	278-287	638.2929	1274.5747	2	-2.74	3.33	33.5	22	Oxidation: 5
A.EIKEMGEMHR.E	278-287	630.2967	1258.5798	2	-0.78	3.62	54.1	37	
A.EIKEMGEMHRELNA.-	278-291	843.9	1685.7865	2	-0.64	2.1	82.1	14	
K.EMGEMHRELNA.-	281-291	658.7891	1315.5649	2	-0.95	1.93	60.1	17	
G.EMHRELNA.-	284-291	500.2408	998.4604	2	6.68	28.16	38	16	