SUPPLEMENTARY INFORMATION

Unfurling of the band 4.1, ezrin, radixin, moesin (FERM) domain of the merlin tumor suppressor

S.D. Yogesha,[1] Andrew J. Sharff,[2] Marco Giovannini,[3] Gerard Bricogne,[2] and Tina Izard[1*]

Supplementary Materials and Methods

Liquid chromatography/mass spectrometry analyses of merlin head:tail crystals

Merlin head:tail crystals were washed twice in reservoir solution (0.2 M ammonium sulfate and 5% PEG 4000) and then dissolved in 10 mM Tris (pH 8). The dissolved crystals were treated with 0.5 mM DTT for 1 hr at 37 ºC, followed by treatment with 2 mM iodoacetamide for 1 hr at ambient temperature, to reduce and carbamidomethylate cysteine residues. The protein was subjected to proteolysis with trypsin (35 ng) and the resulting merlin peptides were analyzed by liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) by our in-house proteomics core. The resulting peptides were first loaded onto a 360 × 100 μm fused silica capillary (PolyMicro Technologies, Phoenix, AZ) pre-column packed with 3 cm of C12 packing material (Jupiter Proteo C12, 4 μm particle size). Salts were removed by washing with 0.1 M acetic acid in 1% acetonitrile and the column was placed in line with a 360 × 100 μm fused silica capillary column packed with 20 cm of C12 material. The peptides were eluted with a gradient consisting of 5% to 55% acetonitrile in 50 min, using a 1,100 series HPLC pump (Agilent Technologies, Santa Clara, CA). A flow rate of 200 nl/min was achieved at the reverse phase column by splitting the flow delivered from the HPLC. Peptides were gradient-eluted and ionized (1.7 kV) by placing a nano-electrospray ionization source (Triversa Nanomate, Advion, Ithaca, NY) into an LTQ-Orbitrap mass spectrometer (Thermo Scientific, San Jose, CA). The LTQ-Orbitrap instrument was operated in a data-dependent mode with dynamic exclusion enabled. The data-dependent method consisted of acquisition of a full scan mass spectrum (m/z 350 – 2000) using the Orbitrap as the analyzer, followed by ten tandem mass spectra (MS/MS) of the ten most abundant ions in the initial full scan. Precursor ions selected for MS/MS were fragmented by collision-activated dissociation (CAD), and the MS/MS scans were acquired using the ion trap as the analyzer. Data were processed using a raw file processing pipeline that extracts tandem mass spectra from the raw data files, which filters the data for spectral quality using the SPEQUAL algorithm,[1] and concatenates the high-quality spectra for database searching via a clustered version of Mascot.[2] Searches were performed against the IPI human protein database. Mascot search results were loaded into Scaffold (Proteome Software, Portland, OR) for statistical analysis followed by manual verification of all peptide assignments. All of the merlin peptides detected are presented in the Supplementary Table.

SUPPLEMENTARY INFORMATION

Supplementary Figure Legends

Supplementary Figure S1. Unfurling of the merlin F2 domain. The F1 subdomain (residues 20-82 and 91-100) is shown in yellow, the F2 subdomain (residues 101-158 and 178-215) is shown in green, and the F3 motif (residues 216-313) is shown in magenta.

(A) Zoomed-in stereo view of the final $2F_o$-F_c electron density map of the F2 α3-α4 loop region. The contour level of the electron density map is 1σ and the resolution is 2.64 Å. Water molecules are shown as a red sphere. Oxygen atoms are in red, nitrogen in blue and carbon in yellow, green, or magenta for the F1-F3 subdomain, respectively. Some residues are labeled.

(B) Zoomed-out view of the merlin dimer looking down the dyad in the same orientation as in panel (A). Some secondary structure elements are labeled with the extension 'b' indicating its location on the F2 subdomain. Residues from the F2 α3-α4 loop are shown in ball-and-stick representation with the remainder of the molecule as a cartoon. The black square indicates the region that is shown in panel (A).

Supplementary Figure S2. Unfurling of the merlin F2 subdomain is likely triggered by the tail domain.

(A) Superposition our 2.64 Å unfurled merlin FERM domain (yellow, green, and magenta) onto the 3 Å full-length moesin crystal structure (white, black, and red; PDB entry 2i1k) and the 2.1 Å truncated moesin crystal structure (grey; PDB entry 2i1j). In the full-length moesin structure, the *C*-terminus (indicated by an arrow that points at moesin residue 320) of the moesin 'A' α-helix (labeled, shown in black, residues 299-320) would have to shift to allow the unfurling of its F2 subdomain.

(B) Superposition of our unfurled merlin structure (orange) onto the moesin head:tail structure (white and black) and the closed, unbound merlin head structure (PDB entry 1h4r; cyan). Both molecules in the asymmetric unit are shown for the moesin head:tail complex and for the closed, unbound merlin FERM structures. The former are very similar with *r.m.s.d.* of 0.14 Å for 2,174 atoms. While the loop succeeding β-strand β6c in our unfurled merlin structure resembles that of the closed moesin head:tail structure (indicated by one arrow on the right), there is further movement in the α-helix α3b and its following loop (indicated by two arrows on the left) in the moesin head:tail structure (green) when compared to the closed, unbound merlin head structure (cyan).

SUPPLEMENTARY INFORMATION

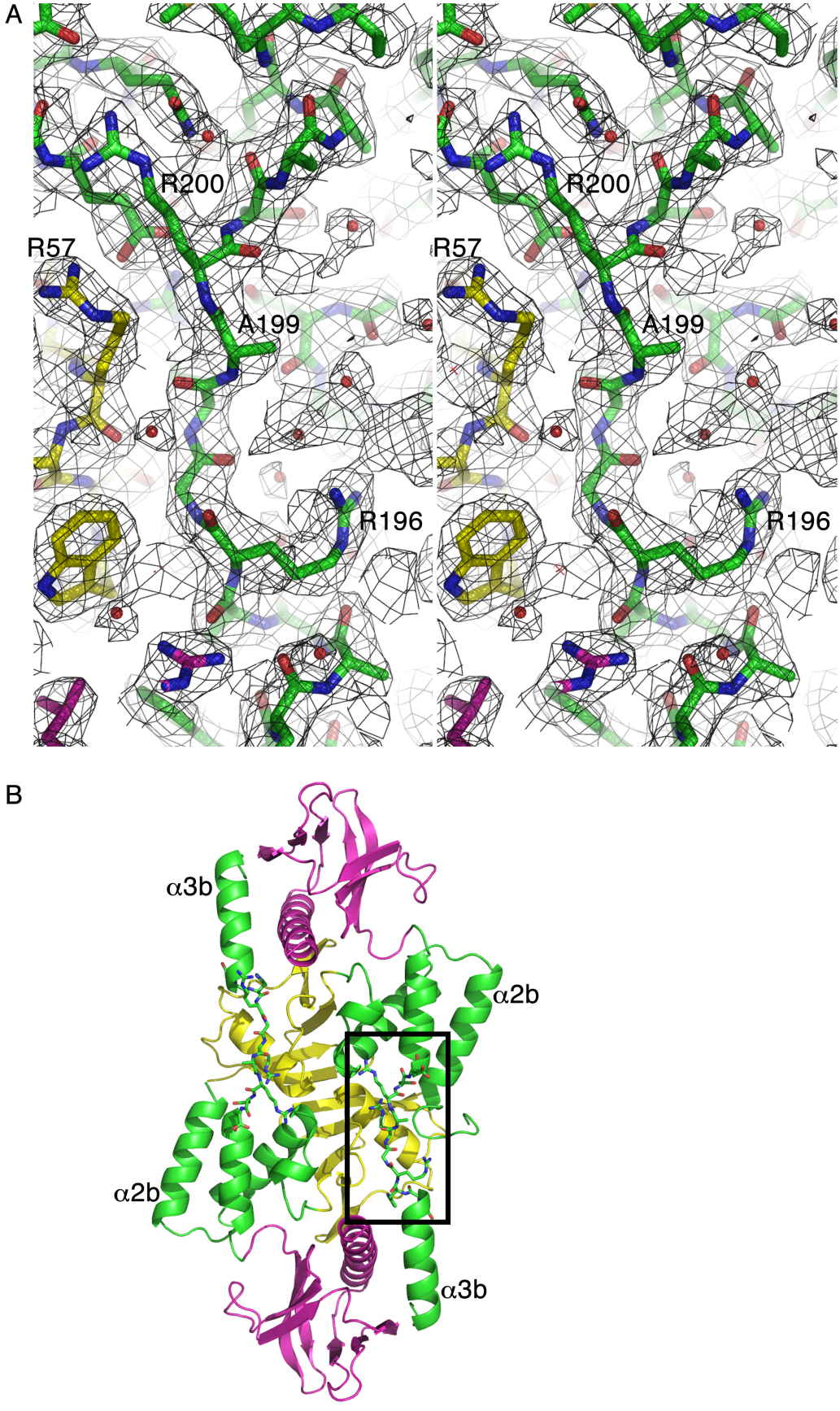
Supplementary Figure S3. Sequence alignment of the FERM domains of moesin versus merlin.

Residues that are identical in the two structures are indicated by an asterisk in the moesin (Moe) sequence (top sequence). The F1, F2, and F3 subdomains are indicated. The four α-helices are highlighted in red (moesin residues 95-112, merlin residues 112-128), yellow (moesin residues 118-135, merlin residues 135-150), green (moesin residues 164-179, merlin residues 181-196), and blue (moesin residues 183-196, merlin residues 202-213), respectively. The α-helix α2b is disordered in our unfurled structure of the merlin head domain. Underlined merlin residues (83-90 and 159-177) indicate disordered regions. A period above the moesin sequence indicates that this residue is involved in the moesin head:tail interface as seen in the crystal structure.[3] Italicized merlin residues in blue font indicate the region that unfurls, which is distinct from any FERM domain structure determined thus far. There are 22 identical residues in the merlin 150-201 region while there are 27 and 38 identical residues in the merlin 98-149 and 202-253 regions, respectively.

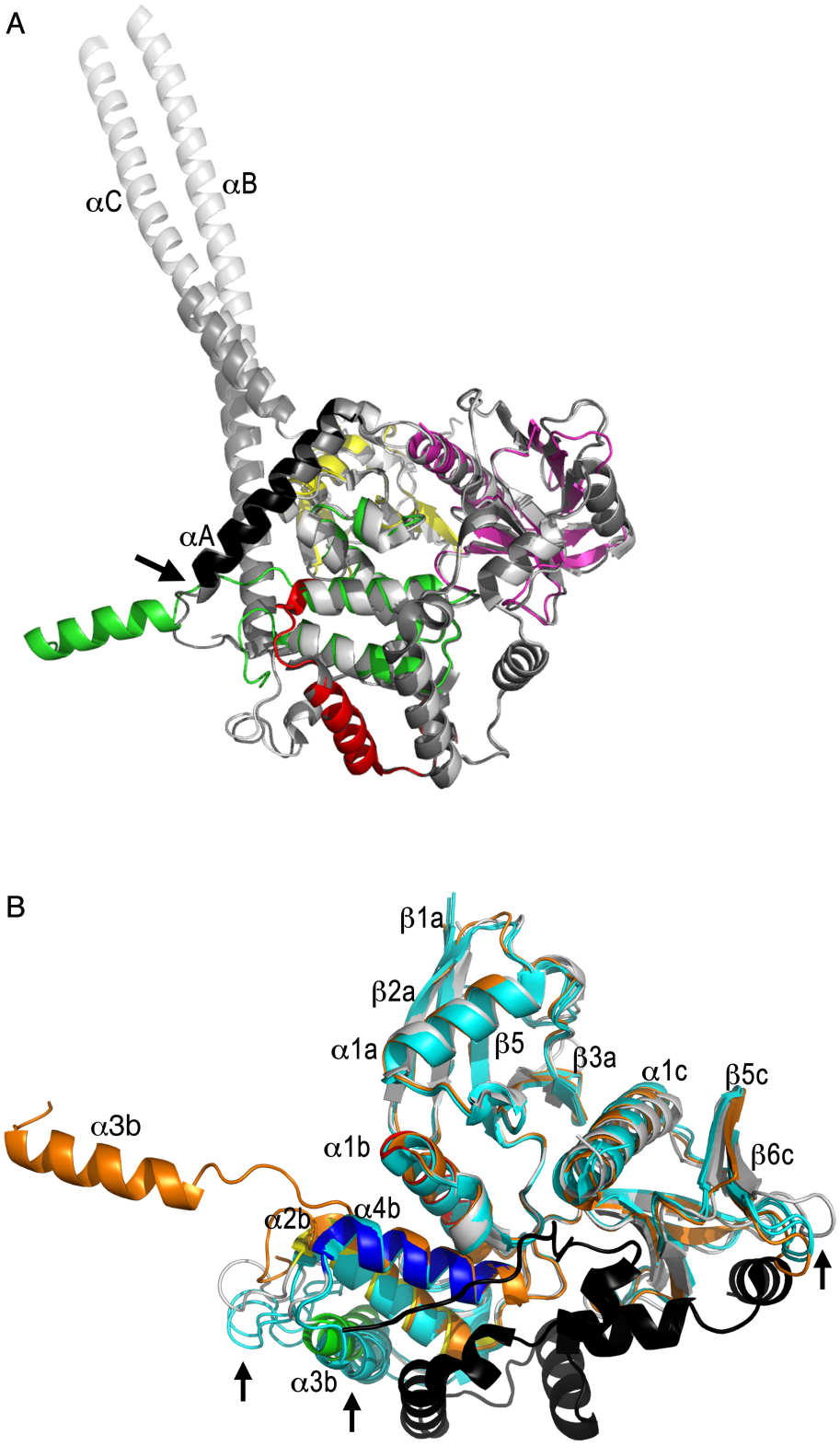
REFERENCES

1. Purvine S, Kolker N, Kolker E (2004) Spectral quality assessment for high-throughput tandem mass spectrometry proteomics. OMICS 8:255-265.
2. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551-3567.
3. Pearson MA, Reczek D, Bretscher A, Karplus PA (2000) Structure of the ERM protein moesin reveals the FERM domain fold masked by an extended actin binding tail domain. Cell 101:259-270.

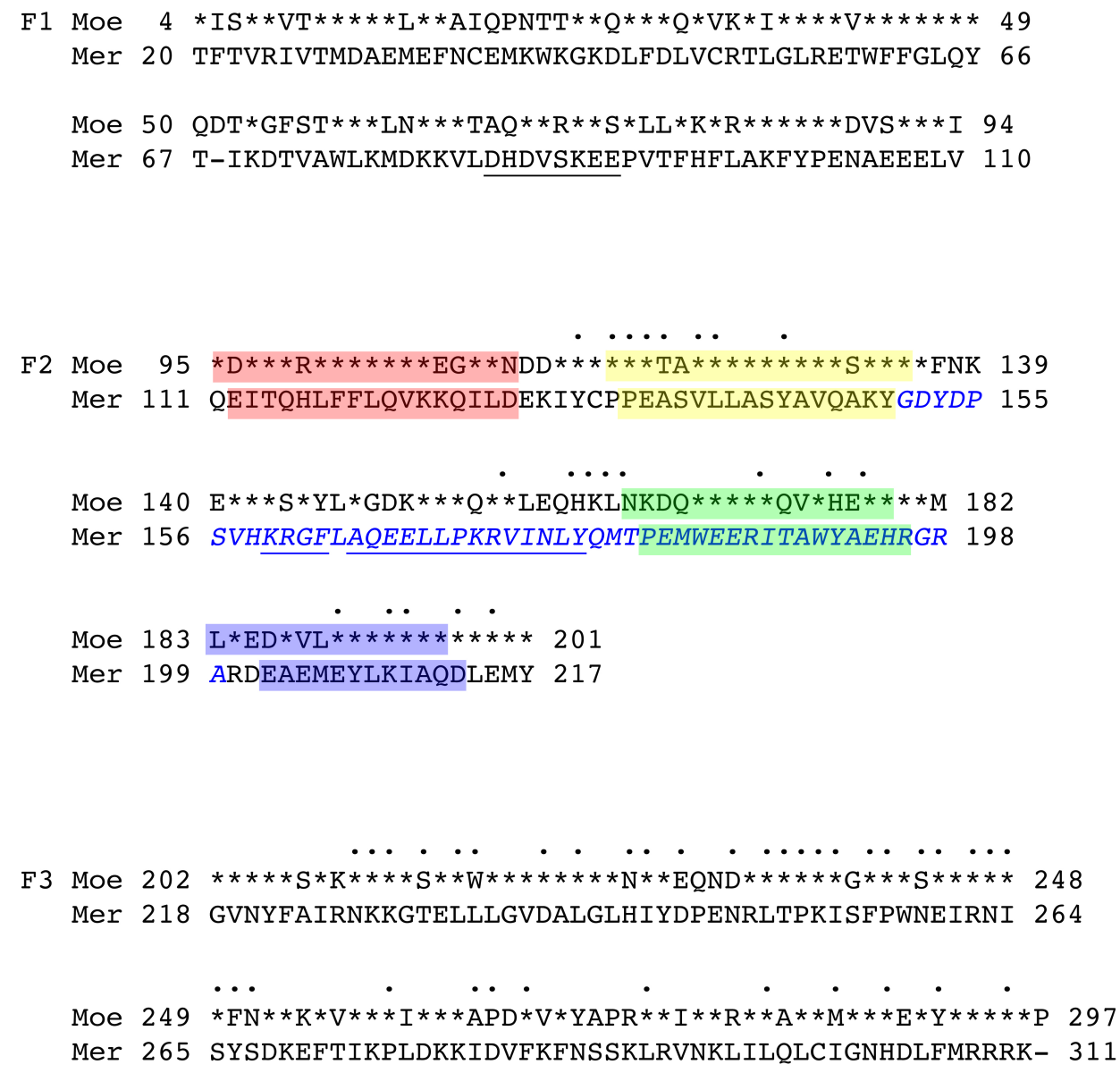
Supplementary Figure S1



Supplementary Figure S2



Supplementary Figure S3



Supplementary Table. Peptides identified from MS/MS. Peptides identified in the mass spectrometric experiment are listed. Amino acids listed as lower case letters indicate that this amino acid is modified: c, carbamidomethylated cysteine and m, oxidized methionine. For simplicity only one peptide was listed for each possible modification since the masses identified are the same. Theoretical masses were calculated and compared to observed masses to determine part per million errors. The table lists all observed masses for each peptide. Note that peptides were identified for both the head and the tail domain of our crystals and thus confirming that our crystals contain both domains, in particular the tail domain, which is disordered in our structure.

		Theoretical (monoisotropic)				Observed (monoisotropic)				Error calculations (PPM)			
Peptide sequence	Residues	MH+1	MH+2	MH+3	MH+4	MH+1	MH+2	MH+3	MH+4	+1	+2	+3	+4
IVTMDAEMEFNcEMK	26-40	1847.7682	924.3878	616.5943	462.6975		924.3881	616.5937			0.32	-0.97	
IVTMDAEMEFNcEmK	26-40	1863.7632	932.3852	621.9259	466.6962		932.3848	621.925			-0.43	-1.45	
IVTmDAEMEFNcEmK	26-40	1879.7581	940.3827	627.2575	470.695		940.3805	627.2562			-2.34	-2.07	
IVTMDAEMEFNcEmKWK	26-42	2177.9374	1089.472	726.6507	545.2398		1089.474	726.6511			1.56	0.55	
IVTmDAEMEFNcEmKWK	26-42	2193.9323	1097.47	731.9823	549.2385		1097.47	731.9813			-0.09	-1.37	
IVTmDAEmEFNcEmK	26-40	1895.753	948.3801	632.5892	474.6937		948.3798	632.5884			-0.32	-1.26	
GKDLFDLVcR	43-52	1222.6249	611.8161	408.2161	306.4117	1222.624	611.8145			-0.41	-2.62		
DLFDLVcR	45-52	1037.5084	519.2579	346.5077		1037.508	519.2554			0.00	-4.81		
KVLDHDVSKEEPVTFHFLAK	80-99	2339.2394	1170.123	780.418	585.5653			780.418	585.5646			0.00	-1.20
VLDHDVSK	81-88	912.4785	456.7429	304.831		912.4783	456.7418	304.8305		-0.22	-2.41	-1.64	
VLDHDVSKEEPVTFHFLAK	81-99	2211.1444	1106.076	737.7197	553.5416		1106.074	737.7186	553.5414		-1.72	-1.49	-0.36
EEPVTFHFLAK	89-99	1317.6838	659.3455	439.8994	330.1764	1317.684	659.3446	439.8979		0.15	-1.36	-3.41	
FYPENAEEELVQEITQHLFFLQVK	100-123	2951.4826	1476.245	984.499	738.6261		1476.242	984.4991	738.6259		-2.10	0.10	-0.27
KQILDEK	124-130	873.504	437.2556			873.5035	437.2545			-0.57	-2.52		
QILDEK	125-130	745.409	373.2082			745.408	373.2071			-1.34	-2.95		
IYcPPEASVLLASYAVQAK	131-149	2080.0783	1040.543	694.031	520.775		1040.54	694.0291			-3.08	-2.74	
YGDYDPSVHK	150-159	1180.5269	590.7671	394.1805		1180.526	590.7658	394.1789		-0.42	-2.20	-4.06	
YGDYDPSVHKR	150-160	1336.628	668.8177	446.2142	334.9125	1336.626	668.8159	446.2123	334.9117	-1.27	-2.69	-4.26	-2.39
GFLAQEELLPK	161-171	1244.6885	622.8479	415.5677	311.9276	1244.688	622.8467	415.5675		-0.08	-1.93	-0.48	
GFLAQEELLPKR	161-172	1400.7896	700.8985	467.6014	350.9529		700.8981	467.6003			-0.57	-2.35	
RVINLYQMTPEMWEER	172-187	2095.0099	1048.009	699.0082	524.5079		1048.01	699.0079	524.5071		1.43	-0.43	-1.53
RVINLYQMTPEmWEER	172-187	2111.0049	1056.006	704.3398	528.5067		1056.007	704.3386			0.38	-1.70	
RVINLYQmTPEmWEER	172-187	2126.9998	1064.004	709.6714	532.5054		1064.004	709.6699			0.28	-2.11	
VINLYQMTPEMWEER	173-187	1938.9088	969.9581	646.9745	485.4827		969.9578	646.9737			-0.31	-1.24	
VINLYQMTPEmWEER	173-187	1954.9037	977.9555	652.3061	489.4814		977.9514	652.304			-4.19	-3.22	

VINLYQmTPEmWEER	173-187	1970.8987	985.953	657.6377	493.4801		985.9514	657.6366			-1.62	-1.67	
ITAWYAEHR	188-196	1146.5691	573.7882	382.8612		1146.567	573.7858	382.8591		-1.92	-4.18	-5.49	
ARDEAEMEYLK	199-209	1354.6307	677.819	452.2151	339.4131		677.8171	452.213			-2.80	-4.64	
ARDEAEmEYLK	199-209	1370.6257	685.8165	457.5467	343.4119	1370.623	685.8146	457.5447		-1.75	-2.77	-4.37	
DEAEMEYLK	201-209	1127.4925	564.2499	376.5024		1127.493	564.2488			0.35	-1.95		
DEAEmEYLK	201-209	1143.4874	572.2474	381.834		1143.488	572.2466			0.52	-1.40		
IAQDLEMYGVNYFAIR	210-225	1902.9418	951.9746	634.9855	476.4909		951.9744	634.985			-0.21	-0.79	
IAQDLEmYGVNYFAIR	210-225	1918.9368	959.972	640.3171	480.4896		959.9689	640.3151			-3.23	-3.12	
KGTELLLGVDALGLHIYDPENR	228-249	2423.2929	1212.15	808.4358	606.5787		1212.151	808.4365	606.5786		0.66	0.87	-0.16
GTELLLGVDALGLHIYDPENR	229-249	2295.1979	1148.103	765.7375	574.5549		1148.104	765.7373	574.5546		0.78	-0.26	-0.52
ISFPWNEIR	254-262	1161.6051	581.3062	387.8732		1161.606	581.3057			0.60	-0.86		
NISYSDKEFTIKPLDK	263-278	1897.9906	949.4989	633.335	475.2531		949.4984	633.34	475.2523		-0.53	7.89	-1.68
EFTIKPLDK	270-278	1090.6143	545.8108	364.2096		1090.617	545.8099	364.2089		2.29	-1.65	-1.92	
EFTIKPLDKK	270-279	1218.7093	609.8583	406.9079	305.4328	1218.707	609.8571	406.9068	305.4322	-1.72	-1.97	-2.70	-1.96
KIDVFK	279-284	749.4556	375.2314			749.455	375.2308			-0.80	-1.60		
IDVFKFNSSK	280-289	1184.631	592.8191	395.5485			592.8181	395.548		-1.69	-1.26		
LILQLcIGNHDLFmR	295-309	1858.9666	929.9869	620.3271	465.4971		929.9875	620.327			0.65	-0.16	
LSmEIEKEK	517-525	1122.5711	561.7892	374.8619			561.7877	374.8612			-2.67	-1.87	
LSmEIEK	517-523	865.4335	433.2204				433.2195				-2.08		
SKHLQEQLNELKTEIEALK	532-550	2251.2292	1126.118	751.0813	563.5628			751.0811	563.5618			-0.27	-1.77
HLQEQLNELKTEIEALK	534-550	2036.1022	1018.555	679.3723	509.781		1018.556	679.3724	509.7802		0.69	0.15	-1.57
TEIEALK	544-550	803.4509	402.2291			803.4506	402.2284			-0.37	-1.74		
ERETALDILHNENSDR	553-568	1911.9155	956.4614	637.9767	478.7343			637.9765	478.7336			-0.31	-1.46
ETALDILHNENSDR	555-568	1626.7718	813.8895	542.9288	407.4484		813.8898	542.928			0.37	-1.47	
VAFFEEL	589-595	854.4294	427.7184			854.4294				0.00			
Peptide sequence	Residues	MH+1	MH+2	MH+3	MH+4	MH+1	MH+2	MH+3	MH+4	+1	+2	+3	+4
		Theoretical (monoisotropic)				Observed (monoisotropic)				Error calculations (PPM)			