

## Supporting Information

# Proteolysis in microfluidic droplets: an approach to interface protein separation and peptide mass spectrometry

Ji Ji, Lei Nie, Liang Qiao, Yixin Li, Liping Guo, Baohong Liu, Pengyuan Yang and Hubert H. Girault

## Experimental

### Reagents and materials

Horse myoglobin (Myo), horse cytochrome *c* (Cyt-*c*), lysozyme *c*, bovine serum albumin (BSA), ammonium acetate, perfluorodecalin (PFD) (purity>95% GC) and trypsin were obtained from Sigma Chemical Co. Inc. (St. Louis, MO). Acetonitrile (ACN, HPLC grade) were purchased from Fisher Scientific (Ottawa, ON, Canada) and used without further purification. Formic acid (FA, analytical reagent, 98%) was purchased from BDH Chemicals, (Toronto, ON, Canada). Deionized water (>18 M $\Omega$ ·cm) was obtained from a Milli-Q Gradient water purification system (Millipore, Bedford, MA). Aquapel was purchased from PPG Industries. The C4 column (0.2\*150  $\mu$ m Magic 300 Å, 5  $\mu$ m) used for reversed-phased chromatography was purchased from Michrom Bioresources Inc. (Auburn, CA).

### Fabrication of the microchip

The microfluidic device was made by conventional soft lithographic techniques.<sup>1</sup> The basic architectures were designed with AutoCAD and transferred to high-resolution photomasks fabricated on transparencies. The negative photoresist SU-8 2035 (Microchem) was spin-coated onto 3-inch silicon wafers which were subsequently

patterned using a MJB4 mask aligner. Development was accomplished by immersion into 1-methoxy-2-propyl acetate (Sigma-Aldrich). A commercially available RTV615 PDMS kit (Momentive) containing the pre-polymer and a cross-linker was used in the recommended ratio of 10:1 (w/w). The mixture was poured on top of the patterned silicon wafers and degassed. The microchannels are 100  $\mu\text{m}$  in depth, 250  $\mu\text{m}$  in width and the length of reaction channels is 46 cm. After curing at 80  $^{\circ}\text{C}$  for 2 hours, the PDMS cast was cut and peeled off from the wafers. Microfluidic channel inlets and outlets were stamped out using a biopsy punch with an outer diameter of 2 mm. The microdevice was assembled by attaching the PDMS cast to a glass slide. The bonding strength was provided by pre-treating the contact surfaces with oxygen plasma for 60 s in the plasma cleaner (PDC-32G, Harrick Plasma). To enable production of water-in-oil drops, we make the device hydrophobic by treating with Aquapel (the commercially available water repellent agent Aquapel (PPG Industries)). This is achieved by flushing Aquapel through the device for a few seconds, flushing with air, and then baking the device at 65  $^{\circ}\text{C}$  for 20 min. Steel less capillary was used as MS emitter and glued onto the outlet of chip using 502 glue. The steel less capillary that was used as emitter was 250  $\mu\text{m}$  in OD and 100  $\mu\text{m}$  in ID.

### **Description of system for protein digestion and MS analysis**

The general design and dimension of the device can be found in **Figure S1**. A typical device presents 50  $\mu\text{m}$  wide channels for droplet formation and a 250  $\mu\text{m}$  wide channel for the later digestion stream. The flow-focusing technique was employed to obtain monodisperse droplets. Protein solutions (such as horse cytochrome c)

(dispersed in different concentrations of ACN solutions) were pumped from inlet A and trypsin solutions (dispersed in 40mM  $\text{NH}_4\text{HCO}_3$ ) were pumped from inlet B with the flow rate of 120  $\mu\text{L}/\text{h}$  respectively. The carrier phase (PFD) was pumped from inlet C with the flow rate of 300  $\mu\text{L}/\text{h}$ . The flow rates of the two phases can be changed for compatible for protein separation with HPLC. The uniform water-in-oil droplet was generated at the cross point I. Subsequently, the droplets travelled through the hydrophobic microchannels (the length can be changed for complete digestion reaction) while the digestion reaction took place inside droplets. The reaction was quenched with formic acid solution (1%, dispersed in 50%ACN solution) which was injected into Port D. In conventional digestion, the reaction time is 6 hours. The temperature of all digestion reactions was controlled at 37 °C. The digested peptides were directly analyzed with ESI-MS/MS or collected from the outlet and analyzed with MALDI-TOF-MS after separation.

### **Analysis of protein mixture by integrated platform**

A three-protein mixture was analyzed by this integrated platform. To achieve good compromise between the resolution for proteins separation and protein digestion, appropriate flow rates were employed. The C4 column was connected to HPLC system (Dionex LC-PACKING Series, USA) and operated at a flow rate of 120 $\mu\text{L}/\text{h}$ . A myoglobin, cytochrome c, lysozyme blend sample (0.13  $\mu\text{g}$  cyto-c, 0.92  $\mu\text{g}$  lyzome, and 0.16  $\mu\text{g}$  myoglobin) was analyzed. The solvent X was 5/95/0.1% ACN/ $\text{H}_2\text{O}$ /FA, the solvent Y was 95/5/0.1% ACN/  $\text{H}_2\text{O}$ /FA. The following separation gradient was set as follows: solvent X 98% (0 min)  $\rightarrow$  95% (3min)  $\rightarrow$  60% (25min)  $\rightarrow$  20% (28min)

→ 95% (35min). The corresponding ACN concentration gradient was: 6.8% (0 min) → 9.5% (3min) → 41% (25min) → 77% (28min) → 9.5% (35min). The corresponding ACN concentration in droplet reactors was 3.4% (0 min) → 4.75% (3min) → 20.5% (25min) → 38.5% (28min) → 4.75% (35min). The effluence was directly injected to microfluidic chip from the Port A and corresponding trypsin solution (2ng/μL) was injected from Port B with flow rate of 120μL/h. After digestion, the modified matrix of 50% ACN containing 1% FA was injected from the inlet D for altering the pH values and the peptides were analyzed with ESI-MS/MS directly. The length of the channel between T junction (channel D – main channel) and emitter tip is 27 mm.

### **Mass Spectrometry**

*MALDI-MS* All the samples were analyzed on an Applied Biosystems 5800 proteomics analyzer (MALDI-TOF MS, Applied Biosystem, USA). 0.5 μL of eluted solution was spotted onto the MALDI plate and dried under ambient conditions. And then another 0.5 μL of CHCA matrix (4 mg/ml α-cyano-4-hydroxycinnamic acid dissolved in 50% aqueous ACN, 0.1% TFA) was added. Positive ion MALDI-TOF mass spectra were acquired. The instrument was operated at an accelerating voltage of 20 kV. A 200 Hz pulsed Nd:YAG laser at 355 nm was used. All mass spectra were obtained with an accumulation of 2000 laser shots under a laser intensity of 4000 instrument units.

*ESI-MS/MS* The LTQ MS (Thermo corporation, USA) was run under full scan mode, the scan range was 350-1800 m/z, the spray voltage was 2.0 kv, the capillary

temperature was 200 °C, the inject time was set 50 ms. For the on-line digestion experiments, the LTQ MS was run under data-dependent mode, the first mass spectrometry stage was set full scan, the mass range was from 350-1800m/z, then the top ten strength ions were selected to do the second stage mass spectrometry, the CID energy was set 35 ev, the dynamic exclusion was used and the repeat duration time was 30 s, all the solvent ions were set in the reject list table and wouldn't be selected to do the MS<sup>2</sup>.

### **Data Analysis**

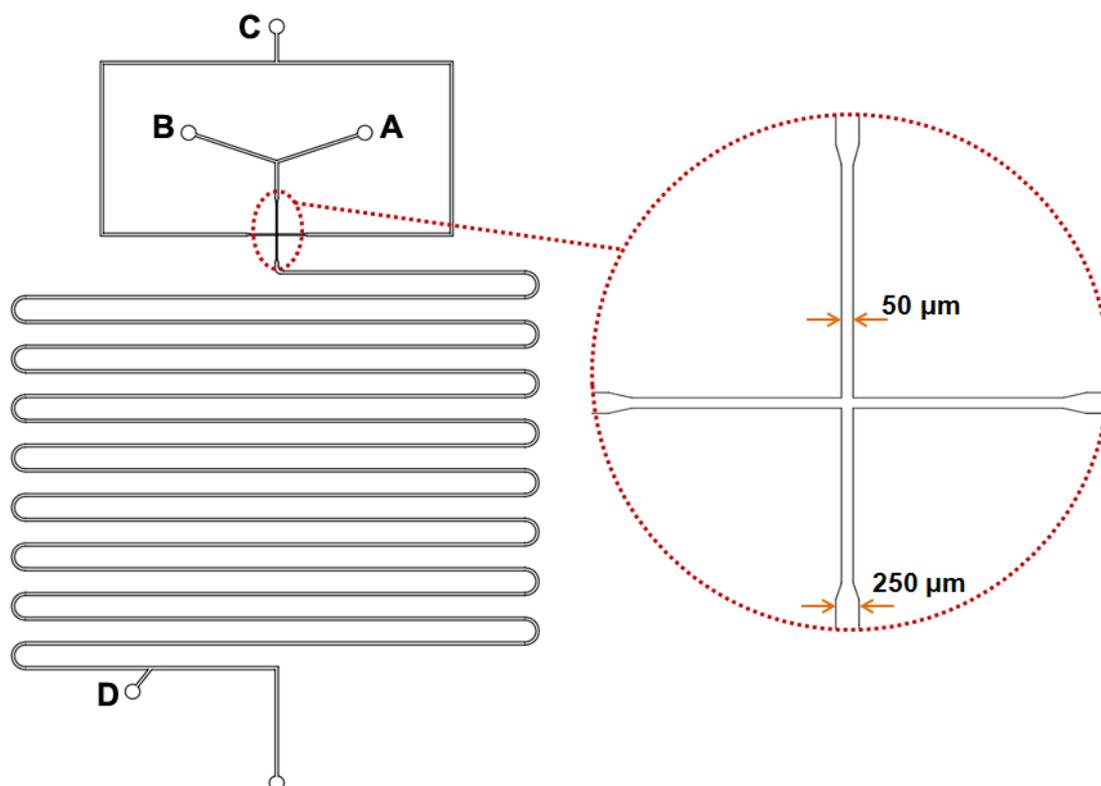
*MALDI analysis* GPS Explorer software from Applied Biosystems with *Mascot* as a search engine and *SwissProt* (version of 050303) as a database were used to identify proteins. All proteins were identified using the peptide fingerprint mass spectra. The peptide mass tolerance was set to 80 ppm and the tandem mass tolerance was set to 0.5 Da.

*ESI-MS/MS analysis* The on-line digestion data was searched in the swissport database using bioworks software. Trypsin was set as enzyme for database searching. Peptides were searched using fully tryptic cleavage constraints, and up to two missed internal sites were allowed for tryptic digestion. The parent ion's tolerance was 2.0 amu. The fragment ion's tolerance's was set 1amu, the missing site was set 2, using B and Y ions to calculated the series, and the modification was set methionine (+15.9994 Da). According to the filtrate criterion of sequest<sup>2</sup>, the identification of peptides were reliable when it's confidence more than 95%. The Xcorr Vs charges was 3.0>(4) > 2.5(+3), > 2.0 (+2), > 1.50 (+1),  $\Delta Cn > 0.1$ .

## Reference

1. D. C. Duffy, J. C. McDonald, O. J. A. Schueller, G. M. Whitesides, *Anal. Chem.* 1998, **70**, 4974-4984.
2. A. Keller, A. I. Nesvizhskii, E. Kolker, *Anal. Chem.* 2002, **74**, 5383 - 5392.

Figure SI-1. The general design and dimension of the device.



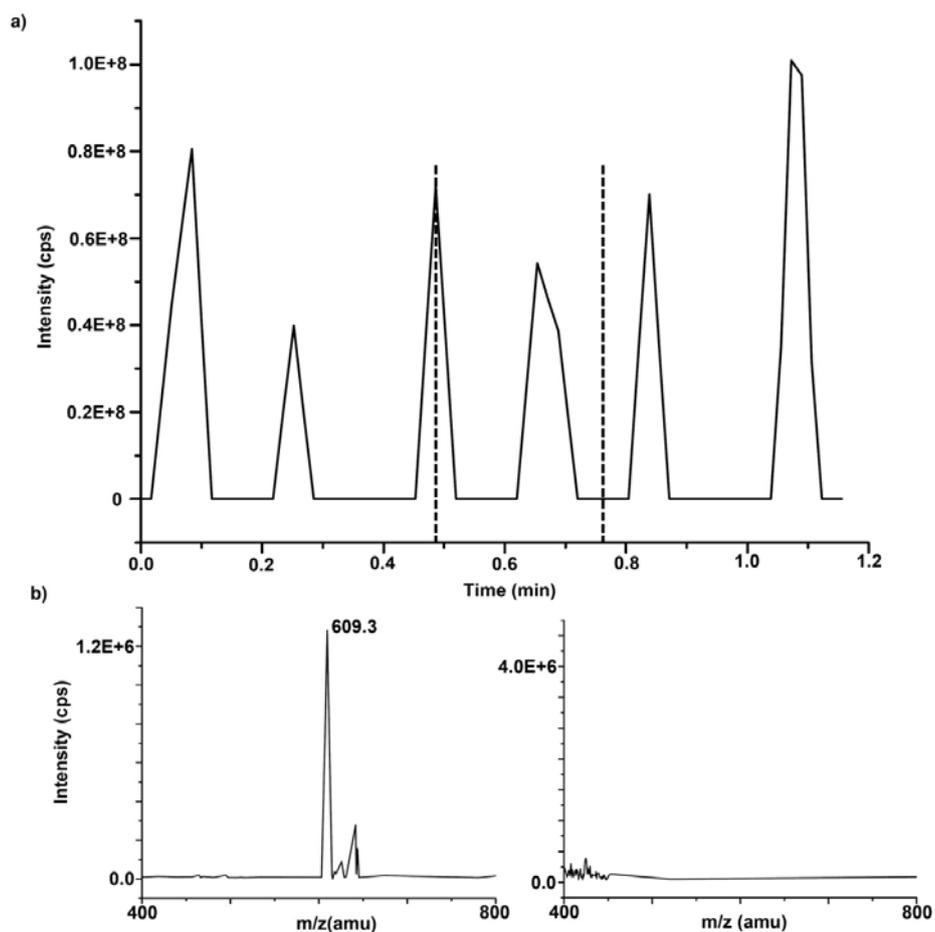


Figure SI-2. a) The total ion current of 100  $\mu\text{g}/\mu\text{L}$  Reserpine sample droplets infused at 240  $\mu\text{L}/\text{h}$  with Octadecafluorodecahydronaphthalene as oil phase; b) The mass spectra obtained from the apex of the peak indicated by dashed line (the left); Mass spectra obtained from the baseline showing the cross-containment between adjacent droplets (the right).

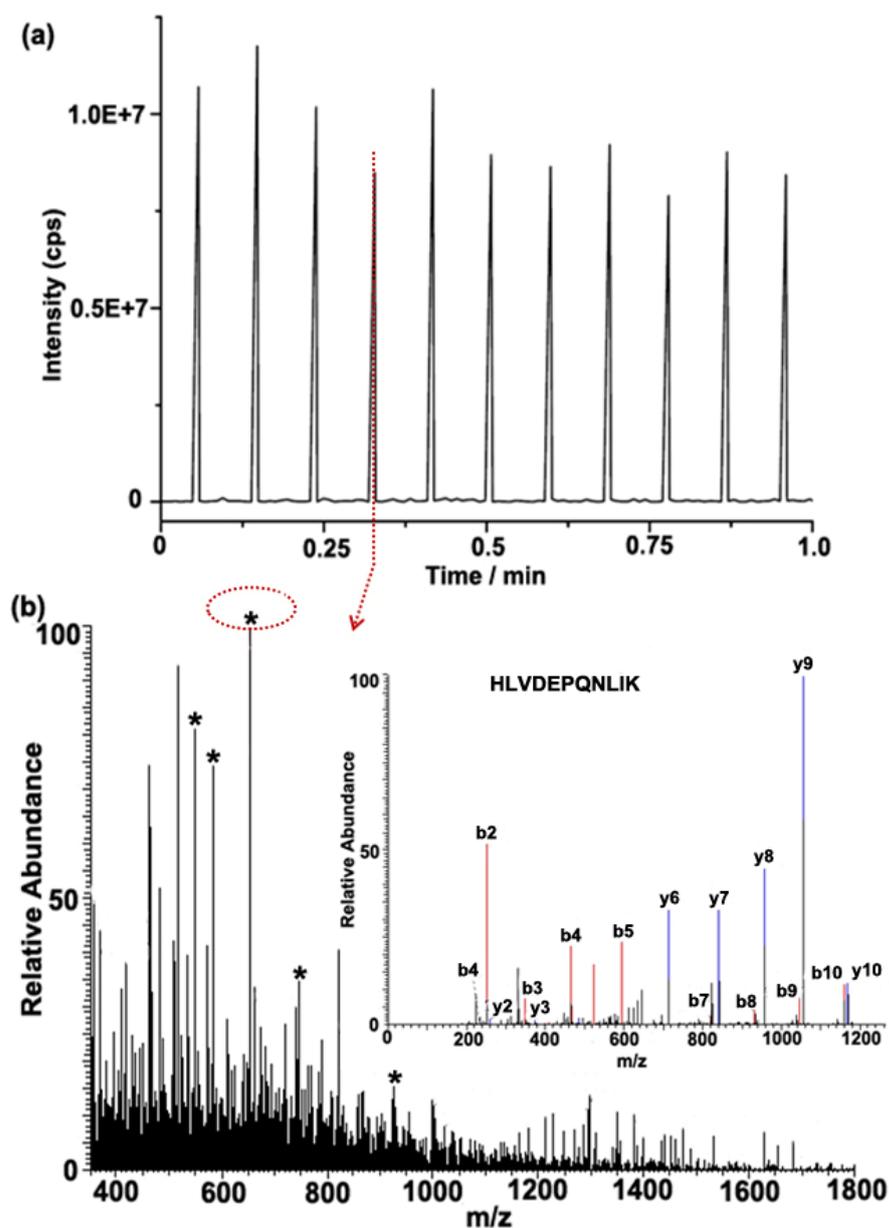


Figure SI-3. a) The base peak chromatography of the on-line protein (100 ng/ $\mu$ L BSA) identification with the microdroplet reactor-ESI-MS/MS; b) The mass spectrum of BSA obtained from the apex of the peak indicated by dashed line in (a); The insert is the MS/MS spectrum of the peak at  $m/z=653.2$  Th with the sequence of HLVDEPQNLK; 100 ng/ $\mu$ L BSA in 50% ACN/49.9% water/0.1% FA at a flow rate of 120  $\mu$ L/h was infused into the droplet-based microchip via inlet A, while the other conditions were same as those in Figure 2. The ESI voltage was 2.0 kV. \* The identified peptides from BSA.

Table SI-1 The detailed information of identified peptide of Myoglobin with microreactor digestion.

Sequence	Observed m/z	Theoretical MH <sup>+</sup>	Charge	$\Delta M$ (Da)	$\Delta M$ (ppm)
K.ALELFRNDIAAK.Y	680.95	1360.76	2	1.13	830.42
K.KGHHEAELKPLAQSHATK.H	660.20	1982.06	3	-1.47	-741.56
K.ASEDLKKHGTVVLTALGGILK.K	716.54	2150.25	3	-0.63	-294.02
A.ELKPLAQSHATK.H	661.54	1322.74	2	0.35	261.67
R.NDIAAKYKELGFQG-	776.39	1553.8	2	-1.01	-647.83
N.VWVGKVEADIAGHGQEVLR.L	692.71	2077.12	3	1.02	491.82
R.LFTGHPETLEKFDK.F	830.04	1661.85	2	-1.78	-1071.57
K.VEADIAGHGQEVLR.L	802.98	1606.85	2	-0.9	-557.43

Table SI-2 The detailed information of identified peptide of Myoglobin with solution digestion.

Sequence	Observed m/z	Theoretical MH <sup>+</sup>	Charge	$\Delta M$ (Da)	$\Delta M$ (ppm)
K.KGHHEAELKPLAQSHATK.H	660.21	1982.06	3	1.43	719.77
K.ASEDLKKHGTVVLTALGGILK.K	716.68	2150.25	3	0.21	99.68
A.ELKPLAQSHATK.H	661.58	1322.74	2	-0.42	-315.50
R.NDIAAKYKELGFQG.-	776.32	1553.80	2	1.16	743.88
N.VWVGKVEADIAGHGQEVLR.L	692.78	2077.12	3	-1.22	-587.88
K.ALELFRNDIAAK.Y	680.90	1360.76	2	-1.04	-765.51
R.LFTGHPETLEKFDK.F	830.04	1661.85	2	1.77	1067.09
K.VEADIAGHGQEVLR.L	802.96	1606.85	2	0.93	581.72
R.LFTGHPETLEK.F	635.69	1271.66	2	0.28	222.57
R.LFTGHPETLEKFDKFK.H	968.15	1937.02	2	0.72	370.02
K.ALELFRNDIAAKYKELGFQG.-	760.42	2283.21	3	1.95	855.46
L.NVWVGKVEADIAGHGQEVLR.L	731.04	2191.16	3	-1.96	-893.67
K.HGTVVLTALGGILK.K	690.42	1378.84	2	-2.00	-1447.53
K.HPGDFGADAQGAMTK.A	750.86	1502.67	2	0.95	631.71
K.HGTVVLTALGGILKK.K	753.84	1506.94	2	-0.75	-495.39

Table SI-3 The detailed information of identified peptides of BSA.

Sequence	Observed	Theoretical	Charge	$\Delta M$	$\Delta M$
	m/z	MH <sup>+</sup>		(Da)	(ppm)
R.DTHKSEIAHR.F	596.94	1193.6	2	0.28	234.58
R.PCFSALTPDETYVPK.A	833.65	1667.8	2	-0.51	-305.79
K.HLVDEPQNLIK.Q	653.23	1305.72	2	0.74	566.74
K.LGEYGFQNALIVR.Y	740.12	1479.80	2	0.43	290.58
R.RHPEYAVSVLLR.L	480.03	1439.81	3	0.29	201.42

Table SI-4 The detailed information of identified peptides of Cytochrome c.

Sequence	Observed	Theoretical	Charge	$\Delta M$	$\Delta M$
	m/z	MH <sup>+</sup>		(Da)	(ppm)
K.IFVQK.C	636.36	636.39	1	-0.03	47.29
K.GITWKEETLMEYLENPKK.Y	736.86	2209.12	3	1.47	665.42
K.TGPNLHGLFGR.K	584.58	1168.62	2	0.54	462.08
K.MIFAGIK.K	779.45	779.45	1	0.0018	2.31
K.TEREDLIAYLKK.A	493.28	1478.82	3	1.01	682.98
K.MIFAGIKK.K	454.04	907.54	2	0.54	595.02
K.GITWKEETLMEYLENPK.K	1040.69	2081.03	2	0.35	168.19
K.TEREDLIAYLK.K	675.15	1350.73	2	-0.44	-325.75
K.TGQAPGFTYTDANK.N	735.55	1470.69	2	0.4	271.98
R.EDLIAYLK.K	964.41	964.53	1	-0.12	-124.41
R.KTGQAPGFTYTDANK.N	799.59	1598.78	2	0.4	250.19
K.KIFVQK.C	381.43	762.49	2	0.36	472.14
R.EDLIAYLKK.A	546.09	1092.63	2	-0.46	-421.00

Table SI-5 The detailed information of identified peptide of Myoglobin.

Sequence	Observed m/z	Theoretical MH <sup>+</sup>	Charge	$\Delta M$ (Da)	$\Delta M$ (ppm)
K.HGTVVLTALGGILKK.K	753.84	1506.94	2	0.75	495.39
K.KGHHEAELKPLAQSHATK.H	660.20	1982.06	3	-1.47	-741.56
K.ASEDLKKHGTVVLTALGGILK.K	716.54	2150.25	3	-0.63	-294.02
K.ALELFRNDIAAKYKELGFQG-	760.42	2283.21	3	-1.95	-856.13
L.NVWGKVEADIAGHGQEVLR.L	731.05	2191.16	3	1.98	902.65
N.VWGKVEADIAGHGQEVLR.L	692.71	2077.12	3	1.02	491.82
K.HGTVVLTALGGILK.K	690.42	1378.84	2	1.99	1447.53
R.LFTGHPETLEKFDK.F	830.04	1661.85	2	-1.78	-1071.57
K.VEADIAGHGQEVLR.L	802.98	1606.85	2	-0.9	-557.43

Table SI-6 The detailed information of identified peptides of lysozyme.

Sequence	Observed	Theoretical	Charge	$\Delta M$	$\Delta M$
	m/z	MH <sup>+</sup>		(Da)	(ppm)
K.AQEMDGFGGYSLANWV	852.61	2558.03	3	-0.19	-74.28
CMAEYES.N					
R.SSSNACNIMCSKLLDENIDDD	987.17	2961.31	3	0.2	67.54
ISCAKR.V					
A.CNIMCSKLLDENIDDDISCAK.R	781.00	2343.03	3	-0.04	-17.07
K.CELAHKLKAQEMDGFGG	856.40	2568.2	3	1.01	393.27
YSLANW.V					
L.ANWVCMAEYESNFNTRAFN	816.90	2452.08	3	-1.38	-562.79
GK.N					