

Manual - HeXicon

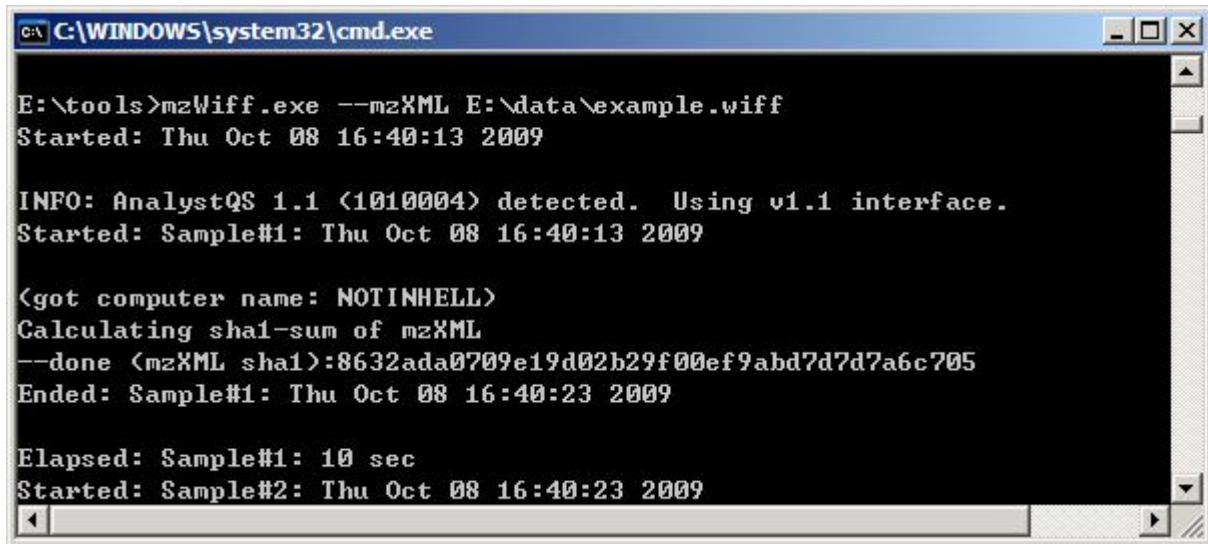
How to install

1. Normally HeXicon does not require any additional package or library. To run HeXicon, simply double click the hexicon.exe.
2. If you encounter an error message of "This application has failed to start because the application configuration is incorrect. ...", please run the vcredist_x86.exe to update the Visual Studio runtime components. This package is also available from the Microsoft Download Center:
<http://www.microsoft.com/downloads/details.aspx?FamilyId=32BC1BEE-A3F9-4C13-9C99-220B62A191EE&displaylang=en>

How to convert data format (.wiff to .mzXML)

1. Download mzWiff:
<http://sourceforge.net/projects/sashimi/files/mzWiff%20%28Analyst%20converter%29/mzWiff-4.3.1/mzWiff-4.3.1.zip/download>
2. Execute the mzWiff.exe in Windows Command Line using the following syntax:
mzWiff --mzXML [full path of the .wiff file]

For example:



```
C:\WINDOWS\system32\cmd.exe

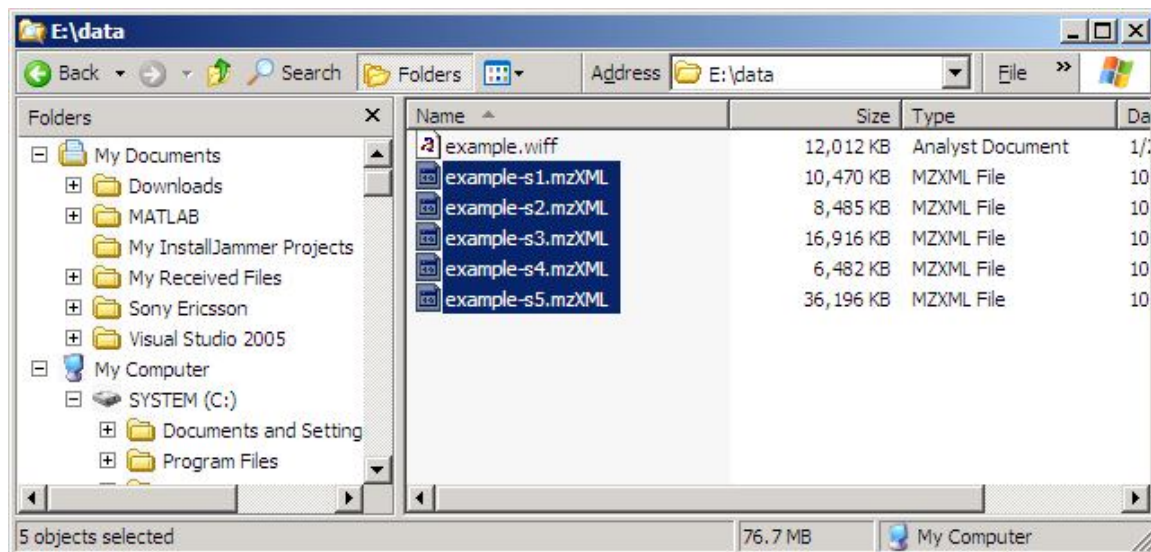
E:\tools>mzWiff.exe --mzXML E:\data\example.wiff
Started: Thu Oct 08 16:40:13 2009

INFO: AnalystQS 1.1 <1010004> detected. Using v1.1 interface.
Started: Sample#1: Thu Oct 08 16:40:13 2009

<got computer name: NOTINHELL>
Calculating sha1-sum of mzXML
--done <mzXML sha1>:8632ada0709e19d02b29f00ef9abd7d7d7a6c705
Ended: Sample#1: Thu Oct 08 16:40:23 2009

Elapsed: Sample#1: 10 sec
Started: Sample#2: Thu Oct 08 16:40:23 2009
```

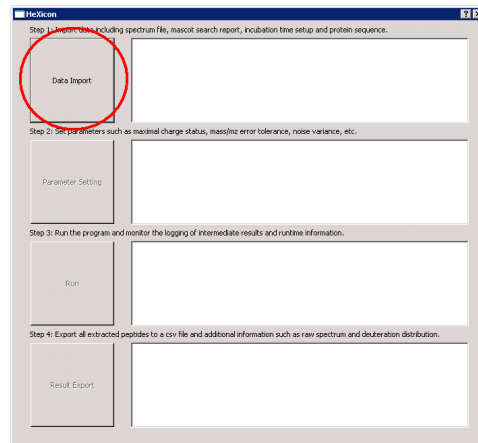
3. When the conversion is finished, the exported mzXML files can be found at the same directory as the source .wiff file.



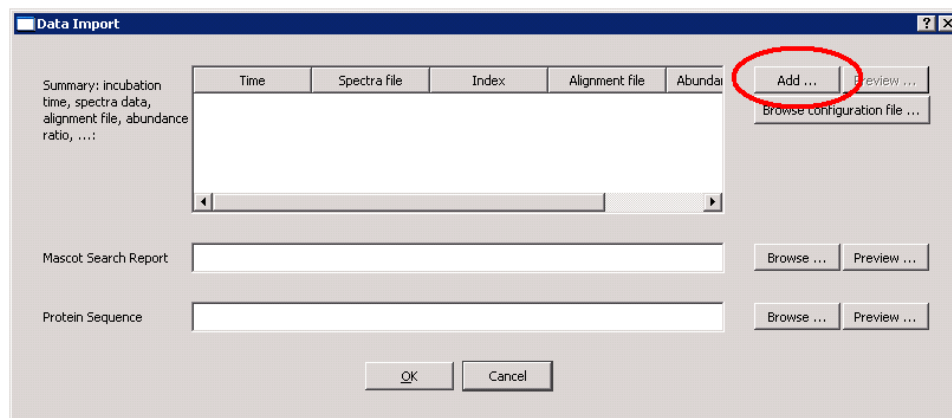
How to use HeXicon

1 Data Import

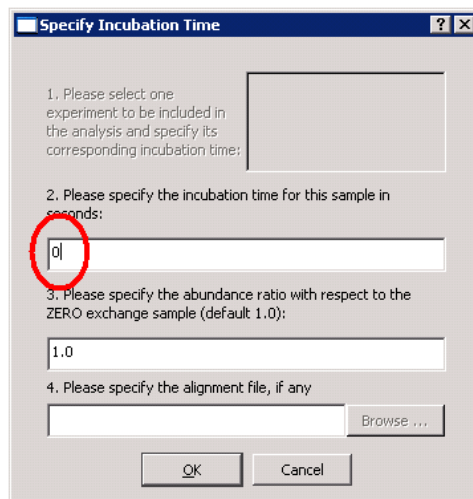
- 1) Execute hexicon.exe to launch the program. Click button “Data Import” to open the “Data Import” dialog.



- 2) To import ms spectrum data, click button “Add ...” and select the target mzXML file.



Specify the corresponding incubation time. The spectrum data from the native sample, i.e. zero exchange, is required.



Repeat this procedure to import the data of all the other incubation times.

- 3) To import MS/MS data, launch your web browser and direct it to the “**Mascot Search Results**” page, e.g.
http://1bcpc1.zmbh.uni-heidelberg.de/mascot/cgi/master_results.pl?file=../data/20080317/F017638.dat

Mascot Search Results

User : XXXXXXXXXX
 Email : XXXXXXXXXX
 Search title : C:\PE Sciex Data\Projects\Default\Data\LC-MS\Chip-Pepsin.wiff (sample number 1)
 MS data file : C:\DOKUME~1\CHRIST~1\LOKALE~1\Temp\masA.tmp
 Database : NCBI nr 20080116 (5853324 sequences; 2017905186 residues)
 Taxonomy : Homo sapiens (human) (199412 sequences)
 Timestamp : 24 Jan 2008 at 11:23:17 GMT
 Protein hits : [gi|56181387](#) STIP1 homology and U-box containing protein 1 [Homo sapiens]
[gi|4506625](#) ribosomal protein L27a [Homo sapiens]
[gi|4885413](#) histidine triad nucleotide binding protein 1 [Homo sapiens]
[gi|34932414](#) non-POU domain containing, octamer-binding [Homo sapiens]
[gi|10436331](#) unnamed protein product [Homo sapiens]
[gi|89041690](#)
[gi|4506439](#) retinoblastoma binding protein 7 [Homo sapiens]
[gi|33870977](#) RNF157 protein [Homo sapiens]
[gi|74762479](#) Cyclin-dependent kinase-like 3 (Serine/threonine protein kinase NKIAMRE)
[gi|134152372](#) CACNA1S protein [Homo sapiens]
[gi|16550104](#) unnamed protein product [Homo sapiens]
[gi|76779832](#) C10orf137 protein [Homo sapiens]
[gi|71480169](#) zinc finger protein 406 isoform TR-ZFAT [Homo sapiens]

Under “**Peptide Summary Report**”, select the list of identified peptides (**WITHOUT the header**) and copy them to the clipboard (Windows, Ctrl+C)

Peptide Summary Report

Format As: Peptide Summary [Help](#)

Significance threshold $p < 0.05$ Max. number of hits 20

Standard scoring ☒ MudPIT scoring ☐ Ions score or expect cut-off 20 Show sub-sets 0

Show pop-ups ☒ Suppress pop-ups ☐ Sort unassigned Decreasing Score Require bold red ☒

Select All Select None Search Selected ☐ **Error tolerant** Archive Report

1. [gi|56181387](#) **Mass:** 34834 **Score:** 1635 **Queries matched:** 59 **emPAI:** 1.51
 STIP1 homology and U-box containing protein 1 [Homo sapiens]
☐ Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Score	Expect	Rank	Peptide
5	404.7020	807.3894	807.3949	-0.0055	0	30	1e-02	5	F.FLGGQCLE
14	411.2276	820.4406	820.4331	0.0075	0	34	1e-02	5	F.FLGGQCLE
19	425.2080	848.4014	848.3884	0.0130	0	34	1e-02	5	F.FLGGQCLE
30	442.2798	882.5450	882.5287	0.0163	0	43	1e-02	5	F.FLGGQCLE
38	468.2229	934.4312	934.4032	0.0280	0	21	1e-02	5	F.FLGGQCLE
39	469.2373	936.4601	936.4375	0.0226	0	23	4.6e+02	5	F.FLGGQCLE.M

Copy
 Select All
 Search Web for "S 404.7020 807...."
 View Selection Source

Open a text file editor, paste the clipboard content into it (Windows, Ctrl+V) and save the file as a .txt file.

2006-08-23_CHIP_calsubset.pir.txt - Notepad

	File	Edit	Format	View	Help
5	404.70	807.39	807.39	-0.01	0
14	411.23	820.44	820.43	0.01	0
19	425.21	848.40	848.39	0.01	0
25	438.23	874.44	874.47	-0.03	0
30	442.28	882.54	882.53	0.02	0
32	450.27	898.53	898.49	0.04	0
38	468.22	934.43	934.40	0.03	0
39	469.24	936.46	936.44	0.02	0
44	481.76	961.51	961.47	0.03	0
49	491.81	981.61	981.60	0.01	0
51	494.73	987.44	987.48	-0.04	0
54	500.28	998.55	998.52	0.03	0
57	502.25	1002.48	1002.46	0.02	0
58	502.25	1002.48	1002.46	0.02	0
59	503.29	1004.56	1004.54	0.02	0

Click "Browse ..." button in "Data Import" dialog and open the .txt file you just saved.

Data Import

Summary: incubation time, spectra data, alignment file, abundance ratio, ...:

	Time	Spectra file	
1	1	E:/data/hx.chip.christian/2006-08-23_CHIP_calsubset-s3.mzXML	0
2	30	E:/data/hx.chip.christian/2006-08-23_CHIP_calsubset-s1.mzXML	0
3	300	E:/data/hx.chip.christian/2006-08-23_CHIP_calsubset-s2.mzXML	0

Mascot Search Report

Protein Sequence

Buttons: Add ..., Preview ..., Browse configuration file ..., Browse ..., Preview ..., Browse ..., Preview ..., OK, Cancel

- To import protein sequence, open a text file editor; paste the protein sequence into it and save the file as a .txt file.

2006-08-23_CHIP_calsubset.sequence.txt - Notepad

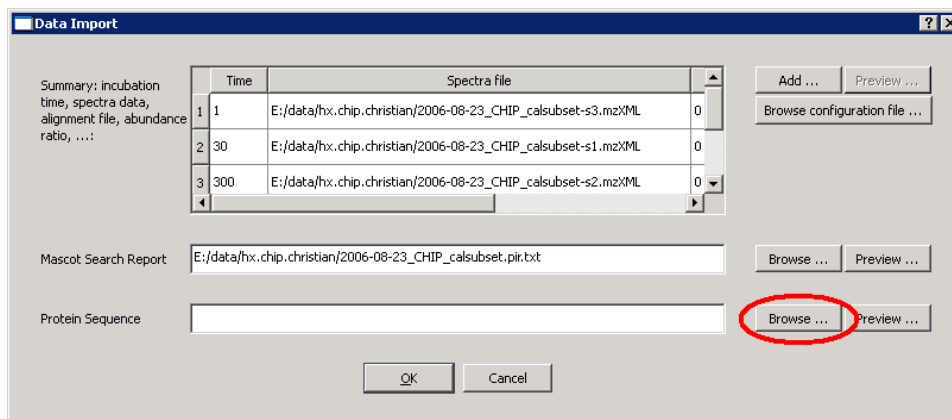
File Edit Format View Help

```

MKGKEEKEGGARLGAGGGSPEKSPSAQELKEQGNRLFVGRKYPEAAAACYGRAIT
RNPLVAVYYTNRALCYLKMQQHEQALADCRRALELDGQSVKAHFFLGQCQLEME
SYDEAIANLQRAYSLAKEQRLNFGDDIPSALRIAKKKRWNSIEERRIHQESLH
SYLSRLIAAERERELEECQRNHEGDEDDSHVRAQQACIEAKHDKYMADMDELFS
QVDEKRKKRDIPDYLCGKISFELMREPCITPSGITYDRKDIEEHLQRVGHFDPV
TRSPLTQEQLIPNLAMKEVIDAFISENGWVEDY

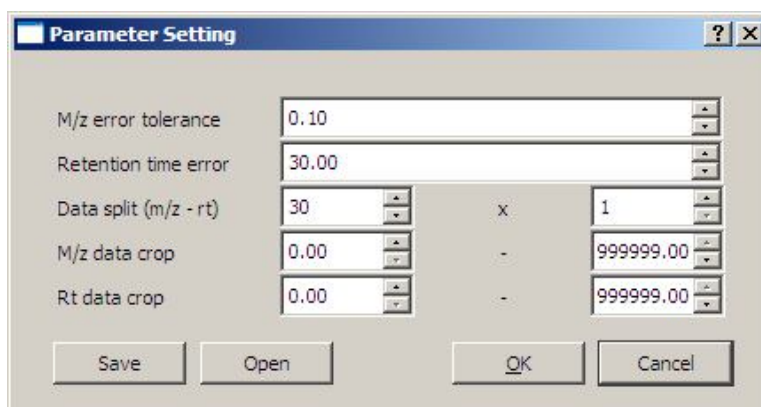
```

Click "Browse ..." button in "Data Import" dialog and open the .txt file which you just saved.



2 Parameter Setting

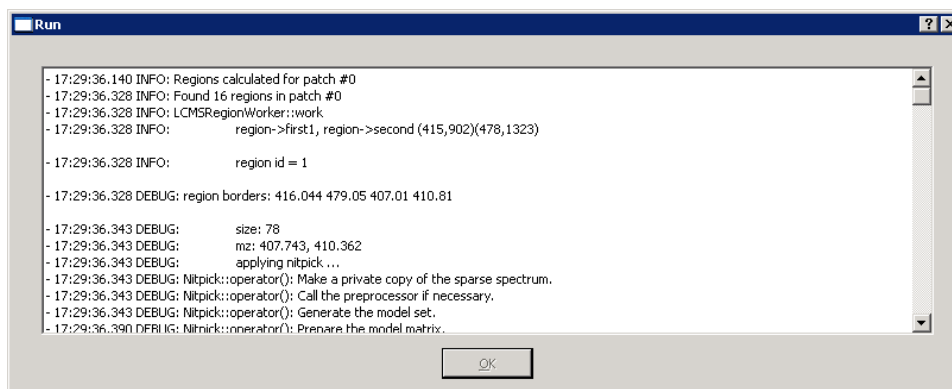
- 1) Click "Parameter Setting" button to open the parameter setting dialog.



- 2) Please refer to **Appendix I** for the detailed specification of parameters.

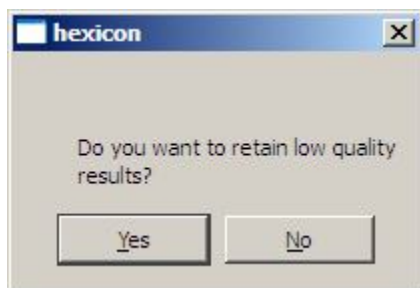
3 Run

- 1) To start the analysis, click "Run" button. The runtime log information will be updated in real time. When the entire analysis procedure is finished, the "OK" button will be enabled.



4 Result Export

- 1) Use the "Export Result" button to export all results to a .csv file. Please refer to **Appendix II** for the specification of the exporting format.
- 2) Select "Yes" in the following dialog to export low quality results and "No" to skip them.



Appendix I

Parameter Specifications

1 M/z error tolerance

HeXicon automatically recovers peptide candidates by matching the picked peaks to sub-sequences of the protein. The matching is determined via the m/z difference. “M/z error tolerance” controls the upper limit of the m/z difference between a picked peak and the theoretical m/z value of a peptide candidate.

2 Retention time window

“Retention time window” defines the retention time range in which HeXicon searches for corresponding signals from different incubation time that belong to the same peptide sequence.

3 Data split (m/z – r/t)

The entire spectrum dataset is too large to fit into the memory of a normal PC. “Data split” defines the grid that separates the original dataset (with overlaps) along both the m/z dimension and the retention time dimension.

4 M/z data crop

“M/z data crop” defines the region along the m/z dimension in which the analysis is performed. All signals outside this region are discarded.

5 Rt data crop

“Rt data crop” defines the region along the retention time dimension in which the analysis is performed. All signals outside this region are discarded.

Appendix II

Result Export Format

Results from the HeXicon analysis are exported into one “.csv” file which includes the following fields:

- 1 Peptide
- 2 Stoichiometry
- 3 Charge
- 4 Monoisotopic Mass
- 5 Monoisotopic Mz
- 6 Protein identification score (e.g. Mascot)
- 7 Number of exchangeable hydrogen
- 8 Peptide length
- 9 Deuteration distribution file name
- 10 Position in the protein sequence
- 11 Estimated quality score
- 12 Average deuteration value over incubation time
- 13 Retention time value over incubation time
- 14 Abundance value over incubation time
- 15 Region id of the corresponding peak groups over incubation time

st11_relGain=0.5,0.0001.split=20x1,maxCharge=5.csv - Microsoft Excel													
Home Insert Page Layout Formulas Data Review View Add-Ins													
Clipboard		Font		Alignment		Number		Styles		Cells		Editing	
Calibri 11		A A				\$ %		Conditional Formatting		Insert		Σ	
Paste		B I U						Format as Table		Delete		Sort & Find & Filter Select	
								Cell Styles		Format			
L1													
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	Peptide	Stoichiom	Charge	MonoMas	MonoMz	Score	Exchangat	Length	Distributi	Position	Quality	1 s	15 s
2	AGIIRTGR	(1)66 (12):	2	842.509	421.254	0	7	8	AGIIRTGR	581<->589	3	422.279	42
3	DASNKAA	(1)45 (12):	1	675.319	675.318	0	6	7	DASNKAA	103<->110	3	676.345	67
4	TAKDYDKAIEL	(1)91 (12):	2	1265.65	632.825	36	10	11	TAKDYDK/	16<->27	3	633.84	63
5	IGYKQNPQAIGQDL	(1)109 (12)	2	1543.8	771.899	53	12	14	IGYKQNP(164<->178	3	772.922	77
6	MSLTADQYKQGNAAFTAKDYI	(1)205 (12)	3	3020.44	1006.81	29	26	27	MSLTADQY	0<->27	3	1007.85	10
7	DAARTKDAEVNNGSSAREIDQL	(1)158 (12)	3	2359.14	786.378	35	21	22	DAARTKD,	487<->509	3	787.39	79
8	IENLKNPKTSEMMKDPQL	(1)166 (12)	3	2243.17	747.721	36	16	19	IENLKNKP	141<->160	3	748.754	75
9	ETLDAARTKDAEVNNGSSAREID	(1)183 (12)	3	2702.31	900.77	38	24	25	ETLDAART	484<->509	3	901.792	90
10	YKARQFDEA	(1)74 (12):	2	1126.54	563.27	39	8	9	YKARQFDI	272<->281	3	564.295	56
11	IEHYNKAWELHKDITYLNNRAA	(1)195 (12)	3	2898.44	966.147	36	23	24	IEHYNKAV	281<->305	3	967.17	97
12	MSLTADQYKQGNAAF	(1)116 (12)	2	1772.8	886.401	29	15	16	MSLTADQY	0<->16	3	887.422	89
13	VDEDDSKIEADKEKAEGNKF	(1)151 (12)	3	2266.06	755.353	42	19	20	VDEDDSKI	252<->272	3	756.379	76
14	IENLKNPKTSEM	(1)114 (12)	2	1530.81	765.403	65	11	13	IENLKNKP	141<->154	3	766.425	77
15	VAKLIGYKQNPQAIGQDL	(1)146 (12)	2	1955.08	977.541	83	16	18	VAKLIGYK	160<->178	3	978.568	98
16	EGKEYFTKSDWPNNAVKAITEM	(1)164 (12)	3	2493.15	831.05	21	19	21	EGKEYFTK	401<->422	3	832.075	83
17	IMQDPVM	(1)60 (12):	1	832.382	832.382	33	5	7	IMQDPVM	542<->549	3	833.401	83
18	FTKSDWPNNAVKA	(1)94 (12):	2	1362.69	681.346	59	10	12	FTKSDWPI	406<->418	3	682.367	68
19	TKLRNAEKLKAEAEA	(1)145 (12)	3	1928.07	642.689	27	16	17	TKLRNAEK	372<->389	3	643.72	64
20	AIMQDPVM	(1)65 (12):	2	903.419	451.709	28	6	8	AIMQDPV	541<->549	3	452.74	45
21	DASNKAAKEGLDQVHRTQQAR	(1)217 (12)	3	3144.6	1048.2	25	27	29	DASNKAA	103<->132	3	1049.23	10
22	IEHYNKAWEL	(1)87 (12):	2	1301.64	650.82	48	9	10	IEHYNKAV	281<->291	3	651.853	65
23	IKRAPEDARGYSNRAAAL	(1)139 (12)	2	1958.04	979.022	25	16	18	IKRAPEDA	422<->440	3	980.049	98
24	VDEDDSKIEADKEKAEGNKFYK	(1)223 (12)	3	3374.59	1124.86	22	28	29	VDEDDSKI	252<->281	3	1125.89	11
25	DEAESNYKKAELDASN	(1)125 (12)	3	1895.87	631.958	0	16	17	DEAESNYK	90<->107	3	633.046	63
26	MGVDI NNA	(1)154 (12):	1	778.335	778.335	21	6	7	MGVDI NNA	191<->198	3	779.359	78
st11_relGain=0.5,0.0001.split=2													
Ready													
Count: 4 100%													