A Parser for mzXML, mzData and mzML files

Bernd Fischer, Steffen Neumann

December 19, 2010

Contents

A short example sequence to read data from a mass spectrometer. First open the file.

```
> library(mzR)
> library(msdata)
> mzxml <- system.file("threonine/threonine_i2_e35_pH_tree.mzXML",
+                      package = "msdata")
> ramp <- rampOpenFile(mzxml)
```

We can obtain different kind of header information.

```
> ramp$getRunInfo()

$scanCount
[1] 55

$lowMZ
[1] 1.58101e-322

$highMZ
[1] 2.371515e-322

$startMZ
[1] 1.513193e-316

$endMZ
[1] 6.928674e-310

$dStartTime
[1] 0.3485

$dEndTime
[1] 390.027

> ramp$getInstrumentInfo()

$manufacturer
[1] "Thermo Scientific"

$model
[1] "LTQ Orbitrap"

$ionisation
[1] "ESI"

$analyzer
[1] "FTMS"
```

```
$detector
[1] "unknown"
```

```
> SH <- ramp$getAllScanHeaderInfo()
> head(SH)
```

```
  seqNum acquisitionNum msLevel peaksCount totIonCurrent retentionTime
1      1              1       1        684     341427000        0.3485
2      2              2       2        432     160473000        5.8561
3      3              3       2        340      58862000       12.5000
4      4              4       3        273      30770400       19.7568
5      5              5       3        238       5291800       26.6056
6      6              6       3        140      14294000       34.0324
  basePeakMZ basePeakIntensity collisionEnergy ionisationEnergy   lowMZ
1   120.0660         211860000               0                0 50.3254
2   120.0660         139169000               0                0 50.4459
3   102.0560          27036600              35                0 50.0658
4   102.0550          26736000               0                0 50.2843
5    56.0497           2188950              35                0 50.1967
6    74.0605          12518700               0                0 50.2254
    highMZ precursorScanNum precursorMZ precursorCharge precursorIntensity
1 298.6730                0      0.0000               0                  0
2 134.1380                0    120.0661               1          210140000
3 134.1430                0    120.0661               1          210140000
4 114.6510                0    102.0600               0            1441880
5 114.9470                0    102.0600               0            1441880
6  84.8802                0     74.0600               0            1556090
  mergedScan mergedResultScanNum mergedResultStartScanNum
1          0                   0                       0
2          0                   0                       0
3          0                   0                       0
4          0                   0                       0
5          0                   0                       0
6          0                   0                       0
  mergedResultEndScanNum
1                      0
2                      0
3                      0
4                      0
5                      0
6                      0
```

Read a single spectrum from the file.

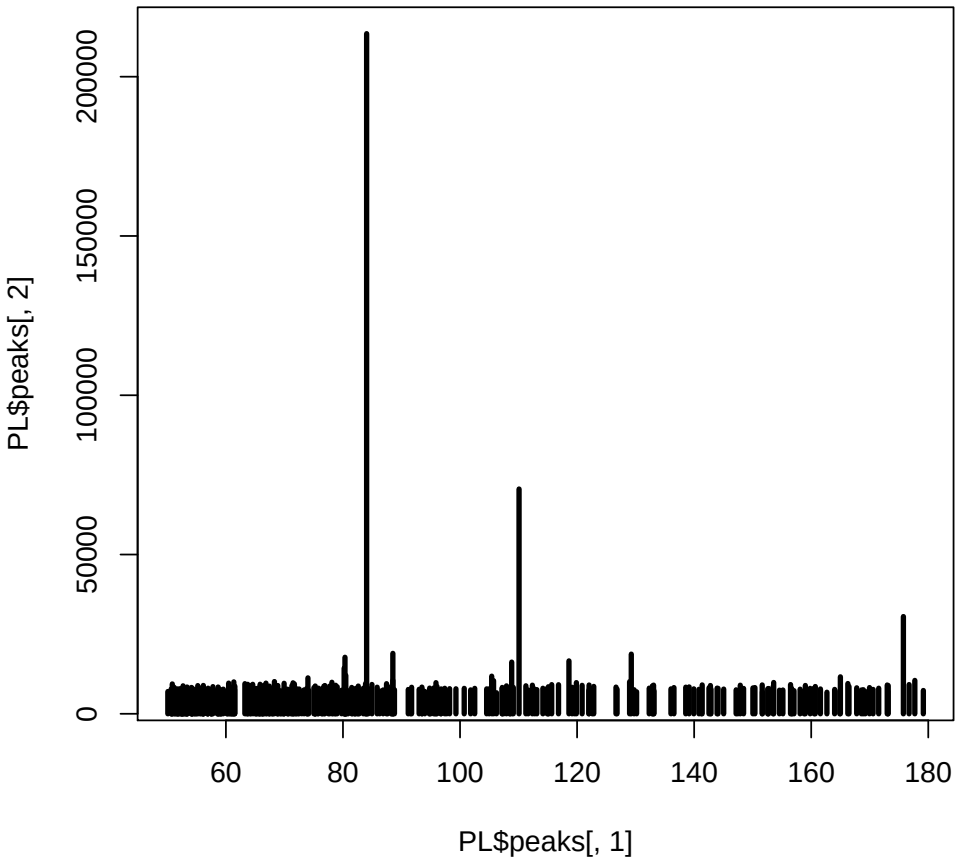
```
> PL <-ramp$getPeakList(10)
> PL$peaksCount
```

```
[1] 317
```

```
> head(PL$peaks)
```

```
         [,1]     [,2]
[1,] 50.08176 6984.858
[2,] 50.62267 7719.419
[3,] 50.70530 7185.290
[4,] 50.73298 7509.140
[5,] 50.83848 9366.624
[6,] 50.88303 8012.808
```

```
> plot(PL$peaks[,1], PL$peaks[,2], type="h", lwd=3)
```



You should close the file when not needed any more. This will release the memory of cached content.

```
> ramp$close()
```