

BIONUMERICS Tutorial: Importing spectrum data: peak lists

1 Aim

Comprehensive tools for the import of spectrum data, both raw spectrum data as processed spectrum data are incorporated into BIONUMERICS. In this tutorial the import of peak lists (= processed spectrum data) is illustrated with a sample data set.

2 Sample data

As an exercise, we will import a set of MALDI-TOF peak lists from different isolates and species. The set can be downloaded from the Applied Maths website: go to https://www.applied-maths.com/download/sample-data and click on "Demo spectra peak lists". When the download is complete, unzip the file.

3 Preparing a sample database

3.1 Creating a new database

- 1. Double-click on the BIONUMERICS icon (Left) on the desktop.
- 2. In the *BIONUMERICS Startup* window, press the Bionumber of the *New database* wizard.
- 3. Enter a database name, e.g. "Demo Spectra".
- 4. Click <*Next*>.

A new dialog box pops up, asking whether to create a new relational database for data storage or to use an existing one.

5. Leave the default option *Create new* enabled and press <*Next*>.

The next dialog asks which database engine should be used for storing data.

6. Select the default option and press < *Finish*>.

The Main window opens with an empty database.

3.2 Creating a spectrum type experiment

Before importing spectrum data, we will first create a spectrum experiment type.

7. In the *Main* window, click on + in the toolbar of the *Experiment types* panel and select **Spec***trum type* from the list (see Figure 1). Press < **OK** >.

Create a new experiment type	?	×
Select what kind of experiment type ye	ou want to	create:
Fingerprint type		^
Character type		
Sequence type		
Matrix type		
Î∠ Trend data type		
A Spectrum type		
≵ Sequence read set type		
🌥 Composito data pat		*
ОК	Ca	ncel

Figure 1: Create a new Spectrum type.

8. Enter a name, for example **Maldi**, leave the units for the horizontal and vertical axis at their defaults and press <*Next*> (see Figure 2).

New spectrum type	?	×
Basic Settings Provide basic information about the experiment type		
Name: Maldi Units Enter the axis unit strings as appropriate for your data. Horizontal axis unit: m/z Vertical axis unit: Intensity		
< Back Next >	Can	icel

Figure 2: Basic settings.

Three predefined preprocessing templates are included: *Preprocessing (Default)*, *Preprocessing (Relaxed)*, and *Preprocessing (Strict)*. The settings of each template can be changed in the *Spectrum Preprocessing* window and saved to the database. Since we will import already preprocessed spectrum data in this tutorial, the choice of a preprocessing template is irrelevant here.

9. Press < *Finish*> to complete the creation of the new spectrum type experiment.

The *Experiment types* panel now lists the spectrum type **Maldi** (see Figure 3).

4 Importing spectra

1. Select *File* > *Import...* (, Ctrl+I) to call the *Import* dialog box (see Figure 4).

କ୍ଷ	·	+ 🖻) ⊗	Ê	∽	1	t	<all experiment="" types=""></all>
#	ŧ	Name		Туре				
] 🔺 🗍	1	Maldi		Spectru	m types			
		manar		opeend	in types			

Figure 3: The Experiment types panel.

In this section we will import mzML files containing peak lists (see 2) in our BIONUMERICS database. mzML is a standardized format for exchange of spectrum type data.

2. Select the *Import mzML data* option, listed under the topic *Spectrum type data* and press <*Import*>.

More peak list import routines can be found under the topic **Spectrum type data**.

Import	? ×
Import data into level 'Raw spectra' Select the kind of data to import:	Import mzIML files and link to new or existing database entries.
	Import Close

Figure 4: The Import mzML data option in the Import dialog box.

- 3. Browse to the folder, select all mzml files in this folder and press <*Open*> and <*Next*> (see Figure 5).
- 4. In the *Import template* wizard page of the wizard, press < *Create new* >.

As an exercise, we will link the file name to the BIONUMERICS Key field.

- 5. Double-click on the last row available in the grid or press < *Edit Destination*>.
- 6. Select *Key* in the *Edit data destination* dialog box (see Figure 6).
- 7. Press <*OK*>.

The grid is updated.

In the *Import template* dialog box, press < *Preview*> and verify the preview of the import (see Figure 7). Close the preview and press < *Next*> and <*Finish*>.

Input	/L file(s) to import	
Select mz	inc me(s) to import	
Select file(s):	C:\Users\Public\D\SpeciesA_isolate1.mzml	
	C:\Users\Public\D\SpeciesA_isolate2.mzml Delete	
	C:\Users\Public\D\SpeciesA_isolate3.mzml	
	C:\Users\Public\D\SpeciesA_isolate4.mzml	
	C:\Users\Public\D\SpeciesA_isolate5.mzml	
	C:\Users\Public\D\SpeciesA_isolate6.mzml	
	C:\Users\Public\D\SpeciesB_isolate1.mzml	
	C:\Users\Public\D\SpeciesB_isolate2.mzml	
	C:\Users\Public\D\SpeciesB_isolate3.mzml	
	C:\Users\Public\D\SpeciesB_isolate4.mzml ¥ 18 file(s), less than 1 Mb	
	To me(s), less than 1 mb	

Figure 5: Select mzml files.

Edit data destination	?	×
 <none></none> Spectrum type Entry info field Spectrum experiments info 	field	
ОК	Can	cel

Figure 6: Select data destination.

- 9. Name the import template (e.g. "Import peak lists") and optionally give it a description. Press < OK >.
- 10. With the new import template highlighted, and the *Maldi* experiment selected (see Figure 8) press <*Next*> to go the next step where an overview of the actions that will be performed during the import is displayed (see Figure 9).
- 11. Press < *Next*> to go to the final step (see Figure 10).

Since we are importing peak lists (= processed peak data) in this tutorial, the option *Input files contain processed peak data* is automatically checked, and no options concerning the preprocessing need to be defined.



When importing raw spectrum data, the option *Input files contain raw spectrum data* is automatically checked, and options concerning the preprocessing can be defined.

12. Keep the option *Input files contain processed peak data* checked for our example files and press <*Finish*>.

The peak lists are imported and linked to the newly created samples. The file names are displayed

revie	w
Nr.	Кеу
1	SpeciesA_isolate1
2	SpeciesA_isolate2
3	SpeciesA_isolate3
4	SpeciesA_isolate4
5	SpeciesA_isolate5
6	SpeciesA_isolate6
7	SpeciesB_isolate1
8	SpeciesB_isolate2
9	SpeciesB_isolate3
10	SpeciesB_isolate4
11	SpeciesB_isolate5
12	SpeciesB_isolate6
13	SpeciesC_isolate1
14	SpeciesC_isolate2
15	SpeciesC_isolate3
16	SpeciesC_isolate4
17	SpeciesC_isolate5
18	SpeciesC_isolate6

Figure 7: Preview.

port spectrum data		? >
Import template Specify how to import data	into the database.	
Import templates:		
<default> Import peak lists</default>	Import peak lists	Create new
import pour loto		Edit
		Preview
		Сору
Experiment type: Maldi	\checkmark	
Experiment type: Maldi	~	

Figure 8: Import template.

in the *Key* field in the *Database entries* panel.

- 13. Click on a green colored dot for one of the entries in the *Experiment presence* panel to open the spectrum in the *Spectrum* window (see Figure 11).
- 14. Click on the spectrum in the *Profiles* panel to make it the active spectrum in the window.
- 15. Click on the *Peak List* panel to display the Peak List table.
- 16. Double-click on a peak in the *Peak List* panel: the peak is highlighted in *Profiles* panel and the view is updated.
- 17. Use the zoom sliders in the *Profiles* panel to obtain the best view.

nport spectrum	data			?	\times
	s orted records to database er c on a cell to get an overview.				
Overview In 'All levels'	✓ create 18 entries	and update 0 entries			
Select modifie	d entries				
		< Back	Next >	Can	cel

Figure 9: Import actions.

Import spectrum data	?	×
Preprocessing Select the preprocessing steps.		
The input file(s) contain processed peak data The input file(s) contain raw spectrum data		
Clip spectrum between and m/z Do preprocessing Predefined Preprocessing (Default) Preprocessing (Strict) Preprocessing (Relaxed)		^
Run the selected workflow in the preprocessing window.		~
< Back Finish	Car	cel

Figure 10: Select preprocessing action.

18. Close the *Spectrum* window with *File* > *Exit*.

Since we have imported peak lists (= processed peak data) no preprocessing is required.

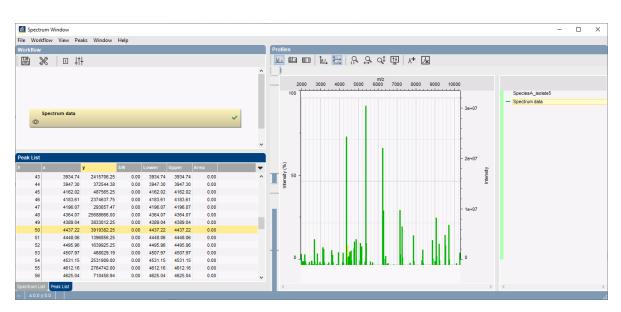


Figure 11: Spectrum with peak list.