

BIONUMERICS Tutorial:

Importing spectrum data in a database with levels

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 we will focus on the import and preprocessing of raw spectrum data in a leveled BIONUMERICS database.

2 Sample data

As an exercise, we will import a set of MALDI-TOF spectra 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 raw spectra". When the download is complete, unzip the file.

3 Preparing a sample database

3.1 Creating a new database

For this example we will be importing the raw spectra to a database with levels. We will start by creating a new, empty database.

- 1. Double-click on the BIONUMERICS icon (
- 2. In the *BIONUMERICS Startup* window, press the 🗟 button to enter 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 *Spectrum type* from the list (see Figure 1). Press < *OK* >.

Create a new experiment type	?	×
Select what kind of experiment type you	want to	create:
Fingerprint type		^
Character type		
Sequence type		
Matrix type		
[†] ∠, Trend data type		
A Spectrum type		
≩ Sequence read set type		
🇠 Composito data pat		*
ОК	Car	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 v Vertical axis unit: Intensity v		
< Back Next >	Can	cel

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 (see 5) and saved to the database.

9. Select the template *Preprocessing (strict)* and press <*Finish*> to complete the creation of the new spectrum type experiment.

The *Experiment types* panel now lists the spectrum type **Maldi** (see Figure 4).

New spectrum type			?	\times
Preprocessing workflow Select a default preprocessing workflow				
Predefined Preprocessing (Default) Preprocessing (Relaxed) Preprocessing (Strict)				~
	< Back	Finish	Ca	ncel

Figure 3: Preprocessing template.

Experiment ty	pes		
<u>କ</u> ୍ରି -	+ 🗗 🛛 🔍	€ ↑	All Experiment types>
#	Name	Туре	•
1 🗆	Maldi	Spectrum types	

Figure 4: The *Experiment types* panel.

3.3 Creating levels

Next, we need to create the correct levels in the database. Database levels are schematically visualized in the *Database design* panel. This panel appears by default as a tab behind the *Entry fields* panel.

When working with levels, it is easier to switch between levels if the *Database design* panel is docked above the *Database entries* panel:

10. Click on the *Database design tab* and - while keeping the mouse button pressed - drag it inside the *Database entries* panel. Drop the floating panel on the top part of the docking guide that appears (see Figure 5).

The Database design panel is now shown above the Database entries panel.

- In the Database design panel, click on "All levels" and select Database > Levels > Add new level... (+). This will show the Level information dialog box. Fill in Species and press < OK >.
- 12. Next, click on the new level "Species" and select *Database* > *Levels* > *Add new level...* (+), fill in *Isolate* and press < *OK* >. Finally, click on the new level "Isolate" and select *Database* > *Levels* > *Add new level...* (+), fill in *Raw spectra* and confirm.

After making the levels, the structure of the database should look like Figure 6.

3.4 Creating information fields

Level-specific information fields can be created during the import of the spectra (see 4), during import of entry information from an external Excel or text file (see tutorial: "Adding entry information")



Figure 5: Docking guide.

Database design				
+ 🖻 🖄 🛛	Ø⇔			
[All levels]	species	▲ Isolate	Raw spectra	^
📥 Maldi	📥 Maldi	📥 Maldi	📥 Maldi	
				~
<				>

Figure 6: Database design after creating levels.

or manually with the menu-items.

To manually add level specific information fields with the menu-items follow these steps:

- Click on the correct level in the Database design panel.
- Select + in the Entry fields panel.
- Enter a name for the new information field and press < OK >.

4 Importing spectra

When importing data (descriptive information and/or experimental data) in a database with levels, it is important to highlight the corresponding level first. Typically, information will most often be imported at the deepest child level (i.e. **Raw spectra** in our example database).

- 1. Make sure the 'Raw spectra' level is highlighted, select *File* > *Import...* (, Ctrl+I), highlight *Spectrum type data* > *Import spectrum data* and press <*Import*> (see Figure 7).
- 2. Browse to the folder, select all files in this folder and press < Open> and < Next>.
- 3. In the *Import template* wizard page of the wizard, press < *Create new*>.

4



Figure 7: The Import tree.

Import spectrum	data	?	×
Input Select spec	ctrum data file(s) to import		
Select file(s):	C:Users\\SpeciesA_lsolate1-1_110131.txt C:Users\\SpeciesA_lsolate1-2_110131.txt C:Users\\SpeciesA_lsolate1-3_110131.txt C:Users\\SpeciesA_lsolate1-4_110131.txt C:Users\\SpeciesA_lsolate1-5_110131.txt C:Users\\SpeciesA_lsolate1-6_110131.txt C:Users\\SpeciesA_lsolate1-6_110131.txt C:Users\\SpeciesA_lsolate1-8_110131.txt C:Users\\SpeciesA_lsolate1-9_110131.txt C:Users\\SpeciesA_lsolate1-9_110131.txt C:Users\\SpeciesA_lsolate1-9_110131.txt S0 file(s), 25 Mb		
	< Back Next >	Ca	incel

Figure 8: Select files.

The only source of information available in the newly created import template is the file name.

- 4. Double-click on the only row available in the grid or press < *Edit Destination*>.
- 5. Select Raw spectra key in the Edit data destination dialog box (see Figure 9).
- 6. Press < OK >.
- 7. Visualize the advanced options for the *Import template* dialog box by clicking on the check box next to **Show advanced options**.
- 8. Press < *Add rule*> to open the *Add data conversion rule* wizard.



Figure 9: Select the destination

9. In the first page of the *Add data conversion rule* wizard select *File* > *Name* (Figure 10) and press <*Next*>.

Add data conversion rule	?	×
Data source Select the data source		
File Path Created Fixed value		
< Back Next >	Can	icel

Figure 10: Adding a new rule to the import template.

- 10. In the second page of the *Add data conversion rule* wizard, select **Isolate key** and press <*Next*> (see Figure 11).
- 11. In the *Data parsing* dialog box, fill in following data parsing string: "[DATA]-*". Press the <*Preview*> button (see Figure 12).
- 12. When the information is parsed correctly press < *Next* > and <*Finish* >.
- 13. Press < *Add rule*> again to open the *Add data conversion rule* wizard.
- 14. In the first page of the Add data conversion rule wizard select **File** > **Name** and press < **Next**>.
- 15. In the second page of the *Add data conversion rule* wizard, select **Species key** and press <*Next*> (see Figure 13).
- 16. In the *Data parsing* dialog box, fill in following data parsing string: "[DATA]_*". Press the <*Preview*> button (see Figure 14).
- 17. When the information is parsed correctly press *<Next>* and *<Finish>*.

Add data conversion rule	?	×
Data destination Select the data destination		
 <none></none> Spectrum type Entry information Raw spectra entry info field Isolate entry info field Species key Species entry info field Spectrum experiments info field 		
< Back Next >	Can	icel

Figure 11: Isolate key.

Add data conversion	ule		?	×
Data parsing Edit the data par	sing			
 Parse componer Regular express 	t: find the component '[DA' ion: match the expression	TA]', use '*' as v and use the su	vildcard bexpres	sion
Data parsing string:	[DATA]-*	· · · · · · · · · · · · · · · · · · ·	~	
Data decoration:	[DATA]	~	·	
Preview Data: Species	A_lsolate1-1_110131	Preview		
Output: Species	A_isolate1			
	< Back	Next >	Can	cel

Figure 12: Parsing string.

- 18. Press < *Add rule*> again to open the *Add data conversion rule* wizard.
- 19. In the first page of the Add data conversion rule wizard select **File** > **Name** and press < **Next**>.
- 20. In the second page of the Add data conversion rule wizard, select Create new under Raw spectra entry info field and press <Next>, specify a name (e.g. Date) and confirm (see Figure 15).
- 21. In the *Data parsing* dialog box, fill in following data parsing string: "*_*_[DATA]". Press the <*Preview*> button (see Figure 16).
- 22. When the information is parsed correctly press < *Next* > and press < *Finish* >.

The grid is updated.

23. In the *Import template* dialog box, press < *Preview*> and verify the preview of the import (see Figure 17). Close the preview and press < *Next*> and < *Finish*>.



Figure 13: Species key.

Add data con	version r	ule		?	×
Data parsi Edit the	ng e data pars	ing			
Parse of Control Par	component r expressi	: find the component on: match the expres	'[DATA]', use '*' as sion and use the s	wildcard ubexpres	l ssion
Data parsin	g string:	[DATA]_*		~	
Data decora	ation:	[DATA]		~	
Preview				_	
Data:	Species/	_lsolate1-1_110131	Preview		
Output:	Species/	4			
		< Pack	Nexts	Can	aal
		< Back	wext >	Can	cei

Figure 14: Species key.

- 24. Name the import template (e.g. "Import to levels") and optionally give it a description. Press < OK >.
- 25. With the new import template highlighted, and the Maldi experiment selected press < Next > to go the next step where an overview of the actions that will be performed during the import is displayed (see Figure 18).
- 26. Press <*Next*>, check *Do preprocessing*, select **Preprocessing template (strict)** as *Preprocessing* option, leave all other settings at default and press <*Finish*>.

The raw spectra will be imported. Depending on the performance of your computer, this may take a few minutes.

Add data conversion rule	?	×
Data destination Select the data destination		
 <none></none> Spectrum type Entry information Raw spectra entry info field Create new> Isolate entry info field Species key Species entry info field Species entry info field Specirum experiments info field 		
< Back Next	> C	ancel

Figure 15: Create new information field.

Add data conversi	on rule			?	×
Data parsing Edit the data	parsing				
 Parse compo Regular expl 	onent: find ression: ma	the component " atch the express	[DATA]", use '** as ion and use the :	s wildcard subexpre	d ssion
Data parsing strir	1g: *_*_	[DATA]		~	
Data decoration:	[DA	TA]		\sim	
Preview	cies∆ Isol	ate1_1_110131	Draviaw	-	
Output: 110	131				
		< Back	Next >	Car	icel

Figure 16: Parsing string.

5 Preprocessing of spectra

Upon completion of the import, the *Spectrum Preprocessing* window will open automatically. If it does not open automatically, you can select all entries on the lowest level "Raw spectra", and use *Analysis* > *Spectrum types* > *Open preprocessing window...* to open the *Spectrum Preprocessing* window.

All the imported spectra are loaded in the *Spectrum Preprocessing* window with the preprocessing template selected in the last step of the *Import spectrum data* wizard (see Figure 19).

Because there is a slight difference in the range of the different spectra, we will be adding a trimming step to this template.

1. With the **Import & Resample** step of the preprocessing template highlighted in the *Workflow* panel, select *Workflow* > *Show flow chart...*. This will open the *Workflow* window for this step of the template.

revie	W				?	×
Nr.	Raw spectra key	Date	Isolate key	Species key		,
1	SpeciesA_Isolate1-1_110131	110131	SpeciesA_Isolate1	SpeciesA		
2	SpeciesA_Isolate1-2_110131	110131	SpeciesA_Isolate1	SpeciesA		
3	SpeciesA_Isolate1-3_110131	110131	SpeciesA_Isolate1	SpeciesA		
4	SpeciesA_Isolate1-4_110131	110131	SpeciesA_Isolate1	SpeciesA		
5	SpeciesA_Isolate1-5_110131	110131	SpeciesA_Isolate1	SpeciesA		
6	SpeciesA_Isolate1-6_110131	110131	SpeciesA_Isolate1	SpeciesA		
7	SpeciesA_Isolate1-7_110131	110131	SpeciesA_Isolate1	SpeciesA		
8	SpeciesA_Isolate1-8_110131	110131	SpeciesA_Isolate1	SpeciesA		
9	SpeciesA_Isolate1-9_110131	110131	SpeciesA_Isolate1	SpeciesA		
10	SpeciesA_Isolate1-10_110131	110131	SpeciesA_Isolate1	SpeciesA		
11	SpeciesA_Isolate2-1_110131	110131	SpeciesA_Isolate2	SpeciesA		
12	SpeciesA_Isolate2-2_110131	110131	SpeciesA_Isolate2	SpeciesA		
13	SpeciesA_Isolate2-3_110131	110131	SpeciesA_Isolate2	SpeciesA		
14	SpeciesA_Isolate2-4_110131	110131	SpeciesA_Isolate2	SpeciesA		
15	SpeciesA_Isolate2-5_110131	110131	SpeciesA_Isolate2	SpeciesA		
16	SpeciesA_Isolate2-6_110131	110131	SpeciesA_Isolate2	SpeciesA		
17	SpeciesA_Isolate2-7_110131	110131	SpeciesA_Isolate2	SpeciesA		
18	SpeciesA_Isolate2-8_110131	110131	SpeciesA_Isolate2	SpeciesA		
19	SpeciesA Isolate2-9 110131	110131	SpeciesA Isolate2	SpeciesA		

Figure 17: Preview.

Database links Link the imported r Double click on a c	ecords to database en cell to get an overview.	tries.		
Overview				
In 'Raw spectra' 🔽 In 'Isolate' 🗹 In 'Species' 📿	create 80 entries create 8 entries create 3 entries	and update 0 entries and update 0 entries and update 0 entries		
Select modified entri	es			

Figure 18: Import actions.

- 2. In the left panel, highlight the operator *Processing* > *X-axis* > *Trimming* and select *File* > *Add operator...* (^{t_2}).
- In the resulting dialog box (see Figure 20), fill in 2000 as X-axis minimum, do not use an upper limit and press < Next>.

The trimming operator is added to the end of the workflow, resulting in the workflow shown in Figure 21.

- 4. After the operator has been added, close the *Workflow* window.
- 5. Select *File* > *Save workflow as template...*, enter a name for the new template (e.g. "Strict with trimming") and press <*OK*>.

Spectru	m Window								– 🗆 X
File Workf	low View Peaks Window I	Help							
Workflow					Profiles				
-				^	hi. di	L 🕅 🗽 🔍 🔍 🗘 🖽 🖈 📈			
	Import & Resample								
*	©		~		_				
					_	5000 10000	15000	20000	^
ĩ	Baseline Subtraction	*			1	+ · · · · · · · · · · · · · · · · · · ·		F 60000	SpeciesA isolate1-1 11013
141						-		40000	- Import & Resample
†1+					1.1	4			
	Compute Noise Smooth	Baseline				the second of the second se	A I	20000	
		i, basenne			1	Jul & Allender Allender Allen Aleren		E 60000	SpeciesA Isolate1-2 11013
	_							40000	- Import & Resample
	A Death Detection	· ·			1.1				
	/ Peak Detection					the sector label of the sector sector is the sector label of the sector sector label of the sector s		20000	
				~	1	Jack Milliell Charles of Carles - Charles - Ch		E 0	SpeciesA isolate1-3 11013
				_				40000	- Import & Resample
Spectrum L	ist					1			
	Key	Date	Species		8	the second of the second se	1	20000	
	SpeciesA_Isolate1-1_110131	110131	SpeciesA	^	1 10	Jack Million Colling Calling Colling			SpeciesA isolate1-4 11013
	SpeciesA_Isolate1-2_110131	110131	SpeciesA		- te			40000	- Import & Resample
	SpeciesA_Isolate1-3_110131	110131	SpeciesA		-				
	SpeciesA_Isolate1-4_110131	110131	SpeciesA			the second se	A	20000	
	SpeciesA_Isolate1-5_110131	110131	SpeciesA		1	Jack Hick Michelle Million Aller Million - Channes -		E 0	Species& Isolate1-5, 11013
	SpeciesA_Isolate1-6_110131	110131	SpeciesA					40000	- Import & Resample
	SpeciesA_Isolate1-7_110131	110131	SpeciesA		1	4		40000	inport a restampte
	SpeciesA_isolate1-8_110131	110131	SpeciesA			the second of the first second		20000	
	SpeciesA_Isolate1-9_110131	110131	SpeciesA		1	Jule M. Alle March Cherry College - Cherry		E 0	Species A. Jeolete 1, 6, 11013
	SpeciesA_isolate1-10_110131	110131	SpeciesA					40000	- Import & Decempio
	SpeciesA_isuidte2-1_110131	110131	SpeciesA			4			inport a restample
	species#_isoid(82-2_110131	10131	SpeciesA	× *		the second of the flate of the second		20000	
- <				'	4	I was a second state of the second state of th		_E 0	×
Spectrum List	Peak List				<			>	< >>
×:0.0 y:									

Figure 19: Preprocessing window after import of spectra.

Parameters Choose para 'Trimming'	meters for algorithm	
X-axis Min:	2000 x-axis unit	
Use upper limit	٥	
X-axis Max:	0 x-axis unit	

Figure 20: Parameters of the trimming operator.

This template will now be available for future import and preprocessing of spectra.

Press the last step of the preprocessing template in the *Workflow* panel, i.e. < *Peak Detection*>, to execute the entire preprocessing workflow. Leave the limit for the signal to noise ratio at default and press < *Next*>.

Depending on the performance of your computer, the execution might take a few minutes.

With the "Peak Detection" step still selected in the *Workflow* panel, the preprocessed spectra are displayed in the *Profiles* panel. The detected peaks are indicated with circles on top of the peaks.

- 7. Click on a spectrum in the *Profiles* panel to make it the active spectrum in the window.
- 8. Click on the *Peak List* panel to display the Peak List table.
- 9. Double-click on a peak in the *Peak List* panel: the peak is highlighted in *Profiles* panel and the view is updated.
- 10. Use the zoom sliders in the *Profiles* panel to obtain the best view.
- 11. Save the results by selecting *File* > *Save spectrum data* (). Close the *Spectrum Preprocessing* window.

The preprocessed spectra are now available in the database and further analysis, such as creating summary spectra or comparing the spectra, can be performed.



Figure 21: Final workflow with trimming operator.