



Running Galaxy workflows using the Clc Main Workbench

Marc Logghe

Galaxy Community Conference 2013

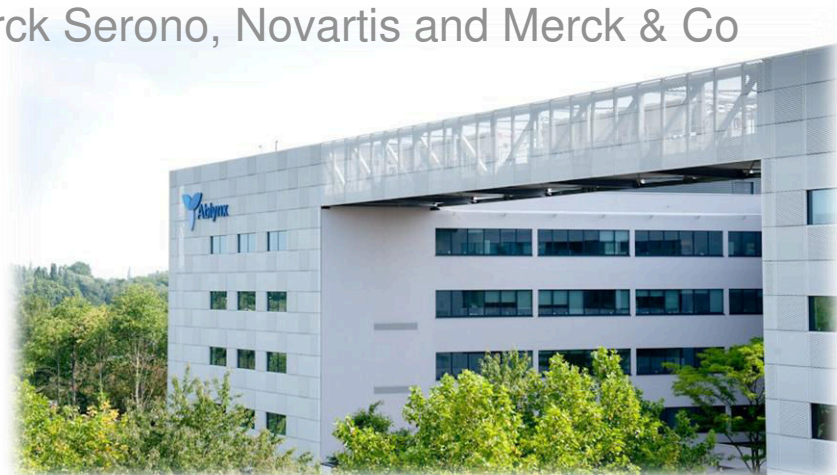
July 2nd

A high-speed photograph of a water splash, showing numerous droplets in mid-air against a blue background. The splash is centered and creates a crown-like shape.

**Nanobodies® -
Inspired by nature**

Ablynx – company overview

- ✔ Drug discovery and development company based in Ghent, Belgium
- ✔ A pioneer in next generation biologics - Nanobodies®
- ✔ Worldwide exclusive rights to commercialise Nanobody products in human healthcare
- ✔ ~25 programmes in the R&D pipeline
- ✔ Two products achieved clinical proof-of-concepts in RA
- ✔ 6 Nanobody products in the clinic - 3 Phase II & 3 Phase I
- ✔ Exclusive rights to >500 patent applications and granted patents
- ✔ Partnerships with Boehringer Ingelheim, Merck Serono, Novartis and Merck & Co
- ✔ Cash at 31 March 2013 of €82.2M
- ✔ >250 employees



What ?

- ✦ Transparently run Galaxy workflows in the Main Workbench of Clc bio

Why ?

✔ Users don't like change

- Yet Another Tool to learn

✔ User friendly

- Stay in comfort zone
- No copy/paste or (manual) export of data needed

How ?

- ✦ Extend functionality of Main Workbench with plugin
 - SDK (<http://www.clcbio.com/customized-solutions/software-developer-kit/>)
- ✦ blend4j (John Chilton)
 - Github (<https://github.com/biologghe/blend4j>)

How ?

Steps executed by plugin

Clc Object parameter (eg. SequenceList)



Fasta Exporter plugin

Export to fasta

Create history

Upload fasta to Galaxy history

Launch Galaxy workflow

Poll

Download Result dataset(s)

Import dataset

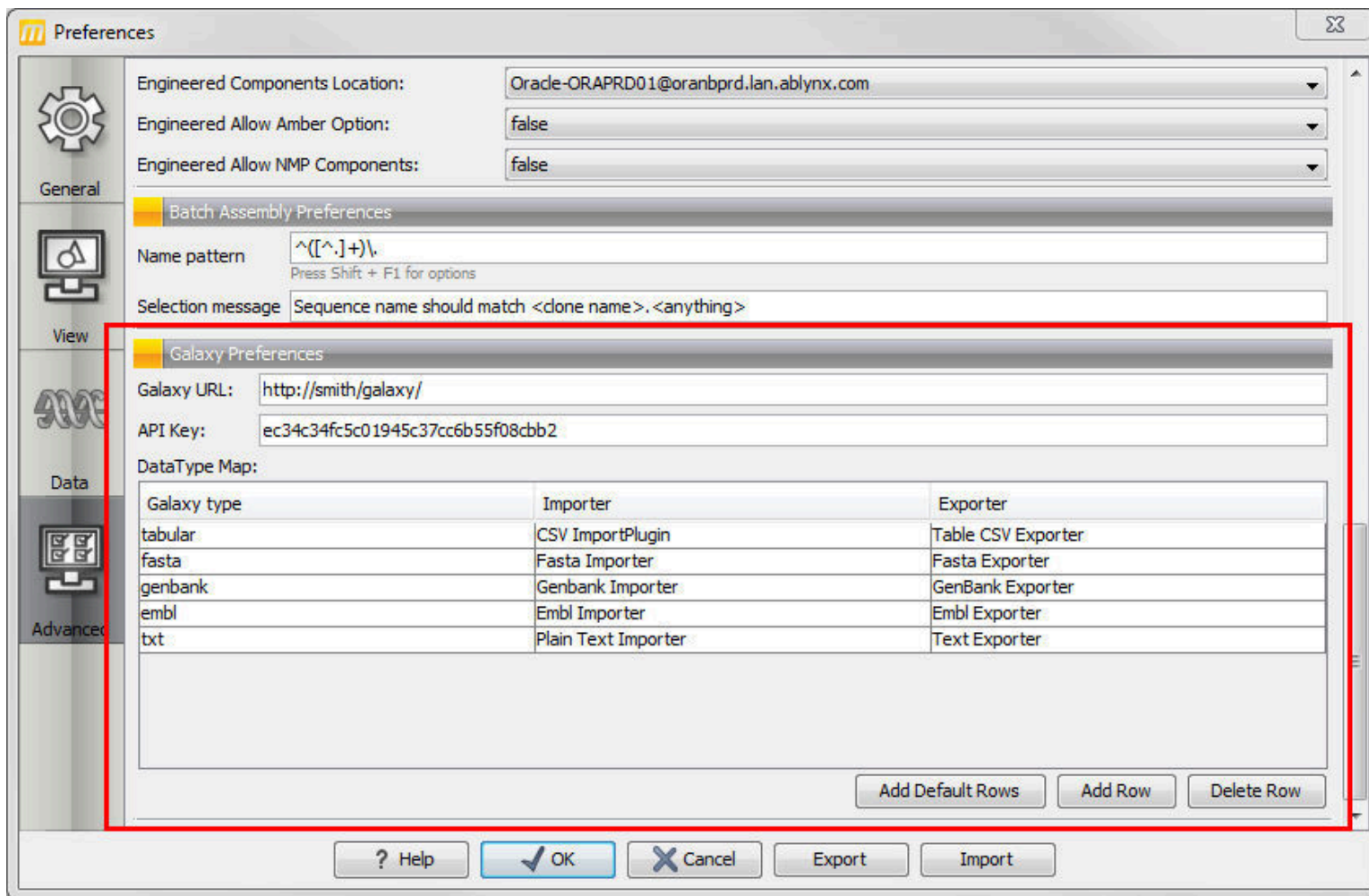
Importer plugin matching Galaxy datatype (set in configuration)



Clc Object(s)

Configuration of Galaxy plugin

Galaxy data type mapping using importer and exporter plugins



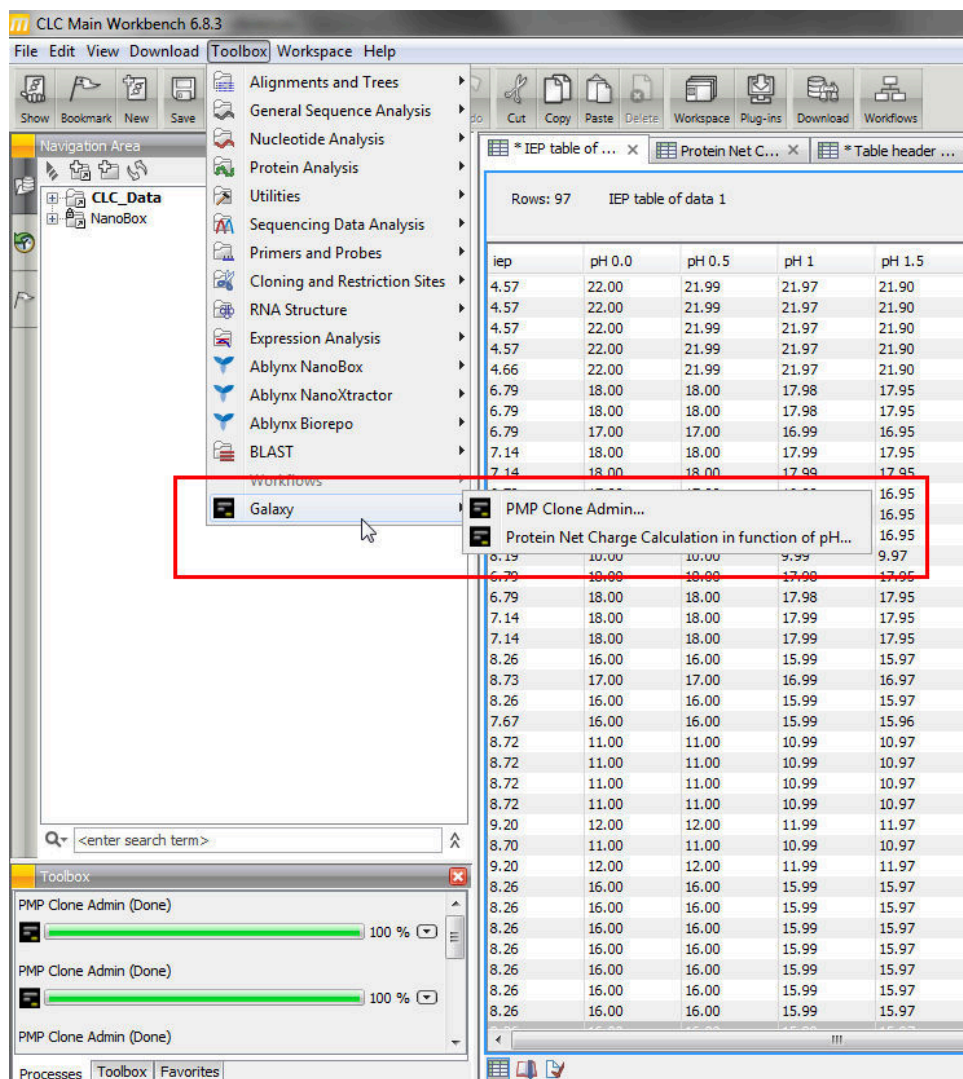
The screenshot shows the 'Preferences' dialog box in Ablynx. The 'Galaxy Preferences' section is highlighted with a red border. It contains the following fields and table:

- Engineered Components Location:** Oracle-ORAPRD01@oranbprd.lan.ablynx.com
- Engineered Allow Amber Option:** false
- Engineered Allow NMP Components:** false
- Batch Assembly Preferences:**
 - Name pattern:** ^([\^.]+)\. (Press Shift + F1 for options)
 - Selection message:** Sequence name should match <done name>. <anything>
- Galaxy Preferences:**
 - Galaxy URL:** http://smith/galaxy/
 - API Key:** ec34c34fc5c01945c37cc6b55f08cbb2
- Data Type Map:**

Galaxy type	Importer	Exporter
tabular	CSV ImportPlugin	Table CSV Exporter
fasta	Fasta Importer	Fasta Exporter
genbank	Genbank Importer	GenBank Exporter
embl	Embl Importer	Embl Exporter
txt	Plain Text Importer	Text Exporter

Buttons at the bottom of the dialog include: ? Help, ✓ OK, ✗ Cancel, Export, and Import.

Galaxy workflows in Toolbox



The screenshot shows the CLC Main Workbench 6.8.3 interface. The 'Toolbox' menu is open, and the 'Galaxy' option is highlighted with a red box. Below it, a sub-menu is visible, also highlighted with a red box, containing the following items:

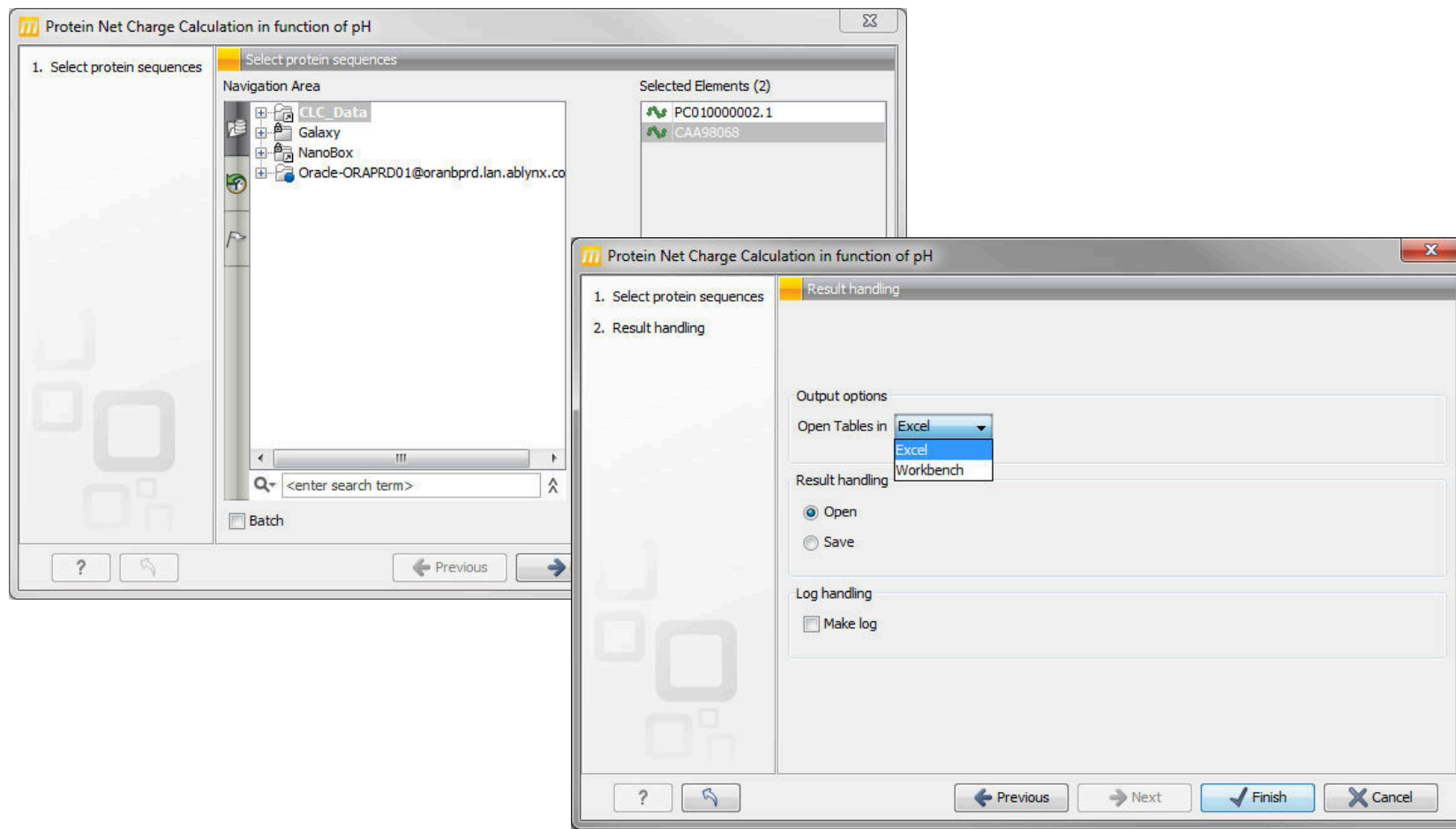
- PMP Clone Admin... 16.95
- Protein Net Charge Calculation in function of pH... 16.95

The main window displays a table titled 'IEP table of data 1' with 97 rows and 5 columns: iep, pH 0.0, pH 0.5, pH 1, and pH 1.5. The table data is as follows:

iep	pH 0.0	pH 0.5	pH 1	pH 1.5
4.57	22.00	21.99	21.97	21.90
4.57	22.00	21.99	21.97	21.90
4.57	22.00	21.99	21.97	21.90
4.57	22.00	21.99	21.97	21.90
4.66	22.00	21.99	21.97	21.90
6.79	18.00	18.00	17.98	17.95
6.79	18.00	18.00	17.98	17.95
6.79	17.00	17.00	16.99	16.95
7.14	18.00	18.00	17.99	17.95
7.14	18.00	18.00	17.99	17.95
8.19	10.00	10.00	9.99	9.97
6.79	18.00	18.00	17.98	17.95
6.79	18.00	18.00	17.98	17.95
7.14	18.00	18.00	17.99	17.95
7.14	18.00	18.00	17.99	17.95
8.26	16.00	16.00	15.99	15.97
8.73	17.00	17.00	16.99	16.97
8.26	16.00	16.00	15.99	15.97
7.67	16.00	16.00	15.99	15.96
8.72	11.00	11.00	10.99	10.97
8.72	11.00	11.00	10.99	10.97
8.72	11.00	11.00	10.99	10.97
8.72	11.00	11.00	10.99	10.97
9.20	12.00	12.00	11.99	11.97
8.70	11.00	11.00	10.99	10.97
9.20	12.00	12.00	11.99	11.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97
8.26	16.00	16.00	15.99	15.97

The 'Toolbox' panel at the bottom shows three instances of 'PMP Clone Admin' with 100% progress bars.

Set input parameters and launch



History is created

Galaxy Analyze Data Workflow Shared Data Visualization Admin Help User Using 497.7 MB

Saved Histories

search history names and tags

[Advanced Search](#)

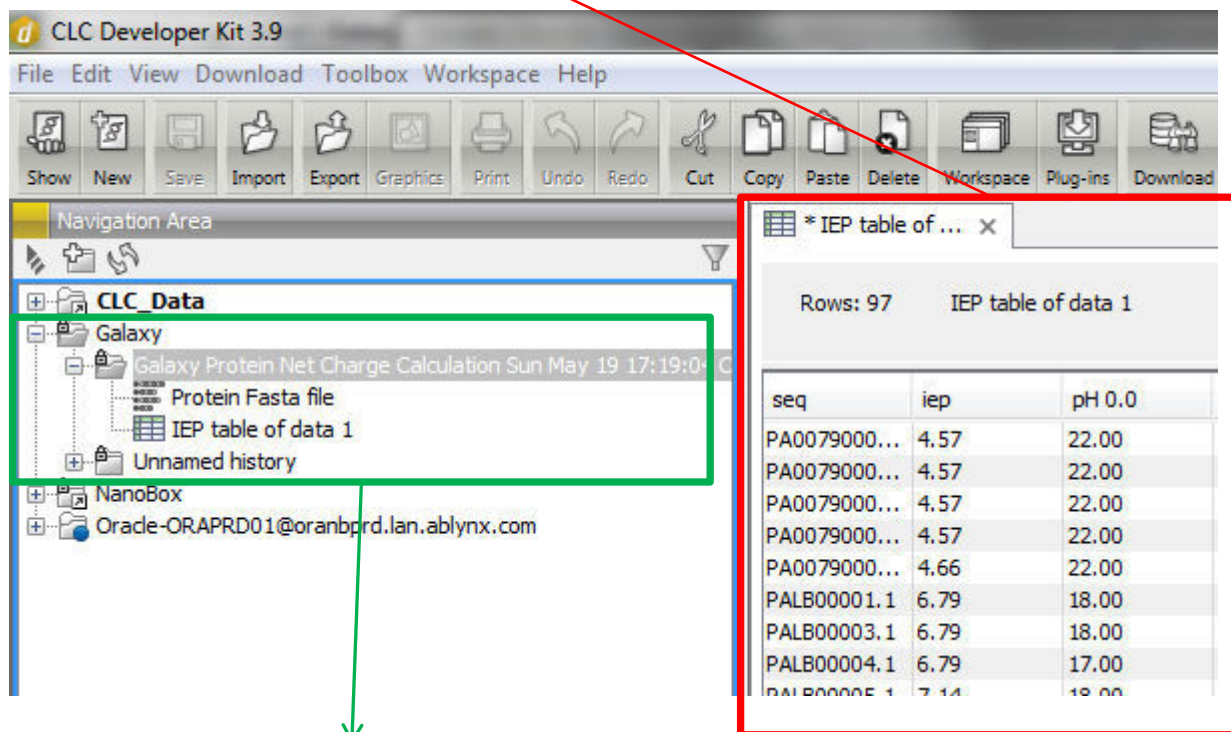
<input type="checkbox"/>	Name	Datasets	Tags	Sharing	Size on Disk	Created	Last Updated ↑	Status
<input type="checkbox"/>	Galaxy Protein Net Charge Calculation Sun May 19 17:19:04 CEST 2013	2	0 Tags		36.7 KB	1 minute ago	1 minute ago	
<input type="checkbox"/>	Unnamed history		0 Tags		0 bytes	9 minutes ago	9 minutes ago	current history

For 0 selected histories:

Histories that have been deleted for more than a time period specified by the Galaxy administrator(s) may be permanently deleted.

Result Dataset import and rendering

- As soon as workflow is finished, result is fetched, imported and rendered



The screenshot shows the CLC Developer Kit 3.9 interface. The main window displays a table titled "* IEP table of ..." with 97 rows. The table has three columns: "seq", "iep", and "pH 0.0". The data rows show various protein sequences and their corresponding isoelectric points (iep) and pI values (pH 0.0).

seq	iep	pH 0.0
PA0079000...	4.57	22.00
PA0079000...	4.57	22.00
PA0079000...	4.57	22.00
PA0079000...	4.57	22.00
PA0079000...	4.57	22.00
PA0079000...	4.66	22.00
PALB00001.1	6.79	18.00
PALB00003.1	6.79	18.00
PALB00004.1	6.79	17.00
PALB00005.1	7.14	18.00

In the navigation area on the left, the "CLC_Data" folder is expanded to show a "Galaxy" folder. Inside the "Galaxy" folder, the "IEP table of data 1" is highlighted with a green box, indicating its location in the workflow history.

- History visible in navigation area as well

Error handling

