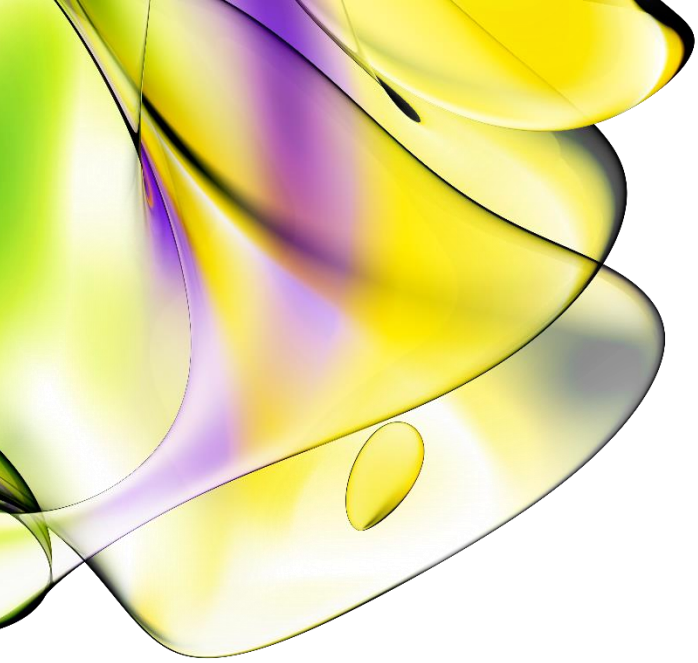


Revvity Signals VitroVivo™ 3.5.0

User Guide

Powered by Spotfire®

Last Updated: March 19, 2024

Table of Contents

1. Introduction

Analyzing scientific data requires a high degree of flexibility to deal with large data volumes in a wide range of experiment types. Integrating this information, visualizing it, and interacting with the data are not trivial tasks.

To address these challenges, we have created the Signals Apps system within Spotfire®. Breaking down complex analyses into small Apps that can be combined into complex Workflows gives the users the required flexibility, while the integration with Spotfire® Automation Services allows the analysis to keep up with the highly automated data generation process.

2. Signals Apps

Signals Apps can be invoked through the **Tools** menu and selecting **Signals Apps** or by clicking on the **Open Signals Spotfire Apps** button.

The Signals Apps screen displays two main tabs (**Workflows** and **Apps**), a logged user indicator (shows whether the user was initially logged or is working offline using the context of the last user) and a **Settings** dropdown (*Cog Wheel* icon) whose options will be explained in other sections of this document.

Both tabs include two reserved categories (Recently used and Favorites). The **Recently used** category shows the latest ten recently used Apps/Workflows and the **Favorites** category shows the favorited Apps/Workflows (using the *Star* icon available when hovering over the App/Workflow card).

Both tabs include a Search component (see Figure 2-3: Workflow Filtering) that allows the App/Workflow cards to be filtered by name or domain/category, thus hiding any irrelevant results, and simplifying navigation.

The **Apps** tab displays the registry of available Apps as shown in Figure 2-1. The Apps are organized by domain.

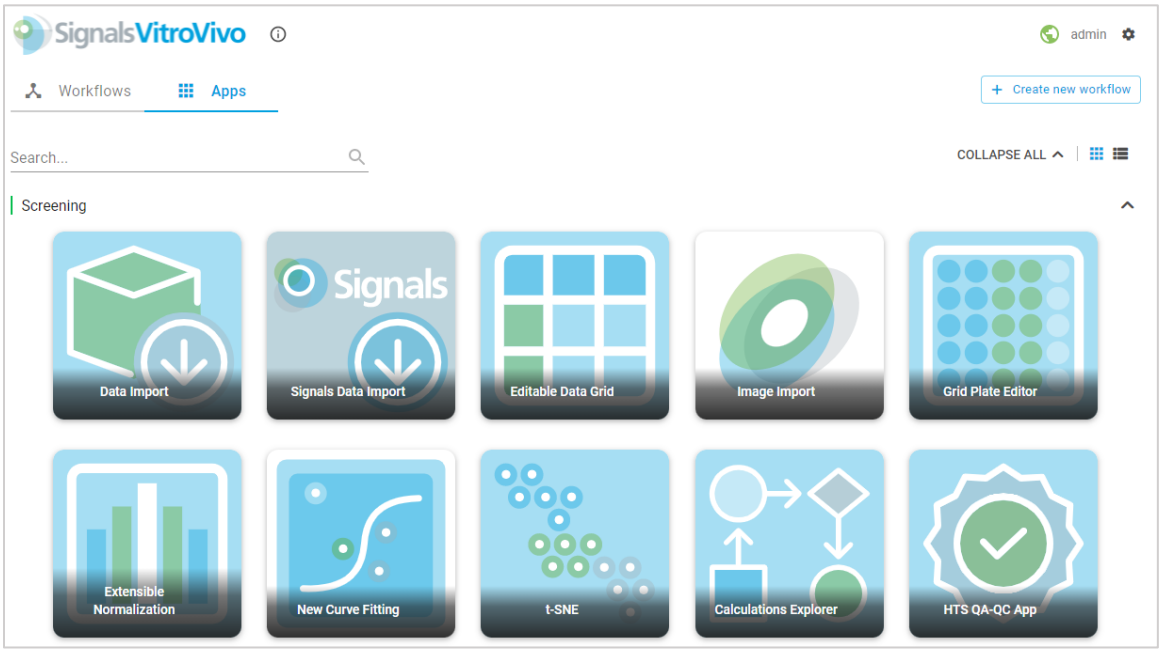


Figure 2-1: User Interface of the Apps Tab

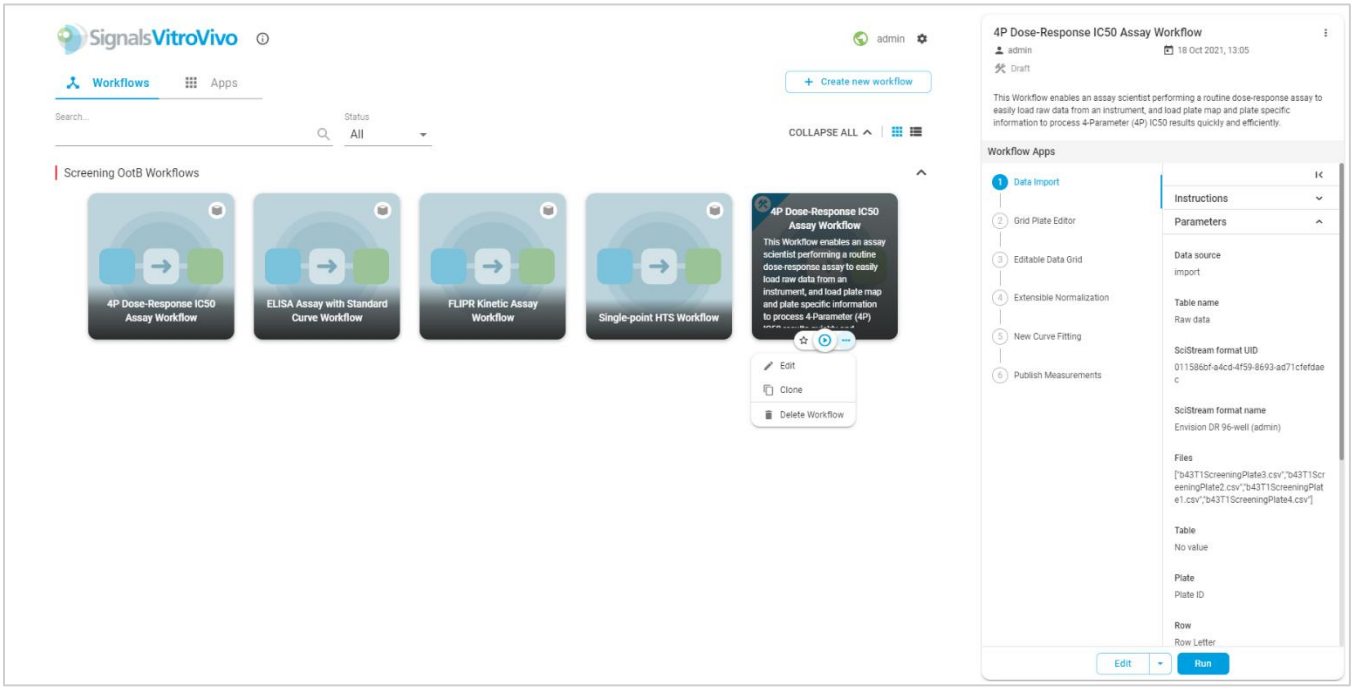


Figure 2-2: User Interface of the Signals Workflow Tab

The **Workflows** tab shows the list of saved Workflows, which can also be filtered by **Status** (**All**, **Draft** or **Published**) using the dropdown available on the left-hand side, as well as a button to create new Workflows and some Workflow specific card interactions to **Run**, **Edit**, **Clone**, or **Delete** Workflows (see Figure 2-2: User Interface of the Signals Workflow Tab and the *App Integrations* section for more details).

The right-hand panel shows the details of the selected Workflow, at the bottom there are buttons to interact with said Workflow. Clicking on a Workflow's card will select it (highlighting the card with a blue ribbon on its upper-left-hand corner) and show its details panel. Creating, cloning, or editing a Workflow will also select it automatically.

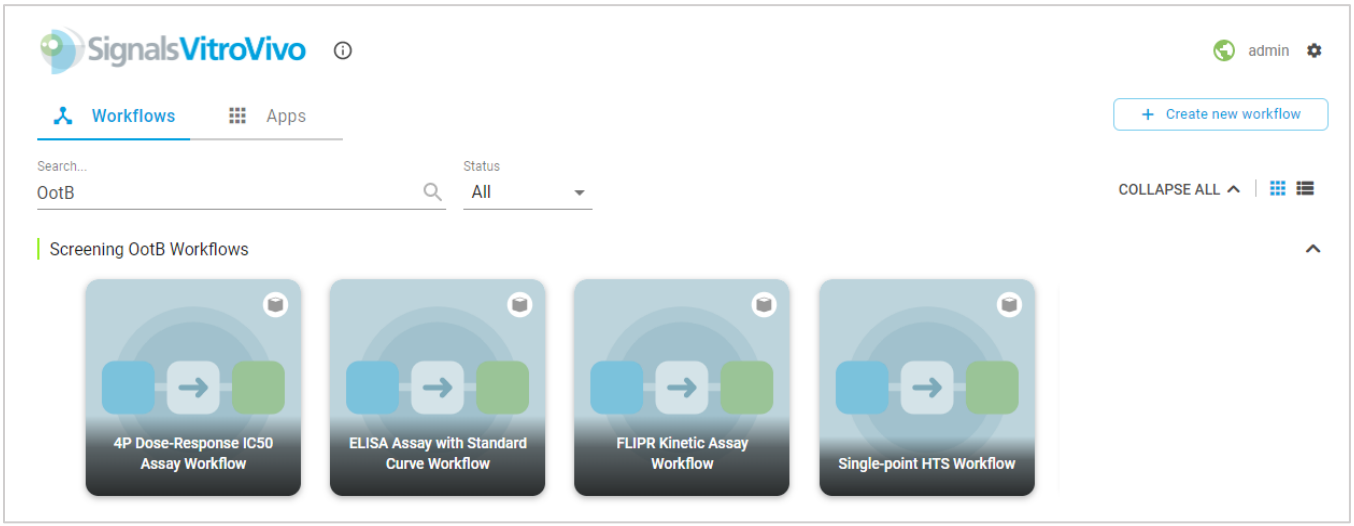


Figure 2-3: Workflow Filtering by Search and Status

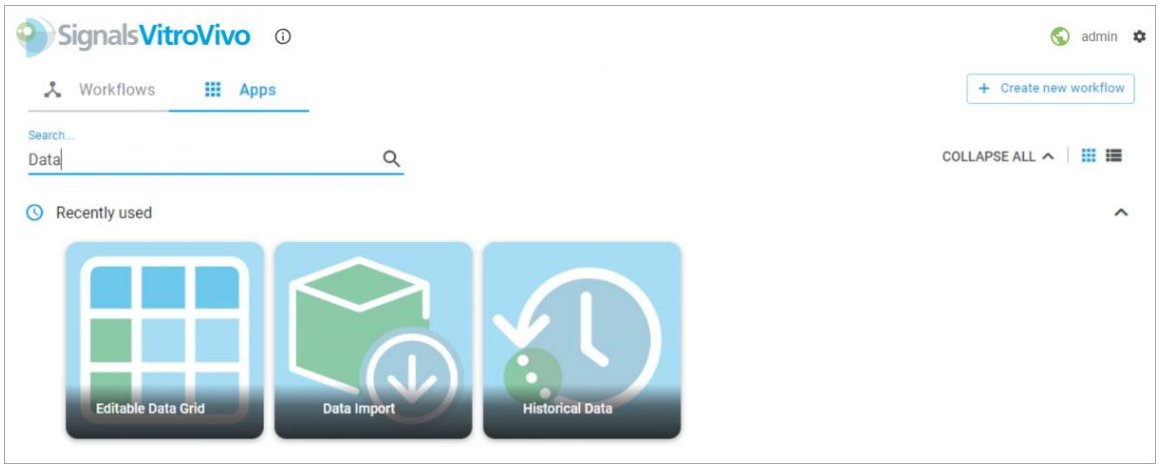
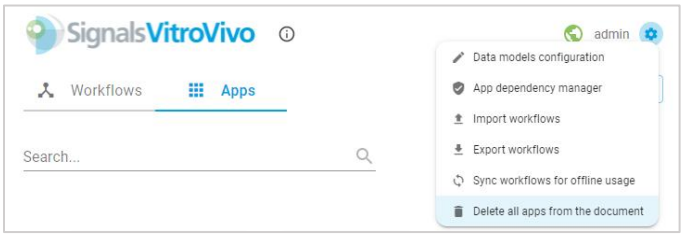


Figure 2-4: Filter Apps by Name

User can quickly clear all Apps from a Spotfire® document by selecting the **Settings** dropdown (*Cog Wheel icon*) and selecting **Delete all apps from the document**. The user is warned that all Apps will be deleted, including associated data tables, where the user must confirm to clear the document. Note that non-App tabs remain unaffected.



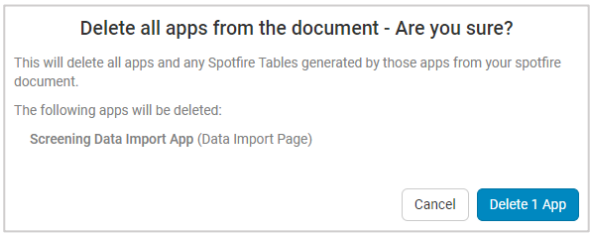


Figure 2-5: Deleting All Apps from a Spotfire® Document

Note: By default, the data table 'Signals Apps' which contains the Signals VitroVivo copyright information is always loaded upon launching the Signals Apps page. This table is created to work around a limitation that prevents Apps from being added when no data table is present in the Spotfire® document. It can be safely deleted once user data has been added to the document.

3. How to Invoke an App

To invoke an App, simply click anywhere on the corresponding App card and it will be added to the document using the default App settings. You can also click on the *Play* icon that appears when hovering over the App to run it.

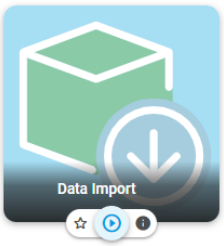


Figure 3-1: App Card

There are two additional options available:

- **Favorites:** The *Star* icon adds the App to the **Favorites** category, displayed at the top of the page.
- **Information:** The *Information* icon provides a preview of the App. (Figure 3-2).

Figure 3-2: App Preview Shown by Selecting Information

4. App Upgrades

An App upgrade will occur when the version of an App in the system is newer than the version of the same App included in a loaded DXP file. Any Apps not included in a Workflow will be updated independently, but Apps that belong to a Workflow will follow the order of the Workflow, so that any data produced and consumed is regenerated consistently.

Before the upgrade is performed, all the App's controls will be disabled, and the user will be able to explore the Workflow until ready. The system will show a notification in the upper-right-hand side of the App with instructions to start the upgrade process (which can be delayed indefinitely if the user deems it necessary). Depending on the situation the message will vary, and the user may have to review or even re-execute the App.

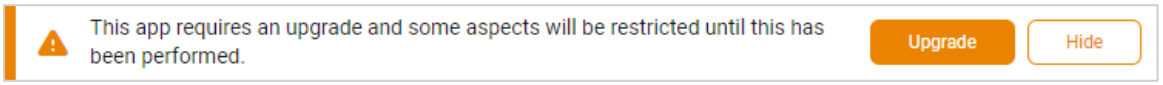


Figure 4-1: Notification on an App that can be immediately upgraded.

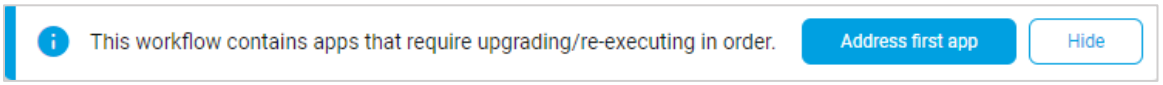


Figure 4-2: Notification on an App that is part of a Workflow with a previous App that requires an upgrade.

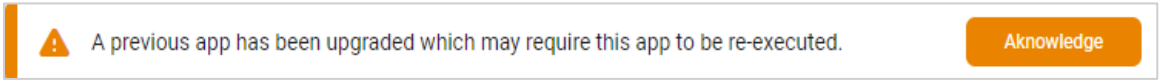


Figure 4-3: Notification on an App that does not require an upgrade but is part of a Workflow with a previous App that was upgraded.

5. App Integrations

Some Apps may provide an integration with an external service such as Signals Notebook™, and this integration may affect how the App is instanced/used.

5.1 Signals Notebook

When the App has a Signals Notebook integration, it might behave differently when a feature is enabled that requires a Signals connection and will prompt to authenticate the Signals server.

Note: The Administrator configures the Signals Notebook server URL and can optionally configure the Signals Notebook Experiment for a specific Spotfire® user group. Refer to the *Revvity Signals VitroVivo Installation Guide* for details on this topic.

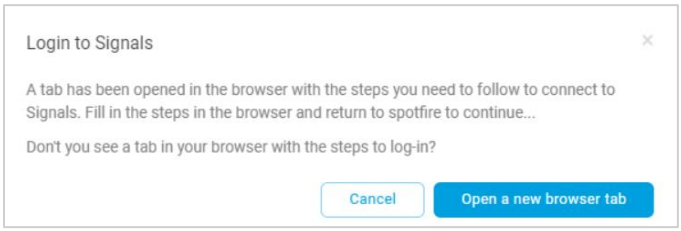


Figure 5-1: Login to Signals Notification

A browser is opened, and the user is prompted to log in to the configured Signals server.

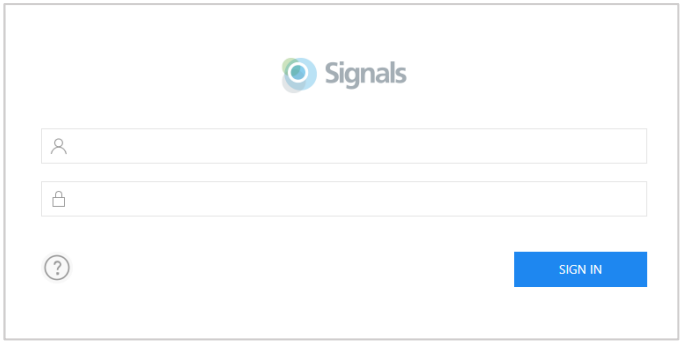


Figure 5-2: Signals Server Login Page

After the user logs in with valid credentials to the Signals server, a form asking for the Signals Notebook Experiment appears (if not configured by the Administrator). The user may optionally filter the available experiments by selecting a Notebook from the corresponding dropdown.

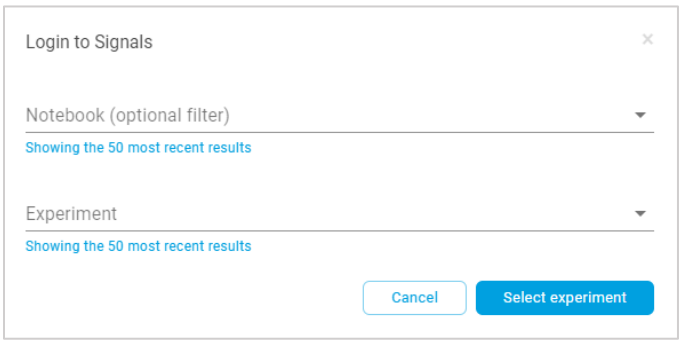


Figure 5-3: Form to Select the Signals Notebook and Experiment

After the user searches and selects the desired experiment and clicks **Select experiment,** the Signals server connection is properly configured, and the data is retrieved.

Note: The user can check the current Signals connection info by selecting the Signals icon in the bottom left-hand corner. The resulting menu lists the connection status and links to the connected server and experiment.

5.2 Opening a DXP with Apps Requiring Signals Notebook

Before the authentication is completed, all the App's controls will be disabled, enabling the user to explore but not alter the Workflow until they are connected to Signals Notebook.

The system will show a message in a notification bar at the top of the App, which will contain instructions to start the connection process.

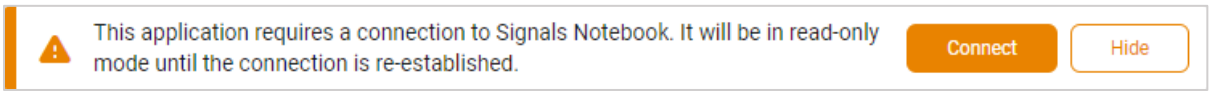


Figure 5-4: Notification Shown on an App that Requires a Signals Notebook Connection

6. Signals VitroVivo Workflows

This section will provide an overview of the Workflow creation process to quickly get new users familiarized with the corresponding Signals VitroVivo features. For detailed information on the specific Screening Domain Apps or the Calculations Explorer please refer to the corresponding sections of this manual.

A Workflow is a formalization of a set of Apps, their order of execution, and their configuration. This concept allows users to string several Apps together, run different types of analysis with the same App, and create and maintain different variants of the same study, etc.

This feature allows users to easily share ideas and collaborate to improve them. To achieve this, a Workflow can have a private visibility (when set to **Draft**) or public (when set to **Published**). Draft Workflows can only be seen by their creator (and thus only they can edit or publish them), whereas published Workflows are visible and editable by all members of the group.

Note: Only the original creator can set a Workflow back to Draft.

Workflows are a key concept within Signals Apps and thus can be accessed directly from the Signals Apps homepage (see Figure 2-1).

Note: Certain Apps may not support the use of multiple instances within the same document.

The Workflows storage must be configured by a Spotfire® administrator. Refer to the *Revvity Signals VitroVivo Installation Guide* for more details.

6.1 Role Based Permissions

Workflow management can be limited to specific users or groups of users by setting up roles using the *Signals Apps Workflows Role - Admin* and *Signals Apps Workflows Role - Author* licenses. By default, everyone has a consumer role, unless any of the aforementioned licenses are granted, in which case, the least restrictive of the granted licenses will apply (thus setting the role).

When the user role does not allow a certain interaction, then the controls in the system will not be present.

Action	Roles		
	Consumer	Author	Admin
See and execute published owned Workflows	✓	✓	✓
See and execute draft owned Workflows	✓	✓	✓
See and execute published unowned Workflows	✓	✓	✓
See and execute draft unowned Workflows	×	×	✓
See and execute Out-of-the-Box Workflows	✓	✓	✓
Add a new Workflow	×	✓	✓
Add a new Out-of-the-Box Workflow	×	×	×
Delete an owned Workflow	×	✓	✓
Delete an Out-of-the-Box Workflow	×	×	×
Edit an owned Workflow	×	✓	✓
Edit an Out-of-the-Box Workflow	×	×	×
Clone an owned Workflow	×	✓	✓
Clone an Out-of-the-Box Workflow	×	✓	✓
Delete an unowned Workflow	×	×	✓
Edit an unowned Workflow	×	×	✓
Clone an unowned Workflow	×	✓	✓

The optimal way of distributing licenses to a large number of users is to create groups and assign the users to the groups. That ensures that all users in a group are granted the same license.

Signals VitroVivo provides a framework in which to create analysis Workflows that can be easily applied to different datasets provided these datasets follow the same structure as the ones used to design the Workflow.

This framework can be used in three different manners that differentiate three clear user roles:

- **Consumer Role:** A user that is not necessarily an expert in the use of Spotfire® and will typically be an expert in the field that will execute the analysis Workflow created by the author and interpret the results. They will not modify the Workflow although they may add additional downstream analysis steps using either Spotfire®, Calculations Explorer, or other analysis Apps. Consumers have access to all published (non-draft) Workflows, but cannot add new Workflows or delete, edit, or clone existing Workflows.
- **Author Role:** An expert Spotfire® user that understands the context of the analysis that needs to be performed and will create the standard analysis Workflows to be used by the Consumers. Authors have access to all published Workflows and their owned draft Workflows. Workflow authors can add new Workflows and are also able to delete and edit their owned Workflows. Lastly, they can clone any Workflow they have access to.
- **Admin Role:** Admins have full access to draft or published Workflows and can create, delete, edit (except Out-of-the-Box Workflows), or clone Workflows regardless of whether they own them.

6.2 Out-of-the-Box Workflows

Out-of-the-Box (OotB) Workflows are included with each installation of Signals VitroVivo and are ready-to-run Workflows designed to introduce users to common assays using examples datasets. They can be identified by the icon  the upper right-hand corner of the Workflow card. OotB Workflows cannot be edited or deleted, but users may clone them to edit the Workflow to suit their needs. See the corresponding *Revvity Signals VitroVivo Quick Start Guide* for full details on how to run these Workflows.

The following OotB Workflows are available for the **Screening** domain:

4P Dose-Response IC50 Assay Workflow

- This Workflow enables an assay scientist performing a routine dose-response assay to easily load raw data from an instrument, and load plate map and plate specific information to process 4-Parameter (4P) IC50 results quickly and efficiently. The results are then published to Signals Inventa using the **Publish Measurements** App. See the *Revvity Signals Inventa User Guide* for full details on this App.

Single-point HTS Workflow

- This Workflow enables an assay scientist performing a single-point high-throughput assay to easily load raw data from an instrument and load plate specific information to process the HTS results quickly and efficiently and select hits to advance to confirmatory screening.

ELISA Assay with Standard Curve Workflow

- This Workflow enables an assay scientist performing a routine ELISA assay to easily load raw data from an instrument, and load plate map and plate specific information to process the ELISA results interpreted from the standard curves quickly and efficiently.

FLIPR Kinetic Assay Workflow

- This Workflow enables an assay scientist performing a routine FLIPR assay to easily load raw data from an instrument, and load plate map and plate specific information to process the FLIPR results from the full kinetic curve, extract key value(s), and calculate a final IC50.

Signals Pipeline Curve Fitting - QA-QC Workflow

- This Workflow enables an automation scientist using Signals Pipelines to easily perform a QA/QC analysis of 4-Parameter (4P) IC50 pipeline results, including import of data from Signals to Spotfire®, QA/QC of curves, and publication of revised results into a Signals Inventa project.

The following OotB Workflow is available for the **SPR** domain:

SPR Kinetic Workflow

- This Workflow enables a scientist performing an SPR Kinetic assay to easily load raw data from a T200 instrument, apply several preprocessing tools, fit sensorgram data to obtain association and dissociation rate constants, and perform selection and classification of SPR analysis results.

6.3 Creating a Workflow

This process will be performed by the author or admin roles.

When creating a Workflow, authors can choose between two different types of building blocks: Signals VitroVivo Apps or Calculations Explorer (CE) Templates.

- **Signals VitroVivo Apps**: These Apps are self-contained blocks that can be used for certain standard analyses such as importing data, normalizing data or adding annotations to existing data tables.
- **Calculations Explorer Templates:** These are a series of calculations and visualizations that are created in an automated manner by either opening the Calculations Explorer App and applying the CE Template directly in the document, or by adding an instance of the Calculations Explorer App to the document where the Calculations Explorer Template has been embedded.

A **VitroVivo Workflow** consists of a series of Apps that are executed consecutively to produce the desired analysis. Standard Apps can be added easily one after another in the document by simply clicking on them in the Signals Apps page and following the steps required in the App to perform the corresponding analysis. In most cases Apps are restricted to one instance per document.

In the case of adding **Calculations Explorer Templates** to the Workflow, this is done by adding an instance of the Calculations Explorer App to the document, and from within the App executing the required CE Template. In this case there are a couple of items to consider:

- This App is not limited to a single instance per document. Multiple instances of the **Calculations Explorer** App can be concatenated to perform steps within a complex Workflow that are not supported by any of the standard existing Apps.
- Each instance of the **Calculations Explorer** App can only execute a single CE Template.
- It is recommended to create each Calculations Explorer Template beforehand from a document containing only the required input data table for the CE Template operations. After doing this, the CE Template can be saved to add to the Workflow within the **Calculations Explorer** App when creating the Apps Workflow. This avoids unintended effects of calculated columns existing in the document that may have been added by other Apps and could unintentionally be included in the Calculations Explorer Template.
- When applying a CE Template, the columns used for matching cannot be calculated columns.

A Workflow can be created using the **+ Create new workflow** button on the top right-hand side of the Signal Apps homepage within the Workflows tab.

Note: When there are no available Workflows, another **Create a workflow** button will appear in the center of the page (Figure 6-1: Signals Apps Workflows Tab with No Available Workflows).

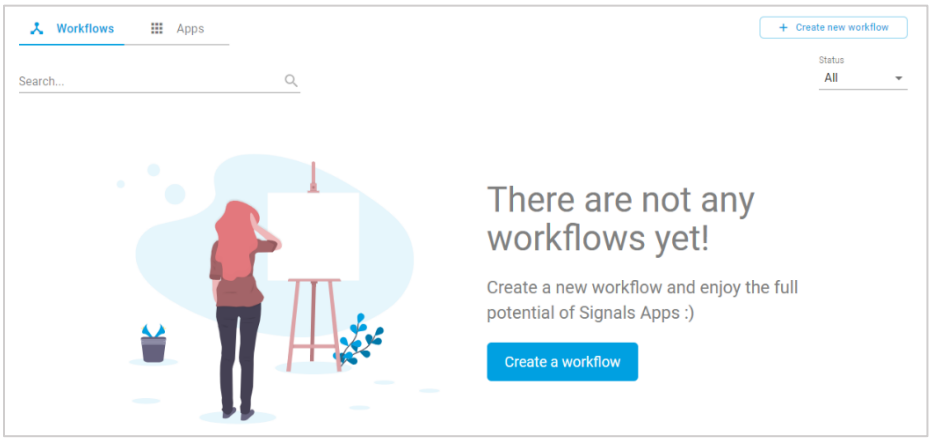


Figure 6-1: Signals Apps Workflows Tab with No Available Workflows

When clicking on either of these buttons the user will be presented with a dialog to choose the Workflow creation method:

- **From Scratch:** Select this option if you would like to create a Workflow in an empty document.
- **From Existing Document:** Select this option if you have a set of Apps already present in the document that you would like to save as a Workflow.

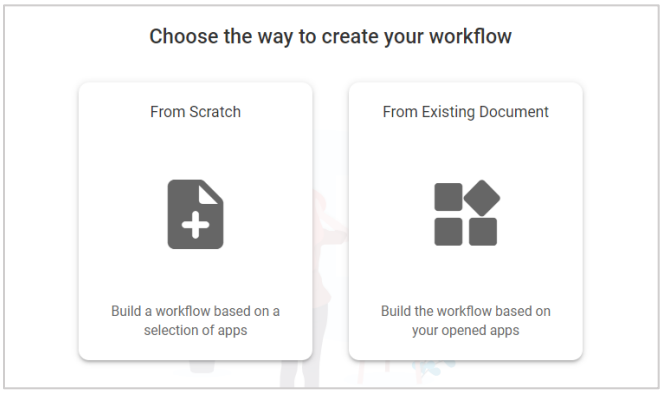


Figure 6-2: Workflow Creation Method

6.3.1 Creating a Workflow from Scratch

After choosing to create a Workflow **From Scratch**, an App selection screen will appear (Figure 6-3).

Apps can be filtered using the **Search** filter and added to the Workflow simply by clicking on them. All added Apps will be shown on the panel available on the right-hand side, which also provides controls to reorder or remove them from the Workflow.

The **Discard** button available at the top right-hand corner of the App selection area cancels the creation process and returns the user to the homepage. Finally, at the bottom of the right-hand panel there are two buttons: **Clear** removes all Apps from the Workflow and the **Configuration** guides the user to the next step.

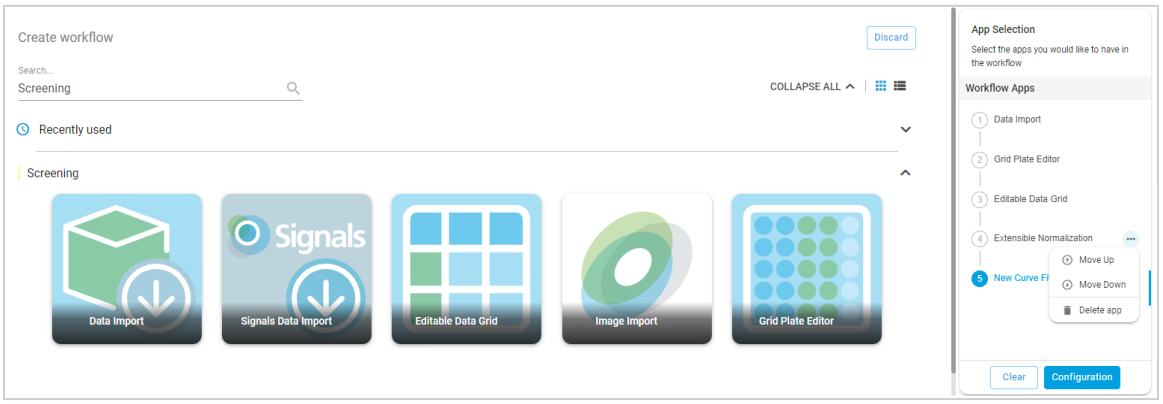


Figure 6-3: Selecting Apps to be Added to a Workflow Created from Scratch

After clicking on **Configuration**, the user will be able to navigate through the Workflow, configuring each App in the process. At the end of the navigation and configuration process, clicking on **Summary & Save** will lead you to the last stage (Figure 6-4).

In the last stage, the user will have to provide a unique name, choose a (new or existing) category, write a description (optional) and choose whether the Workflow will be saved as **Draft** or **Published**. The user can expand the **Instructions** area in the '**Save Workflow**' area to add App specific instructions for each App in the Workflow. Once the user is ready, the blue save button will save the Workflow and the system will return to the homepage. Selecting **Cancel** will return the user to the homepage, discarding the changes made to the Workflow (a confirmation dialog will be shown to ensure nothing is accidentally lost).

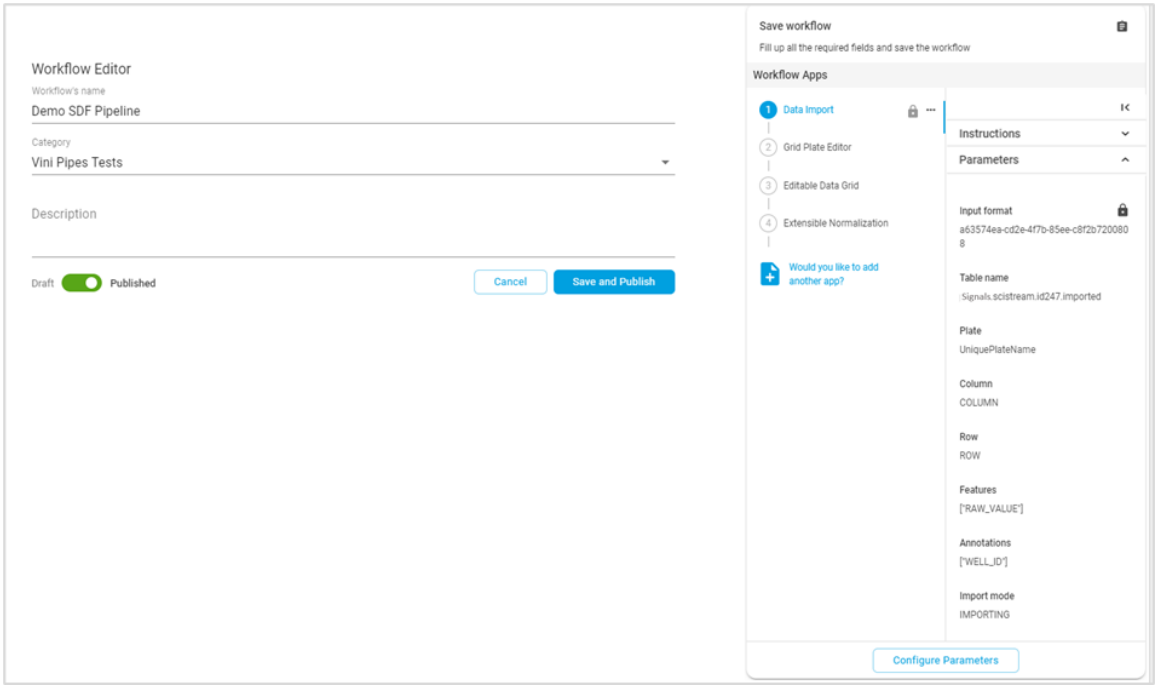


Figure 6-4: Last Step of the Workflow Creation Process

6.3.2 Creating a Workflow from an Existing Document

If the user has created an analysis Workflow in the document using the different Apps, it can be saved in the App store by selecting the Workflows tab and clicking on **Save as workflow**.

After choosing to create a Workflow from an existing document, the last stage of the Workflow creation process will be shown (Figure 6-4).

Just like creating a Workflow from scratch, the user will have to provide a unique name, choose a (new or existing) category, write a description (optional) and choose whether the Workflow will be saved as **Draft** or **Published**. Once the user is ready, the blue save button will save the Workflow and the system will return to the homepage. Selecting **Cancel** will return the user to the homepage, discarding the Workflow (a confirmation dialog will be shown to ensure nothing is accidentally lost).

Note: When creating a Workflow from an existing document, it is not necessary to navigate through the Workflow. The Apps in the document, their configuration and their order are used to create and configure the Workflow.

6.4 Workflow Navigation

When navigating a Workflow, the Workflow navigator panel (Figure 6-5) will be embedded in each App in the lower-left-hand corner and will allow the user to navigate forwards and backwards through the Workflow while configuring and executing the Apps. Since a Workflow can be executed in edition mode (creating from scratch, editing, cloning), in those cases the last App will include a **Summary & Save** button that records the changes made through the execution of the Workflow. During normal executions, for the first and last App, the **Previous** and **Next** buttons will be disabled accordingly.

When moving forwards, the new Apps' pages will be added with all the settings configured to the same values as were used when initially saving the Workflow. This allows the user to easily repeat the same type of analysis on different datasets with a minimum number of clicks.

Note: If a Workflow contains one or more Apps that do not support Web Player execution, then the Workflow will not be supported in Web Player.

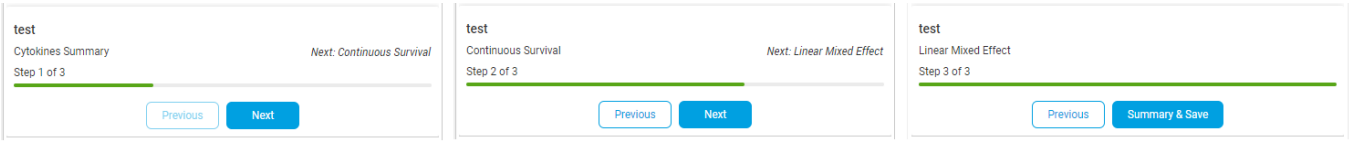


Figure 6-5: Workflow Execution in Edition Mode

If during an execution in edition mode the user attempts to return to the homepage, an overlay will be shown (Figure 6-6) to ensure the user continues the execution or correctly cancels the process (acknowledging the loss of changes).

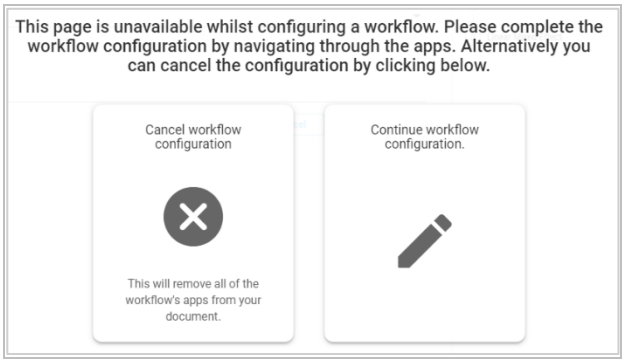


Figure 6-6: Exiting Out of Normal Execution Overlay

6.5 Editing a Workflow

The Workflow edition screen can be accessed through the **More actions** '…' icon on the Workflow card controls, or from the Edit button at the bottom of the Workflow's details panel (see Figure 2-2: User Interface of the Signals Workflow Tab). Note that this option is not available for Out-of-the-Box Workflows.

After choosing to edit a Workflow, the system will navigate to the Workflow edition screen (Figure 6-7).

Note: This screen will be the same as the one shown in the last step of the Workflow creation process.

Similarly, to creating a Workflow from scratch, the user will be able to edit the unique name, choose a (new or existing) category, edit the description and choose whether the Workflow will be saved as **Draft** or **Published**.

The user can expand the **Instructions** area in the '**Save Workflow**' area to add App specific instructions for each App in the Workflow. Once the user is ready, the blue save button will save the Workflow and the system will return to the homepage. Selecting **Cancel** will return the user to the homepage, discarding the changes made to the Workflow (a confirmation dialog will be shown to ensure nothing is accidentally lost).

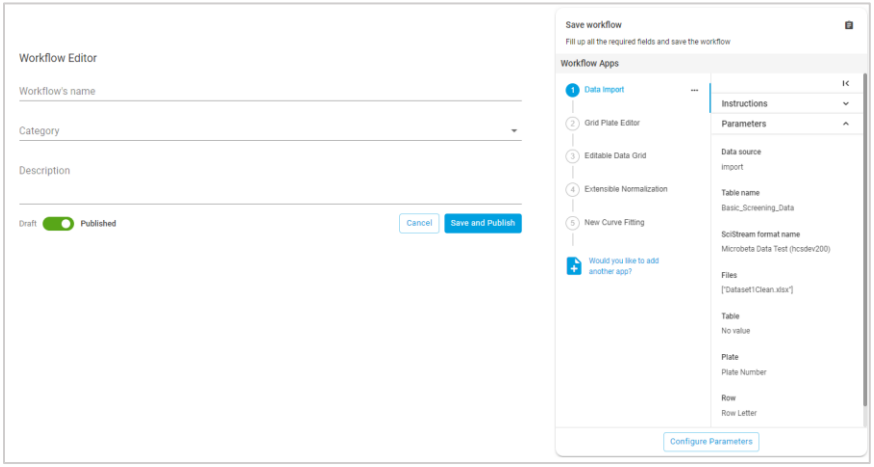


Figure 6-7: Edit/Clone a Workflow Save Screen

The user can click **Configure Parameters** (located at the bottom of the right-hand side panel) to execute the workflow in edition mode and reconfigure the App's parameters. The right-hand side panel also contains other controls that allow the user to further customize the workflow.

6.5.1 Parameter Locking

Parameters can be locked for editing to prevent the modification of their configured values during later executions by clicking on the App, hovering over the desired parameter and clicking on the corresponding *Lock* icon (see Figure 6-7: Edit/Clone a Workflow Save Screen). Note that parameter locking is only supported for SDK based Apps.

The following parameters in the SPR domain Apps cannot be locked:

App	Parameter
SPR Data Import	• Source Instrument
Zeroing	• Start Time • End Time
Cropping	• Start Time • End Time
Multi-Cycle Kinetics	• Start values: ka • Start values: kd • Start values: Rmax • Start values: kt • Start values: RI
Steady State Analysis	• Start of Equilibrium • End of Equilibrium • Analyze replicates separately • Rmax lower limit • Rmax upper limit • Rmin lower limit • Rmin upper limit • Slope lower limit • Slope upper limit
Relative Active Concentrations	• Molecule type

6.6 Cloning a Workflow

The Workflow cloning screen can be accessed through the **More actions** '…' icon on the Workflow card controls, or from the dropdown next to the 'Edit' button at the bottom of the Workflow's details panel (see Figure 2-2: User Interface of the Signals Workflow Tab).

After choosing to clone a Workflow, the system will navigate to the Workflow cloning screen (Figure 6-7: Edit/Clone a Workflow Save Screen).

Just like creating a Workflow from scratch, the user will have to provide a unique name, choose a (new or existing) category, write a description (optional) and choose whether the Workflow will be saved as draft or be published. Since the Workflow is being cloned, these values will be provided by default using the original Workflow as template. Once the user is ready, the blue **Save** button will save the Workflow and the system will return to the homepage. Selecting **Cancel** will return the user to the homepage, discarding the Workflow (a confirmation dialog will be shown to ensure nothing is accidentally lost).

The user can click the **Configure Parameters** button (located at the bottom of the right-hand side panel) to execute the Workflow in edition mode and reconfigure the App's parameters (see Workflow Navigation section).

As expected, once the Workflow is saved, a new cloned Workflow will be available in the homepage.

6.7 Deleting a Workflow

The Workflow deletion button can be accessed through the Workflow card controls (see Figure 2-2: User Interface of the Signals Workflow Tab) and will ask for confirmation to ensure the Workflow is not accidentally deleted. Note that this option is not available for Out-of-the-Box Workflows.

6.8 Running a Workflow

Consumers can access Workflows that have been created and stored in the App store by the authors and use them to perform complex analysis with a few simple clicks.

To use a Signals VitroVivo Apps Workflow the consumer should follow these steps:

- Have a dataset supported by the Workflow available. This would be a dataset that has the equivalent columns/parameters to those that are analyzed in the Workflow.
- Choose the desired Workflow. A Workflow can be executed using the **Execute Workflow** button (*Play* icon) available in the Workflow card controls or using the **Run** button found at the bottom of the Workflow's details panel.
- After clicking on either button, the Workflow will be executed in normal mode (see Workflow Navigation) and any changes made to the App's configuration will only be present in the App's pages corresponding to this specific Workflow execution.
- In the first step of the Workflow the user will need to either input the data in one of the data entry Apps, which will automatically match the required data fields to the Workflow requirements or match the data explicitly with the required fields if the first step of the Workflow is a Calculations Explorer App.
- After the initial data is matched, navigation through the Workflow can be easily done by clicking on the forward and back arrows to move through the different steps, adding when needed any extra input from the user required for the analysis.

6.8.1 Workflow Panel

The Workflow panel allows the user to review all the Workflow's details and perform certain actions.

The Workflow panel can be accessed via the **Show/hide workflow panel** button located on the right-hand side of the Workflow navigator. The Workflow panel will appear on the right-hand side of the screen and its upper part will show the main information of the Workflow and its lower part will contain a list of the Apps that compose the workflow. Please note that:

- The active App is highlighted in blue.
- If any of the Apps' parameters has been locked, the App will indicate it contains locked parameters by also displaying a lock next to it.

The upper part of the Workflow panel will include an ellipsis (…) button with some available actions:

- **Generate report**, which allows you to generate a pdf report of the Workflow execution.
- **Cancel Workflow**, which allows the user to cancel the Workflow execution and close all its Apps' tabs.

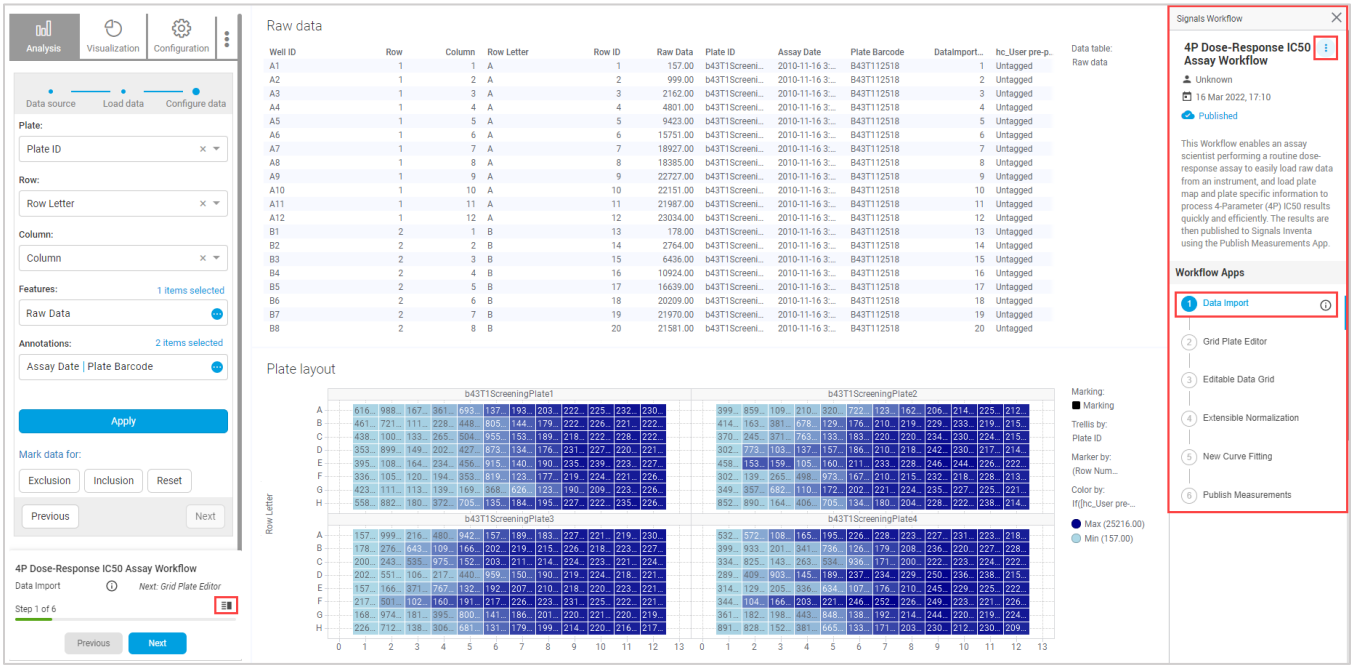


Figure 6-8: Access to the Workflow Panel (via the Button Marked with the Red Box) and its Functionality

6.9 Workflow Reports

The user can configure the report that the Workflow will generate and will be able to choose which Apps and data to include. This can be accessed by clicking on the **Open Report Setup** button found on the top right of the right-hand side panel, located next to **Save workflow** (see Figure 6-7: Edit/Clone a Workflow Save Screen) available during Workflow edition (see section 6.5 Editing a Workflow). Any user can check the report settings in read only mode by opening a Workflow and selecting the ⋮ icon, followed by **Report settings** (see Figure 6-9: Accessing the Report Settings Menu).

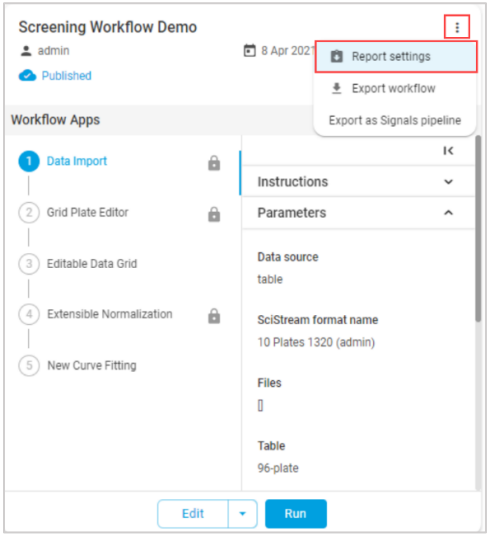


Figure 6-9: Accessing the Report Settings Menu

Figure 6-10: **Report Setup Options**

The report can be generated using the “Generate report” button (see Figure 6-7: Edit/Clone a Workflow Save Screen) from the Workflow panel (see section 6.8.1 Workflow Panel) available during the execution of the Workflow.

6.10 Workflow Instructions

The user can work with the Workflow panel during the Workflow creation or edition phase to configure the instructions to be displayed in the Workflow panel (during use phase) and in the Workflow details sidebar (in the main Workflows page).

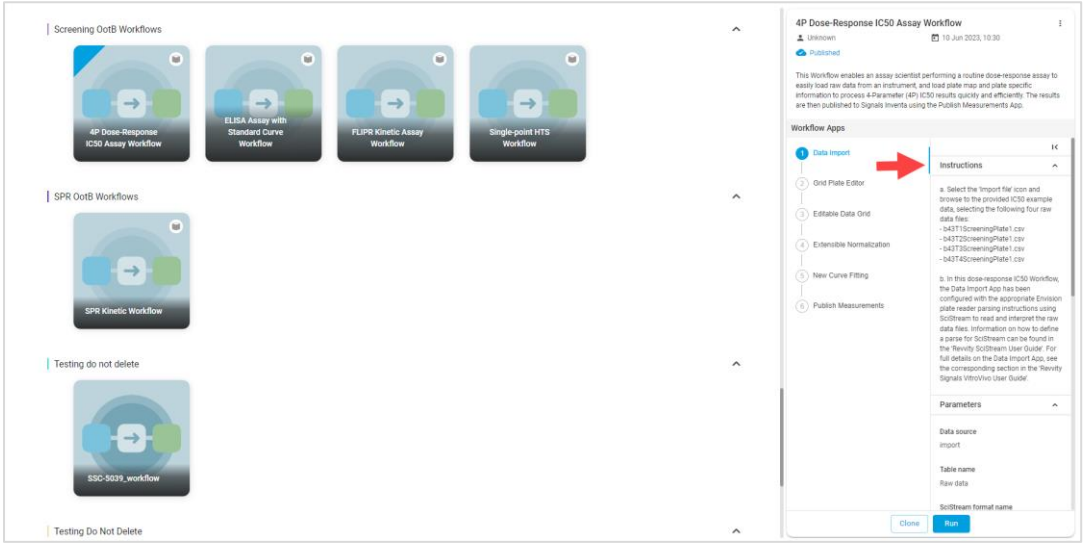


Figure 6-11: **Workflow Instructions (Non-Editable) as Seen by an End User (Highlighted by the Red Arrow)**

A Workflow developer may edit the Workflow instructions by simply opening the Workflow panel, then navigating to the relevant App and finally opening the Workflow panel and expanding the **Instructions** section. Please note that the format entered for the instructions will be used to display them later.

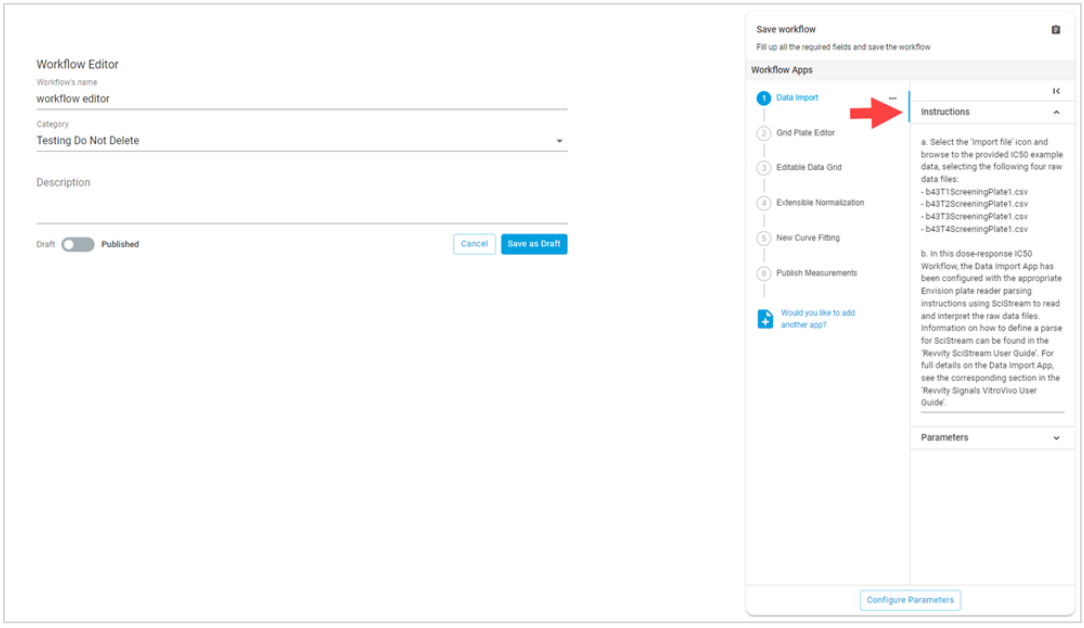


Figure 6-12: **The Editable Workflow Instructions Area (Highlighted by the Red Arrow)**

6.11 Exporting Workflows

To export Workflows, select the **Settings** dropdown (*Cog Wheel icon*) and **Export workflows**. The user may also hover over the lower right-hand corner of a Workflow card to reveal a white circle which will activate selection mode when selected. Note the user can expect the same behavior with the “List mode” view active.

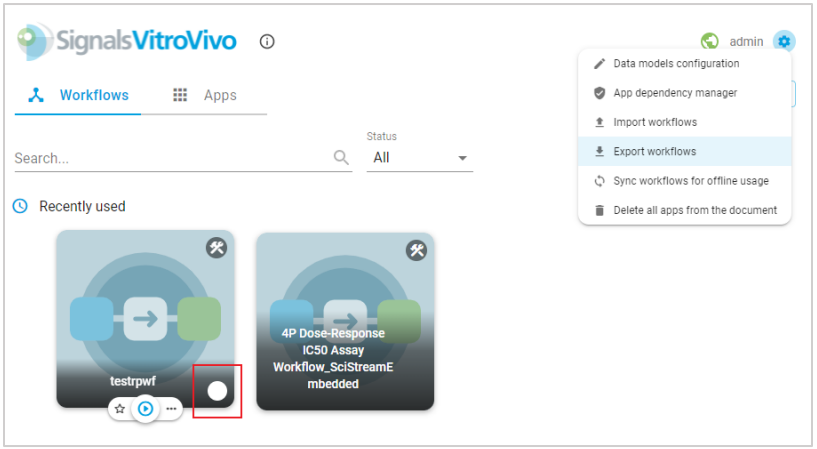


Figure 6-13: **Export Workflows Menu Option**

With the selection mode active, the user can select the Workflow(s) which they would like to export by selecting the corresponding Workflow card. Note the button with the text ‘**2 SELECTED WORKFLOWS’**, which displays the number of currently selected Workflow. Select the X next to the button to close the selection mode.

Additionally, note the checkbox in the lower right-hand corner of the Workflow card which indicates the Workflow has been selected for exporting, and is shown in full color so that it can be easily differentiated from the unselected Workflows.

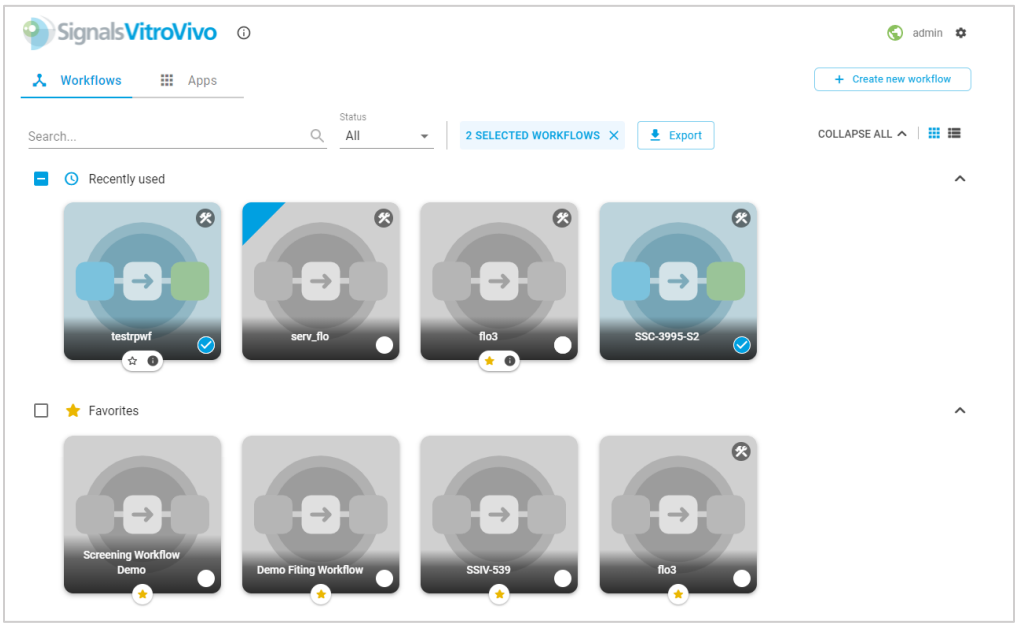


Figure 6-14: **Selection Mode for Exporting Workflows**

If the user wants to see the Workflow details, select the information icon when hovering over a Workflow.

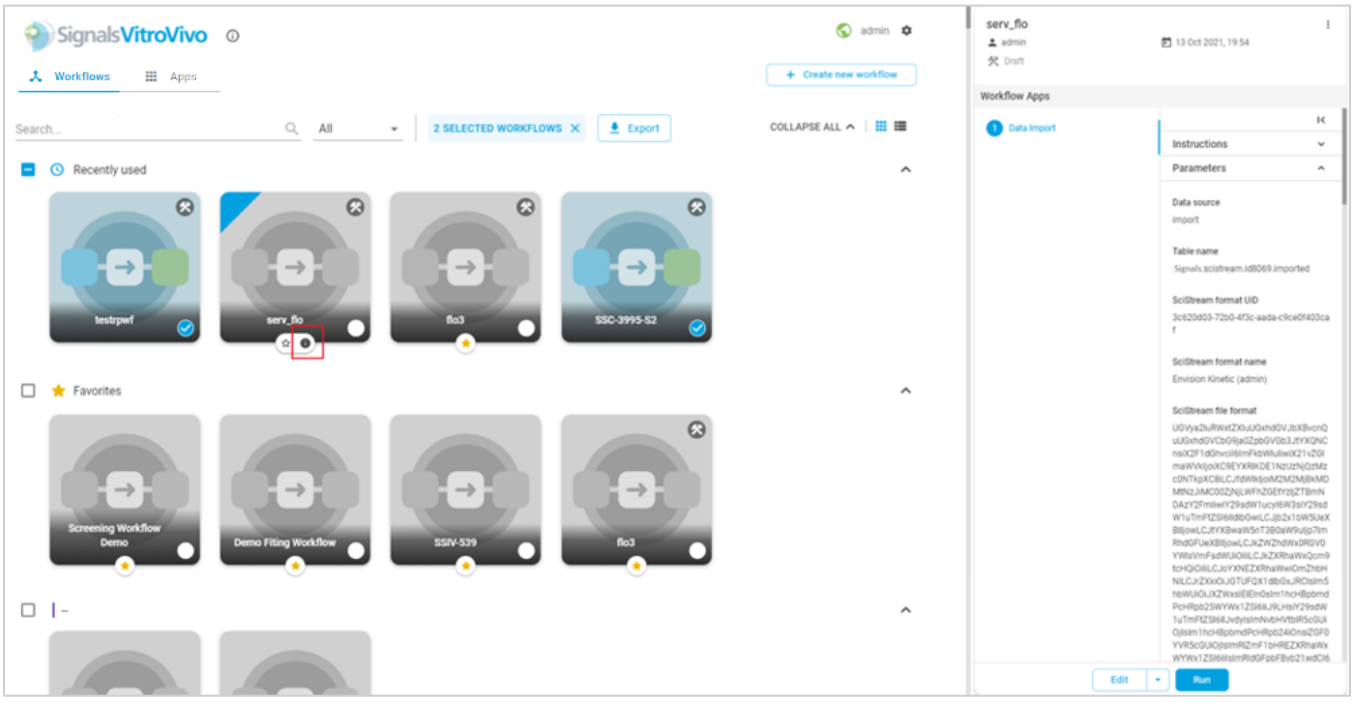


Figure 6-15: **Workflows Details Shown During Exporting**

When the user is ready to export the selected Workflow(s), select **Export** and a popup appears to save the Workflow(s) as a local .zip file with the generic file name “exported-worflows.zip”, by default.

6.12 Importing Workflows

To import Workflows, select the **Settings** dropdown (*Cog Wheel icon*) and **Import workflows**.

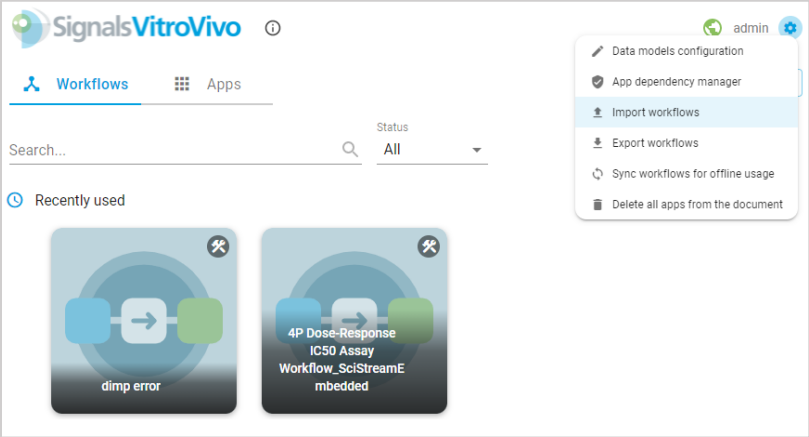


Figure 6-16: **Import Workflows Menu Option**

An **Open File** dialog will appear, prompting users to navigate to and select a .json file with the Workflow definition or a .zip file containing one or more .json files with the Workflow definitions. See section 6.11 Exporting Workflows for details on how to export Workflows.

Select **Open** to import the Workflow(s). When the process is complete, the imported Workflows will appear in the Workflows page.

6.13 Synchronizing Workflows for Offline Usage

If the user often works offline, it is in their best interest to keep the offline cache of Workflows synchronized. To do so, the user will have to select the **Settings** dropdown (*Cog Wheel* icon) located on the top right-hand corner of the Signals Apps page and select **Sync workflows for offline usage**. The system will then connect to the Spotfire® Library and synchronize all available Workflows.

Note: This process might take a while and may be canceled in the displayed progress dialog.

It is important to understand that the system will update the offline version of any Workflow interacted with, but there are certain cases (such as when configuring a system for the first time) that it will be useful to perform a full synchronization.

7. Signals Data Factory for Signals VitroVivo

Signals Data Factory (SDF) is an infrastructural component which provides pipelining capabilities for Signals VitroVivo and data transforming, indexing and query capabilities for Signals Inventa. After the measurement metadata has been correctly defined and measurement data is in Spotfire®, the user may publish the

measurements into Signals Data Factory (manually or automated in bulk) such that the measurement data can be transformed and indexed and then searched using complex queries in Signals Inventa's **Global Search**.

For more information on Signals Data Factory, refer to Appendix B of the *Revvity Signals Inventa User Guide*.

7.1 Exporting Workflows as Pipelines (Beta Feature)

The **Export as Signals Pipeline** Beta feature converts a VitroVivo Workflow into a Signals Pipeline to be run in a spark cluster to assist scientists in the automation of high-throughput screening campaigns involving multiple datasets. Note that this functionality is currently supported only for the OotB **4P Dose-Response IC50 Assay Workflow** as well as clones of this Workflow with minor changes to the settings.

7.1.1 Workflow Selection to Export as a Signals Pipeline

To export a Workflow as a Signals Pipeline:

1. From the **Workflows** tab, select the **4P Dose-Response IC50 Assay Workflow** card and note the side panel on the right-hand side which describes the details of the Workflow.
2. Select the **Export as Signals Pipeline** option, found under the […] ellipsis on the upper right-hand corner of the Workflow Details panel. This executes the Signals Pipeline builder from the set of analytics steps defined by Apps in the selected Workflow.

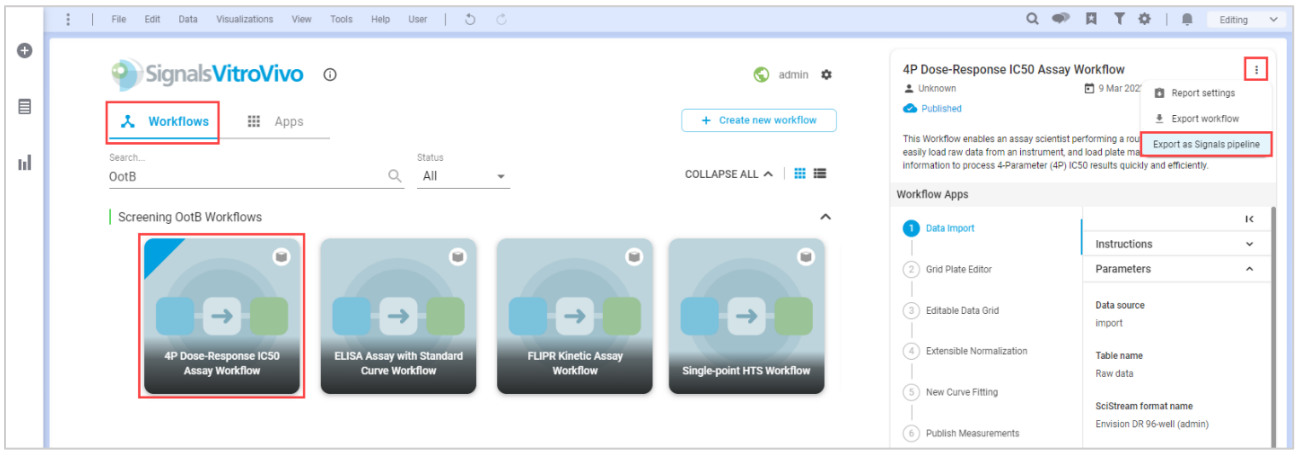


Figure 7-1: **Export as Signals Pipeline Option**

3. Store the zip file generated at any local path. The bundle contains the pieces of information needed to run the configured Workflow as a Signals Pipeline over a given assay dataset. The set of files contained inside the **4P Dose-Response IC50 Assay Workflow - SDF Pipeline.zip** file include:
    - **4P Dose-Response IC50 Assay Workflow - SDF Pipeline.json**: A sequence of analytics steps that execute the 4P Dose-Response IC50 Assay using the Signals Data Factory computing engine, named pipelining, using a spark cluster. Pipeline steps are the counterpart of analysis done by Apps in Spotfire®.
    - **4P Dose-Response IC50 Assay Workflow - SDF Pipeline.spm**: A script listing spm commands to operate in Signal Data Factory, including project and dataset creation, the addition of instrument

and metadata files, pipeline preparation for running, execution of the pipeline, mapping of results, and publishing. This file contains all the references to previous files and should be configured before execution for each screening campaign dataset.

- **automation_script_example.py**: A script example to configure and run the Signals Pipeline bundle.
- **Envision DR 96-well (admin).json**: The SciStream file format describing the format of the data contained within the instrument file.
- **README.md**: A readme file explaining the configuration and execution of this bundle.
- **Standard DR 96-well_PlateDesign.txt**: The plate design in .txt format to be applied to the instrument data.
- **VitroVivo 4P Dose-Response IC50 Assay Workflow Map - SDF Pipeline.map:** The mapping definition file in .json format between the Signals built-in Mtype "*Dose Response – Revvity Signals*" and the 4P curve fitting results data table for publishing the measurements into Signals Data Factory.
- **VitroVivo 4P Dose-Response IC50 Raw Data Map - SDF Pipeline.map:** The mapping definition file in .json format between the Signals built-in Mtype "*Dose Response Well-Level Details – Revvity Signals*" and well-level screening data for publishing the measurements into Signals Data Factory.

7.1.2 Configuration and Execution of a Signals Pipeline

The exported bundle contains the collection of files required to run the **4P Dose-Response IC50 Assay Workflow** using the pipelining engine provided by Signals Data Factory. This bundle is generic and can be used to automatize the analysis of multiple datasets in a screening campaign.

To configure the bundle for a given screening dataset, the user needs to complete empty parameters (arguments between "<>", placeholders) within the provided SPM script, **4P Dose-Response IC50 Assay Workflow - SDF Pipeline.spm**, including the following required information:

- **VitroVivo project name**: Name of the Signals Data Factory project where pipelining results for the 4P Dose-Response IC50 assay will be published.
- **Instrument data file name:** Name to label the instrument file inside the project dataset.
- **SDF dataset name:** Dataset name for a given screening run inside the campaign.
- **Instrument data file path:** Path to local file containing instrument data.
- **Compounds metadata format**: Format of the file containing compound annotations.
- **Compounds metadata file name:** Name to label the annotations file inside the project dataset.
- **Compounds metadata file path:** Path to local file containing annotations to be added to instrument data. Usually compound annotations.

```
# Create Pipeline script

# Create SDF project
project -create -v2  <VitroVivo project name>

# Add Raw file
add-file -name <Instrument data file name> -dataset <SDF dataset name> -scistreamtemplate 'Envision DR 96-well (admin).json' <Instrument data file path>
# Set Dataset for pipeline
set-dataset-pipeline --dataset <SDF dataset name> --pipelineFile '4P Dose-Response IC50 Assay Workflow - SDF Pipeline.json'

# Add Grid Plate Editor metadata
add-file -format tsv -backslashEscape -name 'Standard DR 96-well_PlateDesign.txt' -dataset <SDF dataset name> -metadata 'Standard DR 96-well_PlateDesign.txt'
# Set Grid Plate Editor join step
set-dataset-step-target --step 'Grid Plate Editor' --dataset <SDF dataset name> --targetFiles 'Standard DR 96-well_PlateDesign.txt'

# Persist Grid Plate Editor step
set-dataset-step-persist --dataset <SDF dataset name> --step 'Grid Plate Editor'

# Add Editable Data Grid metadata
add-file -format <Compounds metadata format> -backslashEscape -name <Compounds metadata file name> -dataset <SDF dataset name> -metadata <Compounds metadata file path>
# Set Editable Data Grid join step
set-dataset-step-target --step 'Editable Data Grid' --dataset <SDF dataset name> --targetFiles <Compounds metadata file name>

# Persist Editable Data Grid step
set-dataset-step-persist --dataset <SDF dataset name> --step 'Editable Data Grid'

# Add map
set-dataset-step-map --step 'New Curve Fitting' --dataset <SDF dataset name> --mapFile 'VitroVivo 4P Dose-Response IC50 Assay Workflow Map - SDF Pipeline.map'

# Add map
set-dataset-step-map --step 'Extensible Normalization' --dataset <SDF dataset name> --mapFile 'VitroVivo 4P Dose-Response IC50 Raw Data Map - SDF Pipeline.map'

# Execute Pipeline
execute-dataset-pipeline --dataset <SDF dataset name>
# Import and publish results
load
```

Figure 7-2: **Example Pipeline SPM Script**

Once configured for a given dataset, the SPM script should be executed using the **Signals Project Manager** (spm) tool, provided as part of the Signals Inventa goods. The pre-requisites for pipeline execution include:

1. **SPM tool:** The commands used to run the Signals Pipeline are defined in spm scripting language. Therefore, the spm tool is needed to interpret the commands.
2. **Inventa installation:** The pipeline execution is performed in a Signals Data Factory instance, a component of Signals Inventa. For further details on installation, refer to the *Revvity Signals Inventa Installation Guide*, and for details on the OotB 4P Dose-Response configuration refer to the *Revvity Signals VitroVivo Installation Guide*.

Automation of configuration and execution can be done using scripting. Within the downloaded Signals Pipeline bundle for the **4P Dose-Response IC50 Assay Workflow**, an example python script, *automation_script_example.py*, for configuration and execution is provided. An execution example using the provided data files is available in the documentation folder, **Sample Data > OotBScreening Workflows > 1st Demo - 4P Dose-Response IC50 Assay**.

The Signals Pipeline bundle can be reused as many times as needed with different screening datasets. Each execution should be uniquely identified by Project and Dataset name.

7.1.3 Verify Pipeline Execution in a Signals Data Factory Instance

Below are examples of Signals Data Factory pipeline executions for three screening runs using the Signals Pipeline derived from the OotB **4P Dose-Response IC50 Assay Workflow**.

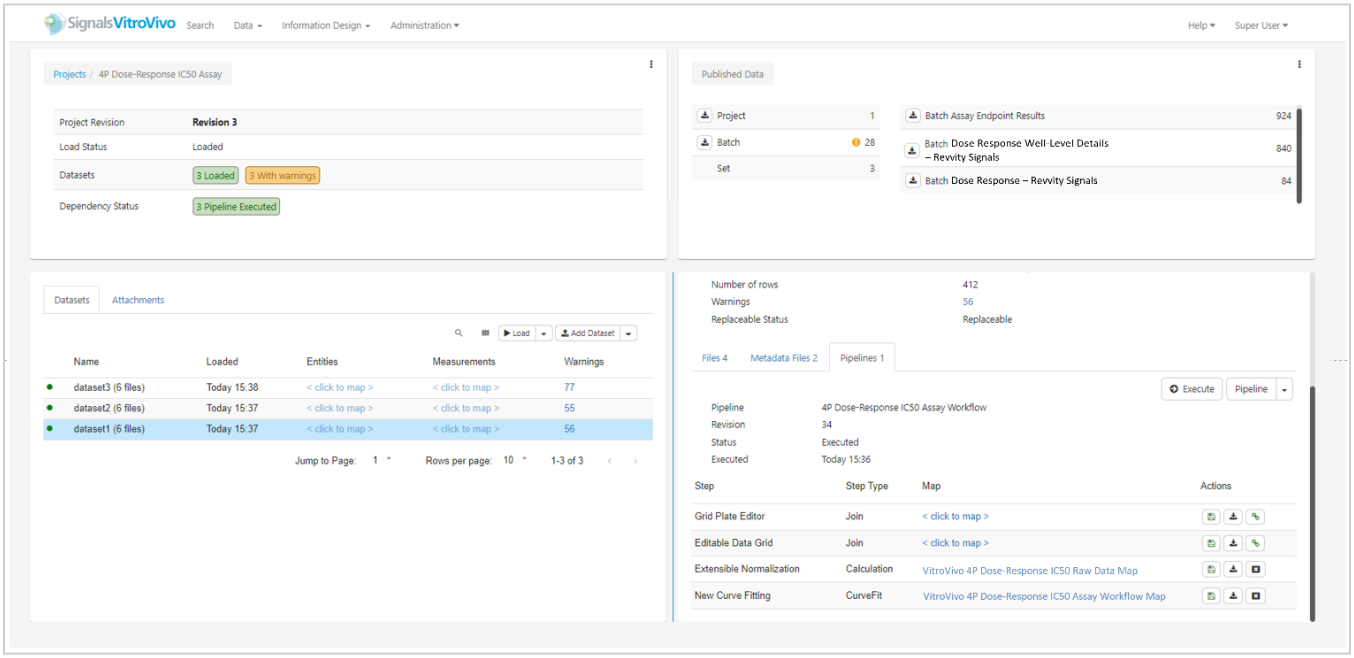


Figure 7-3: **Example Signals Data Factory Pipeline Executions**

8. Sharing DXP Documents Containing Signals Apps in Spotfire® 10.3 and Later

Signals Apps use data functions internally to run analytics. As of Spotfire® 10.3, a security mechanism is present that requires data functions to be trusted before they can be executed. This security mechanism is enabled by default on the Spotfire® Server 10.3 and newer. If there are untrusted data functions in the DXP the user will get a notification.

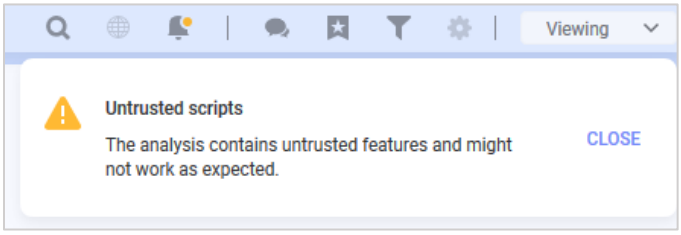


Figure 8-1: **Untrusted Scripts Notification**

The notification details will show a message similar to the following:

> “Untrusted data function (4)
>
> The data function ‘app.test.id1234.run’ is not trusted and could not be executed.
>
> The data function ‘app.test.id1234.run ‘is not trusted.
> You should only trust a data function if you are certain that it comes from a reliable source.”

When a user wants to share a DXP document that contains Apps with data functions, this user must belong to the **Script Author** group in the Spotfire® Server. If the user does not belong to this group the data functions will not be trusted by default by other users, and the shared DXP document will not run the Apps. To bypass this error in Spotfire® Analyst, manually trust the data functions using the **File** > **Manage trust and scripts** menu.

9. Screening Domain Apps

9.1 Introduction

The goal of the Signals VitroVivo Apps is to break down complex analysis Workflows into flexible and small analysis steps allowing users to integrate information, visualize it and interact with it across different domains.

Signals VitroVivo Apps may require additional installation and configuration. For a description of the required configuration, please refer to the *Signals VitroVivo Installation Guide* and the *Signals VitroVivo New User Installation Orientation.* The rest of this section will describe Signals VitroVivo Apps in the **Screening** domain.

9.2 Launching Apps

To launch an App from Spotfire® Analyst, from the main toolbar, navigate to **Tools** > **Signals Apps.** This will open the Signals Apps tab. Alternatively, if this tab is already present in the document, simply navigate to it using the appropriately labelled tab at the bottom of the screen.

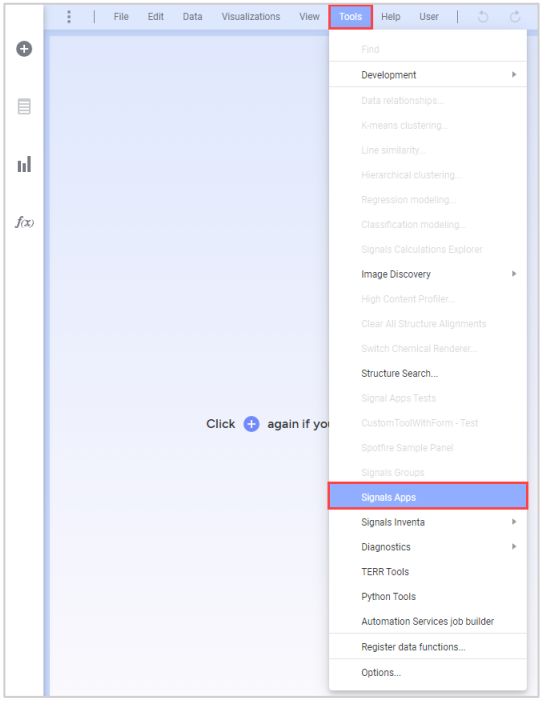


Figure 9-1: Launching Signals Apps

By default, the **Workflows tab** will open upon launch. Navigate to the **Apps tab** to view the available Signals Apps. Scroll down to the **Screening** domain Apps area.

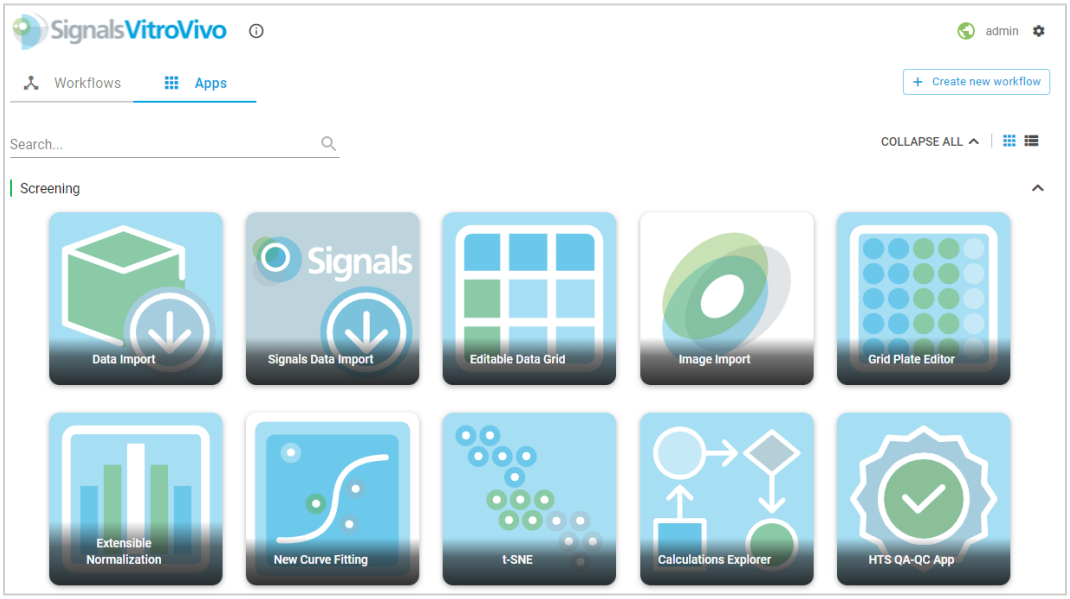


Figure 9-2: Screening Domain Apps

Select the desired App card and a new tab containing the selected App will be added to the document. When hovering over the App card, select the *Star* icon to mark the App as a Favorite or select the *Info* icon to learn more about the App. The *Play* icon will also launch the App.

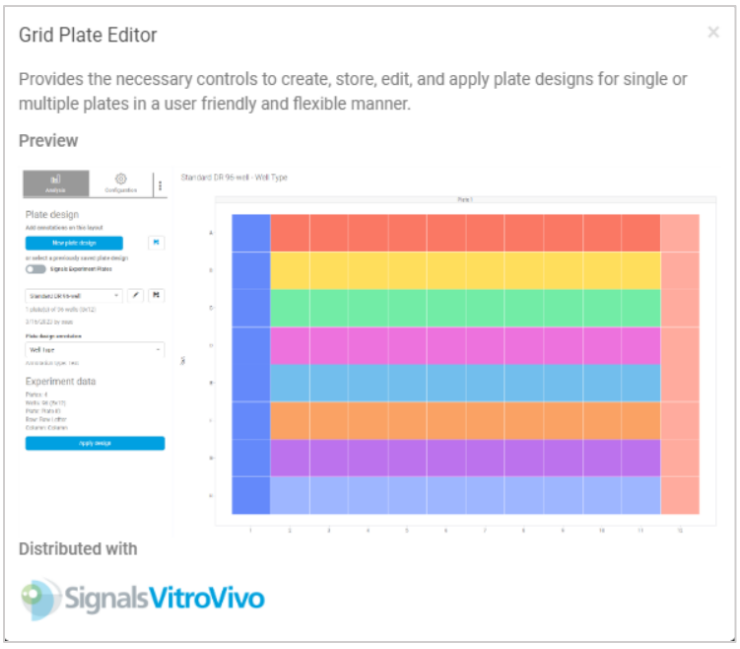
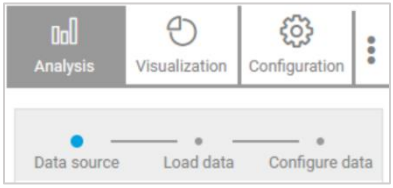


Figure 9-3: App Preview Shown via the *Info* Icon

9.3 App Tab Overview

Most Screening Apps contain a user interface area with three tabs, **“Analysis”, “Visualization”, “Configuration”** and an additional panel **“…”,** displayed in the top-left-hand corner.

- **Analysis:** Contains the main controls for using the App and may contain one or more steps, as shown in the example below, specific to the Data Import App. The current step is highlighted in blue.



- **Visualization (if available):** Provides a toggle to display the **Filter Panel** with a refresh icon. Additional App-specific options may be available and are described, if applicable, in the appropriate Screening App section.

- **Configuration:** Generally, contains a toggle to turn on/off **Guided Mode** (on by default). Additional App-specific options may be available and are described, if applicable, in the appropriate Screening App section.

- **‘…’:** Contains additional information related to the App. Here users can find the “Notification Center”, “About”, “Help” and the ability to “Minimize” the User Interface.

  - **Notification Center:** Provides access to detailed information, warnings, and error messages produced while performing an analysis. Selecting a title will filter the messages by type.
  - **About:** Contains general information about the App, including simplified App name and version.
  - **Help:** When available opens a page with the WebHelp for the App in the user’s browser.
  - **Minimize:** Collapses the UI to provide visualizations more space in the document.

9.4 Data Import App

The **Data Import** App provides a Spotfire® connector to instrument formats in the local SciStream installation to facilitate the importing of data from raw instrument files.

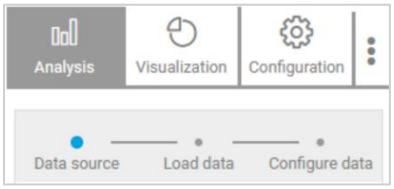
From the Signals Apps page, select the **Data Import** App card. A new tab containing the App will be added to the document and the App will be launched.

Figure 9-4: Data Import App Card

9.4.1 Configuring the Data Import App

9.4.1.1 Analysis Tab

The **Analysis** tab guides the user in preparing, loading, and configuring the data with the following three steps: *Data source*, *Load data* and *Configure data*.



The user has the following controls available from the *Data source* step of the **Analysis** tab**:**

1. From the '**Select Raw Data Source:'** dropdown, select if the data should be loaded from:
    a. An external file via *'File Upload'*
    b. An existing table already present in the document via '*Existing table'*
    c. An existing table in Signals Notebook via '*Signals Experiment'.*
2. Set the name of the table that will contain the imported data. A default table name will pre-populate, and it is recommended that the user edits the table name to be meaningful.

    Note: This option is only available when loading data from an external file or from Signals Notebook.

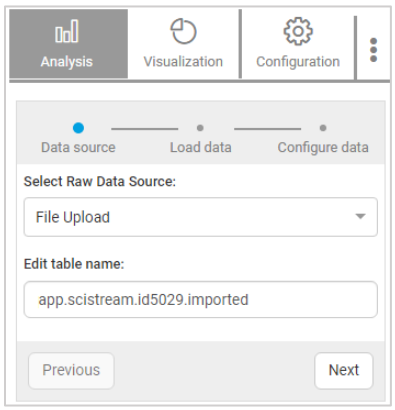


Figure 9-5: Data Source Step in Data Import App for External Files

3. Select **Next** to continue to the *Load data* step and proceed with one of the following options depending on if the data is loaded from an external file, from an existing table in the document, or from Signals Notebook.

In the case of an external file:

a. From the '**Select import format:'** dropdown menu, select the required format. Note under the dropdown selector there are controls that allow the user to manage the formats that are available in the server and locally.

- **Download:** Download the selected format from the server.
- **Upload:** Upload the selected format to the server, replacing it if already present.
- **Delete:** Deletes the remote copy of the selected format from the server.
- **Description:** Displays an overlay with author inputted description of selected format.

Note: Upload and Delete are only available if the user is the author of the selected format.

b. In the '**Imported files:'** section, select the **Import File** button to launch the **Open File** menu and select the desired file(s) (hold Shift to select multiple files) to be loaded from your local drive. The loaded table appears in the visualization area.

c. Select the ***x*** icon next to an imported file to remove it from the box. All data loaded from that file will be removed.

d. Select **Next** to move on to *Configure data* step.

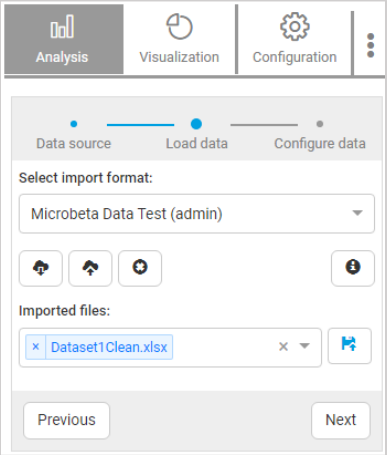


Figure 9-6: Load Data Step in Data Import App for Instrument Files

In the case of an existing table:

a. Select the desired table to be used from the dropdown menu.

b. Select **Next** to move on to the *Configure data* step.

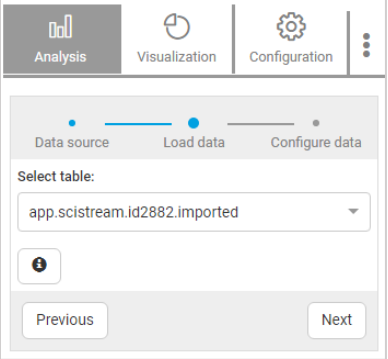


Figure 9-7: Load Data Step in Data Import App for Existing Table

Note: When loading data from an existing table, if the user saves the analysis as a Workflow, the data configuration is not saved, as it is not loaded with any of the formats known to the App. Therefore, the data should be configured for any new data when running the Workflow.

In the case of a Signals Notebook table:

a. From the '**Select import format:'** dropdown menu, select the required format. Note under the dropdown selector there are controls that allow the user to manage the formats that are available in the server and locally.

- **Download:** Download the selected format from the server.
- **Upload:** Upload the selected format to the server, replacing it if already present.
- **Delete:** Deletes the remote copy of the selected format from the server.
- **Description:** Displays an overlay with author inputted description of selected format.

Note: Upload and Delete are only available if the user is the author of the selected format.

   b. Select the ⋯ icon and the checkbox next to the desired data tables in the **Available items** list. Select **Add items > Save** to close the menu.

   c. Select **Next** to move on to the *Configure data* step.

The *Configure data* step allows the user to match the appropriate column header of the data to the corresponding data type for downstream analyses with additional Apps. To configure data:

1. Select the **Plate** column.
2. Select the **Row** column.
3. Select the **Column** column.
4. Select the **Features** column(s) using the ⋯ icon and the checkbox next to the desired feature(s) in the **Available items** list. Select **Add items** > **Save** to close the menu. Note these are the numeric data columns that will be used for the analysis.
5. Select the **Annotation** column(s) using the ⋯ icon and the checkbox next to the desired annotation(s) in the **Available items** list. Select **Add items** > **Save** to close the menu.
6. When connected to Signals Notebook, select the **Experiment properties** using the ⋯ icon and the checkbox next to the desired properties in the **Available items** list. Select **Add items** > **Save** to close the menu. The selected properties will be added as columns to the imported data table.
   a. **Add properties as separate table:** Allows the switching between adding the properties as new columns to the imported table (default) or adding them in a separate table with one row per property.

Once the data has been configured click **Apply** to complete the data import process.

Note: Once applied, the **Mark data for** menu is available for users to exclude or include marked data points. Mark data by selecting it in the visualizations in the right-hand panel and then selecting the corresponding **Exclusion** or **Inclusion** button. The user can optionally reset all excluded data by selecting the **Reset** button.

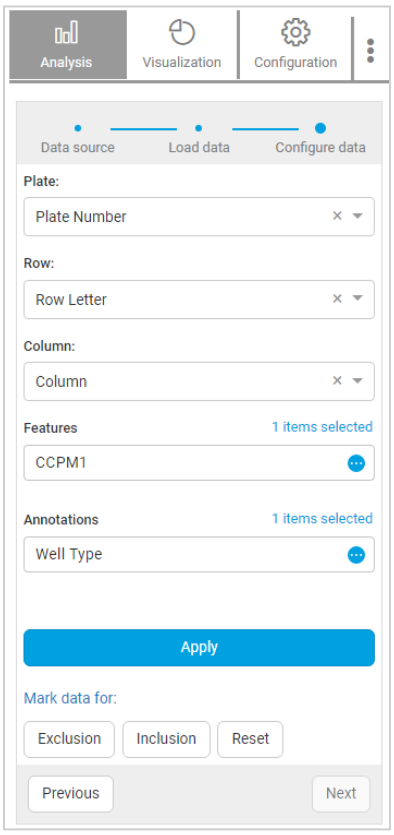
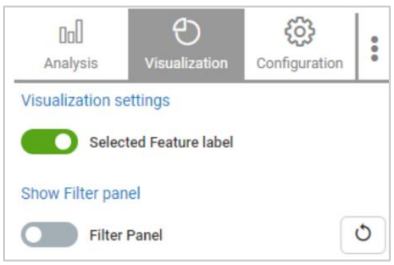


Figure 9-8: Configure Data Step

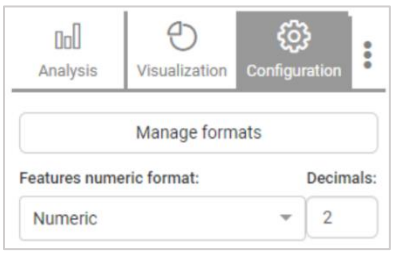
9.4.1.2 Visualization Tab

The **Visualization** tab provides toggles to display the **Selected Feature label** and the **Filter Panel**.



9.4.1.3 Configuration Tab

The **Configuration** tab permits access to the **Manage formats** menu using SciStream and allows the definition of the format for the numeric features.



9.4.2 Data Import App Results

The **Data Import** App provides a tabular representation of the data extracted from the selected file(s) where the columns are tagged with any additional information specified in the *Configure data* step. A preview of the layout of the plate will be shown in the bottom section of the App if the plate layout columns have been defined.

Figure 9-9: Data Import App Main Visualizations

9.4.3 Managing Formats

To access the Manage Formats dialog, select the **Configuration** tab and select **Manage formats**. The full range of functionality is invoked from SciStream. For more details, refer to the *Revvity Signals VitroVivo SciStream User Guide*. The following are the sections of the SciStream User Guide directly related to the functionality brought to the **Data Import App**:

- “3. Importing Instrument Files into Spotfire®”
- “5. Creating a New File Format”
- “6. Pre – defined File Formats”
- “7. Editing a File Format”

- “8. Deleting a File Format”
- “9. Importing and Exporting a File Format”

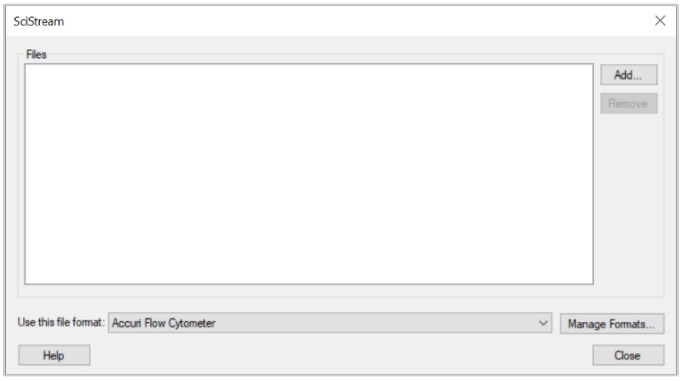


Figure 9-10: Manage SciStream Format

9.4.4 Configuring the Data Import App Library

Refer to the *Revvity Signals VitroVivo Installation Guide* for details on this topic.

9.5 Signals Data Import App

The **Signals Data Import** App imports screening results published in a Signals Data Factory instance into Spotfire® as tables. Note that this App requires Signals Data Factory or Signals Inventa credentials.

From the Signals Apps page, select the **Signals Data Import** App card. A new tab containing the App will be added to the document and the App will be launched.

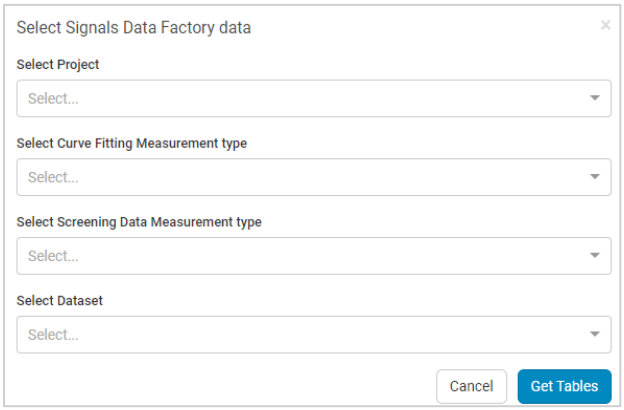
Figure 9-11: Signals Data Import App Card

9.5.1 Configuring the Signals Data Import App

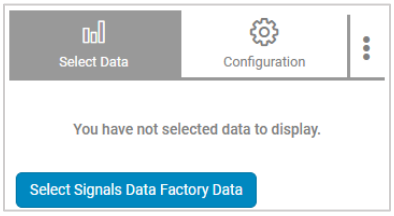
Once the App is open and the user has entered their Signals Data Factory or Signals Inventa login credentials, the **Select Signals Data Factory data** dialog will be displayed. Select from each of the dropdowns the desired data, then select **Get tables**. The required dropdowns are:

- **Project**: the name of the Signals Data Factory project where the pipelining results were published.
- **Curve Fitting Measurement Type**: the measurement type name for the Curve Fit results.
- **Screening Data Measurement Type**: the measurement type name for the well-level screening data.

- **Dataset**: the dataset in the selected Project to perform QA. Only datasets with published measurements for both selected measurement types will be listed.

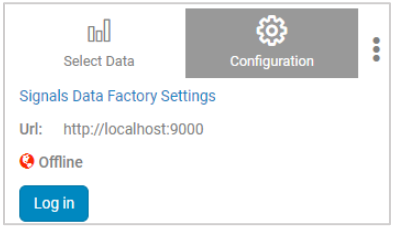


Note: If the user selects **Cancel** a message stating that no data has been selected will be displayed. Select the **Select Signals Data Factory Data** to reopen the data selection dialog.



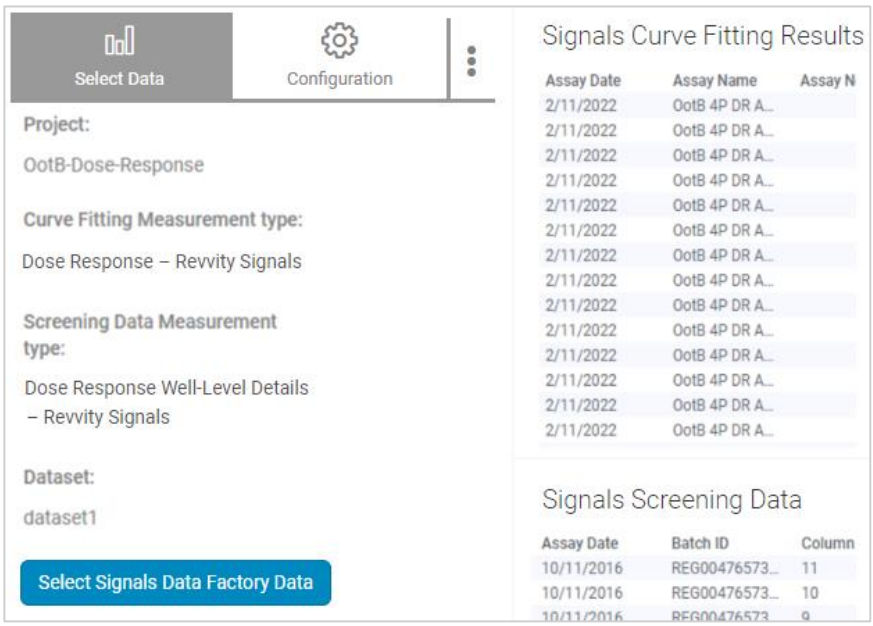
Once the user has selected that data to import, two tabs, **"Select Data", "Configuration"** and an additional panel **"…",** will become enabled in the top-left hand corner.

- **Select Data:** Opens the data selection dialog.
- **Configuration:** Provides information on the status of the Signals Data Factory instance, including Url, connection status, user (when connected), and when offline a 'Log in' button to facilitate connection.



9.5.2 Signals Data Import App Results

The left-hand panel contains the information entered in the data selection dialog and can be edited by selecting the **Select Signals Data Factory Data** button. Two visualizations are generated in the right-hand panel: *Signals Curve Fitting Results* and *Signals Screening Data*.



9.6 Image Import App

The **Image Import** App provides a Spotfire® connector that supports the performance of advanced search and filter features in Image Artist and remote Columbus servers, retrieving the resulting selections directly into Spotfire®.

Note: This App replaces the deprecated **Columbus Navigator** App, and the **Image Import** App supports both Image Artist and Columbus servers.

From the Signals Apps page open the **Image Import** App.

Figure 9-12: Image Import App Card

A new page is created which allows the user to browse data with the **Image Import** App.

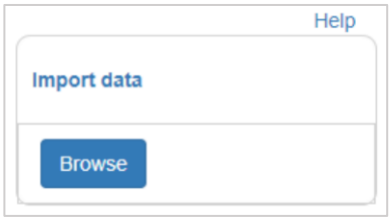


Figure 9-13: Import Data from Image Artist or Columbus

Select **Browse** in the "**Import data**" section to launch the login screen. Afterwards the functionality of the App invokes Image Artist/Columbus. For more information, refer to the Image Artist/Columbus Help from within the application.

9.6.1 Login

To login to an Image Artist or Columbus server:

1. In the **Login – Image Import** dialog, enter your **User name** and **Password** as well as the **URL** of the appropriate server. Ensure no trailing slash is included.

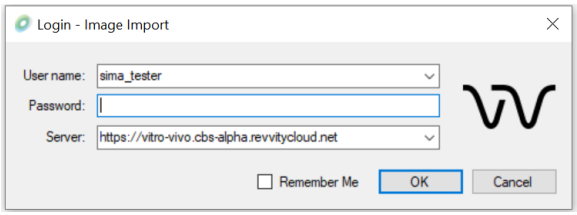


Figure 9-14: Image Import Login Window

2. Check the **Remember Me** checkbox to save the current User name so that the next time the login dialog is opened the user name will populate automatically. Select **OK**.

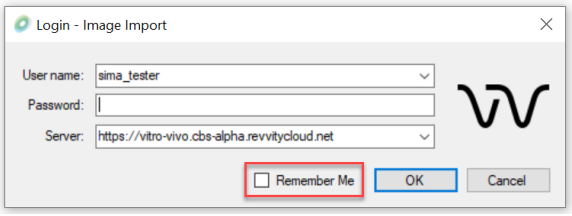


Figure 9-15: Remember Me Checkbox

Note: The password will not be saved for security reasons and will need to be inputted each time.

3. Alternatively, select a remembered user from the **User name** list.

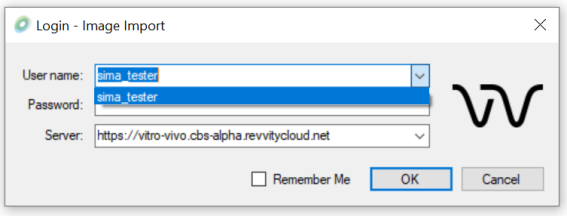
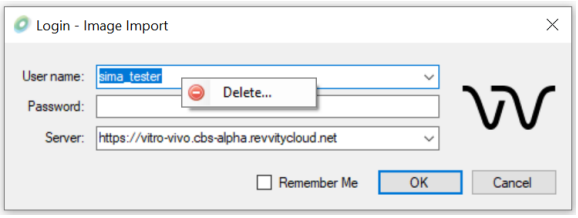
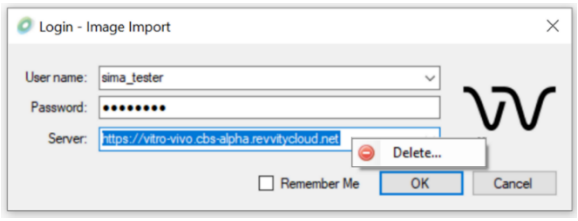


Figure 9-16: User name List

4. Right-clicking the **User name** list will display a **Delete** menu option to delete the user that currently appears in the **User name** box.



5. Once the user has successfully logged in to the selected server, the Server URL will be saved to the **Server** list box. Next time when opening the login dialog, the Server URL that was previously accessed will be populated automatically in the **Server** box. The user may also delete the currently selected URL by right-clicking the **Server** list box and selecting **Delete**.



9.6.2 Filtering the Results

Once logged in, the main **Image Import** interface is displayed. To narrow down the assay definitions and assay results displayed, the filters available on the left-hand side may be used:

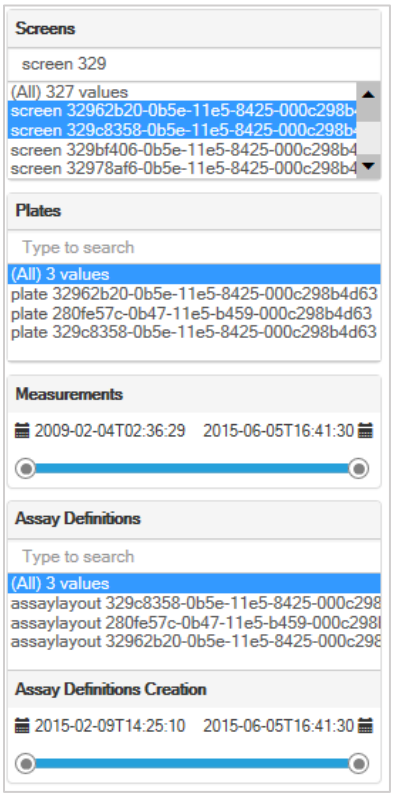
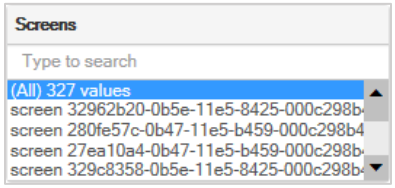


Figure 9-17: Example of Columbus Filters

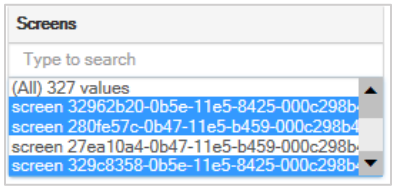
Note: Additional filters not pictured in the image above include Users, Results, and Results Creation.

There are two kinds of filters: a **textbox filter** and a **slide bar**.

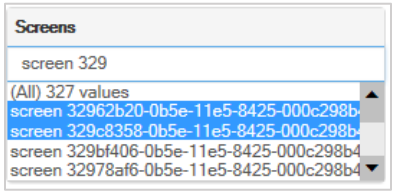
1) For a textbox filter, consider the '**Screens**' example below:



Select one or multiple screens to filter directly by using the **SHIFT** or **CONTROL** keys as shown below:



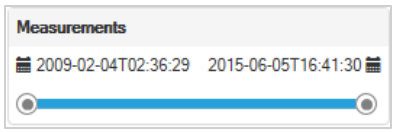
Use the '**Type to search**' textbox to perform a quick search based on what is required:



Note: Select multiple entries by pressing the **SHIFT** or **CTRL** key while performing a quick search.

Note: Upon selecting a certain screen, the other filters such as plates, assay definition, or assay result will change accordingly, however slider bars are not influenced by other filter changes.

2) For a slide bar filter, consider '**Measurements**' as an example.



This control has two pointers: Min and Max. Drag these two pointers left or right to filter data.

9.6.3 Assay Definition and Assay Result

After filtering, the assay definition and assay result will be listed in the bottom-right-hand side for selection.

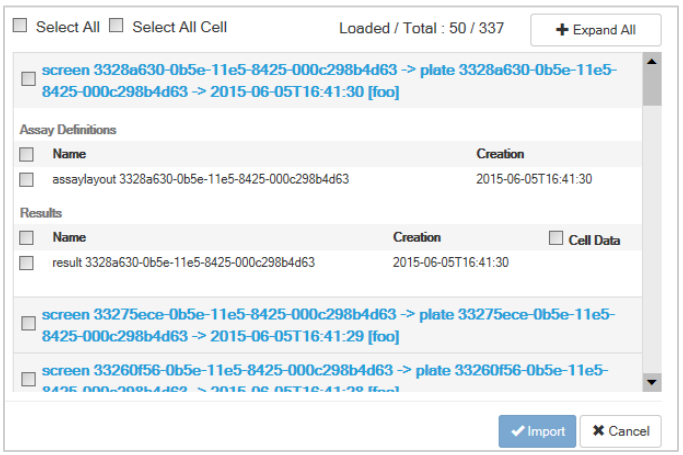
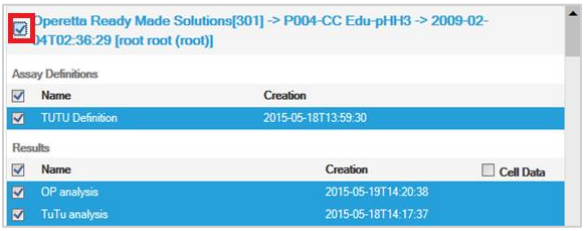


Figure 9-18: Assay Definition and Assay Results Selection

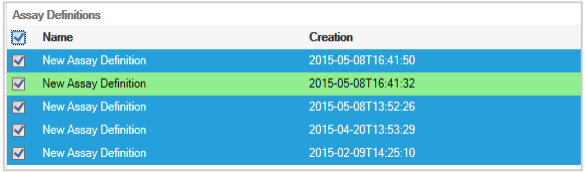
Each item displayed in blue is a measurement and can be expanded and collapsed by selecting the '**Expand All**' button.

- **Select All**: Selects/de-selects all the assay layout definition and result data.
- **Select All Cell**: Selects/de-selects all the cell level data if the result data has cell level data.
- **Loaded/Total**: Shows the total count of matched measurements and the count of currently loaded measurements.
- **Expand All/Collapse All**: Expands/collapses all measurements.

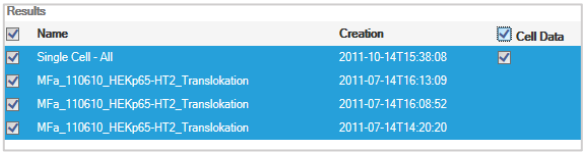
- For certain measurements, select all assay definition and results which belong to the current measurement by selecting the checkbox to its left-hand side.



- For each definition, choose a row by selecting the corresponding checkbox on the left, or select all rows using the checkbox in the top-left-hand corner.



- For each assay result, choose a row by selecting the corresponding checkbox on the left, or choose all rows using the checkbox in the top-left-hand corner. Select all cell data using the checkbox in the top-right-hand corner of the grid.



Scroll down using the scrollbar to load more data.

9.6.4 Image Preview

An image preview is provided in the top-right-hand side when a measurement is selected.

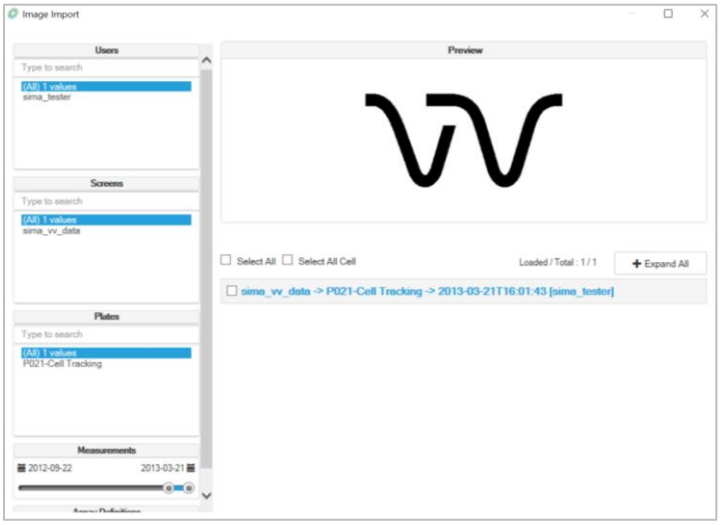


Figure 9-19: Image Preview - Displays the Revvity Logo by Default

9.6.5 Import Data

If any assay definition or result is selected, the **Import** button will be enabled.

Upon selecting the **Import** button:

- If “**Select All**” or “**Select All Cells**” is checked all data is imported including well and cell level data information.
- If “**Select All**” and “**Select All Cells**” are not checked the assay layout, result, and selected cell level data will be imported.

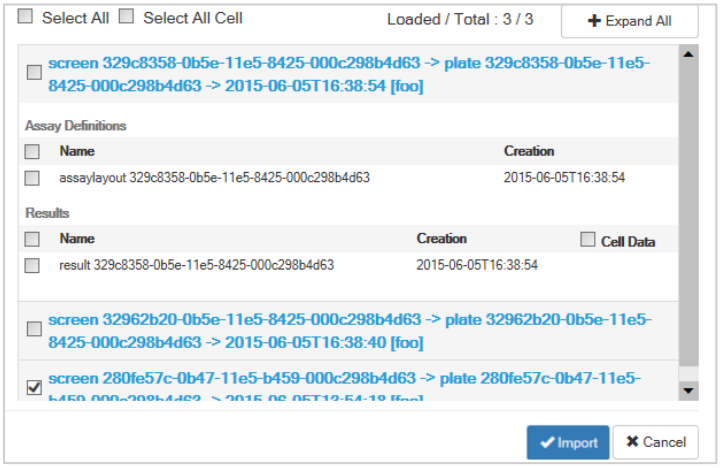


Figure 9-20: Enabled Import Button After Selection

The data is merged based on the definition/result/cell data/PopulationName selected.

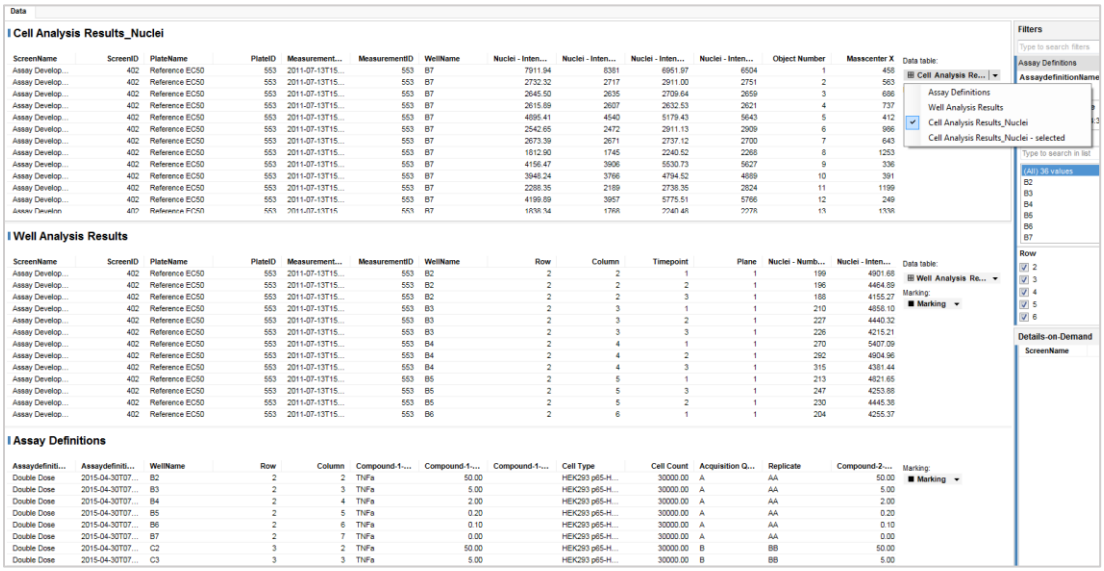


Figure 9-21: Image Import Main Visualization

9.7 Grid Plate Editor App

The **Grid Plate Editor** App provides the necessary controls to create, store, edit, and apply plate designs for single or multiple plates in a user friendly and flexible manner. Plates designs can be stored in the Signals VitroVivo Metastore or a file for future use.

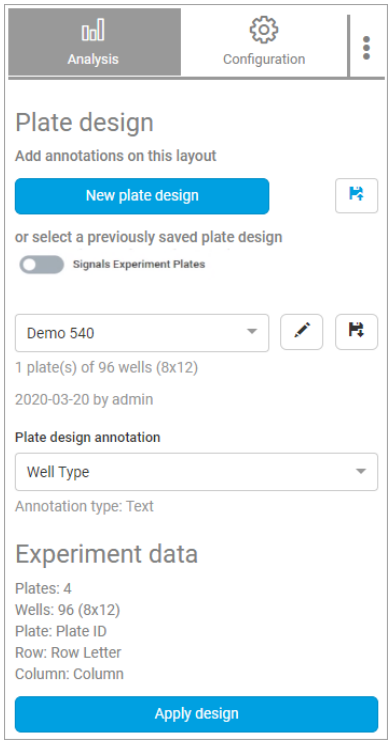
From the Signals Apps page, select the **Grid Plate Editor** App card. A new tab containing the App will be added to the document and the App will be launched.

Figure 9-22: Grid Plate Editor App Card

9.7.1 Configuring the Grid Plate Editor App

9.7.1.1 Analysis Tab

The **Analysis** tab contains a *Plate design* and an *Experiment data* section.



The *Plate design* section contains the necessary elements to create and manipulate a plate design.

- **New plate design:** This button will open the '**Create Plate Layout**' dialog which is used to create a new plate layout as described below.
- **Load plate design:** This icon will open a file browser allowing the upload of a plate design in .json format from the local file system.
- **Signals Experiment Plates**: This toggle allows the use of a plate layout stored in Signals Notebook. When toggled on, a login window will be displayed in the browser if the user is not currently connected to the Notebook.

Note: When connected to Signals Notebook it is possible to load the plate layouts from the Notebook, but it is not possible to modify them or upload a plate layout to the Notebook.

Note: When loading a plate layout from the Notebook the values imported are normalized to mL for volume and g for weight.

- **Previously saved plate design:** This dropdown will display those plate designs available in the Signals VitroVivo Metastore. The selected design determines the dimensions and number of plates in the grid area of the App.
- **Edit plate design:** This icon will load the selected plate design for edition in the App.
- **Save plate design:** This icon will allow the user to save the selected plate design to a .json file in the local file system.
- **Plate design annotation:** This dropdown lists the annotations available in the selected plate design. The one that is selected is the one used to color the grid displayed in the main visualization area of the App.

The *Experiment data* section contains the necessary controls to apply the selected plate design to a table present in the document that has been loaded using either the **Data Import** App or the **Image Import** App. It also provides information on the dimensions and number of plates of the experiment data. This may be useful to ensure the dimensions match those from the annotation to be applied.

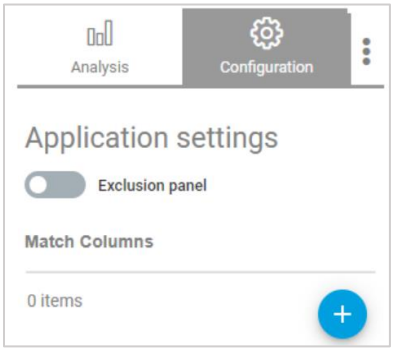
- **Apply design:** This button will apply the plate design to the loaded data table if the dimensions of the plate layout and the data match. When the plate design is applied to the experiment data a notification is displayed. Additionally, if the dimensions are different a warning is raised indicating the dimensions of the data and the layout are incompatible.

After applying the design, the table view for the experiment table with annotations will be updated including the annotations added to the table which will be colored in the same way as in the Plate Layout view.

Note: If there is no screening data in the document, this section will contain a warning indicating the user should add a screening table to apply the design.

9.7.1.2 Configuration Tab

The **Configuration** tab provides a toggle to display the **Exclusion panel** and a **Match Columns** button.



Selecting the toggle will show/hide the **Exclusion panel**. This panel is also accessible from **Tools > Signals Groups.**

The matching component allows the user to match the plate information to the data. If no matching is provided by the user a default matching performed using the ‘row’ and ‘column’ columns from the plate design and the ‘row’ and ‘column’ columns from the loaded dataset.

Within a Workflow, if the user configured and saved customized matching it will be applied, otherwise, the default will be used. If the matching stored in the Workflow is not valid for the new data to which it is being applied an error will be raised and the user will be required to edit the matching to proceed.

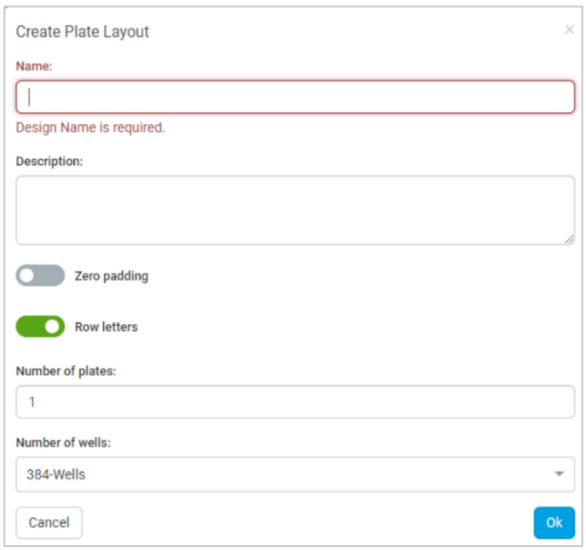
9.7.2 Grid Plate Editor App Results

Once the plate design has been selected or created it can be applied to a specific dataset. When the design is applied, the original table with the additional columns corresponding to the added annotations will be displayed and colored with the corresponding colors as shown in the plate layout representation of the selected annotation.

Figure 9-23: Grid Plate Editor Visualization

9.7.3 Creating a Plate Design

To create a new plate design, select ‘**New plate design’** from the **Analysis** tab. This will open the **Create Plate Layout** dialog menu.

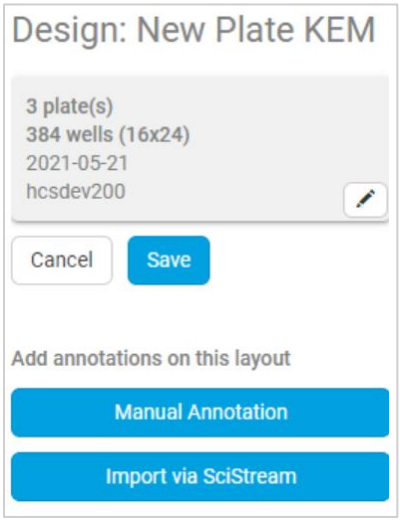


The **Create Plate Layout** dialog contains the following elements:

- **Name:** This textbox sets the plate design name. If the name is already present in Signals VitroVivo Metastore a warning will be displayed indicating the name is already in use.
- **Description:** This textbox allows the user to add an optional description to the plate design.
- **Zero padding:** If this toggle is on, numeric values will be zero padded. Default off.
- **Row letters:** If this toggle is on the rows will be named using letters instead of numbers. Default on.
- **Number of plates:** This textbox allows the user to specify the number of plates to be used in the layout. The default is 1 (only one type of plate). When multiple plates are created for the layout, they will all have the same dimensions and naming constraints (letter or number naming and zero padding).
- **Number of wells:** This dropdown allows the selection of the plate dimensions. These can be a standard plate or a custom plate.
    - Standard dimensions available are 96, 384 and 1536 well plates.
    - Selecting "**Custom**" will allow the user to specify the number of rows and columns.
- **Cancel:** Selecting this button will return the user to the initial screen without changing the document**.**
- **OK:** Selecting this button will take the user to the **Plate design** edition page.

9.7.3.1 Adding Annotations to the Plate Design

Once the user has created a new plate layout the App will present the **Design** menu.

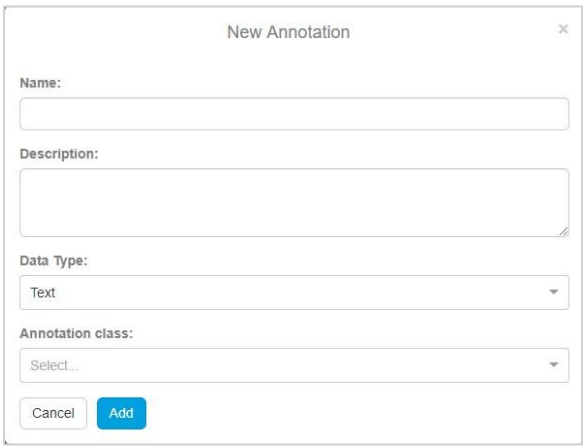


On this menu the user is shown the name and details of the design they are editing with the following controls:

- **Edit:** This icon will open the **Edit Plate Design Metadata** window. Here the user can change the plate design name and description.
- **Cancel:** This button will return to the initial page without performing any changes to the document.
- **Save:** This button will save the plate design to the Metastore.
- **Delete:** This button will only be present after the annotation has been saved to the Metastore. Clicking it will delete the plate design and return to the initial Grid Plate Editor App menu.
- **Add annotations on this layout:**
    - **Manual Annotation:** This button will open the **New Annotation** window that allows the user to add a new annotation to the design.
    - **Import via SciStream:** This button will open the **New Annotations from template** window which allows the user to add a new annotation using an import template stored in the Spotfire® library as described below.
- **Annotations:** This dropdown is only present if an annotation is available in the plate design that is being edited. It allows the user to select any of the existing annotations in the design. The selected annotation will be displayed in the main visualization of the App.
- **Edit:** This button opens the **Update** window for the selected annotation. This window allows the user to change the name and/or description of the annotation as well as providing an option to delete the selected annotation. This button is only present if an annotation is available in the plate design that is being edited.

9.7.3.2 Creating and Editing Annotations

To create a new annotation, the user should select the **Manual Annotation** button. This will open the **New Annotation** UI.

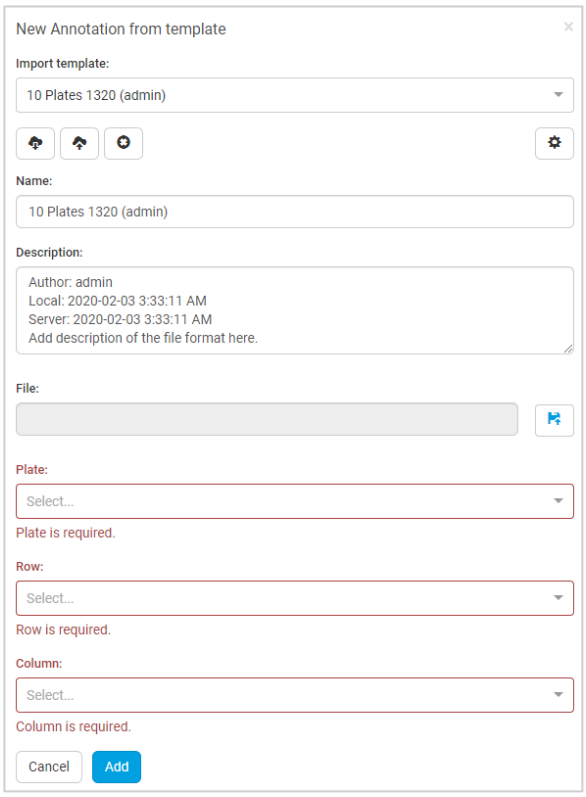


The available options within the **New Annotation** window are:

- **Name:** This will contain the name of the annotation.
- **Description:** An optional description of the annotation.
- **Data Type:** The data type of the annotation. This can be Text, Integer or Real.
- **Annotation class:** This will contain information on the type of annotation in the Screening data model that this annotation belongs to. The currently available annotation classes are:
    - **Replicate:** The annotation indicates a replicate.
    - **Concentration:** The annotation provides concentration information.
    - **Compound Id:** The annotation provides Compound Id information.
    - **Well type:** The annotation indicates if a well corresponds to a control or not, and what type of control (positive or negative).

  The annotation class allows to automatically match some types of information to templates and Apps downstream even if the columns are not named in the same way, thus simplifying the analysis.

To load an annotation using an import template from SciStream the user should select the **Import via SciStream** button. This will open the **New Annotations from template** UI.

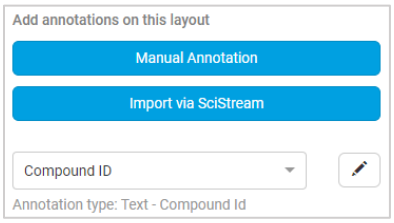


The available options within the **New Annotations from template** window are:

- **Import template:** This dropdown provides a list of the import templates available to the user.
    - **Download** icon**:** Allows the user to download the template if it is not already available locally to use for importing the annotation.
    - **Upload** icon: Allows the user to upload a local template overwriting the existing one in the library. It is only available if the user is the owner of the template selected in the dropdown.
    - **Delete** icon: Allows the user to delete a template in the library. It is only available if the user is the owner of the template selected in the dropdown
    - **Manage Formats** icon**:** Allows the creation of a SciStream import format locally to use for importing the annotation.
- **Name:** Allows the user to input a name for the annotation. It defaults to the SciStream input format name.
- **Description:** An optional description of the annotation. The default is the SciStream input format description.
- **File:** Allows the user to select the file from which to import the annotation using the **'Load plate design'** icon.
- **Plate:** This dropdown contains a list of the annotation file columns and will be used to map the annotation information to the document data. The user should select the column that corresponds to the Plate information.

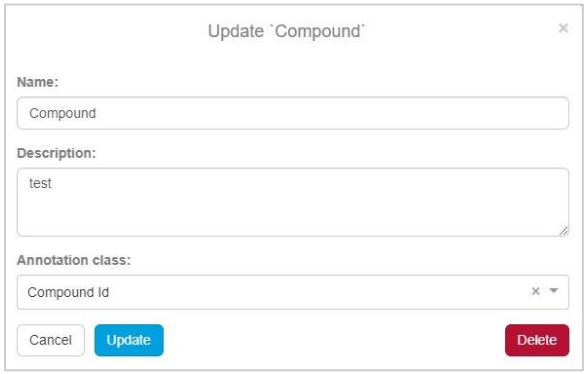
- **Row:** This dropdown contains a list of the annotation file columns and will be used to map the annotation information to the document data. The user should select the column that corresponds to the Row information.
- **Column:** This dropdown contains a list of the annotation file columns and will be used to map the annotation information to the document data. The user should select the column that corresponds to the Column information.
- **Cancel:** Closes the **New Annotations from template** window with no changes.
- **Add:** Adds the annotation information as a new annotation.

When at least one annotation is present in the plate design a dropdown is shown that allows the user to select any of the available annotations and edit them if needed.



The additional controls available when an annotation is present are:

- **Annotation dropdown:** Allows the user to select an annotation from those available in the plate design.
- **Edit or delete annotation:** Allows the user to edit or delete the selected annotation. When clicking on this icon the Update “Annotation” window is displayed:



- The user can edit the **Name**, **Description**, and **Annotation class** of the annotation. Once this is done, selecting **Cancel** will discard any changes, and selecting **Update** will save the changes made. In addition, the **Delete** button will allow the user to delete this annotation from the plate design.

When an annotation is present, the right-hand panel will display information on this annotation. If more than one annotation is present the screen displays the information for the selected annotation.

The upper right-hand panel in the App will display a grid containing an editable representation of one of the plates with the values of the selected annotation present in each cell and the lower right-hand panel will contain a visual representation of all the plates colored according to the values of the selected annotation.

The following controls are available in the right-hand section of the window:

- **Preview:** When selected this will update the lower visualization with any changes that have been performed in the grid available in the upper section.
- **Plate tab:** These tabs allow the user to switch between the different plate types that are present in a single plate design.
- **Plate grid:** The plate grid allows the user to add the annotation values to the cells. It is possible to copy the values of a cell or group of cells by copying and pasting the contents or by selecting the continuous region of cells to be copied and dragging the lower corner of the selection.
- **Plate design overview:** This lower section of the screen provides a visual representation of the annotations over all the plates in the design allowing to easily identify parts that may need to be corrected/updated.

9.8 Editable Data Grid App

The **Editable Data Grid** App provides several important functionalities:

- Ability to insert data from the clipboard that is present in a grid-like structure, such as data copied from Excel.
- Ability to insert data directly from a text file and edit it before adding it to the Spotfire® document.
- Ability to load data from the Spotfire® library when the user is online, and the library preference has been configured.
- Ability to load data from a Signals Experiment.
- Ability to edit the contents of the data loaded from an existing table in the document.

- Data may be added as a **New table**, overwrite an existing table, added as **New columns**, **Replace columns** in the table, or **Replace selected table contents**.

From the Signals Apps page, select the **Editable Data Grid** App card. A new tab containing the App will be added to the document and the App will be launched.

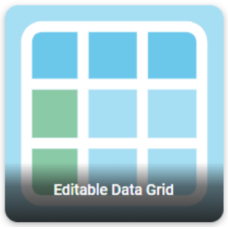
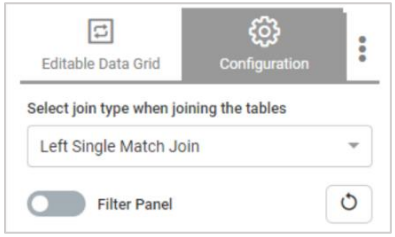


Figure 9-24: Editable Data Grid App Card

9.8.1 Configuring the Editable Data Grid App

The user is presented with an empty data grid and a configuration UI containing two tabs, “**Editable Data Grid**”, "**Configuration**” and an additional panel “…”, is displayed in the top-left hand corner.

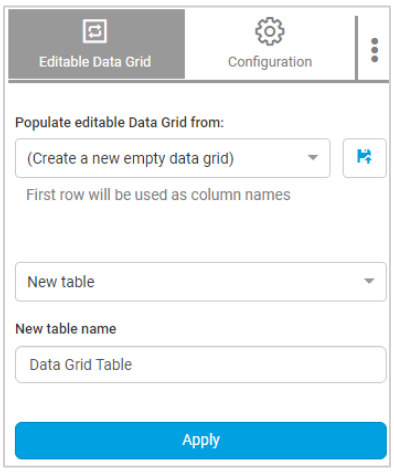
- **Editable Data Grid:** This section contains the main controls that allow the user to load and save the table(s).
- **Configuration**: This section contains a dropdown menu to select a join type between tables (i.e. ‘Left Single Match Join’ and ‘Left Outer Join’). It also contains a toggle to display the **Filters Panel**.



The empty data grid can be used to paste data from the clipboard and will automatically resize to adapt to the size of the pasted data. The data grid always displays an empty row at the end to facilitate the addition of data.

9.8.1.1 Editable Data Grid App Controls

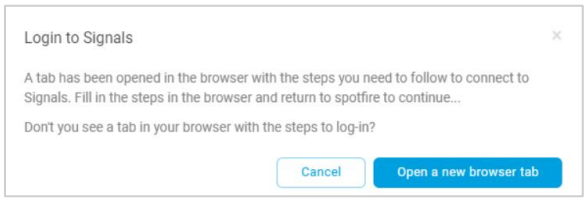
This tab contains the main controls for data grid manipulation.

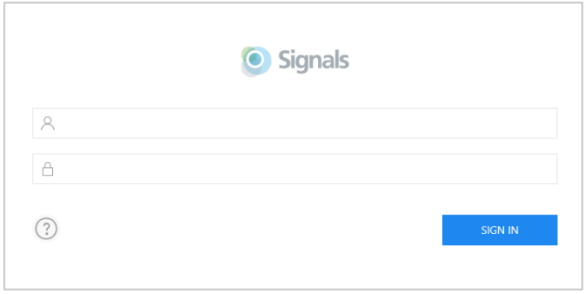


The following controls are available:

- **Populate editable Data Grid from:** This dropdown allows the user to select an existing table in the document, a table saved in the Spotfire® library or an ADT (Administrator Defined Table) saved in a Signals Notebook experiment if opened from within Signals Notebook.
    - **(Create a new empty data grid):** When selected this option will display an empty Data Grid where the user can paste data from the clipboard.
    - **(Spotfire Library):** When selected a new dropdown is shown that will display the existing tables in the configured Spotfire® library folder.
    - **(Signals Experiment):** When selected a new dropdown will be displayed showing the existing ADTs and Signals Notebook materials tables in the experiment from which the document was opened under the **Signals table** dropdown.

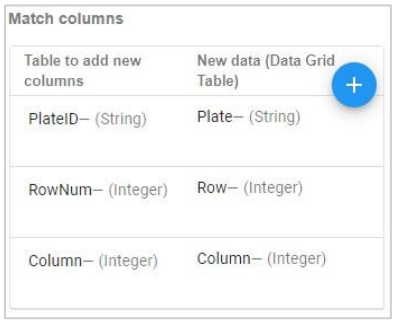
      Note: If the user is not logged in a message will be displayed reminding the user to log in via the web browser.



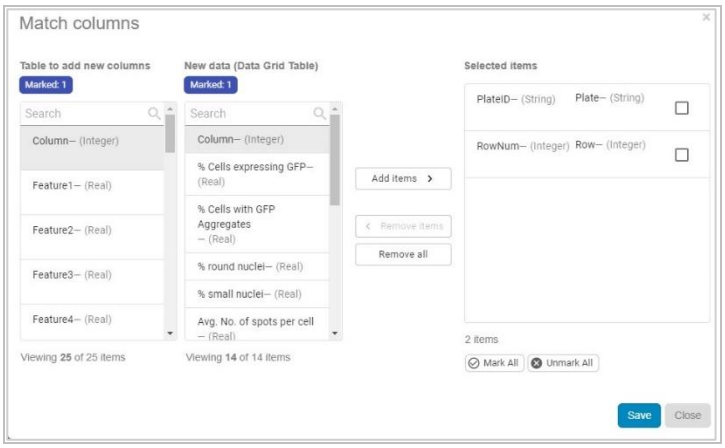


    - **<Table Name>:** In addition to the above options a list of the existing tables in the document are displayed. When one is selected it will be loaded into the data grid

- **Load file…**: This icon allows the population of the grid directly from a file present in the local file system.
- **Save data as**: This dropdown allows the user to select the way in which the grid data will be saved into the current analysis:
    - **New table**: This option will allow the user to add the data as a new table.
        - **New table name:** The name of the new table should be written in this textbox.
        - **Apply:** When the Apply button is selected the new table is added to the document.
    - **New columns:** This option allows the user to add the data from the data grid to an existing table. To do this the user must select the table where the data should be added and configure the matching columns between the data in the grid and the existing table.
        - **Table to add columns:** The user selects from the dropdown the desired table to which the columns will be added to.
        - **Match Columns:** The user configures the matching columns between the columns in the selected table and the columns in the data grid



        - The existing matches, if any, are displayed below the **Match columns** controls. To edit, add or delete matches, click on the *Plus Sign* icon to open the **Match columns** UI where the App matches can be configured.



        The **Match columns** interface has the following controls:

        - **Search boxes:** These will be available for any box containing more than five items. These can be used to easily filter the contents when looking for specific entries.

- **Add items:** Adds the selected pair of elements marked in the **'Table to add new columns**' and the '**New data (Data Grid Table)**' boxes to the '**Selected items'** box.
- **Remove items:** Removes the marked pair from the '**Selected items'** box.
- **Remove all:** Removes all elements from the '**Selected items'** box.
- **Selected items checkboxes:** Used to mark elements in the '**Selected items**' box.
- **Mark All:** Marks all elements in the '**Selected items'** box.
- **Unmark All:** Unmarks all elements in the **'Selected items'** box.
- **Save:** Adds the pairs to the '**Match columns'** list.
- **Close:** Closes the '**Match columns'** UI without changing the matches.
- **Apply:** When the **Apply** button is selected the columns in the data grid that have not been used in the matching will be added to the selected table. The resulting table can be visualized below the data grid with the added columns highlighted.

Note: To add the new columns to the selected table in the document, the combination of matching columns selected must be unique for each row.

- **Replace columns:** When this option is selected the user can add the data from the data grid to an existing table as new columns and remove existing columns from the target table. To do this the user must select the table where the data should be added and configure the matching between the data in the grid and the existing table as well as the columns to replace.
  - **Table to replace columns:** This is the target table where the columns will be added, replacing existing ones.
  - **Match columns:** Here the user can configure the matching between the columns in the selected table and the columns in the data grid in the same manner as described above.
  - **Select columns to remove:** Here the user can select which columns from the target table should be replaced.
- **Replace column values:** When this option is selected the user can add data from the data grid to an existing table replacing existing column values in the target table. To do this the user must select the table where the data should be added and configure the matching between the data in the grid and the existing table, as well as the column values to replace.
  - **Table to replace column values:** This is the target table where the column values will be added, replacing the existing values.
  - **Match columns:** Here the user can configure the matching between the columns in the selected table and the columns in the data grid in the same manner as described above.

Note: The main difference with the previous option (replace columns) is that in the first case the columns are added to the table as new and the columns selected for removal are deleted from the table, whilst in this case the columns are kept in the table and the contents are replaced. This means any existing operations on the columns where the values are replaced are preserved.

- **Replace selected table contents:** When this option is selected, the data in the data grid will replace the contents of the table selected to populate the data grid. This provides an intuitive method to edit the tables present in the document.
  - **Apply:** The selected table is replaced with the content of the grid.

9.8.2 Editable Data Grid App Results

Once the data has been added to the grid, some additional controls may be accessed directly from the grid:

- **Reset grid:** This button will clear the data grid.
- **Column header controls:** The icon provides a single option to clear the contents of a column.
- **Right-click controls:** Currently right-clicking on the data contains the copy option that allows the user to copy the values of the selected cells to the clipboard.

Note: Spotfire® filtering is not integrated into the data grid. However, once the table is saved, if it is loaded in the Editable Data Grid the standard Spotfire® filtering options are available for use and the grid contents will be filtered accordingly.

9.9 Extensible Normalization App

The **Extensible Normalization** App provides functionality to perform inter-plate and intra-plate normalization operations by using a set of predefined normalization algorithms or alternatively using custom normalization templates defined by an author using the **Calculations Explorer**. Support for the custom normalization templates allows authors to define normalization methods different from those provided “out of the box”.

From the Signals Apps page, select the **Extensible Normalization** App card. A new tab containing the App will be added to the document and the App will be launched.

Figure 9-25: Extensible Normalization App Card

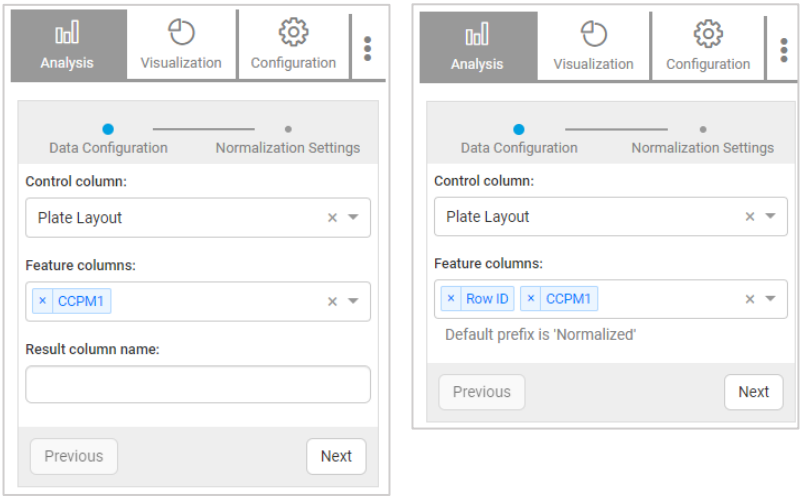
9.9.1 Configuring the Extensible Normalization App

9.9.1.1 Analysis Tab

The **Analysis** tab contains a wizard-like configuration panel where the user can configure the settings in two consecutive steps:

1. **Data Configuration**: This step allows the user to define the:

- **Control column**: the column that contains the information on the control wells.
- **Feature columns**: the Feature(s) to be normalized.
- **Result column name:** an option to provide an alternative name to the results column. If left blank, the normalized data will be added as column(s) in the data table and by default are named with the prefix ‘Normalized’. Note this option is only available if a single feature column has been selected.

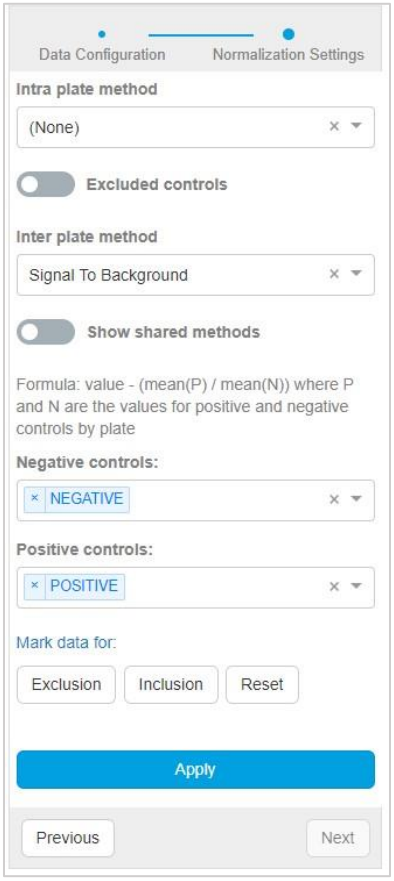


When all the settings are configured select **Next** to proceed to the next step. In case any of the required columns are not selected, a message under the selector informs the user that it requires a valid option before continuing.

2. **Normalization Settings:** This step allows the user to define:
   - **Intra plate method**: A dropdown provides the selection of the method to be used for intra-plate normalization. This normalization is used to correct for spatial effects in the plate such as edge-effects that are sometimes observed in screening experiments. Intra-plate normalization can be performed without inter-plate normalization. The available options are:
     - **(None) -** No intra-plate correction will be performed.
     - **Median polish** - The median polish method will be used to correct for spatial effects in the plate.
   - **Exclude controls:** Toggling this option will exclude the control wells from the calculations when performing the intra-plate correction.
   - **Inter plate method:** This dropdown provides a selection of methods to be used for the normalization between plates. The available “out of the box” options are:
     - (None)
     - Median
     - Normalized percent activation (NPA)
     - Normalized percent inhibition (NPI)
     - Reversed NPI
     - % of negative control
     - Z-score
     - Signal to Background

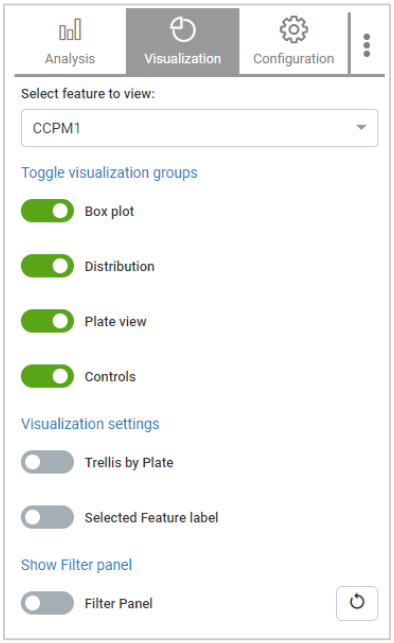
     When the method is selected the formula used in the method is displayed as well as the required Positive and Negative control selection boxes, if any.

- **Show shared methods:** This toggle will retrieve and display in the dropdown menu the Normalization templates in the Signals VitroVivo Metastore that can be used by the App. This option allows users to load custom defined templates that were generated using the **Calculations Explorer** and shared in the Signals VitroVivo Metastore.
- **Negative controls:** This dropdown is available if the selected normalization method requires a negative control. It allows the selection of values from the selected 'Control column' as Negative controls.
- **Positive controls:** This dropdown is available if the selected normalization method requires a positive control. It allows the selection of values from the selected 'Control column' as Positive controls.
- **Mark data for:** Here the user can exclude or include datapoints from the analysis. The options available are:
    - **Exclusion:** All marked data points in any of the visualizations will be excluded from the analysis. This means these excluded points will not be used for the normalization calculations. These can be seen in some visualizations with different color and/or shape.
    - **Inclusion:** Marked points that are set for inclusion will be considered again in the analysis and used in the normalization calculations.
    - **Reset:** Selecting this button will reset any exclusion done and restore the analysis to the original unchanged state.
- **Apply:** This button will apply the normalization to the data.



9.9.1.2 Visualization Tab

The **Visualization** tab contains settings that affect what is displayed in the visualization area.



The available controls are:

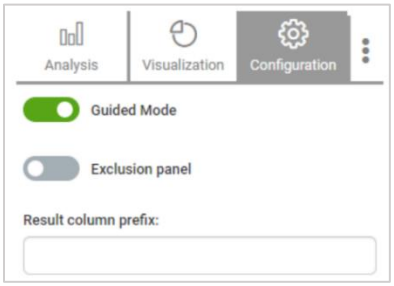
- **Select feature to view:** This dropdown menu will contain a list of features that were included in the normalization. It allows the selection of one to be displayed in the available visualizations.

- **Toggle visualization groups:** Lists the groups of visualizations available in the visualization area with an on/off display toggle. Available options are:
    - **Box plot:** Shows or hides the Box plots.
    - **Distribution:** Shows or hides the Distribution plots.
    - **Plate view:** Shows or hides the plate view heatmap plot.
    - **Controls:** Shows or hides the Control categories plot.

- **Visualization settings**
    - **Trellis by plate:** This toggles on the visualization by plate of the plate view heatmap plot.
    - **Selected Feature label:** This toggle displays the selected feature label, when applicable.

- **Show Filter panel**
    - **Filter Panel:** When activated, this toggle displays the Filter Panel on the right-hand side of the screen. When selecting the reset button next to the toggle, all active filtering will be reset.

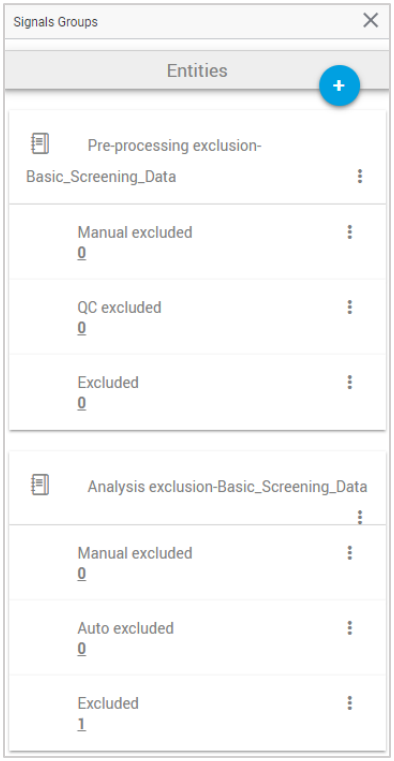
9.9.1.3 Configuration Tab



This section contains the **Guided Mode** toggle to control if the App should follow a given data map or not.

The '**Result column prefix**:' textbox allows the user to specify a prefix to add to the normalized column(s) names with the resulting column being named: "prefix_Feature Name". The default prefix is "Normalized" if left blank.

Additionally, in this section the user can toggle on or off the **Exclusion Panel**, where all the defined exclusion groups are present and will be displayed on the right-hand side of the document. This feature is complementary to the Exclusion buttons present in the **Analysis** tab of the UI. The actions present in the panel allow the user to manage the exclusion from this component, additionally it may be accessed anywhere in the document from **Tools > Signals Groups**.



Currently in Signals VitroVivo by default two groups of Exclusion are available:

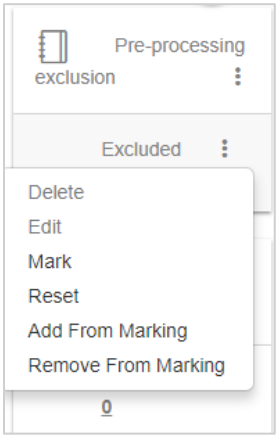
- **Pre-processing Exclusion:** this is used in the Extensible Normalization App and will prevent the Excluded data from being used when Normalizing data and in downstream analysis performed with the Calculations Explorer. It will only be displayed when looking at the raw data and using a different color and shape whenever possible.

- **Analysis Exclusion:** this group is used for exclusion of Normalized data, or any other data in downstream analysis, such as in curve fitting. This data will always be displayed, with different shape and color, however it will not be used in downstream calculation, such as curve fitting.

By clicking the "…" on the right-hand side of the Excluded group name the user accesses the controls to manage this group which are:

- **Mark:** this will mark the excluded data, if any. Additionally, this can be achieved by clicking the count number below the group name;
- **Reset:** this action is the same as the Reset button in the UI, and will bring all data to its original state, meaning there will be no excluded data;
- **Add from Marking:** this action is the same as the Exclude button in the UI. By marking data and clicking this action the user is Excluding data from the analysis;
- **Remove from Marking:** this action is the same as the Include button in the UI. By marking data and clicking this action, if it is excluded data, it will be removed from the Excluded group, thus be considered Included in the analysis.

The **Edit** and **Delete** options are disabled for these default Exclusion groups as they are shared throughout the document and to avoid the QA Exclusion process to malfunction.



9.9.2 Extensible Normalization App Results

When normalization is performed the right-hand side of the screen will be populated with the "after normalization" visualizations, and the visualization settings controls can be used to configure the visualization with the exact information required.

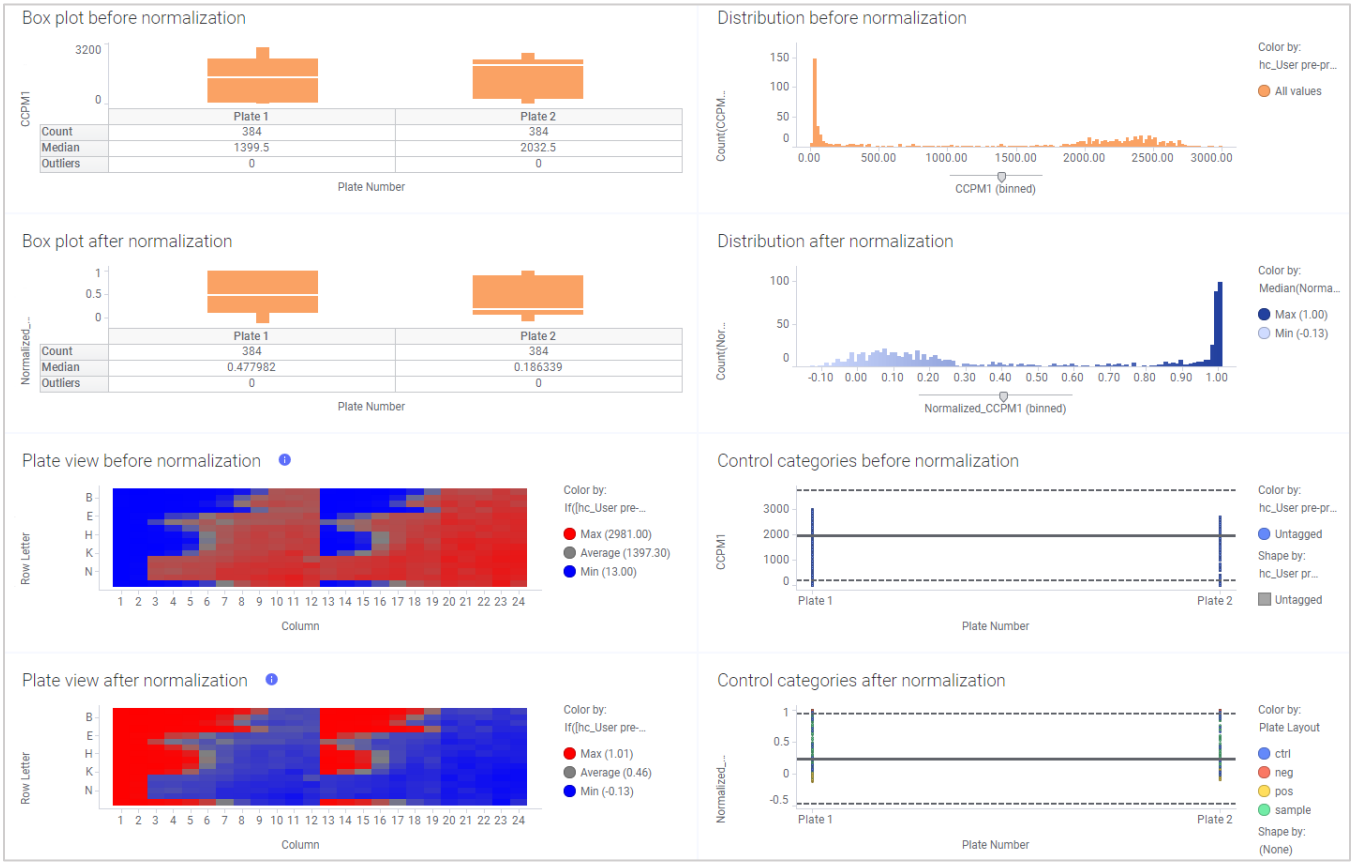


Figure 9-26: Normalization App Visualizations

The visualizations generated by the Normalization App are the following:

- **Box plot before and after normalization:** This boxplot shows for each of the plates the distribution of the raw or normalized values respectively in each of the plates.

- **Distribution before and after normalization:** This visualization displays a histogram with the distribution of the raw and normalized values respectively for the selected feature over all the wells. In the case of the raw values, if any exclusion was applied, those points will appear with a different color than the ones used in the analysis.

- **Plate view before and after normalization:** This visualization displays the plate view before and after the normalization. It shows each well colored using the median of the feature value in that cell. The aim of this visualization is to show edge effects in experiments where the plate layout is the same for all plates. In the case of the raw values, if any exclusion was applied, those points will appear with a different color than the ones used in the analysis.

- **Control categories before and after normalization:** This visualization shows a representation of the distribution of each of the wells before and after normalization colored by the category to which they were assigned: controls or samples. The median and dispersion for each of these categories are also plotted as lines with the same coloring. In the case of the raw values, if any exclusion was applied, those points will appear with a different shape and color than the ones used in the analysis.

9.10 New Curve Fitting App

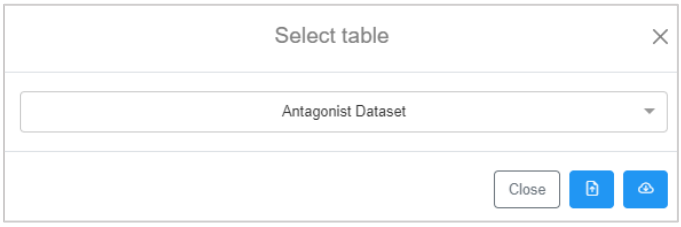
The **New Curve Fitting** App provides a flexible and powerful tool to calculate, review, and report assay curve fitting results using out-of-the-box or custom curve fit equations.

From the Signals Apps page, select the **New Curve Fitting** App card. A new tab containing the App will be added to the document and the App will be launched.

Figure 9-27: New Curve Fitting App Card

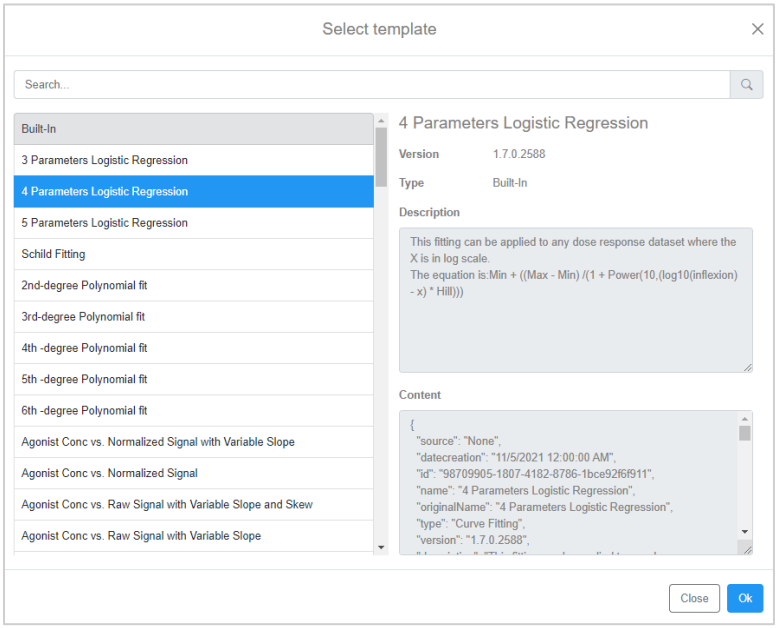
9.10.1 Configuring the New Curve Fitting App

From the left-hand side panel, click the **select** link in the **Apply Template** section, then select the desired data table from the dropdown. The user can then either select the **Load local template** icon or the **Load shared template** icon**.**



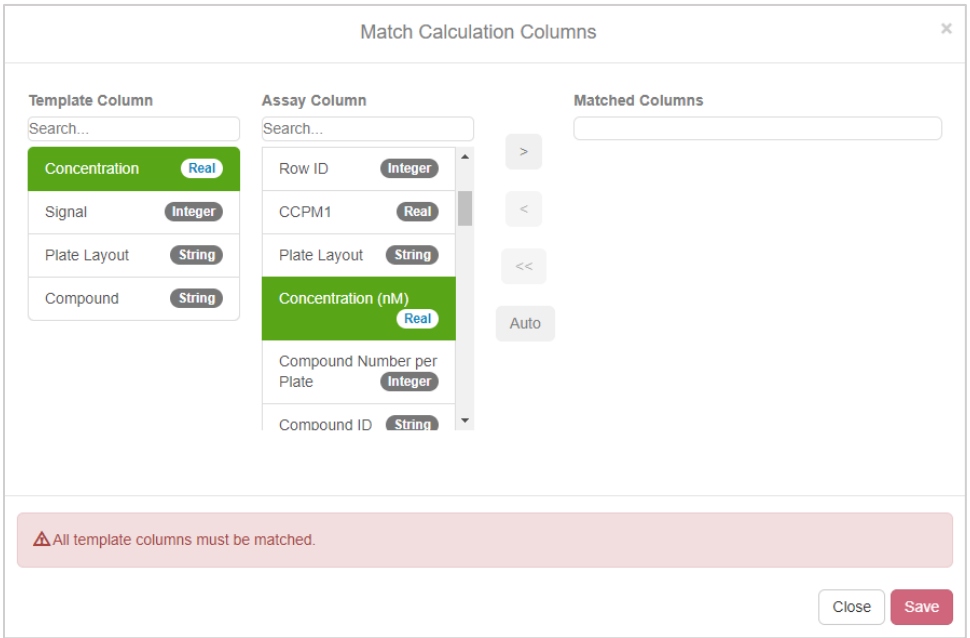
The user is then prompted to select a model from the **Select template** UI, where options are based on if the local template or shared template option was chosen in the previous menu. Once a model is selected the description and contents of the model is displayed.

- **Close:** Selecting this option will close the **Select template** window and allows the user to change the table selection if more than one table is available.
- **OK:** Selecting this option will proceed with the selected curve fitting analysis.



If the user has loaded the data using the **Screening data model** (see 13.2 Appendix Screening Data Model Information), the analysis will be performed, if not, the user is presented with the **Match Calculation Columns** menu, where the user can match the columns required by the fitting with those from the data in the same way as a match would be performed when using the **Calculations Explorer**.

Note: The **New Curve Fitting** App now allows the user to visualize and review curve fit results from pipelining execution, see section Exporting Workflows as Pipelines (Beta Feature). The App automatically detects the presence of tables with pipelining results, loads the CE Template associated with the parent Signals-Pipeline Workflow, and displays/restores curve visualizations for pipelining results as they had been executed locally.

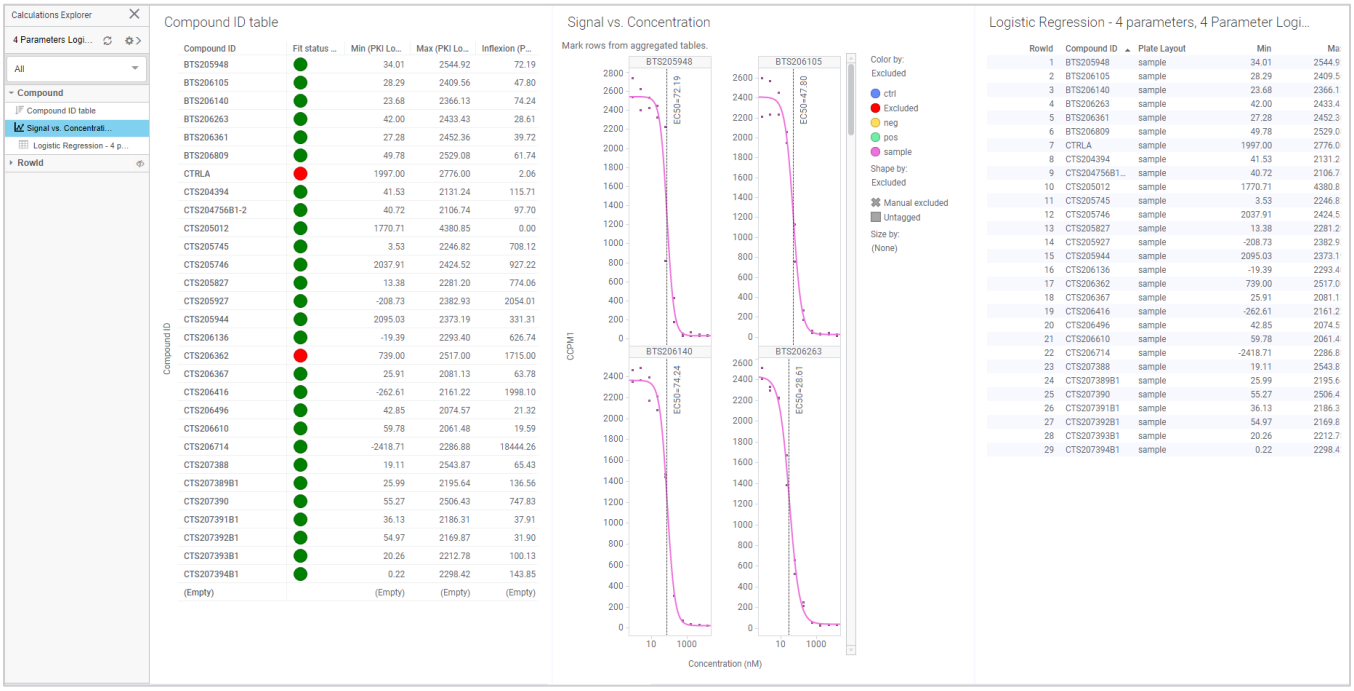


9.10.2 New Curve Fitting App Results

After performing the fitting analysis, the user is presented with the results, these are divided in three sections:

- **Compound table:** This table facilitates the drilldown and navigation through the curve information present in the other section(s). It contains the parameters estimated for each fit. This simplifies the sorting and selection of curves based on any of these parameters.
- **Scatterplot containing the curve fits:** Here the user can visualize the fittings according to the settings used in the selected fitting model. The visualization will be filtered by the selection in the aggregated table and trellised by Compound.
- **Fitting parameters table:** The fitting parameters table is not displayed by default but can be viewed by clicking on the *Crossed Eye* icon by the side of its entry in the Calculations Explorer panel. This table contains the parameters for each of the fittings, including the initial fit settings. The table will be filtered by the selection in the aggregated table.

Once the analysis is performed, the user may navigate through it in the same manner as they would when using any **Calculations Explorer** Template, as described in Calculations Explorer section.



9.11 t-SNE App

The **t-SNE** App allows the analysis of multidimensional data through one of the most recent and popular dimensionality reduction methods. Developed by Laurens van der Maaten and Geoffrey Hinton, t-Distributed Stochastic Neighbor Embedding (t-SNE) [1, 2], is a technique which reduces the dimensionality of a dataset to a lower dimension map according to the probability density of each point. This technique allows the visualization of data in human perceptible dimensions while keeping the intrinsic structure of data. Thus, it is a powerful method to identify hidden structures in data.

The basis of this algorithm is that in a high dimension the density distribution of a single point is approximated to a Normal distribution, whereas in the low dimension map this approximation corresponds to a t-Student distribution. The core of this method is setting this heavy tailed distribution to the map, allowing for the low dimension to compensate the great difference in dimensionality with the high dimensional space. This grants more power to the method as it will provide a higher separation between similar points, thus giving a better notion of the structure inside each group of data. The result is a representation of the natural clusters existent in data in a lower dimension allowing further analysis. The current algorithm in use in this App is the Barnes-Hut implementation of t-SNE, which is the optimized implementation to allow t-SNE to be used in larger datasets more quickly by adding an extra parameter defined in the App as the “Speed/Accuracy ratio”.

From the Signals Apps page, select the **t-SNE** App card. A new tab containing the App will be added to the document and the App will be launched.

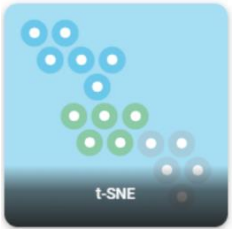


Figure 9-28: t-SNE App Card

9.11.1 Configuring the t-SNE App Data

Once the App has been added to the Workflow but prior to being launched, the input data must be configured. This is done through the **Data configuration** section in the App’s control panel on the left-hand side of the page. The basic version allows only the selection of the features to be used as input.

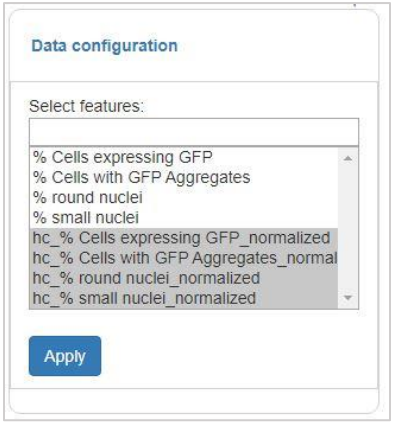


Figure 9-29: Data Configuration Menu for t-SNE

The **t-SNE** App does not require a complex configuration to run. The most straightforward way is to simply select the features of interest and select **Apply**. After the data has been processed, a new visualization will be created.

The **t-SNE** App also allows a more advanced configuration. It is recommended to change this configuration as many times as necessary to explore different results with different values for the parameters. This comes with the nature

of the t-SNE method, which requires this user input, and depending on various properties, such as data size, this parameter exploration is crucial to obtain the best result.

There are three parameters that can be customized in the t-SNE App under **Advanced Settings**:

- **Perplexity (Neighbors)**: Approximated number of neighbors of each point. This will give the App a notion of how data is expected to be distributed and it will try to distribute data accordingly.
- **Speed/Accuracy ratio**: This is a tradeoff between accuracy (low values) and speed (high values). The default value of 0.5 produces robust results and is slightly faster than a traditional t-SNE which could be run by setting this value to 0.
- **PCA as initial approximation**: Checking this option (default) will run a PCA dimensionality reduction before running t-SNE algorithm.

Figure 9-30: Advanced settings menu for t-SNE

After all the settings have been configured proceed to the analysis by selecting **Apply** and the page will be populated with a visualization.

9.11.2 t-SNE App Results

When the **Apply** button is selected, the App is executed and a new visualization is generated. The aim of the t-SNE App is to allow the user to have a notion of the structure of the data and this can be readily visualized through the analysis of the result. This analysis is done by observing different visualizations produced by changing the parameters and verifying how the data tends towards a certain organization. A clear organization in different clusters should be identified with a good choice of the parameters (assuming such structure is present in the data).

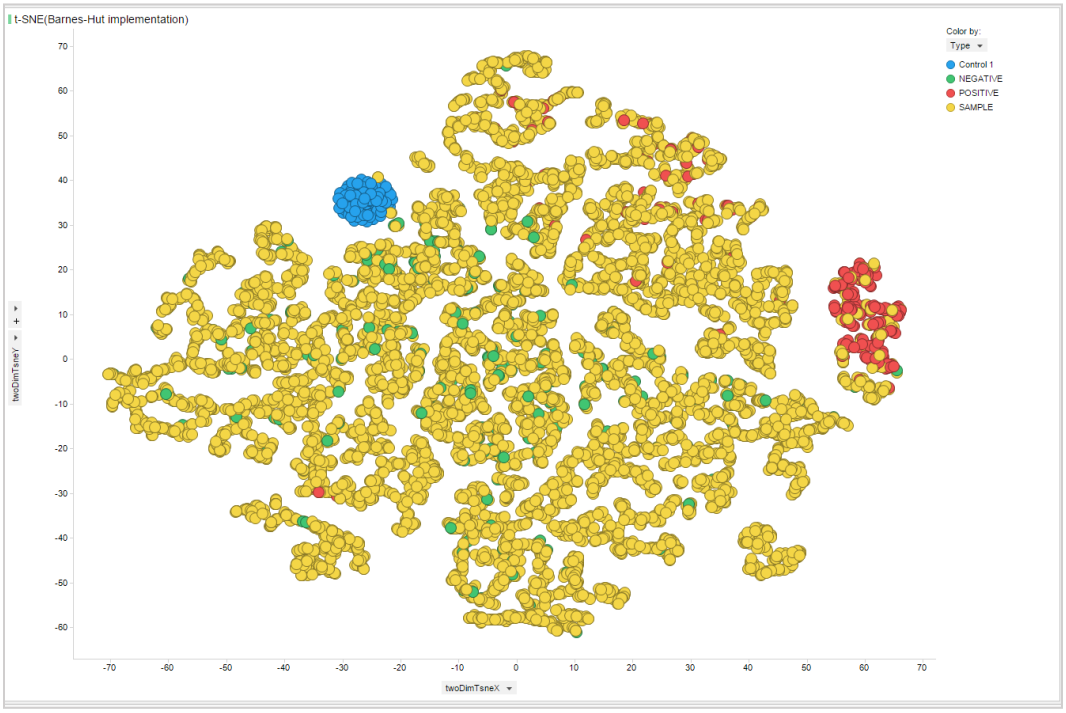


Figure 9-31: 2D Scatterplot of the Data after Running t-SNE

The axes represent the coordinates obtained from running t-SNE. They provide no information on their own, rather the structure of the data is the real goal which is mapped through these coordinates.

9.11.3 t-SNE App Data Exploration

The 2-dimensional representation is not the only possibility to explore data. t-SNE allows the user to observe the distribution in 3-dimensional space. This is easily switched through the **Data exploration** control panel by selecting between 2D and 3D options.

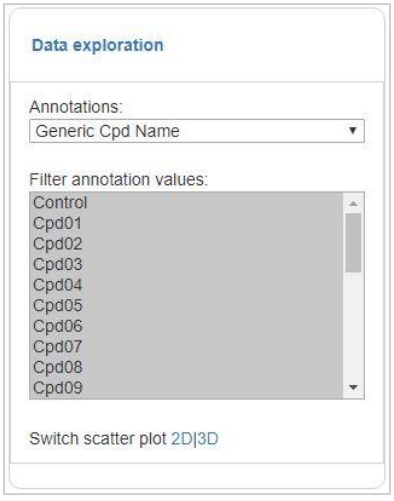


Figure 9-32: The Data Exploration Section Provides Options to Control Coloring, Filtering and the Dimensions of the Result

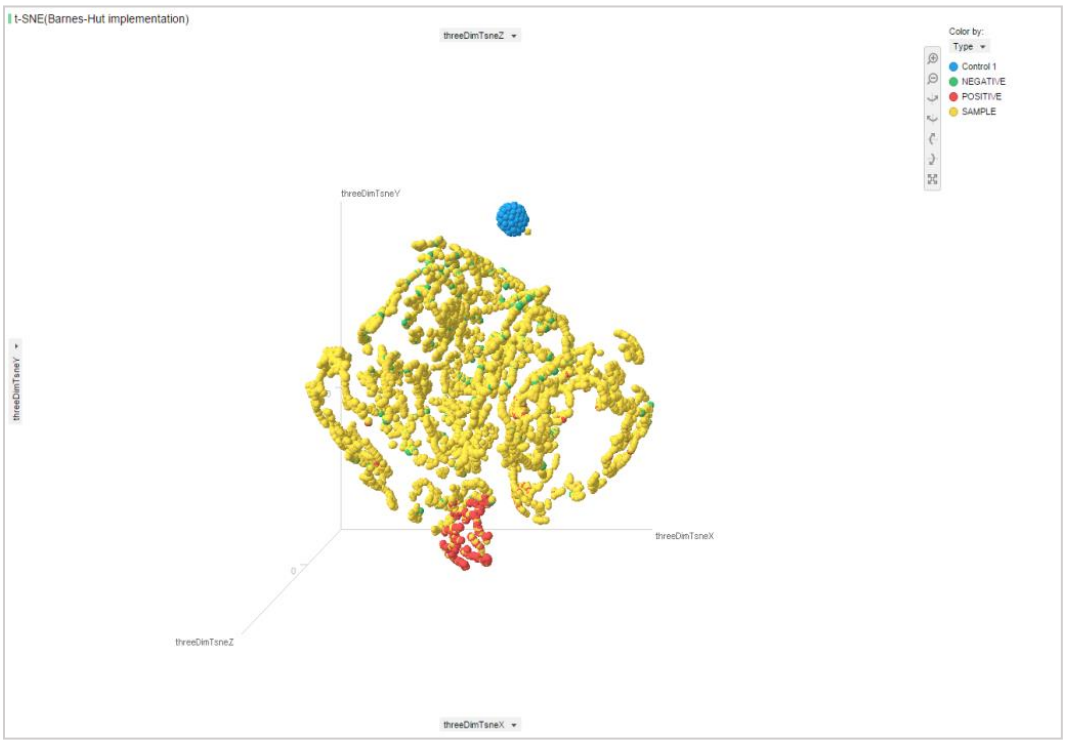


Figure 9-33: 3D Scatterplot Showing the t-SNE Reduction in 3 Dimensions

Additional exploration of the results can be done through filtering and coloring. These controls allow the user to verify how different data classifiers are separated by changing the “Annotations” column. In the example of a simple screening, there are Positive and Negative controls, a neutral Control (“Control 1”) and Samples. In the figures, it is possible to see how well separated the controls are, and how samples are more like either one of the two. This is the actual goal and the power of the t-SNE App.

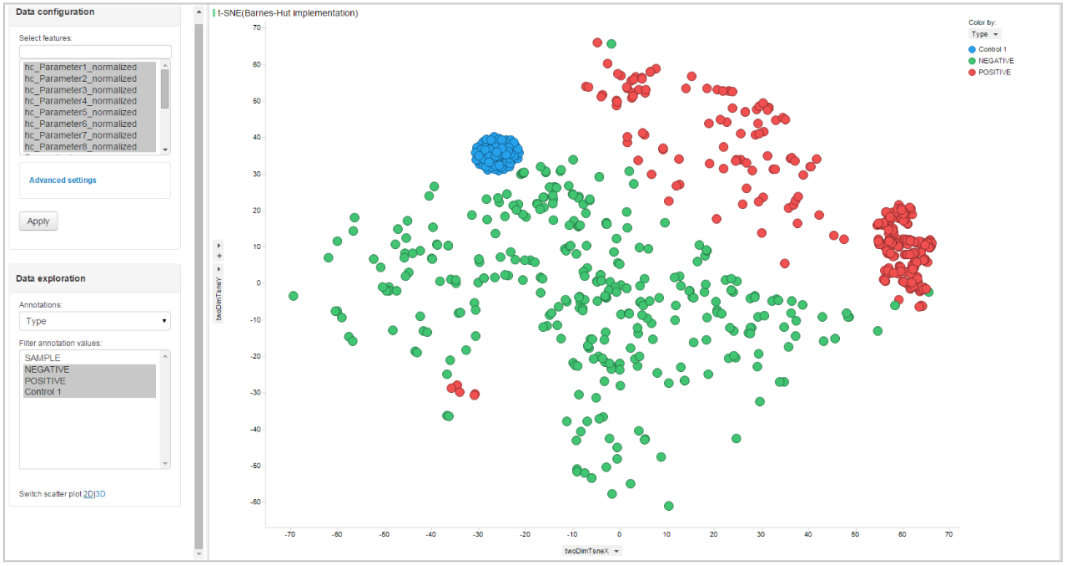


Figure 9-34: t-SNE App Visualizations After Applying where Samples have been Filtered Out to View the Separation of the Controls

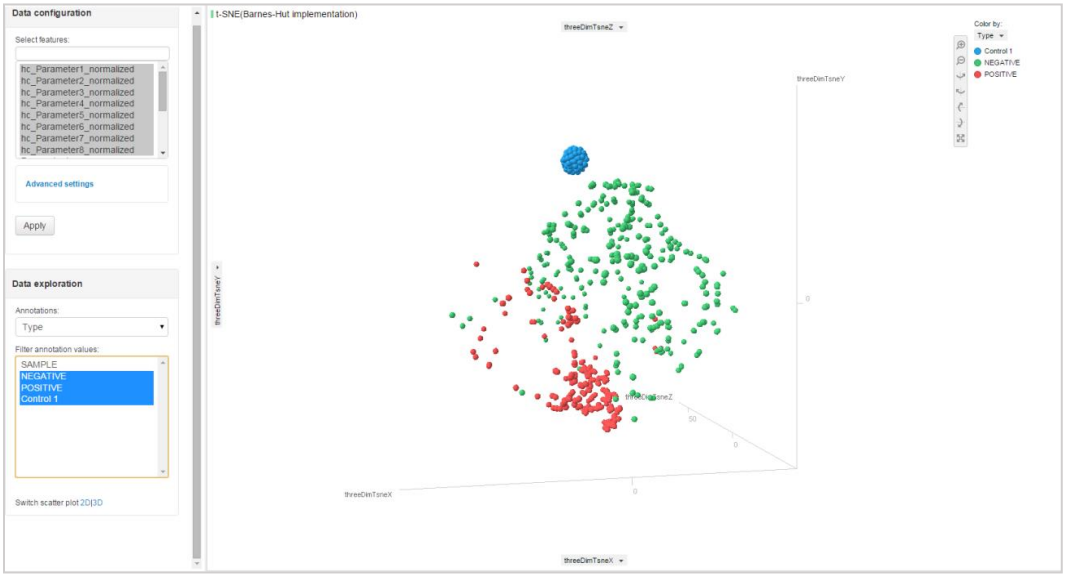


Figure 9-35: t-SNE App Results with a 3D Visualization

Further analysis can be done after t-SNE to identify differences between each group since the coordinates respective to each point are directly mapped to the original data, allowing any type of posterior analysis.

9.11.4 t-SNE App References

[1] L. van der Maaten and G. Hinton. (2008). Visualizing High-Dimensional Data Using t-SNE. Journal of Machine Learning Research, pp. 2579-2605.

[2] L. van der Maaten. (2014). Accelerating t-SNE using Tree-Based Algorithms. Journal of Machine Learning Research, pp. 3221-3245.

9.12 Calculations Explorer App

The **Calculations Explorer** App provides the ability to save and apply a **Calculations Explorer Template** as part of a Workflow to enable the capture and automated replay of transformation, calculations, and visualizations. This simplifies the addition of custom steps to a Workflow created by concatenating different Apps.

Note: A Workflow can contain multiple instances of the Calculations Explorer App and each one of these instances can use a different CE Template.

From the Signals Apps page, select the **Calculations Explorer** App card. A new tab containing the App will be added to the document and the App will be launched.

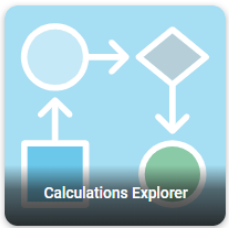
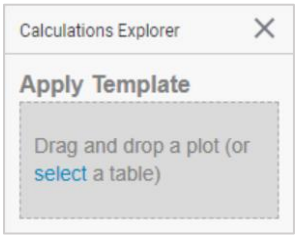


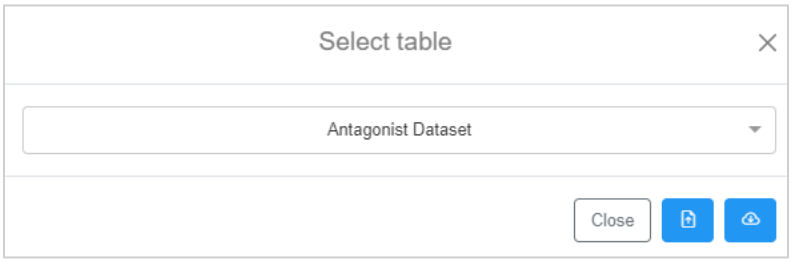
Figure 9-36: Calculations Explorer App Card

9.12.1 Configuring the Calculations Explorer App Data

Once the App has been added to the Workflow, click the **select** link in the **Apply Template** section,



Select the desired data table from the dropdown, then either select the **Load local template** icon or the **Load shared template** icon (see Calculations Explorer section for more details).



Once a CE Template is selected the standard matching interface as described in 10.5.7 Matching the Data to the CE Template section is displayed, allowing the user to match the data to the CE Template required inputs.

In those cases where the data has been loaded using the **Data Import** App, the Calculations Explorer will select the input table by default and open the **Select Template** UI, prompting the user to select from the available CE Templates in the Signals VitroVivo Metastore. If the user wants to select a different table or use a CE Template from a local file, close this screen to get back to the initial page that allows the selection of a table and a CE Template from a local file.

9.12.2 Execution of a Calculations Explorer Template in a Workflow

When a user has saved a CE Template as part of a Workflow using the **Calculations Explorer** App, new data can be automatically matched to the CE Template and will automatically execute when the Apps Workflow reaches the Calculations Explorer App step. If the data cannot be matched automatically an option is provided for the user to perform the matching manually and then apply the CE Template.

A toggle is also provided in the **Details** section of the CE Template that allows the **Autoapply** feature to be switched off.

9.12.3 Calculations Explorer App Results

After the execution of the **Calculations Explorer** App the result will be those provided by the CE Template that was executed.

9.13 High Throughput Screening (HTS) QA-QC App

The **High Throughput Screening (HTS) QA-QC** App allows for the review of plate statistics and QA-QC results for single-point, high-throughput screening.

From the Signals Apps page, select the **HTS QA-QC** App card. A new tab containing the App will be added to the document and the App will be launched.

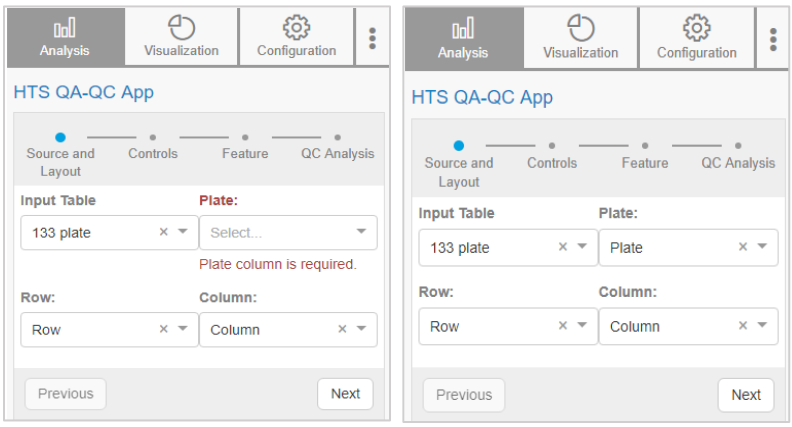
Figure 9-37: HTS QA-QC App Card

9.13.1 Configuring the HTS QA-QC App

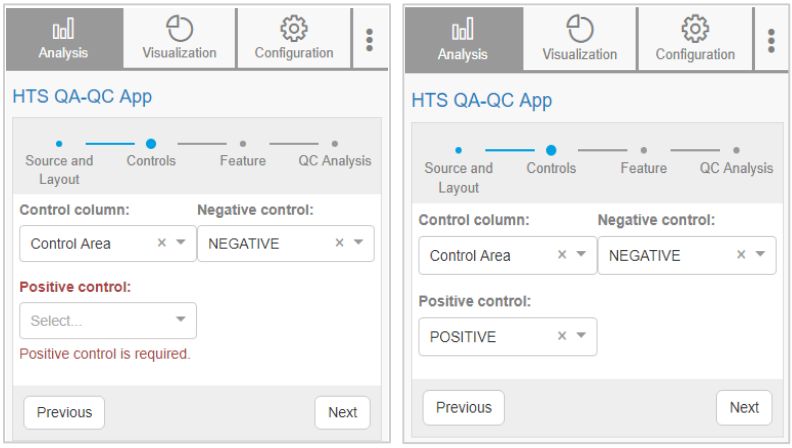
9.13.1.1 Analysis Tab

The **Analysis** tab contains a wizard-like configuration panel where the user can set the settings in four steps:

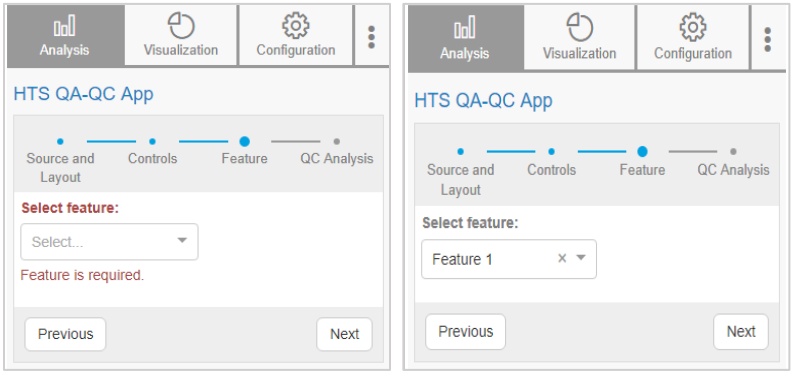
1. **Source and Layout:** The user will set which data table from the ones present in the document to use and the columns that define the plate layout. When all the settings are configured select **Next** to proceed to the next step. In case any of the required columns is not selected, a message under the selector informs the user that it requires a valid option before continuing.



2. **Controls:** In the Controls step, the user can set the column containing the information on the assay controls, followed by the value for each specific control, (i.e. Negative and Positive as shown in the left-hand figure below). As in the previous step, any required selections that are missing will prevent the user from proceeding in the wizard (right-hand figure below). When all selections are done, click **Next** to proceed to next step.



3. **Feature:** This is the last step before the execution of the data function. The user should select the feature of interest to continue. When the **Next** button is clicked the App will perform the QA-QC analysis.



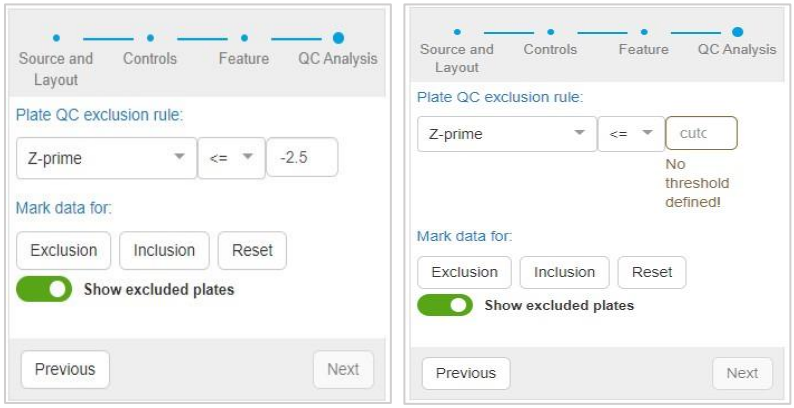
4. **QC Analysis:** In this panel the user has all the tools to perform QC at the plate or well level.
    - **Plate QC exclusion rule**: This section contains two dropdown menus and a textbox:

- The first dropdown allows the selection of the measurement to be used for the analysis.
- The second dropdown allows the user to select the rule to exclude plates. If the selected measurement value for a plate is greater or smaller (according to selection) than the threshold in the third box (the plate) should be excluded.
- The textbox allows the user to input the desired threshold. If the threshold value is empty all plates are automatically included, and the user is warned that no threshold is defined (see figure below).

  Note: The Plate QC is evaluated using all wells without considering the exclusion. This is because the intention is to perform a QC evaluation at the plate level. If exclusion was considered, it could potentially cause bad plates to be accepted if only certain controls are included in the analysis.

- **Mark data for**: This section allows the user to manually exclude or include data points or plates as well as toggling on or off the display of those plates that have been excluded. To mark specific data for exclusion, use the Spotfire® marking and click on the corresponding button.

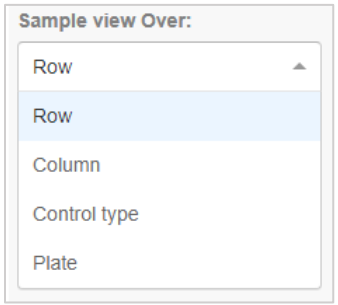
  Note: Plates that are not totally excluded (have less than 90% wells excluded) will appear in the visualization when the excluded plates are hidden colored as “Review plate”.



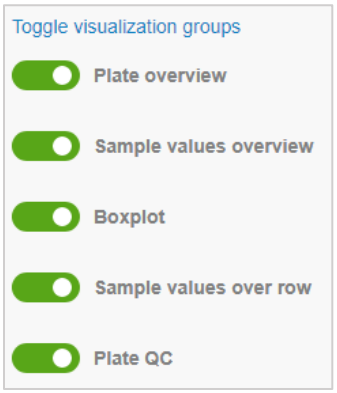
9.13.1.2 Visualization Tab

The **Visualization** tab contains controls to customize the result visualizations:

- **Sample view Over**: This control will determine over which column the X-axis of the “Sample view over X” visualization is computed.



- **Trellis plate overview**: This toggle allows the user to change between a trellised view of all the plates (default) or a view of each individual plate per panel.
- **Toggle visualization groups**: This group of toggles allow the user to hide or show each of the visualizations as better suited for the analysis being performed.



- **Show filter panel**: This allows the user to show or hide the filter panel as well as resetting the filters.
  - The toggle will show/hide the filter panel. This control is set to “off” by default.
  - The reset button will reset any filter applied in the document.

9.13.1.3 Configuration Tab

The third tab in the UI is named the **Configuration** tab. This contains the **Guided Mode** toggle to control if the App should follow a given data map or not and the **Exclusion panel** toggle that will show or hide the Signals Groups panel. This panel can also be accessed from the **Tools > Signals Groups** entry in the menu.

The HTS QA-QC App analysis can be run over data loaded through any of the Data Import Apps, thus using the underlying Screening data model, or use data loaded directly from Spotfire®. When **Guided Mode** is active (default), the App will try to detect which columns correspond to each configuration element, for instance the layout columns (Plate, Row, Column). If **Guided Mode** is off the App will not do any automatic selections. However, for data loaded using the Screening domain Apps with the underlying Screening data model, the App will always know which column or value corresponds to a specific setting.

Note: If the underlying data does not follow the Screening data model, it can be loaded using the **Unguided Mode**, however the exclusion functionality will not be available.

9.13.2 HTS QA-QC App Results

When the analysis is performed the visualization area on the right-hand side of the document will be populated with several visualizations, as seen in the figure below. The controls on the UI section described previously allow the user to perform the analysis and configure the visualizations.

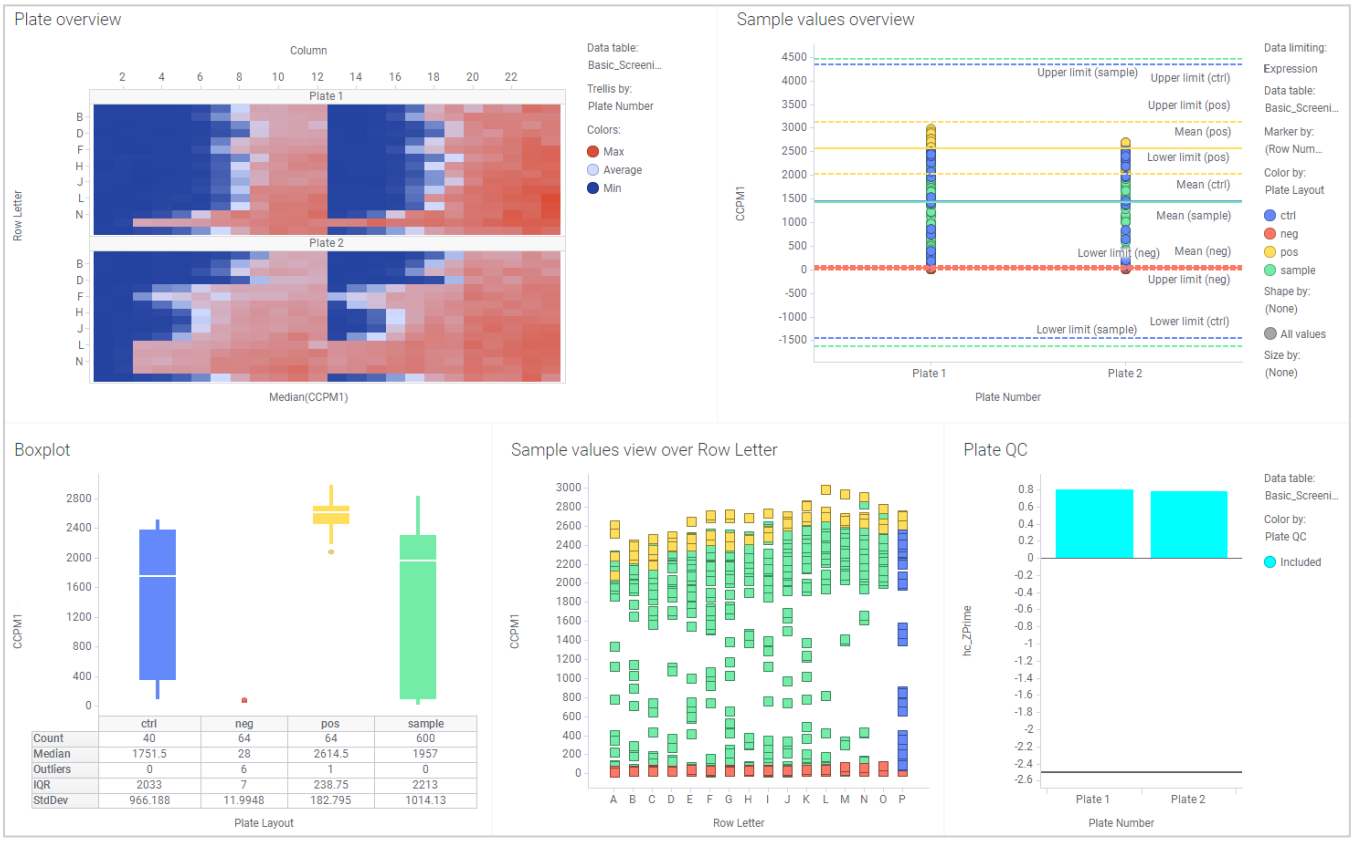


Figure 9-38: HTS QA-QC App Visualizations

The visualizations generated by the **HTS QA-QC** App are the following:

- **Plate overview:** This tiled scatterplot displays a plate view containing the selected feature's distribution across the plates. Each tile corresponds to a specific well value.

- **Sample values overview:** This scatterplot shows the feature distribution across plates, in a view that allows for early detection of outlier values. This is aided by the presence of lines for each of the well types (positive and negative controls as well as the other sample categories present in the selected controls column). The lines shown correspond to:
    - Overall mean for all plates.
    - Upper limit: The average plus three times the standard deviation of the feature distribution.
    - Lower limit: The average minus three times the standard deviation of the feature distribution.

    Any excluded value will be removed and the values in the visualization updated automatically.

- **Boxplot:** Here the feature distribution by control type can be seen together with several statistics calculated for each of the categories in the control column:

- Count
- Median
- Outliers
- IQR
- StdDev

As in the previous visualization, any excluded data points are hidden, and the results are updated automatically.

- **Sample values view over X:** This scatterplot displays the sample values over one of the following categories:
  - Row (default)
  - Column
  - Control Type
  - Plate

  These options are controlled in the **Visualization** tab in the UI. In this visualization excluded data are shown. If the **Show excluded plates** in toggled off excluded points from fully excluded plates will be hidden.

- **Plate QC:** This bar chart displays the QC measurements for all the plates. Through this visualization QC can be performed on the completed plates either by using the UI when setting the exclusion rules or by directly marking plates. It also displays the state of the plate:
  - **Blue:** Plate included/accepted for the analysis;
  - **Red:** Plate excluded/rejected for the analysis;
  - **Orange:** Review plate. The plate contains 10% to 90% excluded wells and should be reviewed.

  Additionally, a horizontal line is present showing the threshold value set for the quality measurement value selected, in the Y-axis. Excluded plates will not be shown in this visualization when the **Show excluded plates** toggle button is off.

9.14 Historical Data App

The **Historical Data** App can be used to pull all historical results for equivalent compounds or compound batches into the current Workflow via Spotfire® Information Links. This allows the user to confirm consistency with historical trends before publishing new data.

From the Signals Apps page, select the **Historical Data** App card. A new tab containing the App will be added to the document and the App will be launched.

Note when the user opens the App within a Workflow, the **Analysis** tab is the initial tab. However, when running outside of a Workflow, the initial tab is the **Configuration** tab so that the user may select the appropriate information link settings.

Figure 9-39: Historical Data App Card

9.14.1 Configuring the Historical Data App

9.14.1.1 Analysis Tab

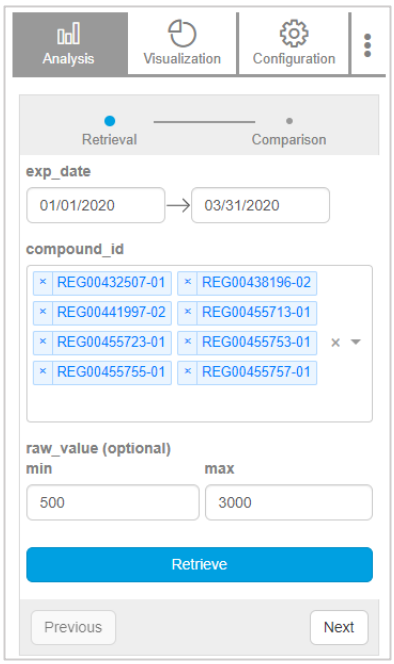
The **Analysis** tab contains a wizard-like configuration panel where the user can configure the settings for 2 different steps:

1. **Retrieval**: This step contains controls that allow the user to configure the input parameters for querying the information link.

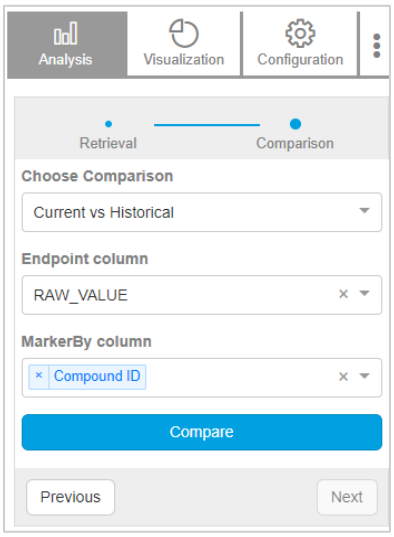
   Note: If the App has not been configured with an information link, this tab has a message “Please select an information link in Configuration”.

   There are three types of inputs:

   - **String inputs**. These types of input parameters can be configured in the Configuration tab to use values from current data or free-entry data.
     - **Current data**: If the user wants to use current data, the parameter must first be matched with a local column. After this is done, the unique values of the local column will be shown in a column value selector.
     - **Free entry**: This mode allows the user to specify any value (not just local values) as input to the query. The user should specify one value per line.
   - **Numeric inputs (optional)**: For this type of input parameter, the user can specify a range (min/max):
     - Both empty => parameter is ignored
     - Min only => include values larger than min
     - Max only => include values less than max
     - Min and Max => include values larger than min and less than max
   - **Date inputs:** Start and stop dates. Includes rows after start date and before stop date. The user can select the dates in the calendar that pops up when clicking on either of the date fields.

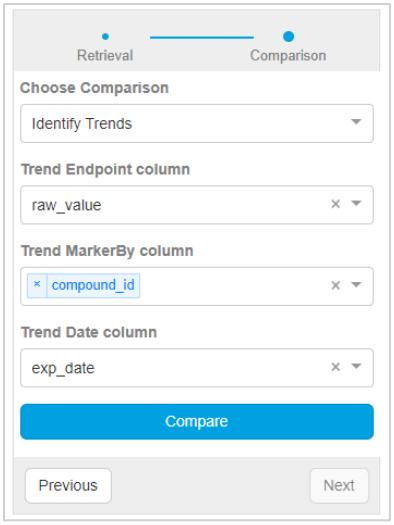


2. **Comparison**: This step allows the user to create two sets of visualizations:
   - **Current vs Historical**. To configure this visualization, the user needs to select an endpoint column and a **MarkerBy** column. Because this visualization compares two sets of data (current and historical), both selections need to be of columns that have been previously matched (in Configuration tab). If not, a warning message will be shown, prompting the user to match the columns. Once configured, this visualization shows the current endpoint values on the X-axis and the historical endpoint values on the Y-axis, showing one marker per value in the **MarkerBy** column.



   - **Identify trends**: To configure this visualization, the user needs to select a **Trend Endpoint column**, **MarkerBy column** and a **Trend Date column** from the historical data. Those columns do not need to be mapped since this visualization involves only one set of data (historical). Once configured, this

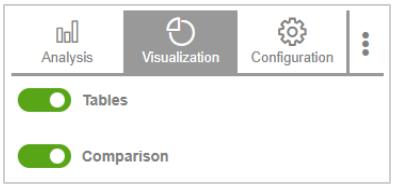
visualization shows the trend with the endpoint value on the Y-axis and the date on the X-axis, showing one marker per value in the **MarkerBy column**.



9.14.1.2 Visualization Tab

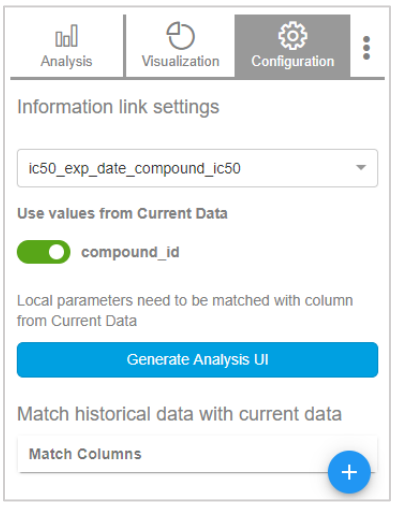
This section contains settings that affect what is displayed in the visualization area. The user can turn off and on groups of visualizations:

- **Tables**: This corresponds to the Historical Data and Current Data table plots.
- **Comparison**: This corresponds to the currently selected comparison plot in the second step of the wizard in the Analysis tab, i.e. Current vs Historical or Identify Trends.

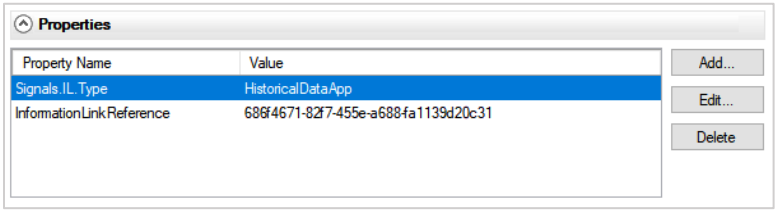


9.14.1.3 Configuration Tab

This section is used during the configuration phase of the App. Once the App has been configured properly as part of a Workflow, these settings will be pre-populated and the user of the Workflow typically will not need to change the settings.

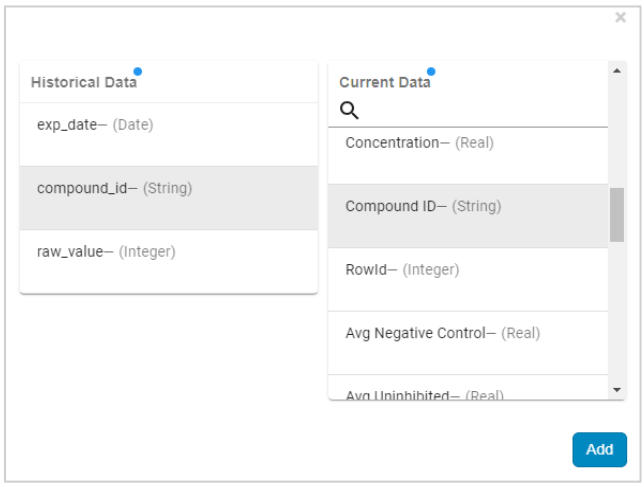


- **Information Link Settings**: Select one of the information links in the connected Spotfire® Server. To appear as an option here, the administrator needs to use the Information Designer (part of Spotfire®) and annotate the information link with a specific property (Signals.IL.Type) and value (HistoricalDataApp):



- **Use values from Current Data**: One switch (default true) will appear per String type parameter in the Information Link.
- **Generate Analysis UI**: When the user clicks this button, the UI in Analysis settings and Use Values from Current Data will be (re)generated and any settings/visualization or previously retrieved data removed.
- **Match historical data with current data**: When the user clicks on the *Plus-sign* icon, a window opens that lets the user match columns from historical and current data. This matching is required for values to appear in the column value selectors of Analysis when configuring the information link settings and for any columns used in the Current vs Analysis visualization.

  Any columns matched must have the same data type, otherwise a small warning triangle will appear in the overview after clicking Add.



9.14.2 Historical Data App Results

The Historical Data App results are shown in the right-hand panel based on the configuration settings selected in the *Retrieval* and *Comparison* steps.

9.15 High Content Profiler App

The **High Content Profiler** App is a Spotfire® add-on that allows the analysis of multivariate and univariate data extracted from instruments and image analysis software. An automated guided Workflow coupled to simple wizards and dialogs enables scientists to perform advanced features such as normalization, selection, classification, profiling, and hit selection in a powerful and user-friendly environment.

From the Signals Apps page open the **High Content Profiler** App. A new tab containing the App will be added to the document and the App will be launched.

Figure 9-40: High Content Profiler App Card

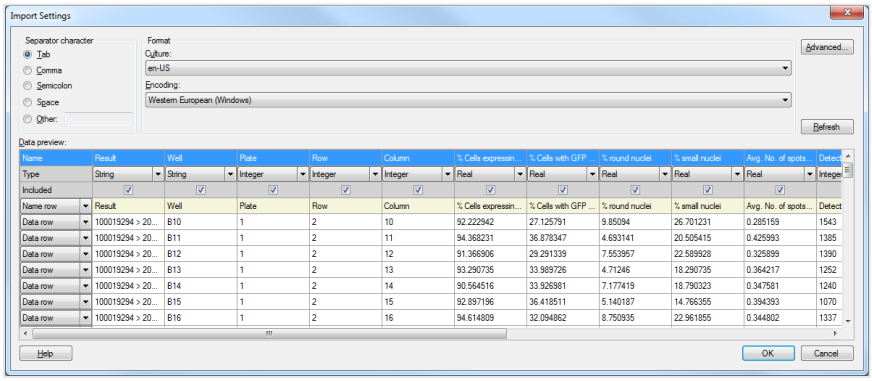
Launching the **High Content Profiler** App first requires a suitable input such as data containing features extracted from images captured on an HCS instrument to be loaded in the active Spotfire® project.

Note: The High Content Profiler is not a standard App, it is a wrapper for the pre-existing High Content Profiler Tool that allows it to be launched from the Signals Apps page. Because of this, this App is not supported as part of an Apps Workflow.

9.15.1 Opening a Project

9.15.1.1 Opening a New Project

When opening a new Spotfire® project, a data table must first be loaded. Select **File > Open > Browse local file** and browse to the desired data file locally and select **Open**. The “Import Settings” dialog appears allowing a preview of the imported data.



The “Import Settings” dialog offers the opportunity to check that the column types automatically set by Spotfire®, like **Integer** or **Real**, correspond to the expected data types of the imported HCS data. If the column data type association is not correct, the column type can be changed manually. Selecting **OK** will load the data set, using the corresponding data types for each column into Spotfire®.

Once the HCS data is loaded, the **High Content Profiler** App can then be launched. Click on the **High Content Profiler** App icon and a new tab containing the **High Content Profiler** App wizard will be added to the document.

Once invoked, the first step in the **High Content Profiler** App allows you to select the Spotfire® input table.

9.15.1.2 Opening an Existing Project

The document resulting from the invocation and execution of the **High Content Profiler** App can be saved as a Spotfire® project document. When opening this saved document, tab panels with the visualizations produced by the automated guided Workflow execution will be immediately available.

9.15.2 Launching an Analysis

You can launch the **High Content Profiler** App through a wizard which guides you through the following steps:

1. Defining the **Experiment type**, **Analysis type** and **Input table**.
2. Defining the analysis layout for **Plates**, **Controls**, **Rows** and **Columns**, when available and selecting the **Features to include** in the analysis.
3. Defining the identifiers of **empty wells** and **negative** and **positive controls** when controls are selected.
4. Defining the **Annotation columns** of the plates and wells.

9.15.2.1 HCP – App Wizard

The **High Content Profiler** App provides a web-based wizard for configuration of the different settings. This wizard contains four main sections together with an advanced settings section.

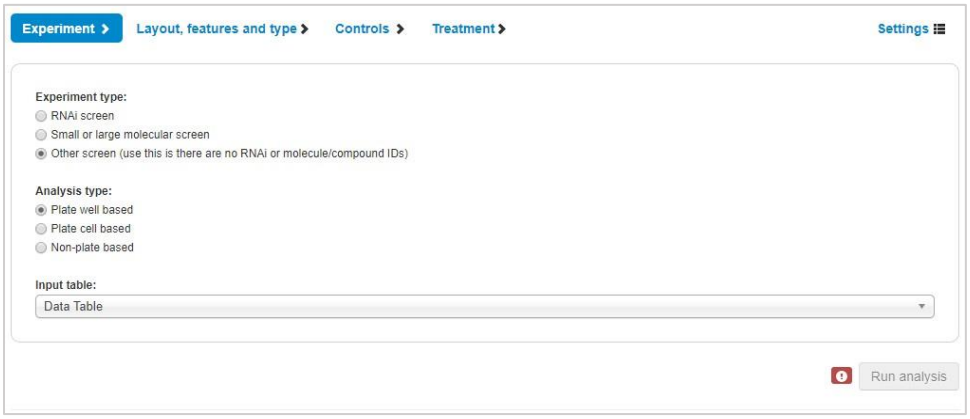
The wizard navigation section provides access to the four main sections that require configuration:

- Experiment
- Layout, features and type
- Controls
- Treatment

The active section at a given moment is marked in blue. The advanced settings section can be accessed from the **Settings** button on the right-hand side of the wizard at any time.

9.15.2.2 HCP – Experiment and Analysis

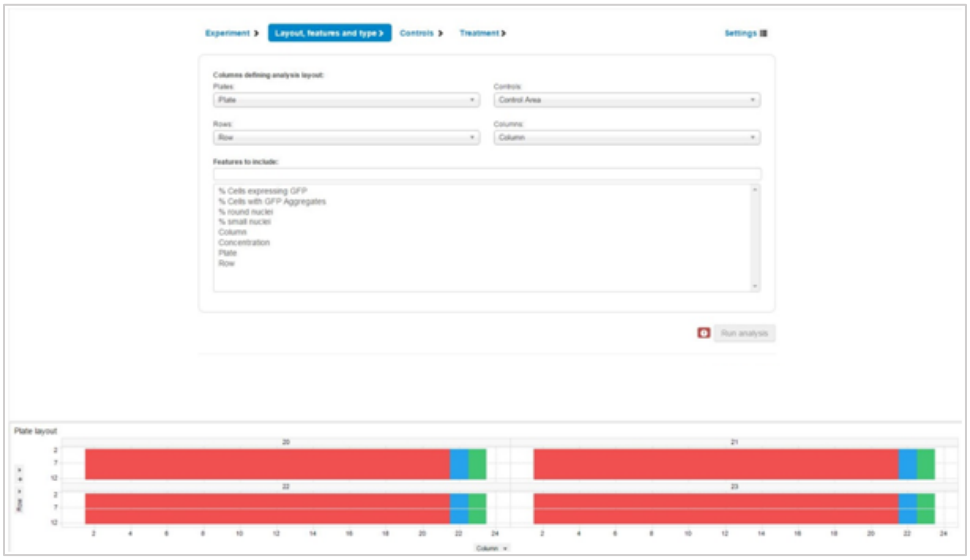
The type of experiment and analysis as well as the data source table can be specified in the first step of the **High Content Profiler** App wizard.



Option	Description
Experiment type	Allows you to specify if the HCS data is part of RNA interference (RNAi) screen, a chemical compounds or large molecule screen or another type of experiment.
Analysis type	Allows the user to select if the analysis is plate and well based, plate and cell based, or non-plate object based.
Input table	From the dropdown menu containing all the tables previously loaded into the current Spotfire® project; select the data table containing the HCS data to analyze. If the input table is not listed in the dropdown menu, cancel the wizard and open the data file from the ‘File’ menu as described above.

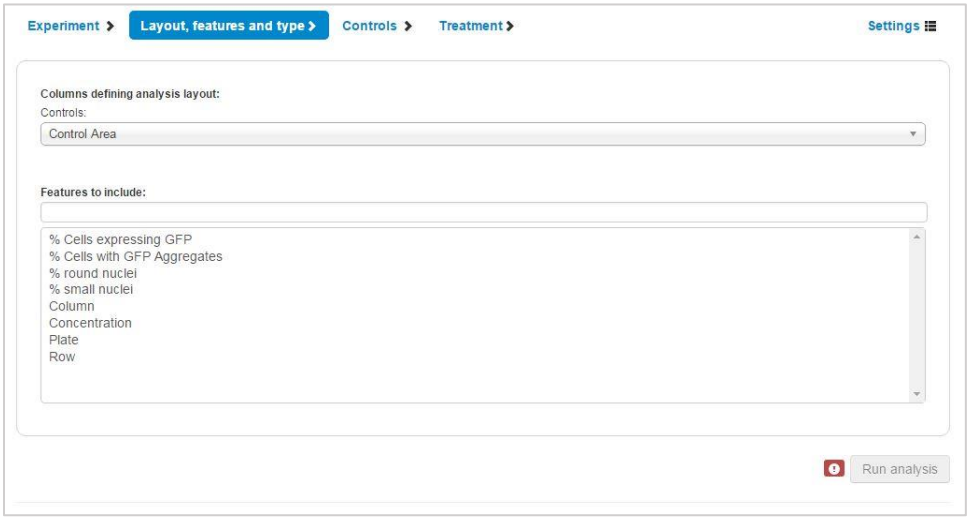
9.15.2.3 HCP – Layout Features and Type

The next step allows you to specify the columns of the data file containing information about the plate's labels and layout, as well as the columns containing the features to be analyzed by the **High Content Profiler** App. The visualization below will update according to the layout columns selected and will be colored according to the selected control column. This visual aid helps avoid the selection of incorrect parameters.



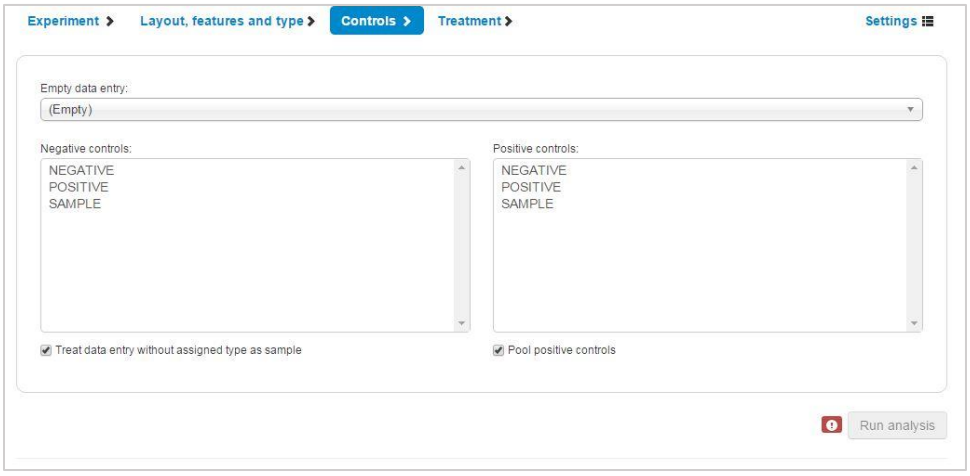
Option	Description
Plate	Indicate which column contains the plate identifier.
Rows	Indicate which column contains the row identifier.
Columns	Indicate which column contains the column identifier.
Controls	Indicate which column contains the control identifier.
Columns defining position for each cell	Indicate which columns contain the X and Y coordinates for the cell positions in the well if available. These selectors are only active if "Plate cell based" analysis type is selected.
Features to include	Allows you to select the columns containing HCS features to include in the analysis of the High Content Profiler. The selected features will be the ones used for the analysis. If a single feature is selected Univariate specific analysis mode will be triggered. This mode is only available for plate-based studies with both positive and negative controls defined. **Note:** Only numerical feature values are supported for analysis.

If the user has selected a non-plate object-based analysis the interface will be slightly different. In this case only the control column and the features to be used in the analysis need to be selected, as there is no layout information required. Also, the layout preview will not be displayed.



9.15.2.4 HCP – Controls

The identifiers of empty wells as well as the positive and the negative controls can be specified in the next step. In addition, if the plates contain technical replicates of the controls, the replicates can be pooled into a single value. If (None) was selected as Control definition in the previous dialog, no control definition is required, and configuration of this section will not be needed.



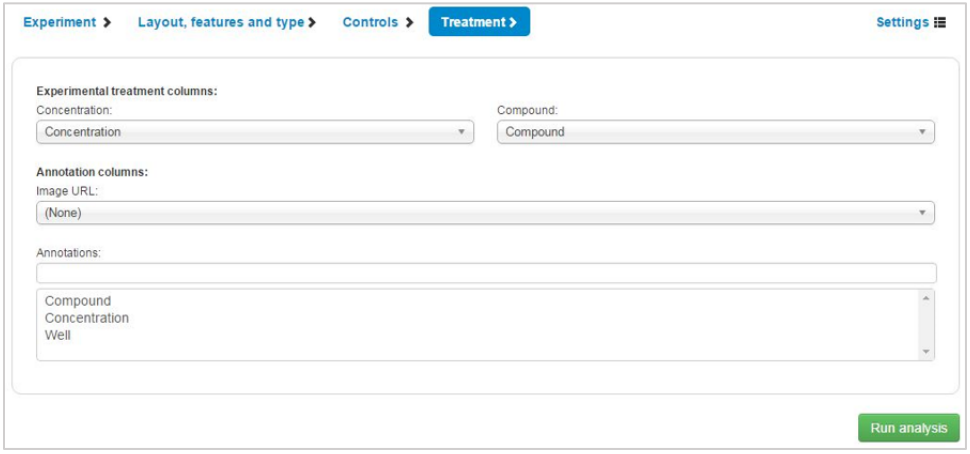
Pooling only affects the feature selection and hit classification.

Option	Description
Empty data entry	Allows you to define the label assigned to empty wells (i.e. wells that do not contain cells by design) which should be ignored in the analysis.

Negative controls	Displays the list of available controls[1]. Allows the selection of those controls that should be considered negative phenotypes or mock treated samples. If more than one negative control type is selected, they will be pooled for the analysis.
Positive controls	Displays the list of available controls[1]. Those controls selected from this list will be considered as positive phenotypes relative to the negative or neutral controls. If more than one control is selected, they will either be treated as the same (pooled) or distinct phenotypes based on the “Pool positive controls” setting.
Pool positive controls	If checked, the positive controls will be treated as replicates of the same control to be pooled as a single value for the feature selection and hit classification. If the option is not selected, each positive control type will remain independent during the feature selection and hit classification steps of the automated workflow.
Treat data entry without assigned type as sample	The data points without a specific type assigned will be treated as samples. If unchecked, the data points without any type specified are not considered during the calculations of normalization, feature selection and classification.

9.15.2.5 HCP – Treatment

The metadata associated with the plates and/or wells from the HCS plates can be used to enrich the visualizations in Spotfire® after the execution of the **High Content Profiler** automated Workflow.



Note: Only the first 1000 available values of the Controls definition column will be displayed. In some cases, when the compound column is selected as Controls definition, more than 1000 values may be present. In this case, the recommended approach would be to create a new calculated column to use as control definition where all compounds except the ones to be used as controls are renamed to sample.

Option	Description
RNAi-ID	Indicates which column contains the identifier of the silencing RNA (siRNA) or siRNA pool tested. This selector is only available if “RNAi screen” is selected.
Compound	Indicates which column contains the identifier of the chemical compound tested. This selector is only available if “Small or large molecule screen” is selected.
Gene	Indicates which column contains the identifier of the gene targeted by the RNAi treatment. This selector is available only if RNAi screen is selected.
Concentration/ RNAi Concentration	Indicates which column contains the concentration of the chemical compound or RNAi tested. This selector is available if “RNAi screen” or “Small or large molecule screen” are selected. **Note:** If more than 5000 compounds are present in the analysis currently only the first 5000 will be available for selection in the dose response page of the results document.
Image URL	Indicates which column contains the URL to the source image from which the features were extracted for each well.
Annotations	Indicate which columns from the HCS data table that were not selected in the previous steps of the wizard should be added to the analysis as annotations.
Analysis settings	Access the advanced settings dialog
Finish	Allows you to launch the data processing steps of the High Content Profiler and to generate the visualization of the results in Spotfire®.

9.15.2.6 Execution of the Workflow

The **High Content Profiler** App Workflow is executed upon clicking the **Run analysis** button available in any of the wizard steps. This button will be disabled until the minimal required parameters for the execution have been correctly configured in the wizard. While the button is disabled a red error provider will be present on the left-hand side of the button and hovering with the mouse over this indicator will provide information on the reason why it is disabled. Once the button is enabled, and the user clicks upon it the Workflow execution will begin. During this Workflow execution, the High Content Profiler will provide information on which step is being executed in the main window. Additionally, detailed information on the analysis steps can be obtained from the “Notifications” dialog.

The “**Notifications**” dialog is available from the *Bell* icon at the top-right-hand corner of the Spotfire® main window.

9.15.2.7 Defining the Analysis Settings

The detailed settings of the **High Content Profiler** App analysis are available from the **Settings** button at the top-right-hand side of the wizard.

Settings

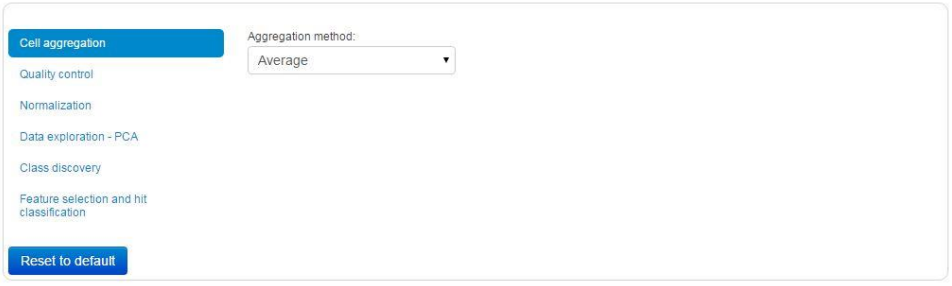
The parameters related to the **Cell aggregation**, **Quality control** (QC), **Normalization**, **Data exploration – PCA**, **Class discovery**, and **Features selection and hits classification can** be changed after selecting any of these topics in the **Settings** dialog.

Tab	Description
Cell aggregation	Allows the selection of the aggregation method to be applied in cell-based analysis. It is only available if “Plate cell based” analysis type is selected.
Quality control	Allows the definition of thresholds for quality control parameters.
Normalization	Allows the selection of normalization methods to be applied to the features.
Data exploration - PCA	Allows the selection of which features should be used for generating principal component analysis (PCA) results.
Class discovery	Allows the configuration of the class discovery parameters.
Feature selection and hit classification	Allows the selection of which features should be used for feature selection and classification, how features should be selected, and which classification method should be used. This option is only available for multivariate analysis (more than one feature is used) where controls have been defined.
Reset to default	Resets all settings in the wizard to the default ones.

The **Settings** dialog is also directly available from the **High Content Profiler** results visualizations in Spotfire®.

Cell Aggregation

This settings panel is only available in single cell plate-based analyses.



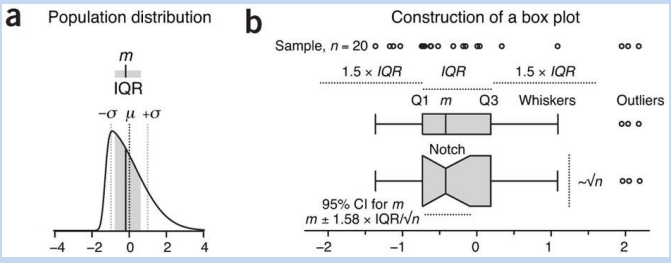
Option	Description
Aggregation method	Allows the user to select the preferred aggregation method. The choices available are: • Average

	• Median • Mode • Geometric mean
Reset to default	Restores the default analysis settings.

Quality Control

Option	Description
Flag wells with … % missing feature values or more	Allows you to change the threshold used to flag wells as missing wells in the quality control output page.
Treat values more than … * IQR away from first or third quartile as outliers	Allows you to define the constant that, multiplied with the inter quartile range (IQR), will define what values will be outliers within each feature. Outliers are those values that are outside the range of [First quartile - Constant * IQR, Third quartile + Constant * IQR].
Reset to default	Restores the default analysis settings.

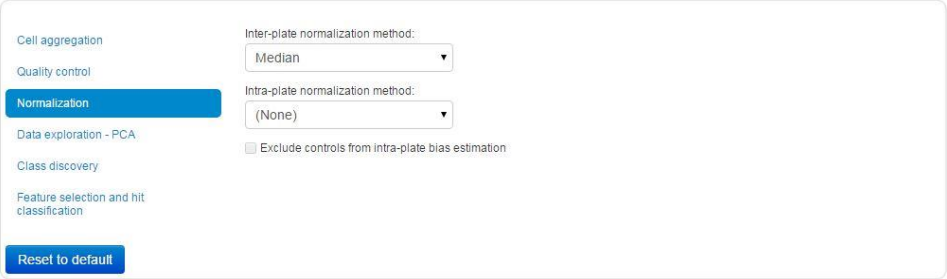
The outliers can be visualized on box plots as follows (taken from Krzywinski& Altman, Nature Methods 11, 119–120 (2014) doi:10.1038/nmeth.2813):



(a) The median (m = −0.19, solid vertical line) and interquartile range (IQR = 1.38, gray shading) are ideal for characterizing asymmetric or irregularly shaped distributions.

The constant used in the High Content Profiler indicates the number of interquartile ranges (IQR) away from First quartile and Third quartile required for a value to be considered as an outlier.

Normalization



Option	Description
Inter-plate normalization method	Allows the selection of the required normalization method. The available methods are: • **None:** Do not normalize feature values. • **Median:** Normalize using the median values of negative controls. Determines first the median of negative controls for each plate and then global median across controls of all plates. Deviations of the per-plate median from the global median are then determined. Finally, values are normalized by subtracting the deviation for a given plate from all values on that plate. This method assumes that the technical biases are mainly additive. • **% of negative control:** For each plate, divide all values by the median value of the negative controls on that plate. Normalized values will be on a scale between 0 and infinity, with 1 denoting the median of negative controls on that plate. • **Normalized percent inhibition (NPI):** This method is suited for antagonist assays. It divides the difference between each measurement and the average of the positive controls on the plate by the difference between the measurements on the positive and negative controls. • **Custom normalization methods:** If the user has registered any custom normalization methods as described in the custom normalization methods section, these methods will also be available here for selection.
Intra-plate normalization method	Allows the selection of the required within-plate normalization method. Within plate normalization can be helpful for removing spatial biases (e.g. row or column effects). The available methods are: • **None:** Do not apply within plate normalization. • **Median polish:** Use median polish (Tukey, J. W. (1977). Exploratory Data Analysis, Reading Massachusetts: Addison-Wesley) to identify and remove row and column biases individually for each plate.

	This option will not be available if the user selects a custom normalization method. However, the user can include this intra-plate normalization within his registered custom data function.
Exclude control wells from intra-plate bias estimation	When using median polish intra-plate normalization, selecting this option allows excluding control wells from bias estimation. This is particularly important if control wells are not distributed randomly or are distributed in ways where specific columns or rows contain preferentially controls of one type.
Reset to default	Restores the default analysis settings.

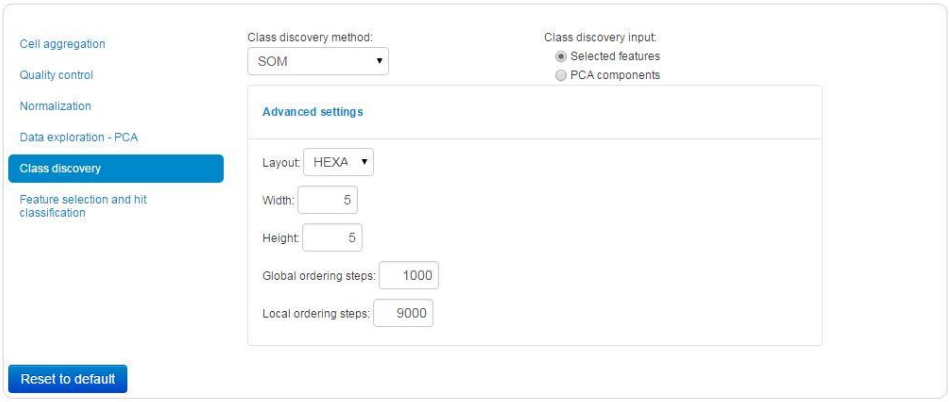
Note: For more information in custom normalization refer to:

Custom Inter-Plate Normalization Methods.

Data Exploration - Principal Component Analysis (PCA)

Option	Description
Imputation method	Allows the selection of the method for imputing missing values in the data set. The available methods are: • **Replace with constant value:** Replaces all missing values with the constant value defined in the text field on the right-hand side. • **Feature average:** Replace missing values with the average of all values of the same feature. • **Well average across features:** Replace missing values with the average of all feature values of the same well. • **Average of K-nearest neighbors (KNN):** For each well with missing values first identify the set of *k* most similar wells. Then use the average of values from the same feature in that set to replace the missing value.
Reset to default	Restores the default analysis settings.

Class Discovery



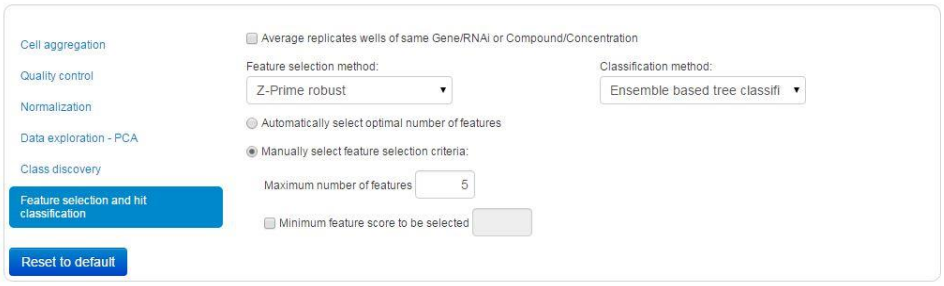
Option	Description
Class discovery method	Displays the methods available for class discovery, currently SOM or None.
Class discovery input	Allows the user to select what to use as input to the class discovery method selected, this can be: • The selected features, in which case the input will be the scaled normalized features. • The PCA components.
Advanced settings	This will open a configuration menu for the selected method.
Reset to default	Restores the default analysis settings.

SOM Advanced Settings

Option	Description
Layout	Provides the choice of the grid layout for the SOM training, this can be: • Hexagonal (HEXA) meaning each of the nodes has 6 neighbors. • Rectangular (RECT) where each of the nodes will have 8 neighbors.
Width	Width of the SOM grid in cells (1-50).
Height	Height of the SOM grid in cells (1-50).
Global ordering steps	The number of steps that will be used for the initial training of the SOM (1-1000).
Local ordering steps	The number of steps to be used in the fine tuning of the SOM (1-10000).

Feature Selection and Classification

This settings panel is not available in Univariate specific analysis mode or when no control column has been selected.



Option	Description
Average replicate wells of same Gene/RNAi or Compound/Concentration	If selected, this option will average replicates defined by identical combinations of RNAi/Gene identifiers or Compound identifier/Concentration for feature selection and classification. Wells of types selected as positive or negative controls will never be averaged.
Feature selection method	Allows the definition of the method used to select the most relevant features for discrimination of positive and negative controls. The available methods are: • None – use all features • Ensemble based tree classifier • MRMR • T-score • Z-prime score • Robust Z-prime score
Classification method	Allows the selection of the method used to assign class labels corresponding to Positive and Negative controls to wells with unknown type. The available methods are: • Ensemble based tree classifier • Robust Mahalanobis distance classifier
Automatically select optimal number of features	If selected, the optimal number of features to be used for classification will automatically be determined by searching groups of 1 to the total number of features (or 50, whatever is smaller) of the features with highest ranking using the selected feature selection method. **Note**: that depending on the data set size this might take some time.

Manually define feature selection criteria	If selected, allows you to manually define criteria for selecting features to be used for classification using the feature scores and ranks after applying the Feature Selection method.
Maximum number of features	Maximum number of top ranked features to be selected for classification.
Minimum feature score to be selected	If selected, allows the definition of a threshold on the feature scores according to the feature selection method below which features will never be selected. For example, for Z-prime scores, features with scores below 0 often are not considered relevant.
Reset to default	Restores the default analysis settings.

9.15.3 Workflow Results

The execution of the **High Content Profiler** App automatically adds a set of tabs to the current Spotfire® project.

The available tabs depend on the type of analysis.

- Map and settings
- QC cell density (Plate cell-based analysis only)
- QC cells distribution (Plate cell-based analysis only)
- Cell Features overview (Plate cell-based analysis only)
- Cell gating (Cell/object-based analysis only)
- Cell QC Overview (Non-plate object-based analysis only)
- QC overview (Plate based analyses only)
- Plate layout (Plate based analyses only)
- Plate values (Plate based analyses only)
- Plate layout editor (Plate based analysis only)
- Annotation overview (Non-plate object-based analyses only)
- Feature overview (Plate based analyses only)
- Exploration of data structure
- Class definition
- Feature selection and classification (If adequate controls are defined)
- Hit detection (If adequate controls are defined)
- Dose response (If a concentration column is defined)

In the case of Univariate analysis, which is available only for plate-based analyses where both positive and negative controls are defined and a single feature is selected, there are certain changes in the tabs displayed:

- Feature overview page is not available
- Feature selection and classification page is not available
- Hit detection page is not available
- Univariate hit detection page is available

The first tabs allow the evaluation of the quality of the experiment and explore the main characteristics of the dataset. In all these tabs, the “hc_” prefix indicates values calculated in the High Content Profiler automated Workflow. This

will also be true for any of the new tables generated by the Workflow, they will all be preceded by a “hc_” which makes them easily identifiable.

A detailed description of each of these tabs is provided.

9.15.3.1 Map and Settings Page

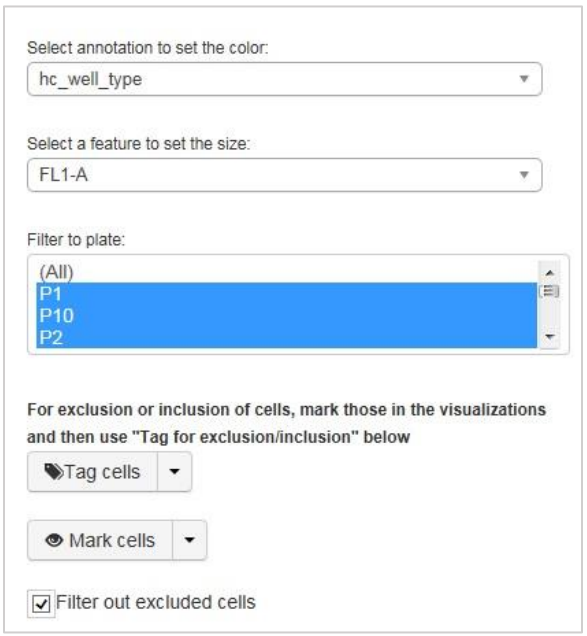
This page provides an overview of the available visualizations in the executed Workflow. The thumbnails can be clicked on to navigate to the desired tab. On the upper part of the visualization the buttons “**Workflow Map**” and “**Settings**” can be used to alternate between the map visualization and the Settings visualization described above.

9.15.3.2 QC Cell Density Page

This page provides information on the number of cells in each of the selected wells, as well as a visual representation of their distribution within the wells.

The user interface provides the following controls:

- **Select annotation to set the color**: This dropdown menu provides the option to select the annotation to use for coloring the visual representation of cells and wells.
- **Select a feature to set the size**: Allows the selection of the feature to use to size the cell representations in the Cell location overview visualization.
- **Filter to plate**: Provides filtering controls to filter to select specific plates for visualization.
- **Tagging controls**.
- **Filter out excluded cells**: This checkbox controls the visualization of the excluded samples.



Cells tagged for exclusion using the tagging controls will not be included in the aggregation of the cells by well used in the downstream analysis. This is regardless of their presence or not in the visualizations (according to the “Filter out excluded cells” checkbox).

The QC cell density page provides the following visualizations:

- **Plate layout overview:** This panel allows the user to select specific wells for which to display the details in the Cell location and Cell summary panels.
- **Cell location overview:** This panel is displayed only if the X and Y coordinates for the cell positioning have been provided. It shows a trellis containing a representation of the cells in the wells selected in the “Plate layout overview” panel. The color and sizing of the cell representations is defined by the features selected in the user interface.

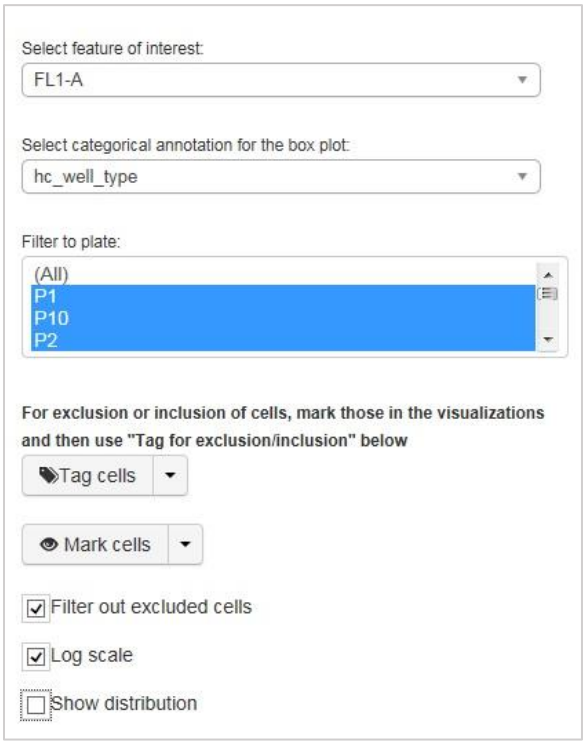
- **Cell summary:** This table contains a summary of the wells selected in the “Plate layout overview” panel.

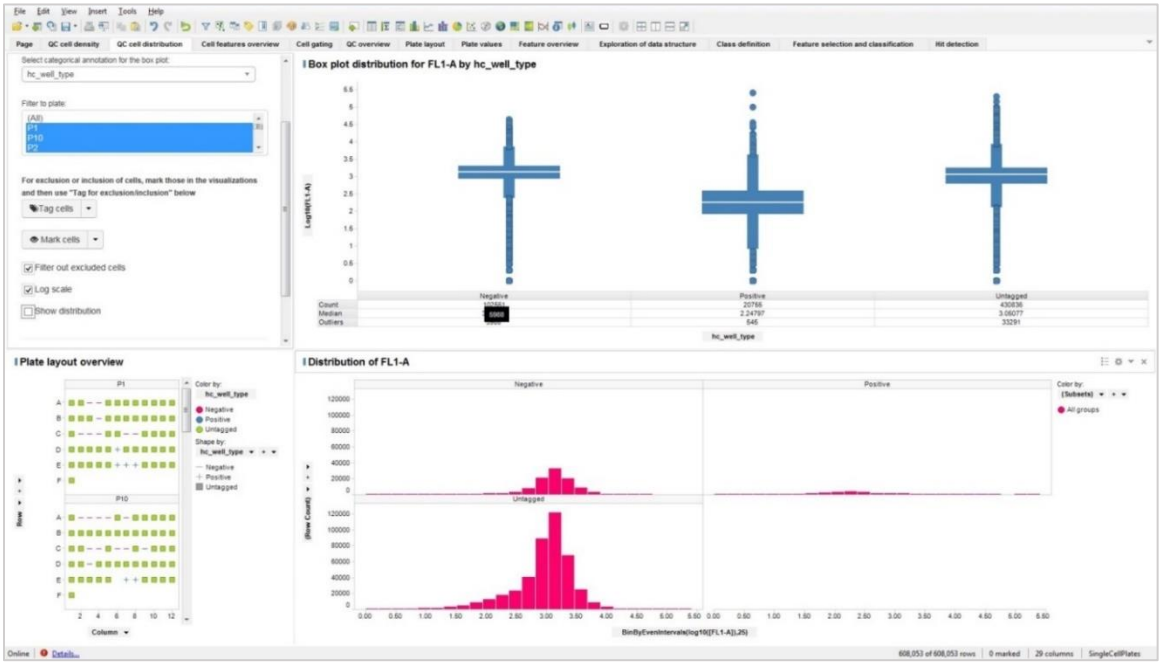
9.15.3.3 QC Cell Distribution Page

This page provides information on the distribution of the selected feature values over the different cells and facilitates the identification of outliers for selected features to exclude cells from downstream analysis.

The user interface provides the following controls:

- **Select feature of interest:** Allows the selection of the feature to evaluate.
- **Select annotation to compare:** This allows the selection of a categorical annotation to compare the distribution of the feature across the samples with the different categories in the selected annotation, in both the boxplot and bar chart visualizations.
- **Filter to plate:** Provides filtering controls to filter to select specific plates for visualization Exclusion controls.
- **Filter out excluded cells:** This checkbox determines if those cells tagged for exclusion will be visible or not in the different cell level visualizations.
- **Log scale:** Determines if the scale of the features in the visualization is represented as linear (default) or log.
- **Show distribution:** When checked, this option overlays the distribution of the feature in the form of a vertical histogram over the boxplot.





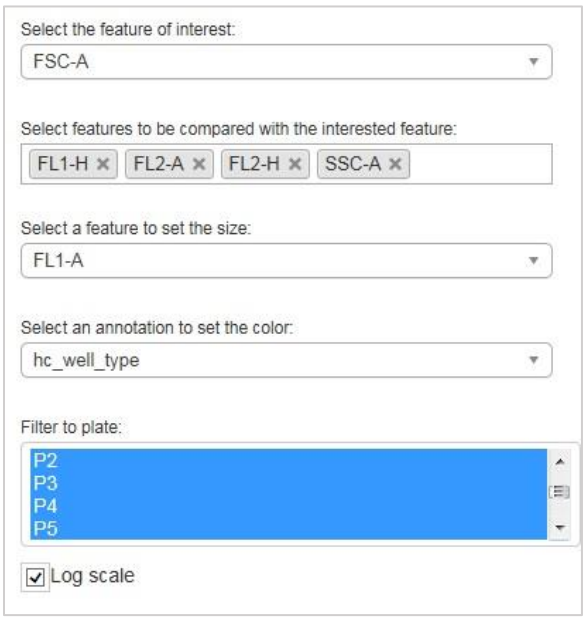
The QC cell distribution page provides the following visualizations:

- **Plate layout overview:** This panel allows the user to select specific wells for which to display the details in the Distribution plot.
- **Box plot distribution:** This panel provides a box plot visualization of the selected feature of interest separated by the selected annotation if any.
- **Distribution:** This panel provides a bar chart visualization of the selected feature of interest separated by the selected annotation if any. In this bar chart the distribution of the data selected on the plate layout visualization can be seen against the overall distribution of the data.

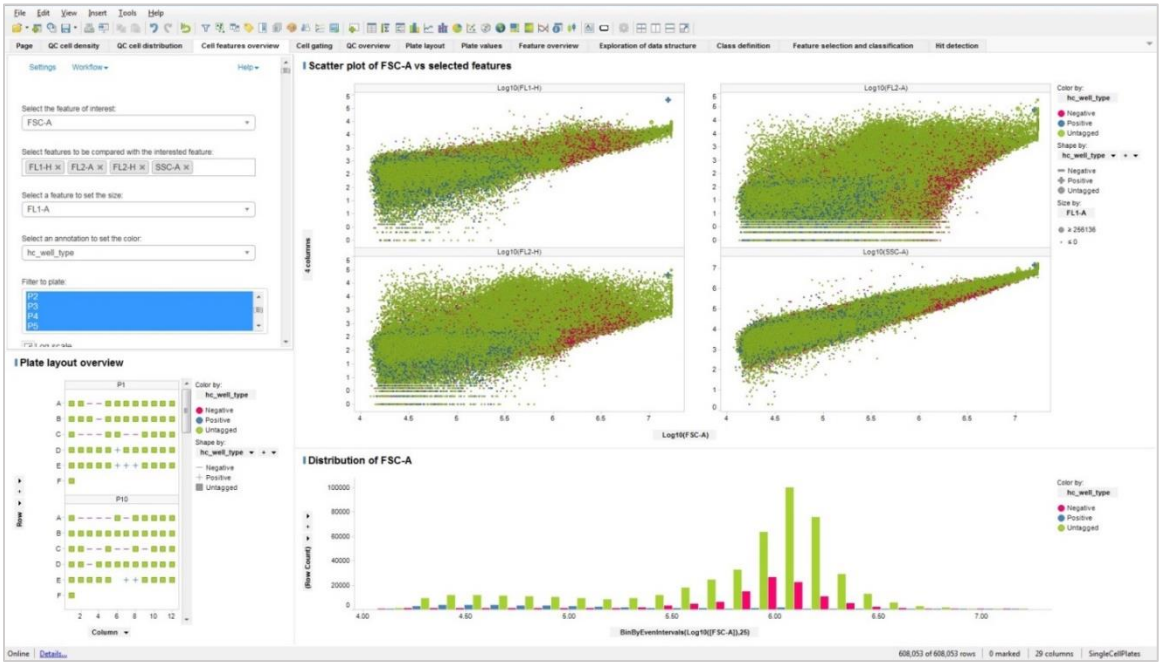
9.15.3.4 Cell Features Overview Page

This page provides visualizations that allow the comparison of different features as well as identify outlying samples and regions of the feature distribution that are of interest for the downstream analysis.

The available options in the user interface are:



- **Select feature of interest:** Allows the user to select the feature to display in the X-axis of the scatterplot and in the bar chart.
- **Select secondary features to compare:** Provides the option to select the features to be compared with the main feature of interest in a trellis of scatterplots.
- **Select a feature to set the size:** Allows the user to select the features that will be used to size the datapoints of the scatterplots.
- **Select an annotation to set the color:** Will determine the color with which the points in the scatterplot are plotted as well as the colors of the bar plots.
- **Filter to plates:** Provides filtering controls to filter to select specific plates for visualization.
- **Log scale:** When checked, the visualizations will display the features in log 10 scale (default is unchecked).

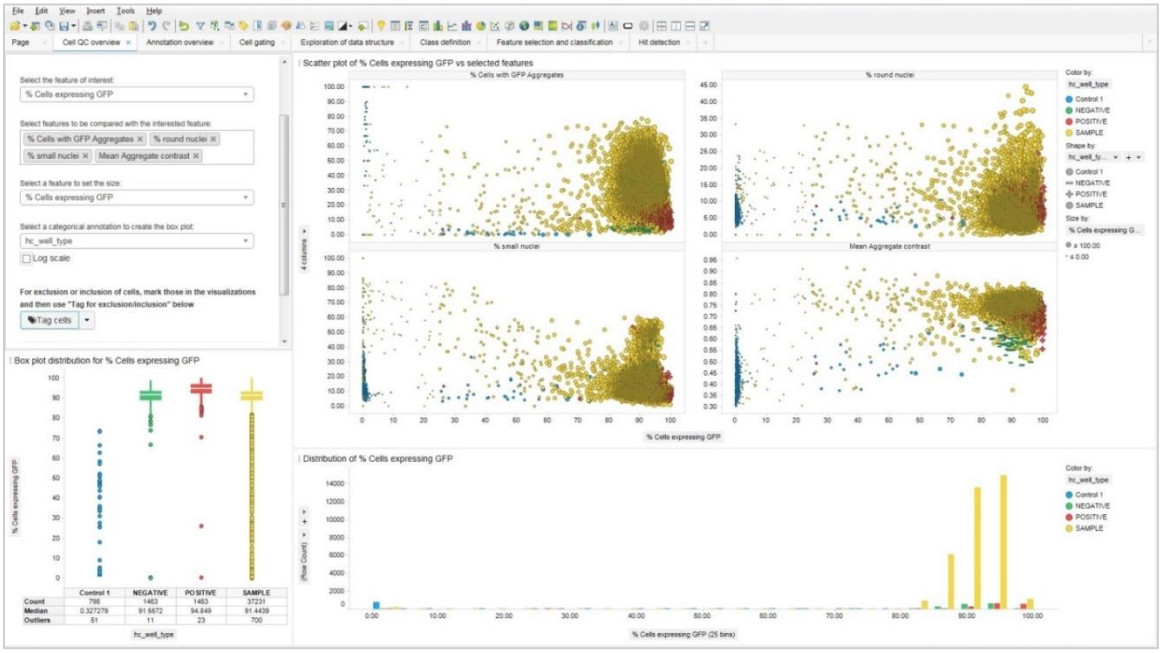


The features overview contains the following visualizations:

- **Plate layout overview:** This panel allows the user to select specific wells to mark in the Scatterplot and Distribution visualizations.
- **Scatterplot:** This plot shows the distribution of the feature of interest in the X-axis against all the selected secondary features in the Y represented as al trellis. Color and sizing are defined by the selections made in the user interface controls.
- **Distribution:** This plot shows the distributions of the feature of interest as a bar chart where each of the categories present in the selected annotation are shown in parallel.

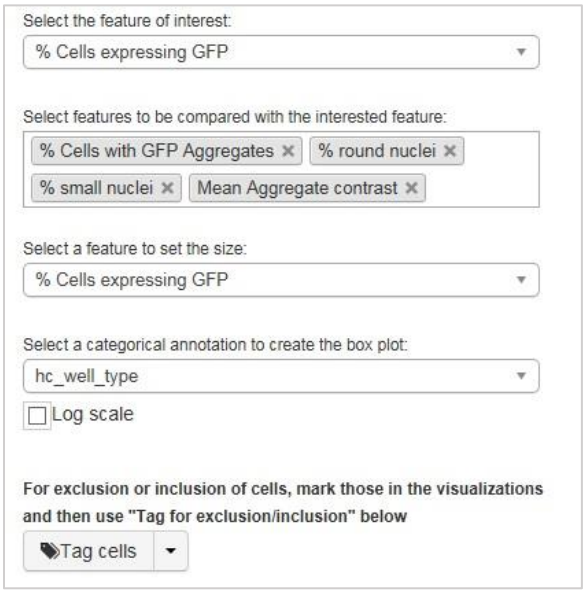
9.15.3.5 Cell QC Overview Page

The Cell QC Overview page appears only when using Single cell non-plate-based workflow. It provides information on the distribution of a selected feature of interest compared to other available features as well as differences in the feature values distribution between different groups within the analysis.



The Cell QC overview contains the following visualizations:

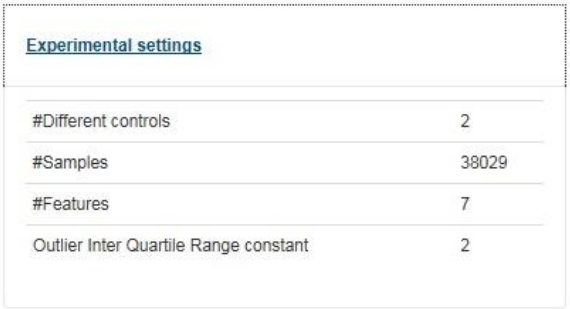
- **Box plot distribution of the feature of interest:** This panel provides a boxplot representation of the feature values divided by the categories present in the annotation selected in the user interface controls.
- **Scatterplot of the feature of interest:** This plot shows the distribution of the feature of interest in the X-axis against all the selected secondary features in the Y represented as al trellis. Color and sizing are defined by the selections made in the user interface controls.
- **Distribution:** This plot shows the distributions of the feature of interest as a bar chart where each of the categories present in the selected annotation are shown in parallel.



The html control panel allows the configuration of the visualizations using the following controls:

- **Select feature of interest:** Defines the main feature to be examined, this feature will be the one displayed in the X-axis of all the scatterplots as well as in the box plot and bar charts.
- **Select secondary features to compare:** This searchable list box allows additional features to be compared to the main feature using the scatterplot representation.
- **Select feature to set the size:** This dropdown menu allows the selection of an additional feature to use to define the size of the points in the scatterplot.
- **Select a categorical annotation to create the boxplot:** This dropdown menu allows the user to select a categorical annotation column (including the column selected for controls) to display the distribution of the data split over the different categories in the boxplot and bar charts, the scatterplots are also colored according to this selection.
- **Log scale check box:** Checking this box will change the scale of all the representations to log scale.

Additionally, the tagging controls as described in Cell / Well and Feature Tagging section are available for the exclusion of cells that appear as outliers.



The cell QC overview contains also information on the Experimental settings including the number of controls, samples and features used in the analysis.

9.15.3.6 Annotation Overview Page

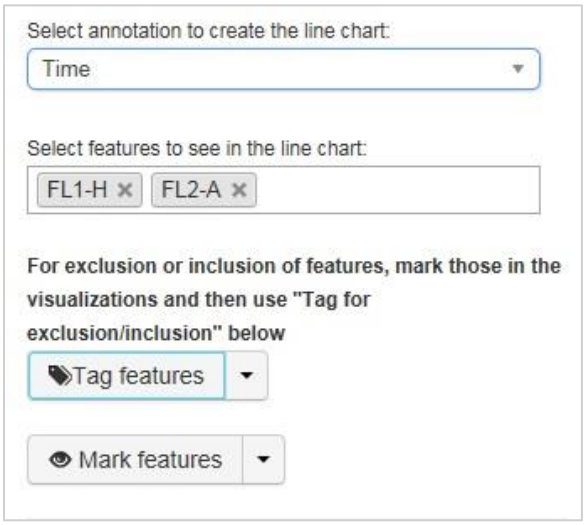
This page is only available in the Non-plate based single cell analysis workflow. It provides visualizations to explore the behavior of the different features over selected annotations. This allows the features to be plotted over time or event number in the case of data collected over a period, in order to detect any biases that may have occurred.

The Annotation overview page contains the following visualizations:

- A line chart visualization of the selected features against the annotation of interest.
- A summary table of the features used in the analysis.

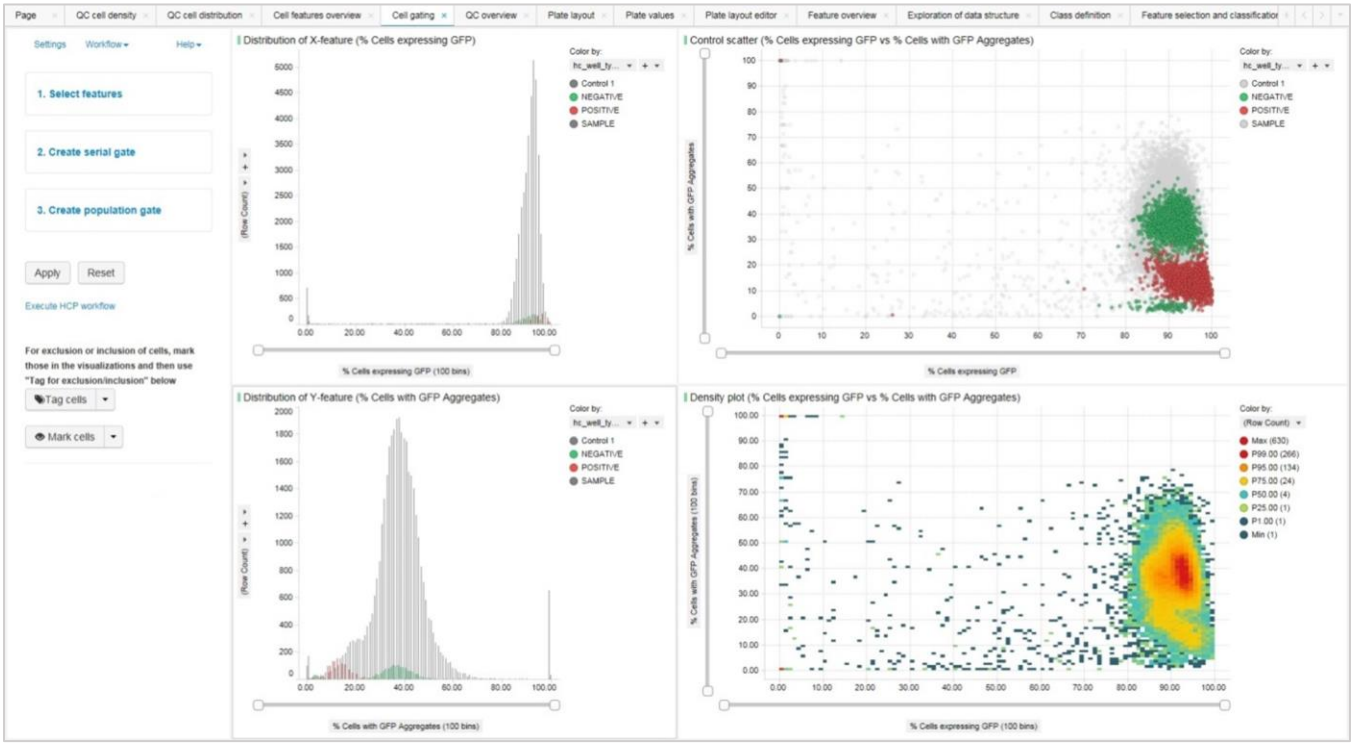
The user interface in this page contains the following controls:

- **Select annotation of interest:** This allows the selection of one of the columns that have been added as annotations to the workflow to plot the features of interest against it.
- **Select features of interest:** This allows the users to select the features to plot against the selected annotation.
- **Feature exclusion controls:** This allows the user to exclude from the analysis those features that may show some problem by selecting them in the features overview table. It is also possible to include these features back in by using the tag feature for inclusion option of the tagging dropdown menu.



9.15.3.7 Cell Gating Page

The gating page allows the user to perform different types of gating, filtering and tagging to easily select the cell population of interest for the downstream analysis as well as define different cell sub-populations.



Gating Page

The gating page visualization provides four different visualizations:

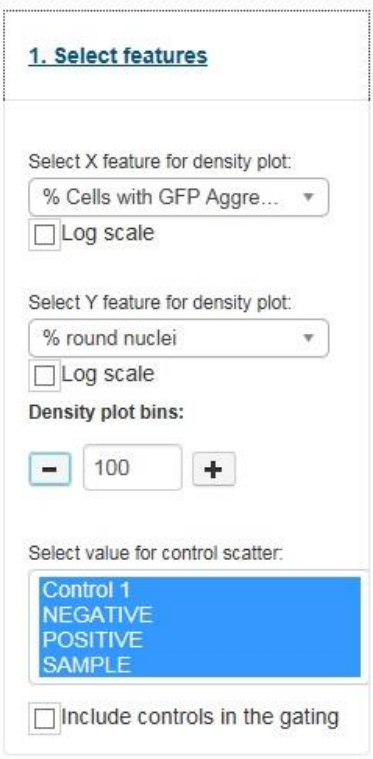
- **Distribution of X:** A bar chart representing the distribution of the feature values for the feature selected in the X-axis.
- **Distribution of Y:** A bar chart representing the distribution of the feature values for the feature selected in the Y-axis.
- **Control Scatter:** A Scatterplot representing the distribution of all datapoints on the feature space defined by the features selected in the X and Y axes. In this plot those datapoints that are defined as positive or negative controls are colored and the rest of the samples are not. This allows the user to easily identify the distribution of the control samples with respect to the overall distribution of the data for the selected pair of features.
- **Density plot:** A density plot representation of the distribution of all datapoints in the feature space defined by the selected X and Y axes.

The gating page user interface provides four different sections that allow to easily explore the data in the best way to highlight the individual cell and population characteristics and create the different gates and populations required for the downstream analysis. The different sections are:

- Select features
- Create serial gate
- Create population gate

These sections are described in more detail below. Additionally, the user also has the exclusion tagging controls that allow the user to exclude specific cells from the analysis. These controls allow the user to exclude specific samples from the analysis.

Select Features



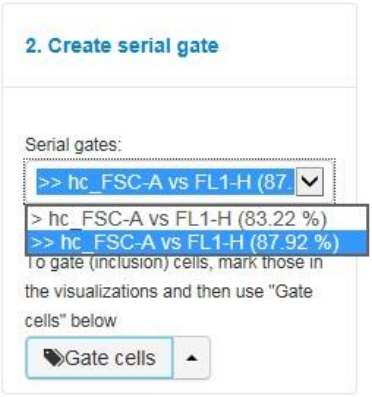
The select features user interface controls contain the following options:

- **Select X feature**: allows to select the feature to display in the X-axis and Distribution of X bar chart.
- **Select Y feature**: allows to select the feature to display in the Y-axis and distribution of Y bar chart.
- **Log scale checkboxes**: Determines the scale in which each of the axes is displayed:
    - Unchecked linear (default)
    - Checked log scale
- **Number of bins**: This allows the user to indicate the number of intervals (bins) into which the data should be grouped to generate the density plot and the bar charts. The default is 100.
- **Select groups to display**: determines what groups of samples are displayed from those present tin the controls column. The display of the different sample types as defined by the selected control column can be displayed or not to focus on different properties of the data.
- **Include controls in gating**: This checkbox determines if the controls are filtered out by the gating process or not. The default (unchecked) behavior is that the controls are not filtered out by the gating.

Create Serial Gate

Create serial gate user interface controls contain the following options:

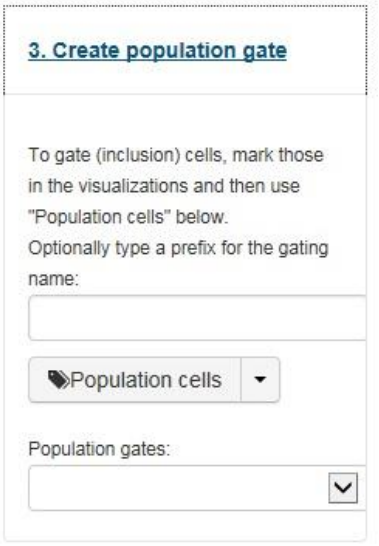
- **Serial gates**: Contains a dropdown list of the defined serial gates.
- Gate cells dropdown menu contains options to create new gates or recover previous gating stages:
    - **Add new gate**: This option will create a new serial gate. A new entry will be added to the Serial gates' dropdown menu and the selection in this menu will move to this new gate. The data will be filtered to those cells that pass the gate, except for the control samples that will be filtered or not depending on the state of the include controls in gating checkbox.
    - **Recover selected**: This option will recover the selected gate in the Serial gates' dropdown. Any later gates that had been defined will be removed as well as any population gates that may have been defined in the next section.



The gates that have been defined will be displayed in the serial gates' dropdown menu. The information displayed in these gates is the order, represented by the number of ">" before the name. The parameters used to create the gate and the fraction of cells that passed the gate calculated over the number of cells present before the gating steps.

This means that for the first gate this is the fraction of total cells in the experiment that pass the gate, and for the second and successive gates it is the fraction of cells that passed the previous gate and passed the one selected.

Create Population Gate



The create population gate user interface controls contains the following options:

- **Prefix text box**: This text box allows the user to add a prefix to the population gate generated. The default gate name generated will be a combination of the features used to define it.
- **Population cells dropdown menu**: This control is used to define different regions in the data corresponding to different populations. In this case the data is not filtered, but the samples are assigned to the corresponding population. This makes it easy to identify the data from a specific cell population and to obtain statistics on these different populations.
    - Add new population creates a new population gate with the selected samples and adds the corresponding entry to the population gates dropdown menu.
    - Remove selected removes the population gate selected in the Population gates menu.
- **Population gates dropdown**: This dropdown contains the list of defined population gates. Selecting any of the gates in the list will mark the samples belonging to this population in the visualization.

The naming of the population gates follows the same logic as that of the serial gates, although in this case there is only one level, so it is not necessary to indicate if there is the first or second gate. Also, population gates can only be defined over the last level of serial gating, and a warning will be displayed if population gates have been defined and a new level of serial gating is attempted, as this will erase all defined population gates.

When a population gate is selected in the population gates dropdown menu, the cells included in this gate are marked in the visualization.

Additional Gating Options

This section provides the user with the necessary controls to re-execute the workflow on the gated population.

- **Apply:** Clicking on this button will execute the analysis on the gates defined, meaning a new column will be created for each of the gates defined in the gating process.
- **Reset:** Clicking this button will reset all defined gates, both serial gates and population gates will be removed.

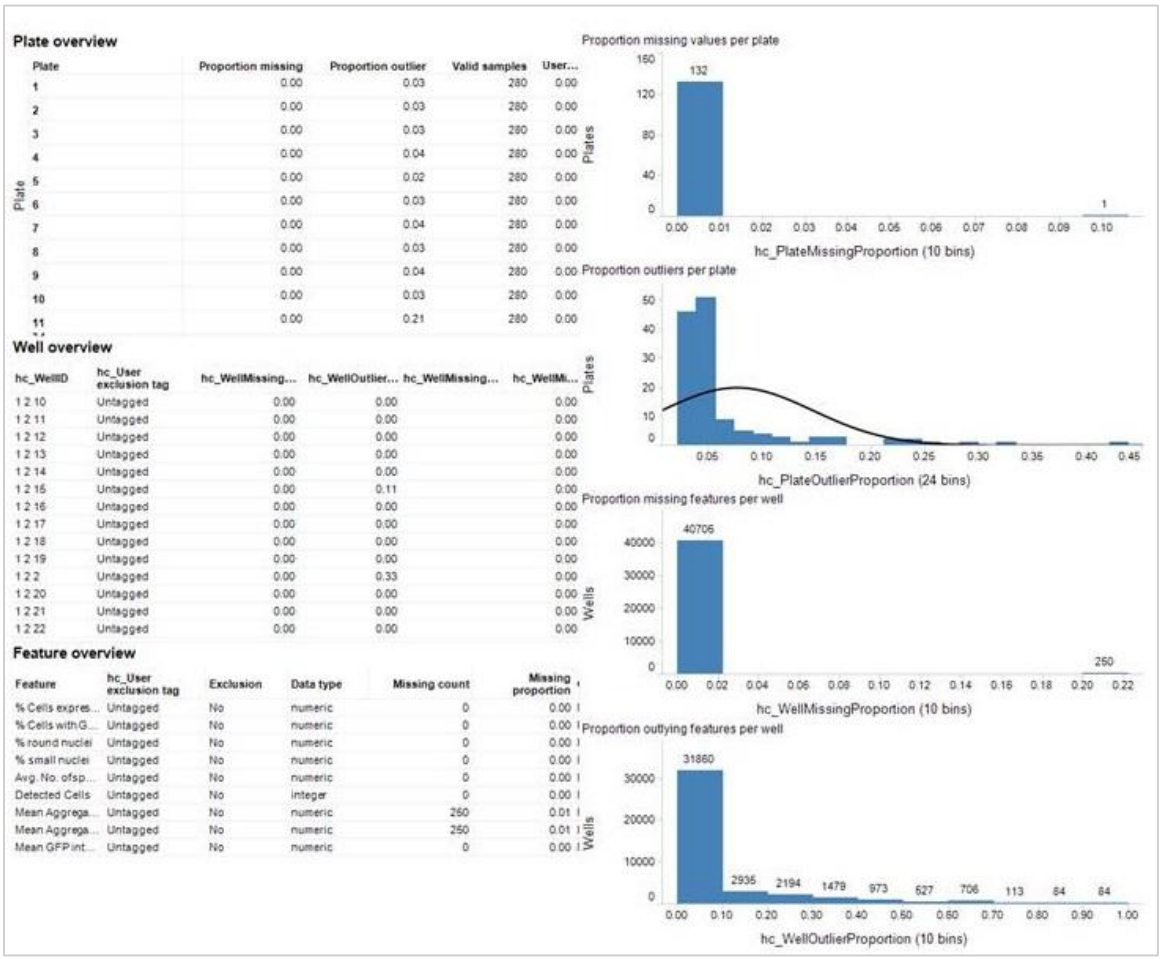
- **Execute HCP Workflow:** Clicking on this link will open the main workflow dialog using as input the result from the gating after aggregation.
- **Tagging controls**: Allow for the control of the inclusion and exclusion of specific cells.

9.15.3.8 QC Overview Page

This Quality Control summary page is present in all plate-based data analyses. It displays the most important characteristics of the loaded data and constitutes a starting point to drill down into individual plates or wells. The automated QC visualizations generated by the High Content Profiler are aimed at speeding up the identification of problematic plates or wells in an experiment and pointing the scientists to the plates and wells that must be checked. These controls allow checking for batch effects or biases that would otherwise be difficult to identify.

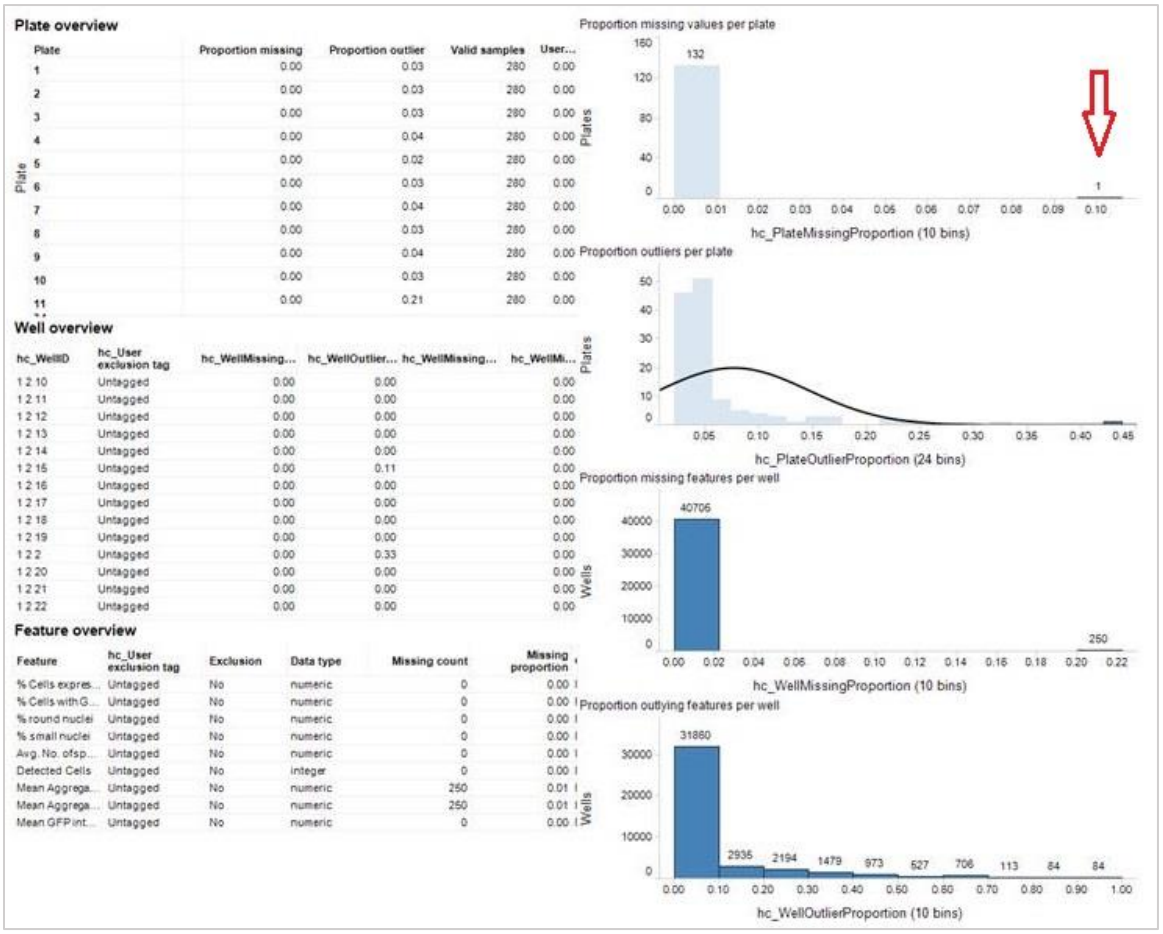
The QC overview pages answer the following basic questions at the level of Plates, Wells and Features:

- Are there missing values? If yes, then where are they located?
- Are there outliers? If yes, which plates, wells and features are outliers?

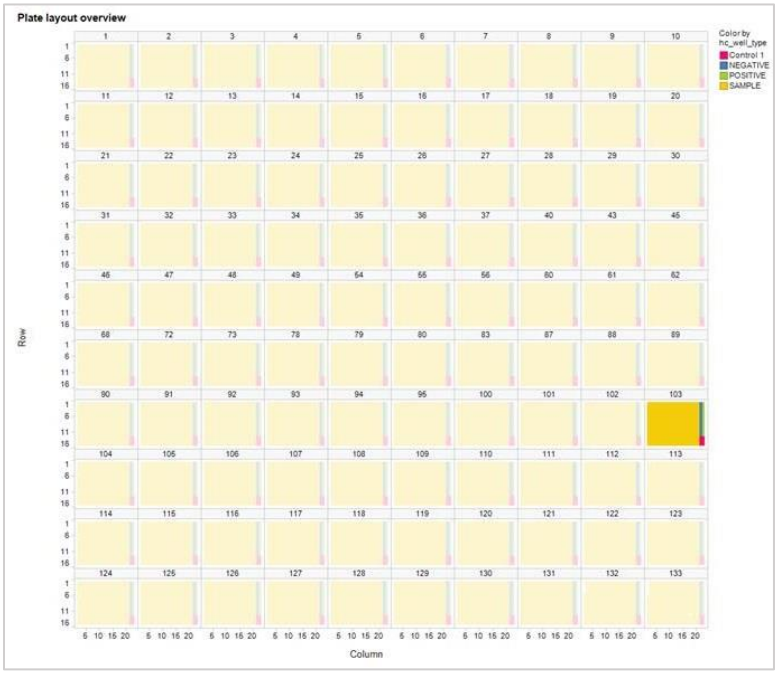


The QC overview page is a starting point to drill down into the QC issues in an experiment at the well level. For instance, by selecting the plate with the highest proportion of outliers in the summary bar chart (refer to the example

below with the plate highlighted with the red arrow); the selection will be cascaded to all the visualizations of the same page, as well as the relevant visualizations in the following pages:



This allows to rapidly focus on the HCS plate where irregular events are located. In this example, the plate with the highest proportion of outliers will be immediately ready for inspection in the Plate layout page:



QC Overview Navigation Panel

The navigation panel contains the following specific elements:

- The Experimental settings box shows relevant information regarding the imported data set and selected QC relevant parameters. Apart from the number and layout of plates, number of features, number of controls, number of samples it also reports threshold used for flagging missing wells and defining outliers.

- The lower part contains dropdown menus for adding, removing and resetting exclusion tags to wells and plates or to features as well as for visualizing those wells that are excluded. Tagged wells/plates or features will not be used in analysis. In addition, the “Filter out excluded wells” checkbox will filter out all the wells that are tagged for exclusion from the visualizations.

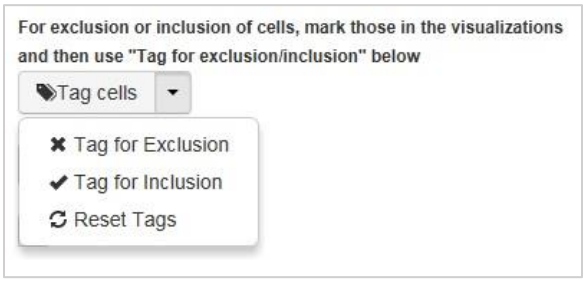
To tag wells:

1. Mark wells or complete plates in any of the visualization of QC related result page
2. Click on the Tag wells dropdown menu and select tag for exclusion to exclude the selected wells, tag for inclusion to remove the exclusion flag from previously excluded flags and reincorporate them to the analysis or reset tags to remove all exclusion flags. Detailed information on well and feature tagging is provided below.

9.15.3.9 Cell / Well and Feature Tagging

Cell Tagging

The **Tag cells** dropdown menu is available in the QC cell density, QC cell distribution, Cell gating and Cell QC overview pages. This menu contains tools that allow the user to exclude certain cells from the analysis based on the information provided in the different tables and visualizations. Whenever these tags are used, on one or more cells, a refresh icon will appear in each of the tables and visualizations that would require updating to reflect the modified results after the changes in the wells included in the analysis.

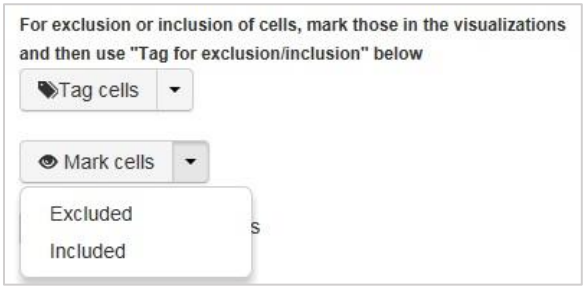


The options available in this menu are:

- **Tag for Exclusion:** Choosing this option will mark the tagged cell for exclusion, meaning that the cell will not be used in the analysis when it is refreshed.
- **Tag for Inclusion:** This option will include a previously excluded cell.
- **Reset Tags:** This option will reset the tagging to the original settings where all cells are included in the analysis.

These options allow the user to experiment with the analysis results obtained after excluding different sets of cells that may be outliers or behave in a different manner form most of the other cells.

Mark Cells



The **Mark cell** option will allow the user to highlight those wells that have been tagged for exclusion or those that will be included in the analysis using the corresponding options as shown in the figure above. To use this effectively the “Filter out excluded cells” checkbox should be unchecked, because if this checkbox is checked, the entries tagged for exclusion will not be visible.

Well Tagging

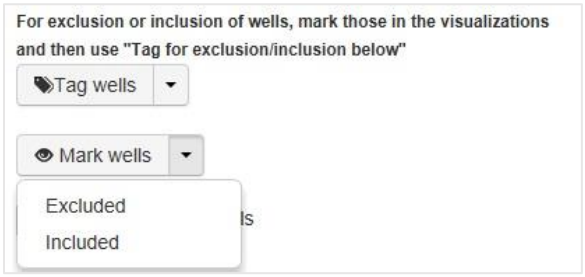
The **Tag wells** dropdown menu is available in the QC overview, Plate layout, Feature overview and Dose response pages. This menu contains tools that allow the user to exclude certain wells from the analysis based on the information provided in the different tables and visualizations. Whenever these tags are used, on one or more wells, a refresh icon will appear in each of the tables and visualizations that would require updating to reflect the modified results after the changes in the wells included in the analysis.

The options available in this menu are:

- **Tag for Exclusion:** Choosing this option will mark the tagged well for exclusion, meaning that the well will not be used in the analysis when it is refreshed.
- **Tag for Inclusion:** This option will include a previously excluded well.
- **Reset Tags:** This option will reset the tagging to the original settings where all wells are included in the analysis.

These options allow the user to experiment with the analysis results obtained after excluding different sets of wells that may be outliers or behave in a different manner form most of the other wells.

Mark Wells



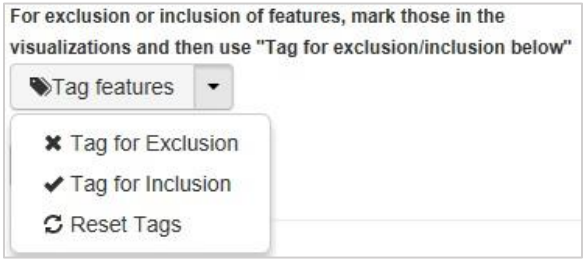
The **Mark well** option will allow the user to highlight those wells that have been tagged for exclusion or those that will be included in the analysis using the corresponding options as shown in the figure above. To use this effectively the “Filter out excluded wells” checkbox should be unchecked, because if this checkbox is checked, the entries tagged for exclusion will not be visible.

Mark, Tag and Exclude Data Points

These controls are equivalent to the Marking and tagging controls described above. They are present in the Exploration of data structure page and will affect wells or cells depending on the type of analysis that has been performed (plate-based or non-plate based).

Feature Tagging

The Tag features dropdown menu is available in the QC overview page and the Annotation overview page (Non-plate-based analysis). As in the case of the Tag wells described above, this menu contains tools that allow the user to exclude features from the analysis based on the information from the different tables and visualizations. Whenever these tags are used, on one or more features a refresh icon will appear in each of the tables and visualizations that would require updating to reflect the changes in the features included in the analysis.

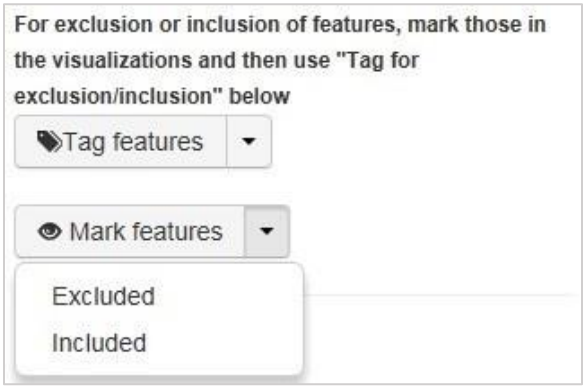


The options available in this menu are:

- **Tag for Exclusion:** Choosing this option will tag the selected features for exclusion, meaning they will not be used in the analysis when it is refreshed.
- **Tag for Inclusion:** This option will include a previously excluded feature.
- **Reset Tags:** This option will reset the tagging to the original settings where all features are included in the analysis.

These options allow the user to experiment with the analysis results obtained after excluding different sets of features that may be outliers or behave in a different manner form most of the other wells.

Mark Features



The **Mark features** option will allow the user to highlight those features that have been tagged for exclusion or those that will be included in the analysis using the corresponding options as shown in the example above.

QC Overview Visualizations

Visualizations on the QC overview page provide an overview on data state, completeness and possible outliers. This overview is extremely valuable in large experiments, since it allows to quickly point out issues in very large experiments and, subsequently, to use the following pages to drill down and focus on each of the events to control.

The visualizations are available as:

- Tables - allows reordering the information by clicking on the columns' headers.
- Bar Charts - to easily select extreme event.

Marking elements allows to inspect the same element on the following QC related pages or to tag them for exclusion.

Plate Overview

Displays a table with plate-centric quality metrics

Column	Description
Plate	Plate identifier.
Proportion missing	Proportion of missing values on that plate, across all features and wells on that plate.
Proportion outlier	Proportion of feature values flagged as outlier in the corresponding feature, across all features and wells on that plate.
Valid samples	Number of samples on that plate which have at least one non-missing value in any of the features.
User excluded wells count	Number of wells on that plate tagged for exclusion.

Well Overview

Displays a table with well centric quality metrics.

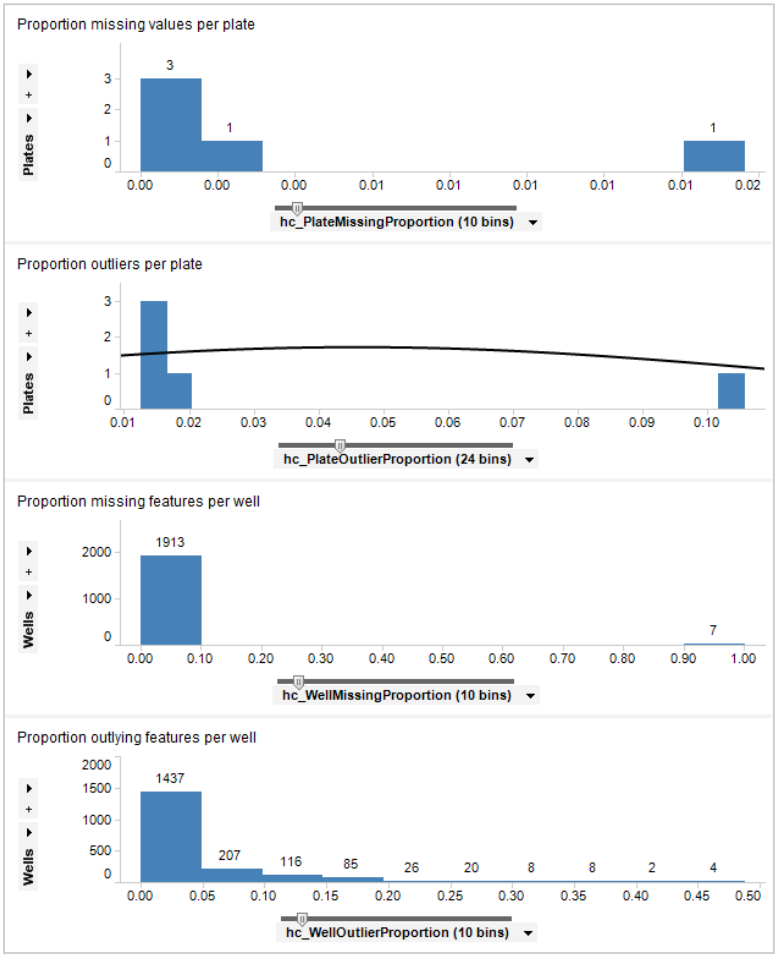
Column	Description
Well ID	Well identifier. Combination of Plate identifier + Row identifier + Column identifier.
User exclusion tag	Specifies that the selected well is to be excluded from analysis. 'Untagged' wells will be included in the analysis. Use the corresponding buttons in the navigation panel on the left-hand side to change the tags for marked wells.
Well Missing Proportion	Proportion of features in that well with missing value.
Well Outlier Proportion	Proportion of feature values that are considered outliers in that feature when applying the threshold as configured in the Analysis settings dialog.
Well Missing flag	Flag will be True if the percentage of missing values in that well is above the threshold defined in the Analysis settings dialog and False otherwise.
Well Missing Count	Number of features in that well with missing value.

Feature Overview

Displays a table with feature-centric quality metrics.

Column	Description
Feature	Feature name as defined in imported table.
User exclusion tag	Specifies that the selected feature is to be excluded from analysis. 'Untagged' features will be included in the analysis. Use the corresponding buttons in the navigation panel on the left-hand side to change the tags for marked features.
Exclusion	Flag indicating if that well will be excluded from further analysis based on non-numeric data type, complete absence of valid values or constant value.
Data type	Data type as inferred from values found for that feature. Data types supported for analysis are numeric and integer columns. All other data types will be treated as missing values and ignored in analysis.
Missing count	Number of wells with missing values in that feature.
Missing proportion	Proportion of wells with missing values in that feature.
Constant	Flag indicating if the values of that feature are essentially constant or not.

Bar charts for proportion of missing values and outliers



Bar charts on the right-hand side of the page show (top to bottom) frequency distributions of:

- Proportion of missing values per plate
- Proportion of outliers per plate
- Proportion of missing features per well and
- Proportion of outlying features per well

Those visualizations can be used to identify plates or wells with values on the far tail of the distribution which could indicate wells or plates with extreme values that should be inspected.

9.15.3.10 Plate Layout Page

Displays information on the plates and well annotations. The navigation column on the left-hand side allows a set of actions defining the overview through:

- **Annotations dropdown** - Allows selecting and coloring each plate based on the annotations associated to the wells.
- **Filter to Plate -** Allows for the selection of the plates that will be displayed in the page, ranging from one plate to all plates in the analysis (100 plates are displayed in the example figure below). If more than 100

plates are displayed, they will be split into pages with 100 plates each to maintain the individual plate visualizations of a sufficient size.

9.15.3.11 Plate Values Page

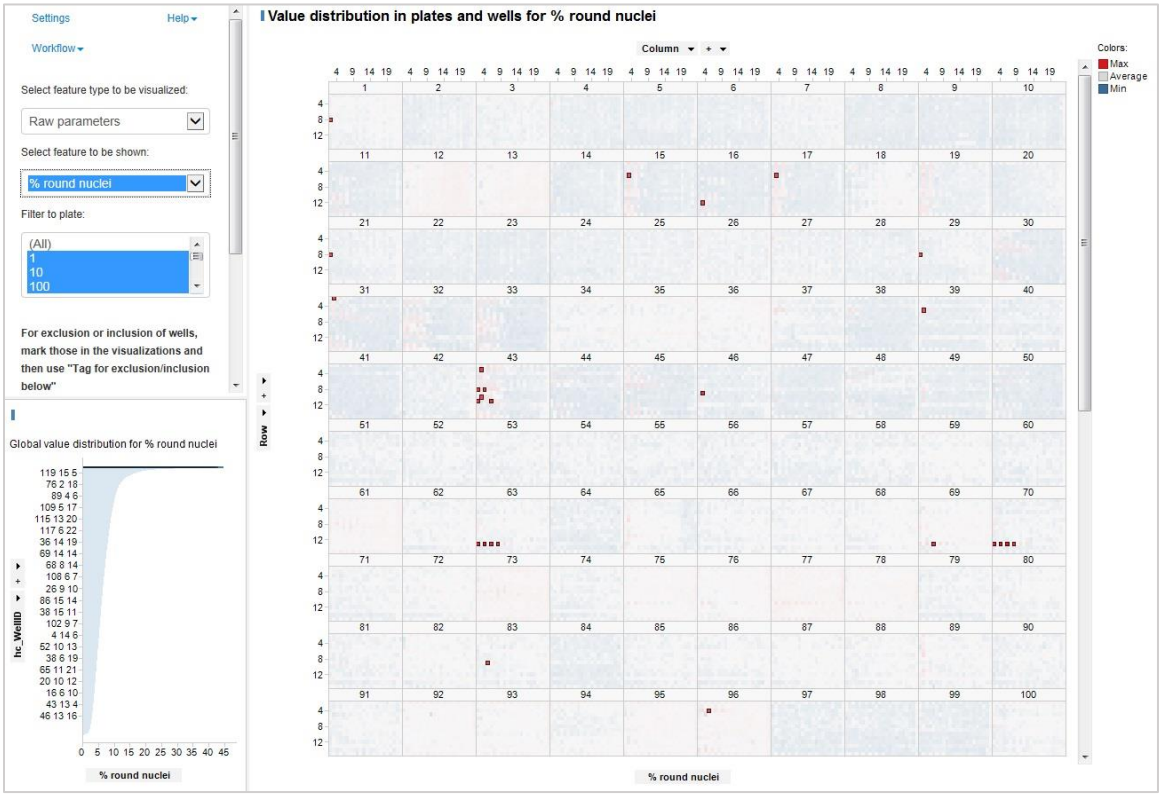
This page allows the exploration of the distribution in the plates of each of the HCS features extracted. The navigation column on the left-hand side of the page allows a set of actions defining the overview through:

- **Parameter selectors**

Allows the selection and coloring of each plate based on a given feature (raw, normalized or Z-scored).

- **Distribution histogram**

Allows the marking and coloring of the plates containing the selected feature in a given range of values as selected by the user (in the example below the 28 highest values for the selected parameter are shown in red).

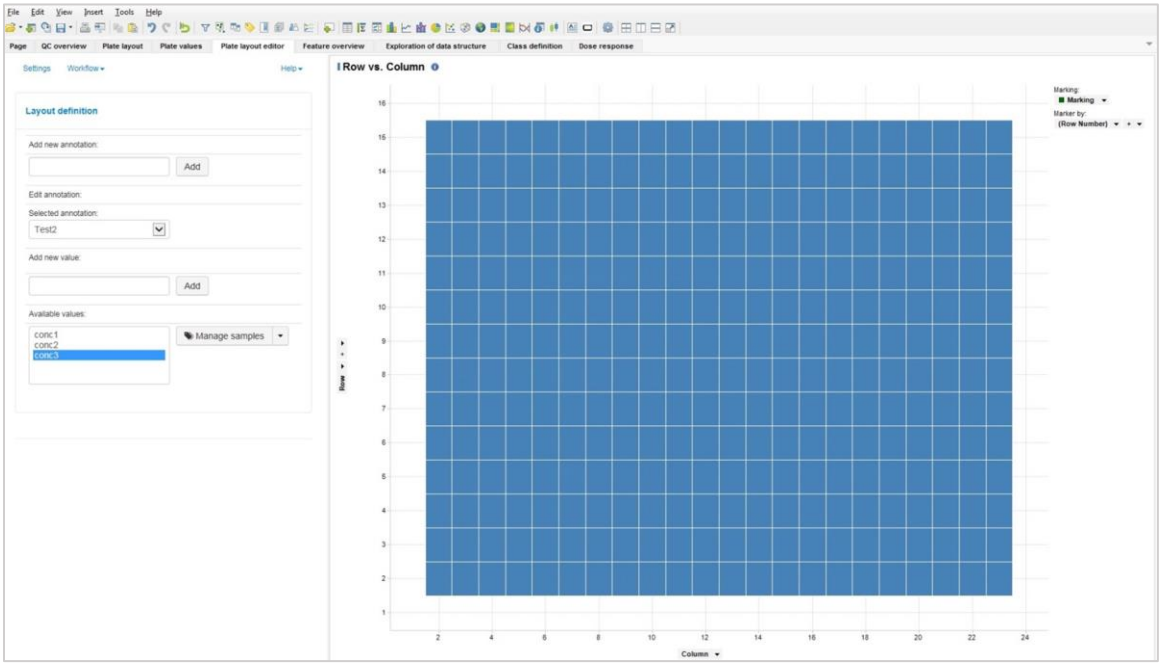


The selections and markings performed on this page are automatically reflected in the other pages of the project.

9.15.3.12 Plate Layout Editor

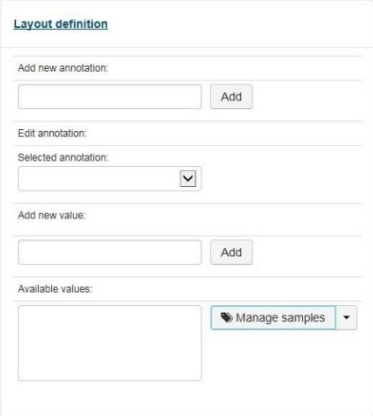
The plate layout editor allows the user to add new annotations to the wells in plate-based analyses.

The page provides a representation of the plate layout where groups of wells can be selected by the user to provide annotations for them using the controls available in the user interface.



The user interface contains the following controls:

- **Add new annotation:** This option allows to introduce a name for a new annotation column that will be added to the workflow by clicking on Add.
- **Edit annotation:** This section contains several parts.
    - **Selected annotation:** This option provides a dropdown menu that allows to select between the different annotations that have been defined.
    - **Add new value:** A textbox control that allows for the creation of new groups within the selected annotation by introducing the name of the group in the box and clicking on Add.
    - **Available values:** A textbox that allows the selection of the different annotation values we want to assign to the selected wells. This assignment is done by using the manage samples dropdown menu.
    - **Manage samples:** A dropdown menu that facilitates the managing of the classes defined in the current classification:
        - **Assign:** Assigns currently selected annotation value to the selected wells.
        - **Remove:** Removes the currently selected annotation value from the selected wells.
        - **Mark:** Marks the wells that have tagged with the currently selected annotation.



9.15.3.13 Feature Overview Page

This page allows the exploration of the distribution across plates of each of the HCS features extracted. The navigation column on the left-hand side of the page allows a set of actions defining the overview through:

- **Feature type selector**

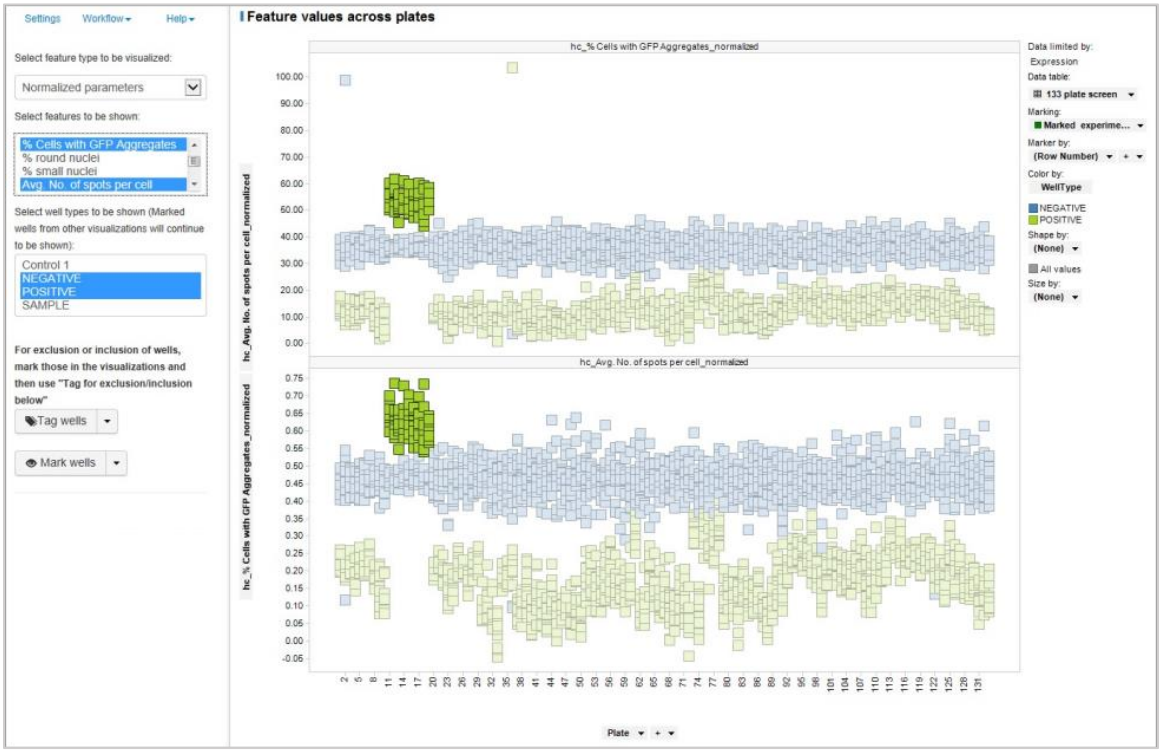
Allows the selection of a feature type to view from the different feature types available (raw, normalized or Z-scored).

- **Feature filter selector**

Allows the specification of which features of the selected type will be plotted across the plates.

- **Well type selector**

Allows the visualization of the features for specific well types across the plates (e.g. in the example below, the positive control appears to contain two subpopulations for a given pair of features).



This page supports the discovery of plate layout effect or batch effects. In the example shown above, the shapes of the data points reveal for the same positive control differences that could be a strong batch effect or sub-populations of positive control depending on the positive control being located on different plates.

9.15.3.14 Exploration of Data Structure Page

This page allows the exploration of the dataset through a 3D Principal Component Analysis (PCA) visualization complemented by a set of visualization options:

- **Switch scatterplot 2D|3D**

When at least three features are used this option allows to switch the visualization between a 3D scatterplot (default) and a 2D scatterplot.

- **Loadings Plot**

Provides a visualization of the Loadings information of the PCA.

- **Explained variance**

Provides a visualization of the Variance information explained by each of the PCA components.

- **Identify Batch effects:**

Provides a visualization of the PCA calculated on the plate controls, which can reveal batch effects between plates.

- **Color well by:**

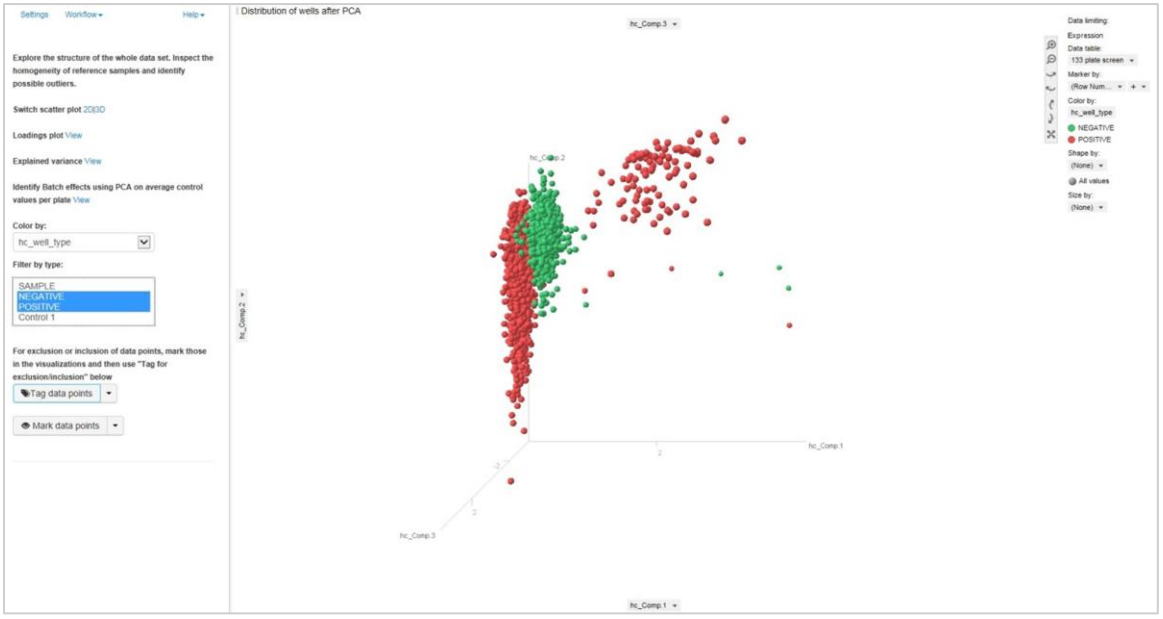
Allows you to choose the data coloring scheme (e.g. using the type of sample: whether it is a control or the test population).

- **Filter by well type:**

Allows you to focus on the class to visualize (e.g. only the positive control, as shown in the example below, where the location on certain plates of the positive controls determines their distribution on the PCA).

- **Tagging controls:**

Provides access to the tagging controls described earlier that will allow the user to exclude problematic data.



This page helps explore data distribution relative to controls or subgroups of the data and take decisions about the analysis accordingly.

Note: The sign assignment in the PCA is arbitrary, which means that with the same data it is possible to obtain equivalent PCA solutions where the component signs are inverted. This particularity of certain types of modern data analysis methods which include PCA is known and has no influence on the results of the analyses performed in the High Content Profiler. However, it is worth mentioning here, as this type of sign flip may often occur after the removal of outliers and re-execution of the workflow. This procedure will be done frequently in the initial exploratory analysis, and the sign change could be surprising for those analysts who are not familiar with the arbitrariness of the sign solution. More information on sign ambiguity on this type of methods can be found in (E. Rasmus Bro and T. Kolda. Resolving the sign ambiguity in the singular value decomposition Journal of Chemometrics, 22(2):135-140, 2008).

Note: In those cases where the number of features used in the analysis is less than three in place of the 3D scatterplot, a 2D scatterplot representation will be used for representing analyses done over two different features and a bar plot representation of the distribution of the features over the different categories of the annotation selected for coloring will be displayed when a single feature is used.

9.15.3.15 Class Definition

The main goal of this page is to provide the capacity to identify different groups of samples from the data that may have different behaviors. This can identify subgroups within the data or different classes of response within the analysis.

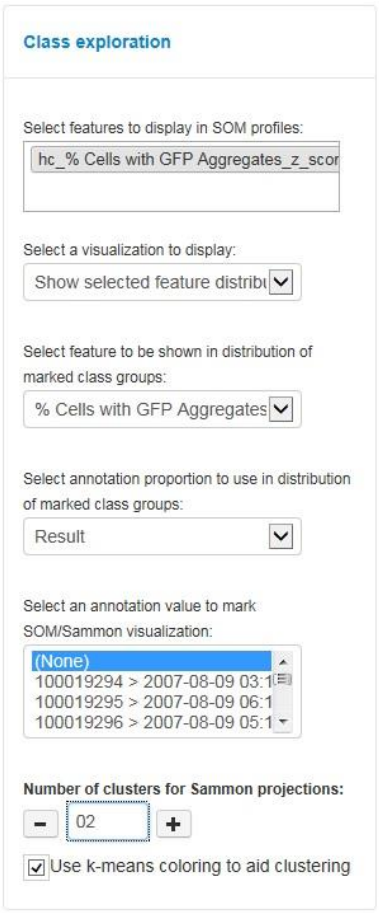
To identify these classes a self-organizing map (SOM) is used that will perform an unsupervised clustering of the data into a predefined number of cells this number of cells can be configured by the user in the advanced settings panel. From this initial classification the user can select the cells showing similar profiles to create the groups of interest using the class definition assignment tool provided in the user interface. This tool allows the creation of new groups and classes and the assignment of the different samples to them.

Note: The SOM execution is not deterministic; this means that two consecutive re-executions will not necessarily provide the same groupings, although if the groups existing in the data are sufficiently distinct, they will be similar. The user interface provides two sections with controls for Class exploration and controls for defining new classes (Class discovery). Once the new classes are defined, these can be used within the workflow, as the column type for example, by re-executing the workflow and selecting this new column.

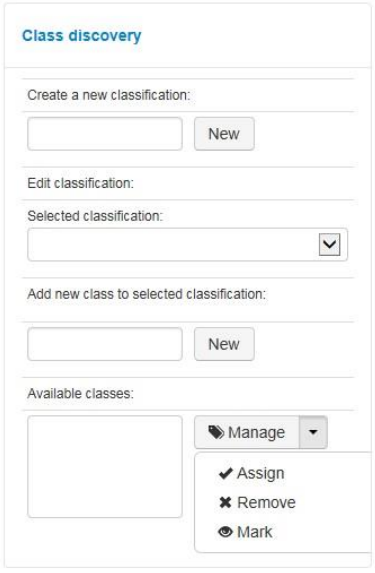
Class Exploration

The controls available in the class exploration user interface section are:

- Select features to display in SOM profiles: This allows for the selection of the features that will be displayed in the X-axis of the selected profile visualization as well as in the SOM grouping visualization. This control can help select the features that provide the most relevant profile for differentiating between classes.
- Select a visualization to display: Allows to select between the following drill down visualizations:
    - Distribution of the selected feature: This will show the overall distribution of the selected feature as a bar chart with the distribution of the selected data represented in parallel.
    - Show selected annotation proportion: Displays the proportion of the selected annotation categories as a bar chart with the distribution of the selected data represented in parallel.
    - Show selected feature across selected annotation: Displays a line chart representing the profile of the selected feature over the selected annotation category against the profile over the same feature of the selected data from the annotation.
- Select feature to be shown in distribution of marked class groups: Defines the feature to be used in the distribution of marked class groups.
- Select annotation proportion to use in distribution of marked class groups: Defines the annotation to be used to set the proportions to be used in the distribution of marked class groups. This dropdown will also include the control column when available.
- Select an annotation value to mark SOM/Sammon visualization: Allows to easily mark the selected annotation categories in the SOM and Sammon visualizations.
- Number of clusters for Sammon projections: This option allows defining the number of clusters that are expected using a number between 2 and 10, and the profiles for the OM will be clustered using K-means clustering with the specified settings to identify a clustering with this number of groups.
- Use k-means coloring to aid clustering: Checking this checkbox will color the points in the Sammon according to the k-means group to which they have been assigned. This may aid in the clustering definition.

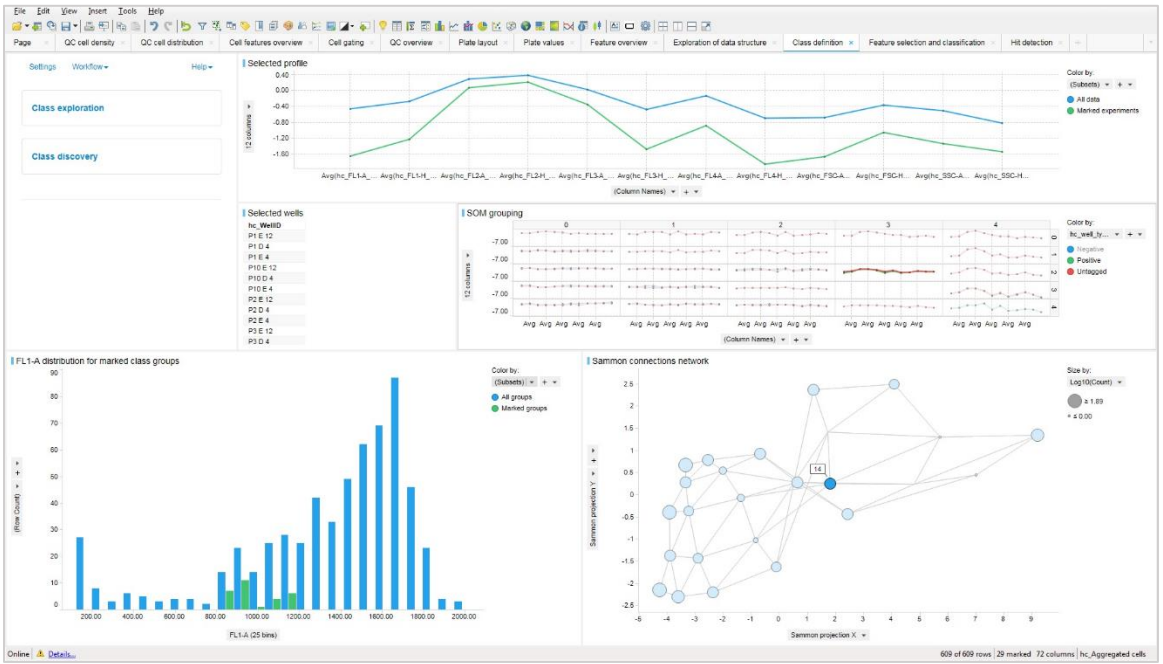


Class Discovery Panel



The class discovery html control panel contains the necessary controls to define the classes.

- **Create a new classification:** This option allows introducing a name for a new classification that will be added to the workflow by clicking on new.
- **Edit classification:** This section contains several parts.
    - **Selected classification:** This option provides a dropdown menu that allows to select between the different classifications that have been defined.
    - **Add new class to selected classification:** A textbox control that allows for the creation of new groups within the selected classification by introducing the name of the group in the box and clicking on new.
    - **Available classes:** A textbox that allows the selection of the different classes to which we want to assign the selected nodes by clicking assign, or which we want to highlight in the visualization by clicking on Mark.
    - **Manage:** A dropdown menu that facilitates the managing of the classes defined in the current classification:
        - **Assign:** Assigns the marked data to the currently selected class.
        - **Remove:** Removes the marked data from the currently selected class.
        - **Mark:** Marks the data belonging to the currently selected class in the visualizations.



The Class definition page contains the following visualizations:

- **Selected profile:** The selected profile visualization will display the overall profile of the samples across all features against the profile across all features of the selected samples. The wells containing the selected samples will be displayed also in the selected wells table.
- **SOM grouping:** Provides a visual representation of the cells used in the self-organizing map with each of the trellis panels representing one cell of the SOM. The obtained groups are represented on the SOM grouping visualization where the average profile of each well across the features can be observed

represented on the SOM grid and grouped by sample class if available. If a cell in the SOM grid has no sample assigned to it, it will appear empty in this visualization and will show no samples in the Sammon connections visualization either.

- **Details visualization:** The plot represented here depends on the selection chosen in the “Select a visualization to display” dropdown menu. The three available options are:
    - **Distribution of the selected feature:** This will show the overall distribution of the selected feature as a bar chart with the distribution of the selected data represented in parallel.
    - **Show selected annotation proportion:** Displays the proportion of the selected annotation categories as a bar chart with the distribution of the selected data represented in parallel.
    - **Show selected feature across selected annotation:** Displays a line chart representing the profile of the selected feature over the selected annotation category against the profile over the same feature of the selected data from the annotation.
- **Sammon connections network:** This visualization provides a visual aid to the interpretation of the SOM by displaying those SOM nodes that are more similar as closer in the Sammon bidimensional representation, as well as representing the structure of the network by showing the neighbors of each node in the SOM grid by joining them with a line. The number of samples assigned to each of the node is represented by the size of the node.

Note: When performing changes in the analysis that require the re-execution of some of the step, ensure the Sammon representation is also updated by clicking on the update icon displayed on the plot when it requires an update.

- **Selected wells table:** Provides a list of the wells in the selected data.

9.15.3.16 Univariate Hit Detection

This page provides the user with the visualizations needed for the analysis and selection of hits when a single feature has been selected for the analysis. In this case there is no point in performing the feature selection, as there is only one feature, and the common measurements used in high throughput screening can be used with no need of applying more complex methods that take into account the interactions between different features.

To best perform the analysis in these cases the univariate hit detection page contains all the necessary tools to identify those plates where the separation between controls is best as well as the necessary functions to remove outliers and identify the best hits from the remaining samples.

To perform the sample selection, the user can determine if the quality of the plates is sufficient by using different measurements that evaluate how good the separation between the positive and negative controls in each plate is. The measurements available for this are:

- Signal to background
- SSMD

$$\hat{\beta} = \frac{\bar{X}_P - \bar{X}_N}{\sqrt{s_P^2 + s_N^2}}.$$

Where X_P and X_N are the average of signal of the positive and negative controls in the plate respectively and S_P and S_N the Standard deviations.

- Z-prime
- Z-prime robust. This is calculated the same way as Z-prime, but using the median and the MAD instead of the mean and the standard deviation

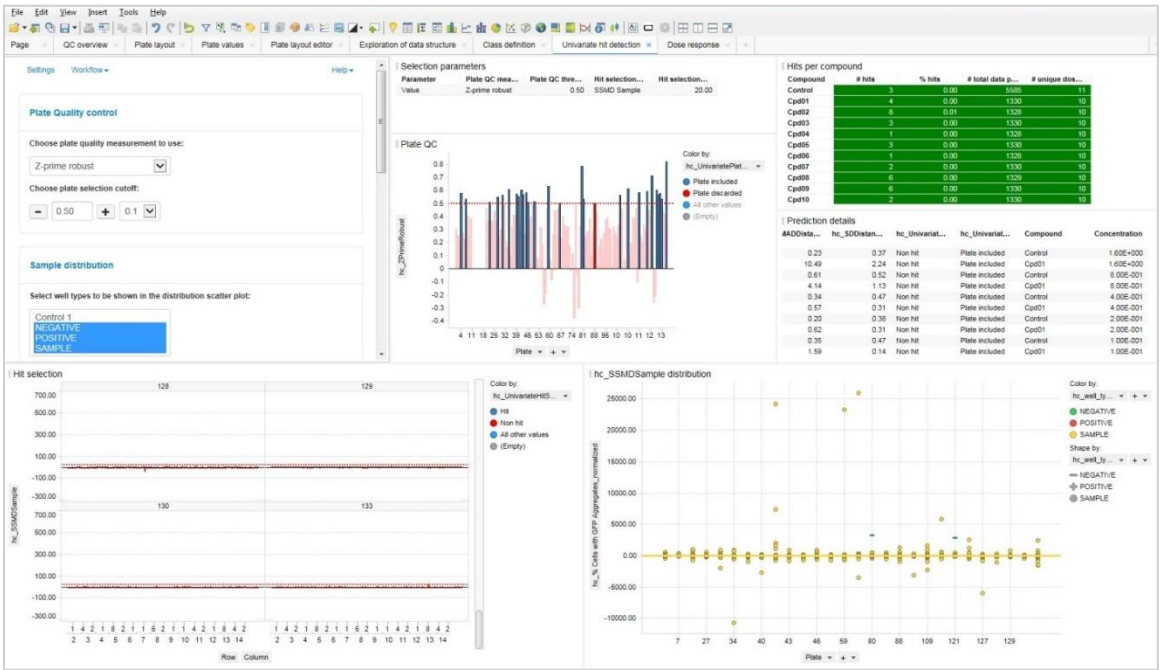
Once the plates of interest are selected, determining if the samples are hits are not can be achieved using several different measurements:

- Raw values of the feature
- Normalized feature values, this normalization is according to the selected normalization method
- Z-scored normalized parameter. Uses the normalized parameter but scaled using Z-score
- SSMD: The formula used for calculating the SSMD at the sample level is:

$$\mathrm{SSMD} = \frac{X_i - \bar{X}_N}{s_N \sqrt{2(n_N - 1)/K}},$$

Where Xi is the measurement for the sample n is the number of negative controls and K is nN-2.48.

- Fold Change when compared to the negative control
- The number of Standard deviations away from the mean of the negative controls
- The number of MADs away from the median of the negative controls



This page contains the following visualizations:

- **Selection parameters table:** This table contains the selected plate QC measurement and the threshold set, as well as the hit selection measurement with the numeric threshold to which it has been set.

- **Plate QC bar chart:** This visualization will contain for each of the plates the value of the calculated per plate statistic that has been selected for the plate QC. Depending on the user defined threshold, the plate bars will be colored in red (below the threshold) or blue (above the threshold).The selected QC measurement will also determine the range of optimal values.
- **Feature distribution scatterplot:** The samples belonging to the selected plates in the Plate QC chart are displayed here, allowing the user to identify possible outliers, as well as examine the distribution of the different types of samples contained in the column selected for control. The central value (mean or median) as well as the dispersion can also be overlaid on this plot for each of these categories using the available controls.
- **Hit selection bar chart:** For each of the selected plates, the per sample value of the selected measurement to use for Hit Selection is displayed. Colored also according to the threshold that has been set.
- **Hits per gene/compound table:** Information on the hits will be aggregated by Gene or Compound, if that information is available and the corresponding workflow type is selected. If not, the table will not be present. This table contains the gene/compound details including the number of hits if more than one replicate is available.

Column	Description
Gene/Compound	Gene or Compound identifier, depending on selected workflow type.
# hits	Number of hits found for this gene or compound.
% hits	Proportion of hits among all replicate wells available for this compound.
Med projection score [positive class name]	Median projection score across all wells available for this compound. If there is more than one positive control class, there will be one column for each of them.
# total RNAi / concentrations	Total number of wells for this compound.
# unique RNAi/concentrations	Number of unique RNAi or concentration levels found for this gene or compound.

- **Prediction details:** This table will contain the individual sample details, for the selected compounds/genes if available, including the values for the different statistics used to determine if it is a hit and the classification according to the selected measurements and threshold. The available measurements in this table are:

Column	Description
WellID	Well identifier, Plate ID + Row ID + Column ID.

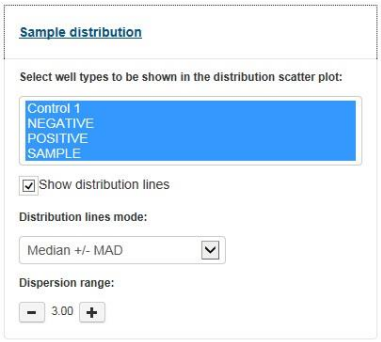
Gene/Compound	Identifier of gene or compound, if available.
RNAi	Identifier of RNAi if available.
Concentration	Concentration level, if available.
Type	Well type as specified by the user.
Predicted class	Class assignation as determined by the classification model.
Projection score [positive]	Projection score for this well. If there is more than one positive control class, there will be one column for each of them.
SSMD	SSMD measurement for the sample. (Strictly standardized mean difference).
Fold change	Fold change with respect to the negative control.
SD away from neg means	Number of Standard deviations away from the negative control.
MADs away from negative median	Number of MADs away from the positive control.

The user interface contains four different sections as described below:

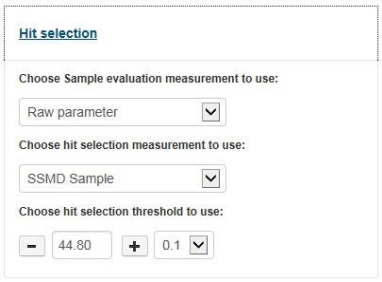
- **Plate quality control:**
  - **Choose quality measure to use**: Allows the selection of a plate level quality measurement to select those plates where the separation between the negative and positive controls is adequate.
  - **Choose cutoff**: Allows the user to set the threshold for the selected measurement to use to select those plates of interest.

- **Sample distribution:**
  - **Select well type to be shown in the distribution scatterplot:** Allows the user to select which sample types will be visualized from those types present in the control column selected for the analysis.

- **Show distribution lines checkbox:** When checked (default) the distribution lines are displayed over the scatterplot distribution.
- **Distribution lines mode:** Allows the selection of the distribution lines to display, either mean and StdDev or Median and MAD.
- **Dispersion range:** Allows the user to select the distance in number of StdDev or MAD away from the Mean/Median where the dispersion lines will be plotted in the scatterplot.



- Hit selection
  - **Choose sample evaluation measurement to use:** This selector determines the measurement displayed in the distribution scatterplot.
  - **Choose hit selection measurement to use:** This allows the user to select between the different hit selection measurements available.
  - **Choose hit selection threshold:** Allows the user to determine the threshold to be used to determine if a sample is a hit.



- **Tagging controls:** These controls perform the same functionality as has been described previously allowing the user to exclude specific wells from the analysis if they are considered outliers.

9.15.3.17 Feature Selection

The main goal of this section of **High Content Profiler** is to provide the capacity to select the most important set of features that can be used to separate the distinct classes or phenotypes. In this section, the capability of the High Content Profiler to support feature selection and hit classification is described.

As many data analysis techniques were originally not designed to cope with large amounts of irrelevant features, feature selection has become crucial in High Content Screening analysis. There are different goals for feature selection as a prior step before analysis. Likely, the most important goal is to avoid the potential over fitting of data

to the model, which affects their generalization and performance. Another important consideration in feature selection is what it is known as “Curse of dimensionality”. This term refers to the fact that when we deal with many features or dimensions, the available data points or samples becomes very sparse in this multidimensional space and therefore it is very difficult to understand the structure of the data or to classify them, unless we have a huge amount of data points (data samples would look very dissimilar to each other). In other words, the amount of data needed to obtain statistically relevant results grows exponentially with the number of features.

The main assumption of feature selection is that the original image data set contains redundant and irrelevant features that should be eliminated before doing any further analysis like classification and hit stratification.

There are two general approaches to feature selection:

- **Filters** - Filter type methods are essentially pre-processing steps applied to the data. Features are selected based on their relevance to discriminate the classes.
- **Wrappers** - Wrapper methods on the other hand, use a classification method to evaluate the usefulness of the features in providing an accurate classification.

High Content Profiler provides the several approaches to select the features before creating a classification model or ranking the hits:

- *Minimum Redundancy Maximum Relevance (MRMR):* This is a filter method that demands the features to be maximally dissimilar to each other, for example by maximizing their mutual Euclidean distances or minimizing their pairwise correlations. These minimum redundancy criteria are combined by the maximum relevance criteria such as maximal mutual information with the classes or phenotypes.
- *Ensemble based tree classifier variable importance: Ensemble based tree classifier (Breiman 2001)* is a classification method based on the ensemble of decision trees classifiers. The rationale of this method is the use of hundreds of randomly generated trees that are properly combined to improve the classification performance. The main principle behind this method is that a group of weak learners can come together to form a strong learner. Ensemble based tree classifiers can rank the importance of variables by evaluating the implication of each feature in the classification error.

Z prime factor: The Z-prime factor (Zhang JH, Chung TD, Oldenburg KR. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen. 1999;4(2):67-73) is defined in terms of the means and standard deviations of both the positive (p) and negative (n) controls. The Z’-factor is defined as:

$$\text{Z-factor} = 1 - \frac{3(\sigma_p + \sigma_n)}{|\mu_p - \mu_n|}.$$

Z-prime factor Interpretation:

1.0	Ideal
0.5 - 1.0	An excellent assay
0 - 0.5	A marginal assay
<0	Overlap between positive and negative controls

Note: For High content screening assays often values above 0 are considered sufficient (Bray and Carpenter, Advanced Assay guidelines, 2013).

Z' factor robust:

This factor is like the original Z prime factor, the only difference being that the mean is replaced by the median and the standard deviation by the median absolute deviation in order to increase robustness against outliers.

T-statistic: The t-statistic is defined as:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2}}}$$

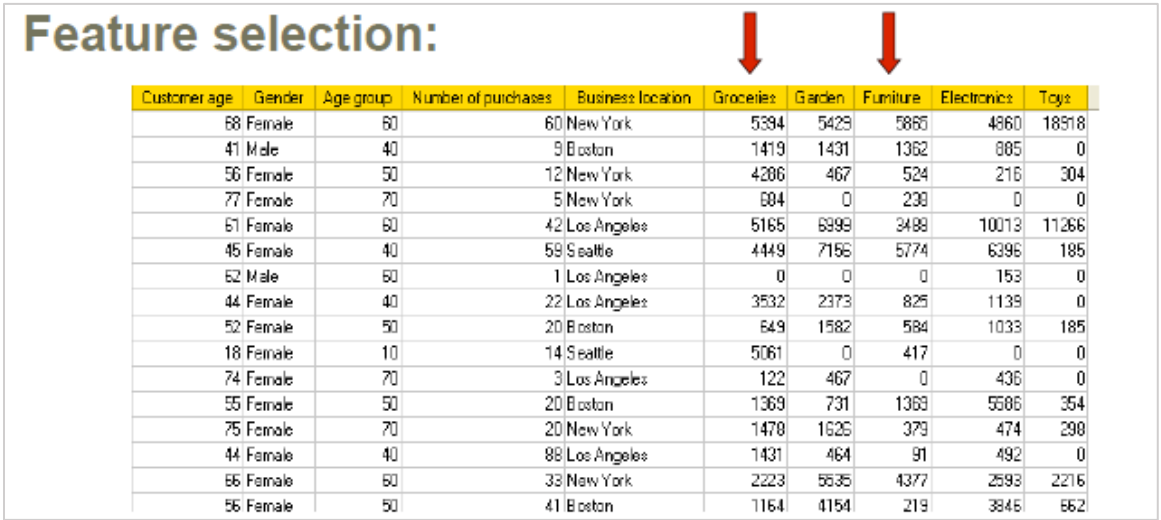
The absolute value of the t-statistic will be reported.

The combination of a filter and a wrapper approach make High Content Profiler robust in almost any situation of dirty and noisy high dimensional data.

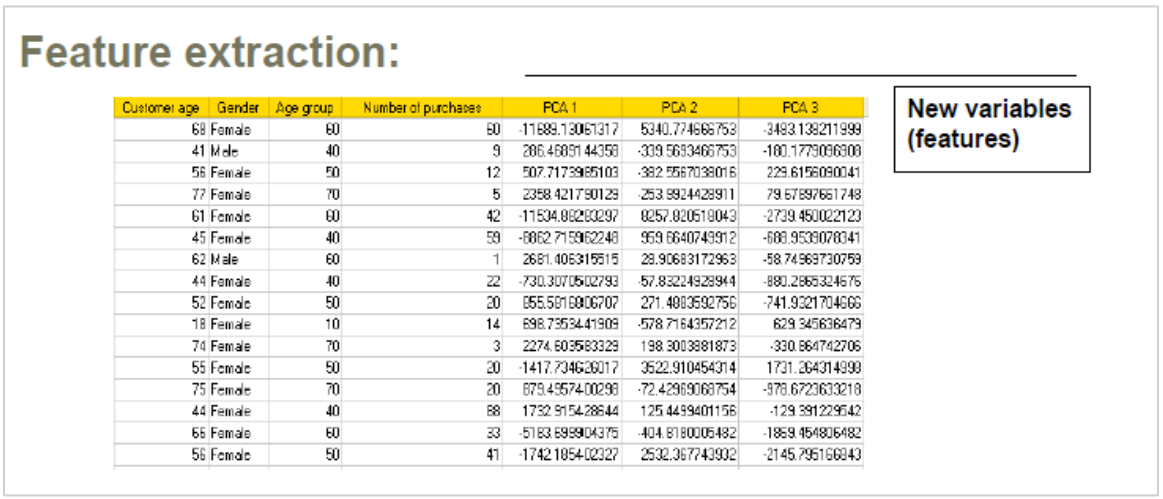
9.15.3.18 Feature Selection versus Feature Extraction

There is an important difference between feature selection and feature extraction that needs to be clarified in this manual for further reference. While feature selection selects the best and most discriminative features among the set of original ones, feature extraction methods do not select any features. They are transformative: they apply a transformation to the data to project it into a new feature space with lower dimensionality. PCA is the best example of this. Mapping and transformation of the original space into a lower one has lots of advantages since the new set of features is usually uncorrelated or independent. The main drawback it has is related to its interpretability. During the transformation process the interpretation of the new set of features becomes very difficult or impossible in some cases.

High Content Profiler use PCA to transform the data for visualization, QC and exploration. However, it can also be used as a feature extraction process to relate to classification. The following figures clarify these concepts.

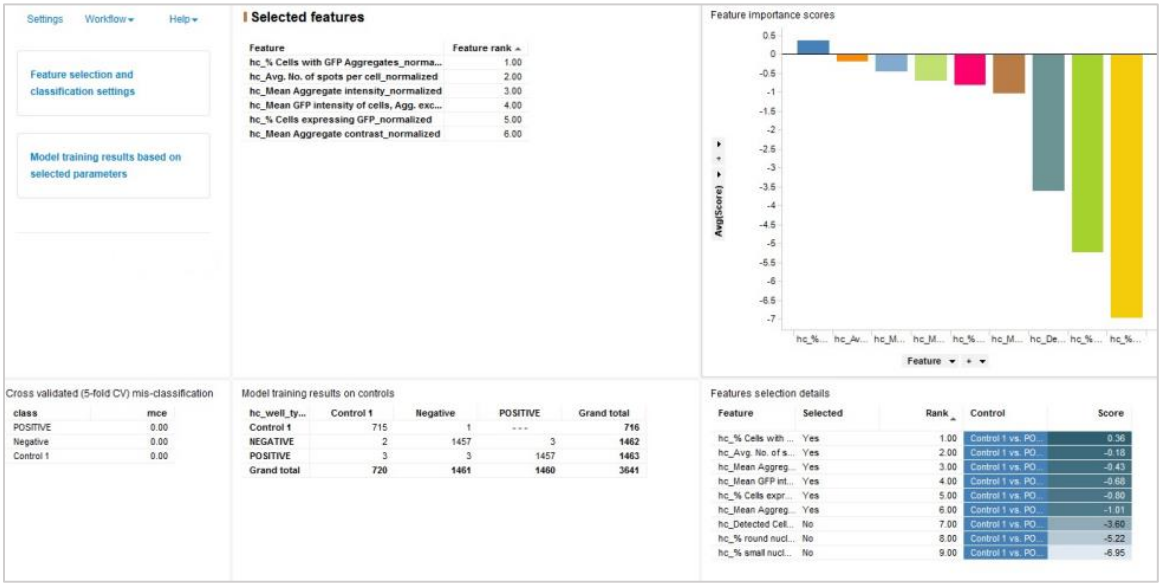


Customer age	Gender	Age group	Number of purchases	Business location	Groceries	Garden	Furniture	Electronics	Toys
68	Female	60	60	New York	5394	5429	5865	4860	18518
41	Male	40	9	Boston	1419	1431	1362	885	0
56	Female	50	12	New York	4286	467	524	216	304
77	Female	70	5	New York	694	0	239	0	0
61	Female	60	42	Los Angeles	5165	6999	3488	10013	11266
45	Female	40	59	Seattle	4449	7156	5774	6396	185
52	Male	50	1	Los Angeles	0	0	0	153	0
44	Female	40	22	Los Angeles	3532	2373	825	1139	0
52	Female	50	20	Boston	649	1582	584	1033	185
18	Female	10	14	Seattle	5061	0	417	0	0
74	Female	70	3	Los Angeles	122	467	0	436	0
55	Female	50	20	Boston	1369	731	1369	5586	354
75	Female	70	20	New York	1478	1625	379	474	298
44	Female	40	88	Los Angeles	1431	464	91	492	0
66	Female	60	33	New York	2223	5535	4377	2593	2216
56	Female	50	41	Boston	1164	4154	219	3845	662



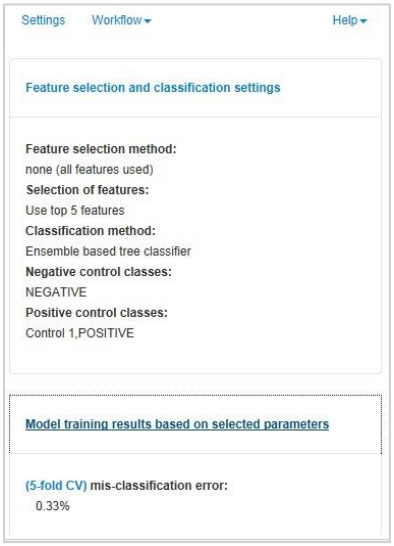
Customer age	Gender	Age group	Number of purchases	PCA 1	PCA 2	PCA 3
68	Female	60	60	-11688.13061317	5340.774666753	-3493.138211999
41	Male	40	9	286.46891 44358	-339.5693466753	-180.1779096308
56	Female	50	12	507.7173965103	-382.5567038016	229.6156090041
77	Female	70	5	2358.421790129	-253.9924428911	79.67897661748
61	Female	60	42	-11534.86293297	9257.820518043	-2739.450022123
45	Female	40	59	-6862.715962248	959.6640749912	-688.9539078041
62	Male	60	1	2681.405315515	28.90683172963	-58.74969730759
44	Female	40	22	-730.3070502793	-57.83224928944	-880.2865324676
52	Female	50	20	855.5818906707	271.4803592756	-741.9321704666
18	Female	10	14	696.7353441909	-578.7164357212	629.345636479
74	Female	70	3	2274.603583329	198.3003881873	-330.864742706
55	Female	50	20	-1417.734626017	3522.910454314	1731.264314999
75	Female	70	20	879.49574 00298	-72.42969068754	-976.6723633218
44	Female	40	88	1732.9154 28644	125.4499401156	-129.391229542
66	Female	60	33	-5183.6989 04375	-404.8180005482	-1869.454806482
56	Female	50	41	-1742.1854 02327	2532.367743932	-2145.795166843

The following visualization is available on the Feature selection and classification page:



Navigation Panel

The navigation panel on the left-hand side provides the following information:



The upper part of the panel lists the methods and parameters used for feature selection and classification and the control classes that were used to train the model.

The lower part shows the global misclassification error estimate computed when applying the model to held out samples not used for building the model. This estimate describes the classification error to be expected when applying a model using the same methods and parameters to a data set with similar characteristics. If the samples to be predicted have value distributions somewhat like those in controls used for building the model, the misclassification estimate can be used to assess the quality of the current model. The hold out samples are either selected using 5-fold cross validation or, in the case of Ensemble based tree classification without feature selection, using out of bag samples (i.e. samples not used for building an individual tree).

Visualizations

Selected features table

This table displays the list of features selected for classification, ordered by importance according to the feature score as computed by the selected feature selection method.

Cross validated (5-fold CV) miss-classification table

This table displays the misclassification error from hold out samples (using either cross validation or out of bag samples) individually for each class to be predicted. Model training usually tries to minimize the global misclassification error, not the individual ones. Strong imbalances in size of control samples therefore might lead to models biased towards the larger class.

Classification metrics

This table provides the user with several classification metrics that evaluate how well the feature classification performs.

Binary analysis

- Matthew's correlation coefficient: This measurement is regarded as one of the best ways to describe the confusion matrix with a single number. It considers true and false positives and negatives and is generally regarded as a balanced measure which can be used even if the classes are of very different sizes. The MCC reflects the correlation between the observed and predicted binary classifications. The value provided

ranges from −1 to +1. A coefficient of +1 represents a perfect prediction, 0 no better than random prediction and −1 indicates total disagreement between prediction and observation. It can be calculated using the following formula:

$$\mathrm{MCC} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Where TP: true positives, TN: true negatives, FP: false positives and FN: false negatives

- Sensitivity: The sensitivity or true positive rate reflects the effectiveness of a classifier to identify positive labels. It evaluates the number of true positives against all positives:

$$TPR = TP/P = TP/(TP + FN)$$

Where TPR: True positive rate, TP: true positives, P: Positives and FN: false negatives

- Specificity: This measure evaluates the effectiveness of a classifier to correctly identify negative labels. It evaluates the number of true negatives against all negatives

$$SPC = TN/N = TN/(FP + TN)$$

Where TN: True negative, N: negatives, FP: false positives and TN: true negatives

Multiclass Classification

- Sensitivity per class: This would be the same as for binary classification applied to each class separately, as opposed to in the binary classification where it was calculated only for positive controls.
- Specificity per class: This would be the same as for binary classification applied to each class separately, as opposed to in the binary classification where it was calculated only for positive controls.
- Sensitivity with macro average: This is an average of the per class sensitivity. This metric is invariant to each class size.
- Sensitivity with micro average: This is a cumulative calculation of sensitivity. This metric is not invariant to each class size and thus it favors bigger classes.
- Specificity with macro average: This provides an average of the per class specificity. This metric is invariant to each class size.
- Specificity with micro average: This provides a cumulative calculation of specificity. This metric is not invariant to each class size and thus it favors bigger classes.

Model Training Results on Controls Cross Table

This cross table displays the predictions on hold-out samples (i.e. those not used to build the model, either from cross validation or out of bag samples) vs. the known class labels on the model training set, i.e. those wells belonging to any of the selected controls. This information allows to both assess the general performance of the model and to identify control wells that are misclassified. The latter group of samples might include outlier or mislabeled control wells. It is therefore recommended to inspect misclassified samples using the QC visualizations and exclude them from analysis if appropriate.

Feature Importance Scores Bar Chart

This bar chart shows the distribution of feature importance scores as generated by the used feature selection method. Highest values generally indicate higher importance for discrimination of positive and negative classes.

Feature Selection Details Table

This table contains details on the feature selection results. The following columns are available:

Column	Description
Feature	Name of the feature
Selected	If the feature was selected for classification
Rank	Rank of the feature according to its feature score
Control	Description of the control classes used to select the features
Score	Feature score as determined by the feature selection method used

Rows without provided feature score, rank and selection information corresponding to features excluded from feature selection and classification because of increased number of missing values. Those features are excluded to generate a feature matrix with the minimum number of missing values while maintaining the maximum possible number of features. All rows (wells) containing at least one missing values in the remaining matrix are excluded from feature selection and classification model learning.

Note: When more than one class is present in the analysis and the positive controls for these classes are not pooled, the score displayed in the Feature selection details table is the median of the feature scores obtained from the feature selection method used over all classes. A feature that is selected because the score it obtains discriminating between the negative control and one of the positive classes passes the threshold defined in the analysis settings but may not be good at discriminating between the negative and positive controls in other classes. In this case the score it obtains when discriminating these other classes will not pass the threshold and the median for this feature may be above this threshold although the specific score for the class comparison in which the feature was selected passed the cutoff.

Note: The ranking of the features is calculated based on the score assigned to the feature. These scores for each feature are calculated when the classifier is built during the cross validation however if the features are very similar repeated scores can be returned by the classifier. In these cases where tied scores occur, the provided Rank is the mean of the ranks that would have been assigned to the tied features. This means that it is possible to have several features with the same rank.

9.15.3.19 Hit Detection

Once the feature selection process is completed, High Content Profiler automatically classifies the samples into the preselected classes using the selected features in the previous step. Classification of hits is carried out using different classification approaches.

9.15.3.20 Classification Models

Ensemble based tree classifier: As described in the feature selection section, this method is a classification method based on the ensemble of decision trees classifiers. The rationale of this method is the use of hundreds of random generated trees that are properly combined to improve the classification performance. The main principle behind this method is that a group of weak learners can come together to form a strong learner. This approach is very robust in high dimensional datasets and performs feature selection simultaneously to improve the classification accuracy.

Robust Mahalanobis distance classifier: This classifier is based on the Mahalanobis distance metric between multidimensional data points (samples) from the mean of a group of values (controls) which it is defined as:

$$D_M(x) = \sqrt{(x-\mu)^T S^{-1}(x-\mu)}.$$

Where $x = (x_1, x_2, x_3, \ldots, x_N)^T$ are the multivariate random variables,
$\mu = (\mu_1, \mu_2, \mu_3, \ldots, \mu_N)^T$ defines the mean of the controls and *S* is the covariance matrix.

For each individual sample, its Mahalanobis distances in multidimensional space are calculated from the center of the positive and negative controls. A decision tree is then trained to identify the optimal cutoffs for classifying samples in positive and negative classes based on its Mahalanobis distances from the centers of positive and negative controls. Finally, this decision tree is applied to the same Mahalanobis distances of samples to be predicted.

The algorithm by default tries to use a robust method (Minimum Covariance Determinant, Rouseeuw 1999) to estimate the covariance matrix in order to limit influence of outlier samples in controls. If for any reason that method fails, the empirical covariance estimator will be used as a fallback.

One Dimensional Projection of Hits

Classification assigns categorical labels to each sample to be predicted. Assuming that the trajectory from negative to positive controls in the feature space of informative features is a continuum that describes the similarity of samples to either of both controls, we also might be interested in a continuous measure for that similarity. Here we provide a measure that describes similarity with a given positive control phenotype on a continuous scale. We define a reference vector that starts at the center of negative controls and points towards the center of a given positive control. For each sample in the data set, we use scalar projection to project a vector from the negative center to that sample on the reference vector. The length of that projection has value 0 if the sample is located at the negative center and 1 if the sample is located at the center of the positive control. Values below 0 denote samples even farther away from positives than the negative control center while values above 1 denote samples that are farther away from the negative control center than positive controls.

The projection score uses the features selected by feature selection for classification. Projection scores computed on the full feature set without feature selection might not be meaningful, especially if there are many non-informative features. Also note that the scale of that the relationship of the score with phenotypic similarity is not guaranteed to be linear.

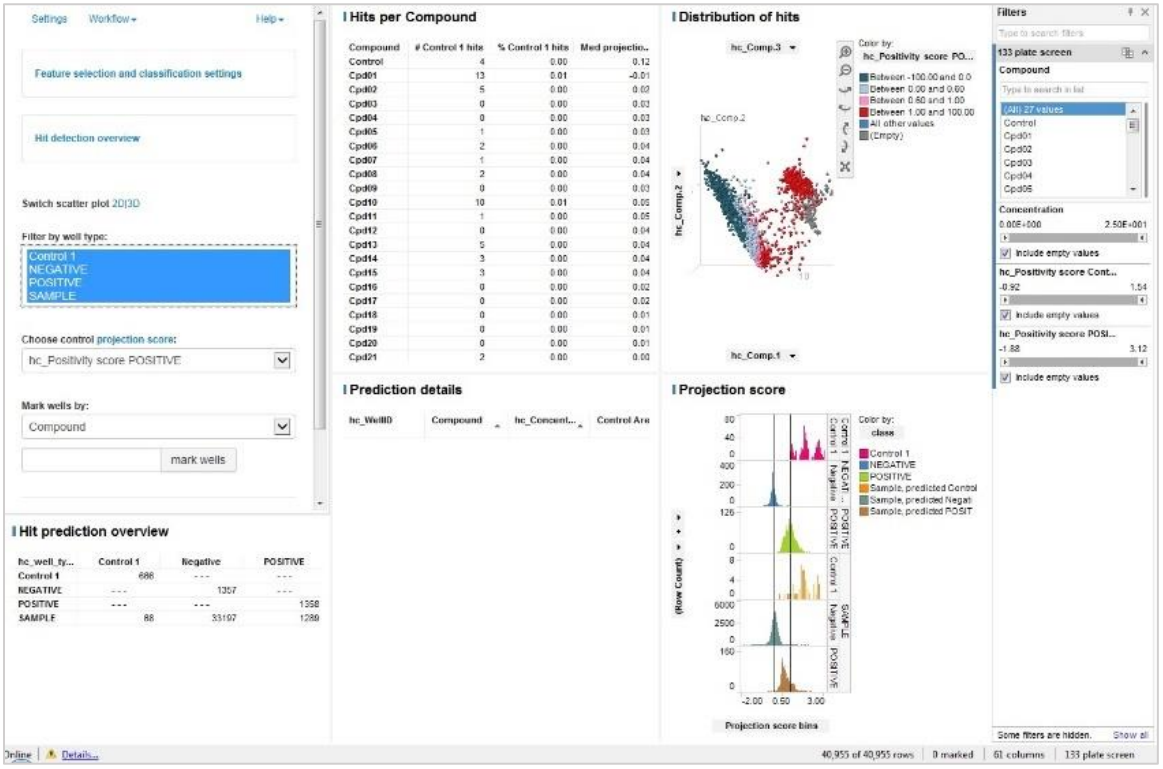
Hits Prediction and Scoring

Class labels based on positive and negative controls will be assigned to all samples that do not contain missing values in selected features. If there is more than one positive control and pooling is not enabled, separate classes are assigned for each positive control type, using their original names. If there is only a single positive control type,

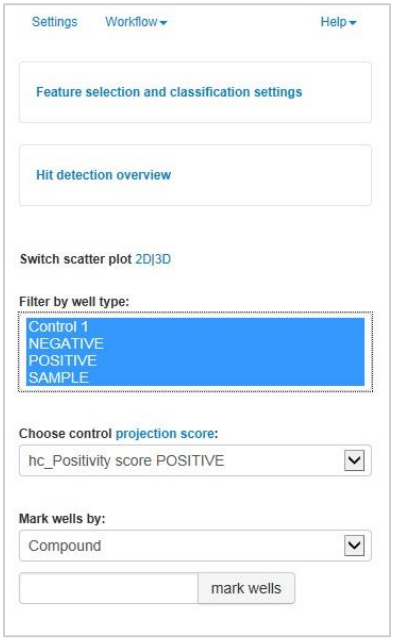
“Positive” is assigned for hits and “Negative” for non-hits, otherwise the original positive control names are assigned for hits. In addition to the class label, a Projection score value for each of the positive control types (if pooling is not selected) or for the combined positive controls is assigned as described above.

Hits will be aggregated by Gene or Compound, if that information is available and the corresponding Workflow type is selected.

Visualizations for Hit Detection



Navigation Panel



The navigation panel provides the following information:

- **Feature selection and classification settings:** Shows the methods and parameters used for feature selection and classification.
- **Hit detection overview:** Shows the cross-validated misclassification error estimate from the training set (i.e. control samples) and number of hits (i.e. classifications belonging to any of the positive classes) on the complete data set, i.e. including positive controls).
- **Switch scatter plot:** Allows switching the Principal component analysis scatter plot from 3D to 2D and back, allowing the display of the scatter plot also within the Spotfire® Web Player.
- **Filter by well type:** Allows you to select well types to be shown in scatter plot and Projection score bar chart.
- **Choose control projection score:** If there is more than one positive control and pooling of positive controls has not been selected, a separate projection score will be computed for each positive control type. Select here which one should be used in Projection score bar chart and scatter plot.
- **Mark wells by:** Allows marking genes in the Hits per gene/Hits per compound table, Prediction details table and in plots on this page based on gene/RNAi/compound identifiers and additional well annotations provided by the user.

Visualizations

Hit prediction overview cross table

This cross table shows an overview on classification results for all samples and controls in the data set based on the final model learned on the training set in relation to its well type.

Note: For control wells (but not samples), the model was learned on the same data as used for predictions and is therefore expected to show signs of over-fitting (i.e. lower classification error than to be expected on unknown samples).

Hits per gene/Hits per compound summary table

This table displays the following information:

Column	Description
Gene/Compound	Gene or Compound identifier, depending on selected workflow type
# hits	Number of hits found for this gene or compound
% hits	Proportion of hits among all replicate wells available for this compound
Med projection score [positive class name]	Median projection score across all wells available for this compound. If there is more than one positive control class, there will be one column for each of them
# total RNAi / concentrations	Total number of wells for this compound
# unique RNAi/concentrations	Number of unique RNAi or concentration levels found for this gene or compound

Prediction Details Table

This table displays the following information:

Column	Description
WellID	Well identifier, Plate ID + Row ID + Column ID
Gene/Compound	Identifier of gene or compound, if available
RNAi/Concentration	Identifier of RNAi or concentration level, if available
Type	Well type as specified by the user
Predicted_class	Class assignation as determined by the classification model
Projection score [positive]	Projection score for this well. If there is more than one positive control class, there will be one column for each of them
Rowcol	Coordinates of the sample on the plate

Distribution of Hits Scatter Plot

This scatter plot represents the distribution of hits in a projection of all data by the Principal component analysis on all features. Color is defined by the value of the selected Projection score column (Red: near the positive controls, Blue: near the negative controls).

Projection Score Bar Chart (Histogram)

This bar chart indicates the distribution of projection score values across well types and predicted classes. In general cases, one would expect a good agreement between the value of the projection score and the classification result, with clear separations of predicted classes. This plot also allows identifying wells where projection score and classification do not match, which might indicate outlier status or misclassifications.

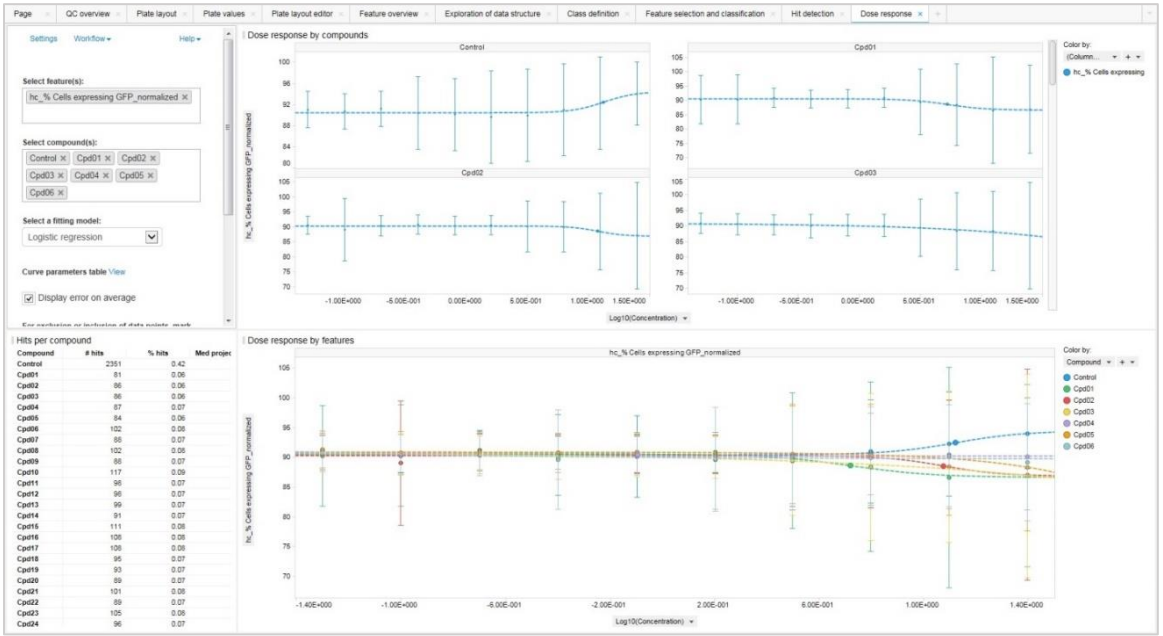
Filters

Filters are available for limiting the information displayed to selected Gene/Compounds or RNAi/concentration levels and for any of the projection scores.

9.15.3.21 Dose Response

Dose response curves are one of the most common methods used to select hits in compound screening analyses where a group of samples are exposed to different concentrations of a set of compounds to evaluate the response produced by these compounds.

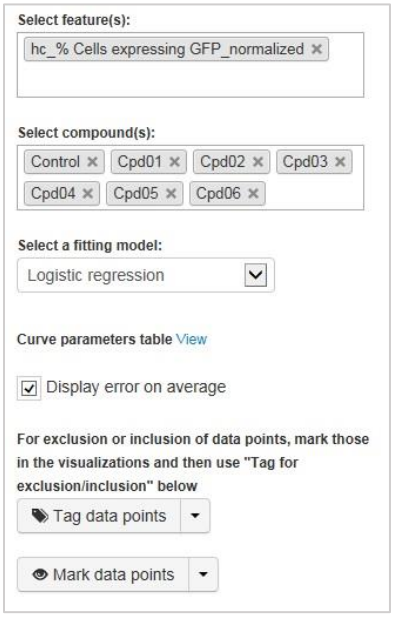
This visualization will be available for RNAi screens and Small or large molecule screens where a concentration column has been provided to the wizard.



The Dose response page provides the following visualizations:

- **Dose response by compound:** This visualization shows a scatterplot representation of the compound concentration vs. the normalized value of the selected feature trellised by compound. This allows to easily evaluate the effect of each individual compound on the selected features.

- **Dose response by feature:** This visualization shows a scatterplot representation of the compound concentration vs. the normalized value of the selected feature trellised by feature. This allows to easily to easily evaluating the behavior of each feature in response to the selected compounds.
- **Hits per gene/compound table:** Information on the hits will be aggregated by Gene or Compound, if that information is available and the corresponding workflow type is selected. If not, the table will not be present. This table contains the gene/compound details including the number of hits if more than one replicate is available. The compounds can be selected in this table to visualize the dose response curves corresponding to each of them.



The user interface in the dose response page contains the following controls:

- **Select feature(s):** Allows the user to select the features for which he would like to calculate the curves.
- **Select compound(s):** Allows the user to select the features for which he would like to calculate the curves.
- **Select a fitting model:** Provides a dropdown menu that allows the user to select the type of curve fitting model to use for the dose response curve calculation. The available options are:
    - (None): No line is shown
    - Polynomial curve 4: Level 4 polynomial curve fit
    - Polynomial curve 5: Level 5 polynomial curve fit
    - Logistic regression
    - Straight line
    - Gaussian curve
    - Exponential curve
- **Curve parameters table view:** This link displays a table containing the parameters calculated for the fitted curves.
- **Display error on average checkbox:** When checked (default) the average of the feature value for the same compound and concentration is displayed together with the error bars that represent the dispersion. When unchecked, the individual data points are displayed with no error bars.

- **Tagging controls:** These controls allow the user to exclude data points from the analysis by selecting them from the analysis.

Note: Both the Select compound and the Hits per compound table control the compounds displayed in the selected features and selected compounds visualizations, both options are provided for the convenience of the user, however care should be taken to avoid filtering out all the data when combining the filters.

9.15.3.22 Custom Inter-Plate Normalization Methods

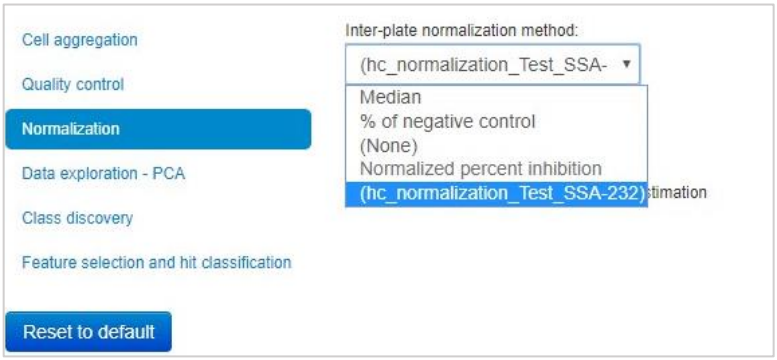
This option allows those users with more familiarity with Spotfire® to create their own custom data functions that would be registered in the Spotfire® server and use these data functions in the normalization step in place of those provided by the workflow.

To do this the data functions must be created using the template TERR data function provided in Appendix C and should be registered following the standard procedure for registering custom data functions in Spotfire® with the following additional restrictions:

- The data function must have only three input parameters which are the following:
  - A non-optional Table parameter called "Input_table".
  - A non-optional Value parameter called "neg_controls".
  - A Value parameter called "pos_controls".
- The data function must have only one Table output parameter called "out".
- The name of the registered data function should contain the prefix “hc_normalization”.

Once this is done the data function will be available for use in the normalization settings menu.

To select this custom normalization method in the HCP-Treatment form of the wizard step, select analysis settings Normalization. The custom normalization method will be available in the normalization methods drop down menu as shown in the figure.



9.15.4 Appendix A - Supported Input File Formats

The **High Content Profiler** tool can be used on a table present in Spotfire®. This table should have the information for the different features as columns and the information on the different Genes/Compounds as rows. Depending on the type of workflow selected, in addition to the feature columns, other columns may also be required:

- Gene/compound ID (RNAi and Compound screens)

- Plate (Plate based analyses)
- Row (Plate based analyses)
- Column (Plate based analyses)
- Sample type (Univariate analysis)

9.15.5 Appendix B - Main Table Information

When the High Content Profiler is executed it will add several columns to the input table or to the hc_Aggregated_cells table, which is created in the case of plate-based analyses from cell level data. This table contains the data from the aggregation of the cell level information at the well level, with a single aggregated entry provided for each of the wells containing cell information, the columns mentioned below may be present depending on the selected workflow and analysis. Each of these columns will be prefixed with “hc_”. Although most of them have been mentioned in relation to those visualizations in which they appear, below is a description of each of these columns.

Column	Description
Hc_User exclusion	Specifies that the selected well is to be excluded from analysis. ‘Untagged’ wells will be included in the analysis. Use the corresponding buttons in the navigation panel on the left-hand side to change the tags for marked wells.
Hc_Gating tag	Specifies that the selected well is to be included or excluded from the analysis.
Hc_Combined exclusion tags	Contains a combination of the hc_User exclusion and hc_Gating tag columns that will indicate which are the final cell level data to be aggregated in the well level table.
Hc_percent objects	This column contains the percentage of the total cells contained in each of the wells.
HC_Concentration	This column contains the concentration information if present.
Hc_WellID	Well identifier. Combination of Plate identifier + Row identifier + Column identifier.
Hc_WellMissingProportion	Proportion of features in that well with missing value.
Hc_WellMissingCnt	Number of features in that well with missing value.
Hc_WellMissingFlag	Flag will be True if the percentage of missing values in that well is above the threshold defined in the Analysis settings dialog and False otherwise.
Hc_PlateMissingCnt	Number of features in the plate with missing value.
Hc_PlateMissingProportion	Proportion of features in the plate with missing value.

Hc_well_type	Column containing the type of sample present in the well.
Hc_WellMissingType	Boolean value indicating if the well is missing.
Hc_PlateHasSamples	Boolean value indicating if the well has a sample.
Hc_PlateValidSamples	Boolean value indicating if the plate contains valid samples.
Hc_<featureX>_normalized	Columns containing the normalized values for each of the features. For features selected for passthrough in normalization, the value provided is the same as the raw value.
Hc_<featureX>_z_score	Columns containing the z-score scaled normalized values for each of the features. For features selected for passthrough in normalization, the value provided is the same as the raw value.
Hc_WellOutlierCount	Number of feature values that are considered outliers in that feature when applying the threshold as configured in the Analysis settings dialog.
Hc_WellOUtlierProportion	Proportion of feature values that are considered outliers in that feature when applying the threshold as configured in the Analysis settings dialog.
Hc_PlateOutlierCount	Number of feature values flagged as outlier in the corresponding feature, across all features and wells on that plate.
Hc_plateOutlierProportion	Proportion of feature values flagged as outlier in the corresponding feature, across all features and wells on that plate.
Hc_CompX	Components of the PCA.
Hc_SOM_X	X coordinate of the SOM cell to which the sample is assigned.
HC_SOM_Y	Y coordinate of the SOM cell to which the sample is assigned.
Hc_pdcv	Class of the sample used in cross validation.
Hc_Positivity score <Ctrl>	Positivity score as measured relative to the relevant control.
Hc_predicted_class	Predicted class.
Hc_SigToNoise	Contains the Signal to noise calculated for the plate. It is present only in the case of univariate analysis.
Hc_SSMD	Contains the SSMD calculated for the plate. It is present only in the case of univariate analysis.

Hc_ZPrime	Contains the Z-Prime calculated for the plate. It is present only in the case of univariate analysis.
Hc_ZPrimeRobust	Contains a robust version of the Z-Prime calculated for the plate. This Z-Prime is calculated using Median and MAD. It is present only in the case of univariate analysis.
Hc_FoldChange	Contains the Fold Change calculated for the well. It is present only in the case of univariate analysis.
Hc_SSMDSample	Contains the SSMD calculated for the well. It is present only in the case of univariate analysis.
Hc_SDDistanceToNegatives	Contains the Distance to the negative controls calculated for the well in terms of Standard deviations from the negative controls mean. It is present only in the case of univariate analysis.
Hc_MADDistanceToNegatives	Contains the Distance to the negative controls calculated for the well in terms of MADs from the negative controls median. It is present only in the case of univariate analysis.
Hc_UnivariatePlateQCClassification	Contains plate level QC information that indicates if a plate is within the QC threshold defined by the user.
Hc_UnivariateHitSelection	Contains well level QC information that indicates if a well is within the QC threshold defined by the user.

9.15.6 Appendix C - Custom Data Function Contract File

9.15.6.1 Custom Data Function Contract File

The custom data function contract file is a template for a data function that must be followed to use the custom normalization feature within the HCS workflow.

```
################################################
# Mi-nimum requirements:
# Input table with the structure:
# Col 1 Plate
# Col 2 Row
# Col 3 Column
# Col 4 Type (Control area column)
# Col 5-n Parameters
# Neg control list Strings used for neg control
# Pos control list Strings used for pos control

################################
# INSERT CUSTOM FUNCTIONS HERE #
################################

CUSTOM FUNCTIONS

############################
# END OF CUSTOM FUNCTIONS #
############################
```

```
##############################################
# INSERT CUSTOM FUNCTION CALLS AND CODE HERE #
##############################################
# The custom code should generate one normalized column per feature raw feature in
# the input
# The name of the normalized features should be the same name as the input feature.
# The returned table should be in the same format as the input table as described
# above
# Col 1 Plate
# Col 2 Row
# Col 3 Column
# Col 4 Type (Control area column)
# Col 5 to n features


CUSTOM CODE

###################
# END CUSTOM CODE #
###################

colnames(normalized_data) <- paste(colnames(normalized_data),"_normalized",sep="")
```

[1] Bray MA, Carpenter A. Advanced Assay Development Guidelines for Image-Based High Content Screening and Analysis. 2013 Mar 01. In: Sittampalam GS, Gal-Edd N, Arkin M, Auld D, Austin C, Bejcek B, Glicksman M, Inglese J, Lemmon V, Li Z, McGee J, McManus O, Minor L, Napper A, Riss T, Trask OJ, Weidner J, editors. Assay Guidance Manual [Internet]. Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2004-. Available from http://www.ncbi.nlm.nih.gov/books/NBK126174/.

[2] Rousseuw P, Van Driessen K. A Fast Algorithm for the Minimum Covariance Determinant Estimator, 1999, American Statistical Association and the American Society for Quality, TECHNOMETRICS.

[3] Breiman, Leo (2001). "Random Forests". Machine Learning 45 (1): 5–32.

9.16 Image Discovery

Image Discovery enables rendering and visualization of scientific images directly within Signals VitroVivo. Image Discovery provides connection to an image repository such as Columbus, Harmony, Image Artist, or other image sources such as file shares. Using Image Discovery, Signals VitroVivo can render images in labels, tool tips, tables, axis labels, as well as several new image visualization and settings panels and tools. Imaging scientists have total control over their images including channel and color selection, adjustments of brightness, contrast, and gamma values for each channel, and much more. Furthermore, Image Discovery empowers scientists to immediately correlate and validate results and insights obtained from Signals VitroVivo by inspecting the corresponding images on the same platform.

Feature List for Image Discovery:

- Access and render images stored in Columbus system directly within **High Content Profiler** and Spotfire®.
- Access and render images stored in Image Artist system directly within **High Content Profiler** and Spotfire®.
- Access and render images stored in various image repositories including Harmony database and other file shares directly within High Content Profiler and Spotfire®.
- Define custom rules or definitions to access and render images stored in other image repositories (such as GE InCell image files).
- Render raw images (multi-channel, multi-field, multi-timepoint, multi-plane) in High Content Profiler with controls over channel selection, color selection, brightness, gamma, and histogram values.
- Brand new Image Settings panel to adjust and preview rendering settings.
- Brand new Thumbnails panel to view images side by side and channel by channel.
- Full image viewer with a view finder to zoom in and out as well as panning capabilities.
- Supports visualization of images from multiple screens or projects.
- Ability to use the image rendering engine to display images with Tooltip, Label, Axis, columns within a data table and Details-on-Demand, and more.
- Automatically crop and render single cell images/thumbnails to support flow cytometry-like analysis workflows.
- Local caching of image data to optimize visualization performance.
- Visualize images in the Consumer (web) client.

9.16.1 Setting Preferences

Preferences for **Image Discovery** can be set to define the default values that are used when a data file is opened, and images are retrieved from an image source and displayed in Spotfire®. Image columns can be inserted into the current data set, and an image renderer is automatically set on these columns.

Through the Spotfire® Administration Manager, you can see if the image renderer is automatically set on the image column by default or modify the preference.

The following preferences can be set for Image Discovery through the Spotfire® Administration Manager:

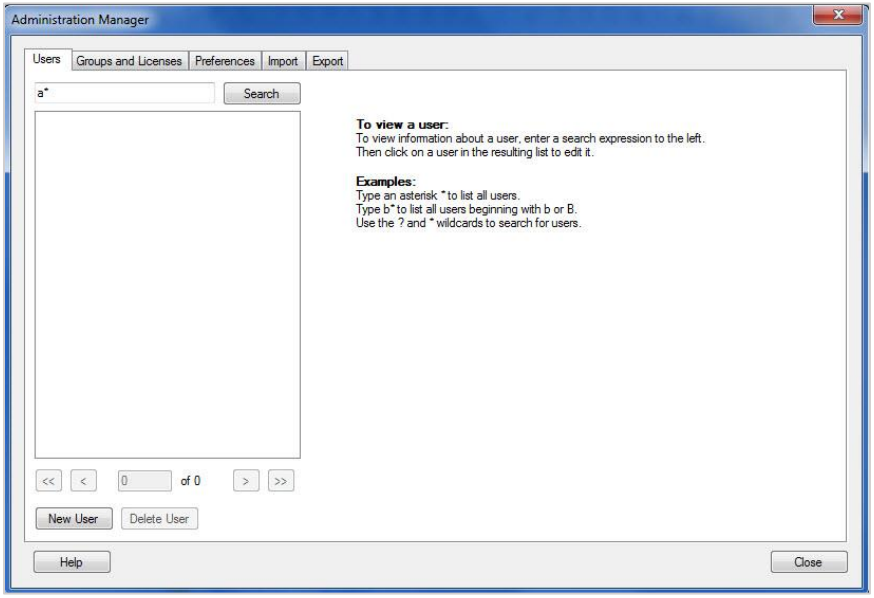
Preference	Description	Default Value
AutoSetImagesRenderer	Can automatically set an image renderer after inserting a new image column. Otherwise, Spotfire® will set a default Text renderer for the image column.	True
Rules	Refer to Groups Rules.	

Note: Changes to the preferences will take effect after restarting Spotfire®.

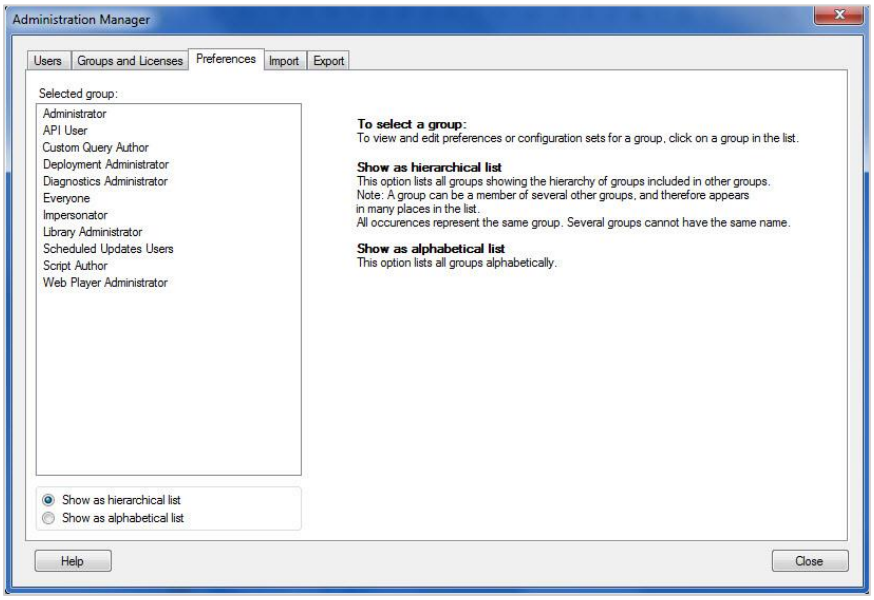
The following section describes how to set the AutoSetImagesRenderer preference through the Administration Manager in Spotfire®. Refer to Group Rules section for details on configuring the Rules preference.

To set AutoSetImagesRenderer preference:

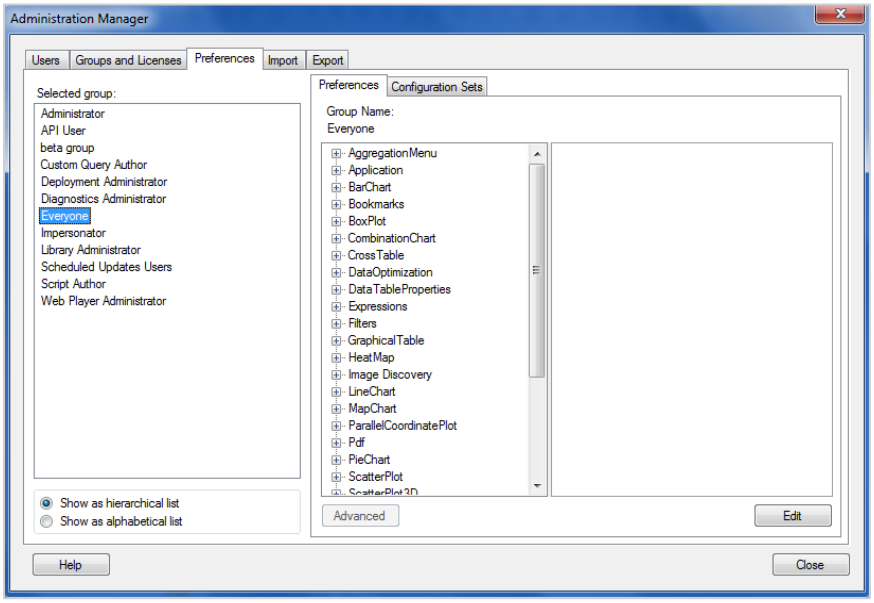
1. Open the Spotfire® client, and logon as a Spotfire® Administrator. From the **Tools** menu, select the **Administration manager…** sub-menu item. The **Administration Manager** window opens.



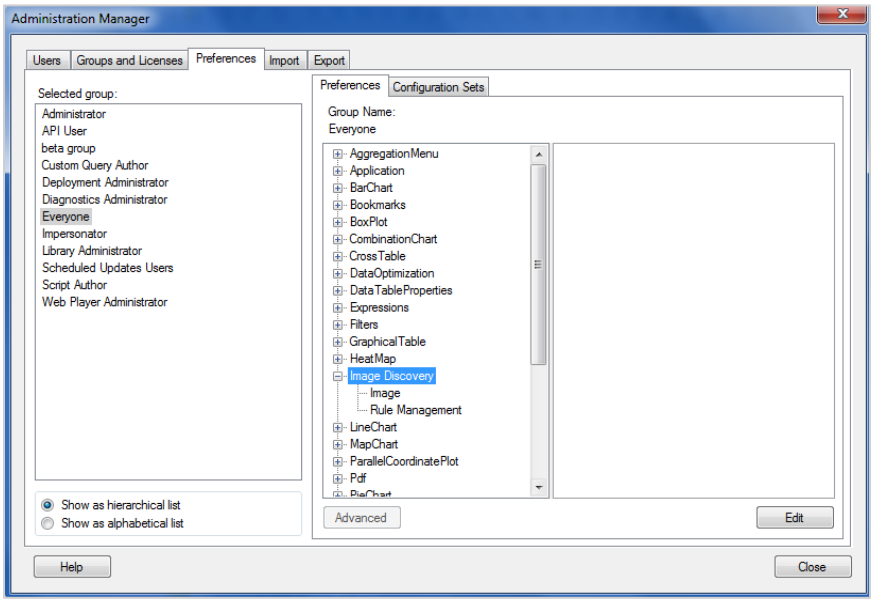
2. Click on the **Preferences** tab.



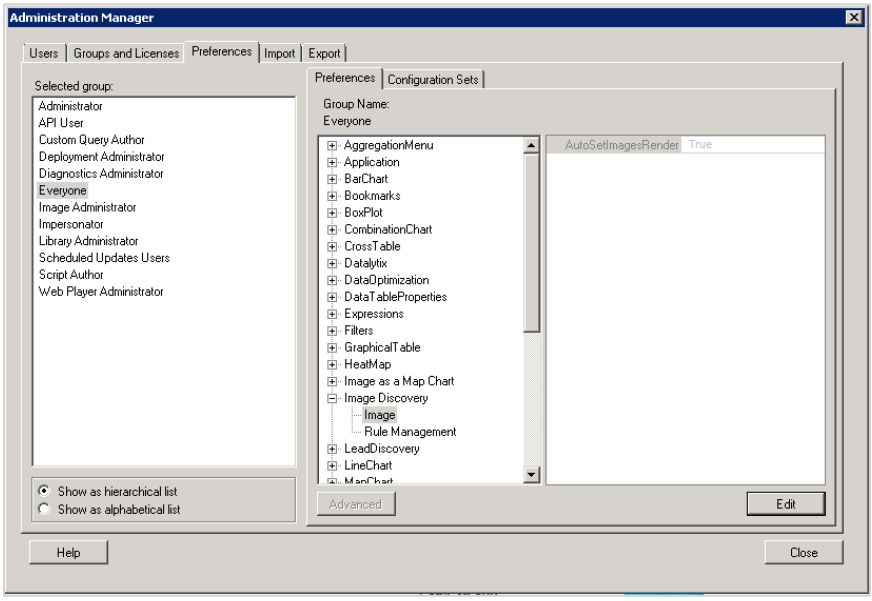
3. Select the **Group Name** to which the preferences should be applied.



4. Select the **Preferences** tab in the right-hand panel. Expand the **Image Discovery** category in the tree.

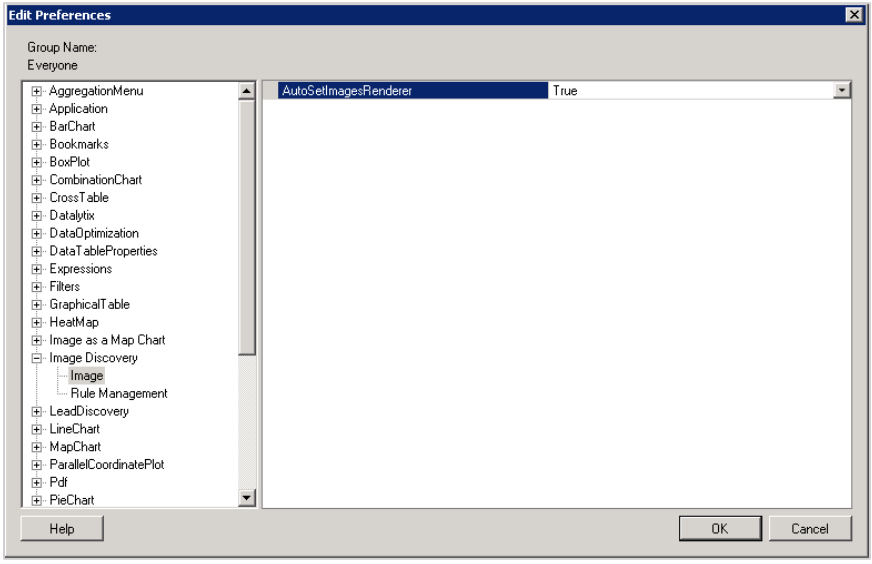


5. Select the **Image** sub-category in the tree.



6. Click on the **Edit** button.

The **Edit Preferences** dialog opens.



7. Set the **AutoSetImagesRenderer** parameter preference (True or False).
8. Click **OK** to save and close the dialog.
9. Click **Close** to exit the Administration Manager.

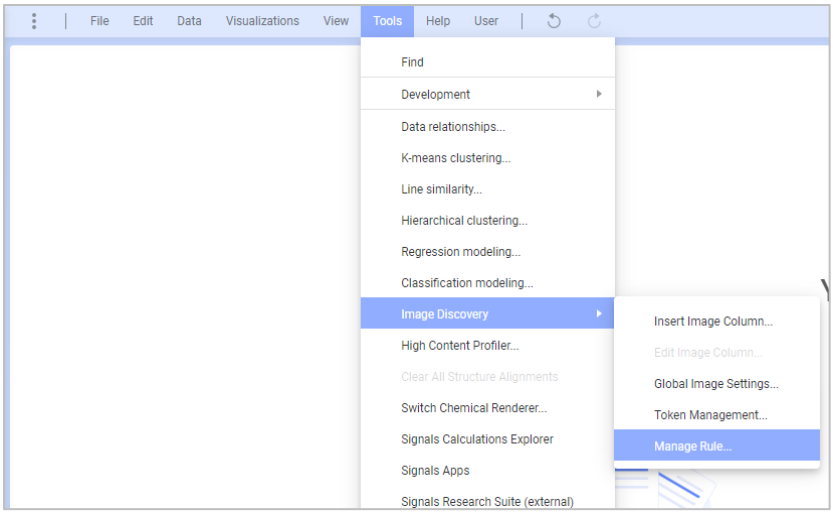
9.16.2 Rule Management

A rule is a predefined set of field mapping/value processing methods to help users retrieve images from the image source and display them in the data table. There are three types of rules: predefined rules, group rules and user

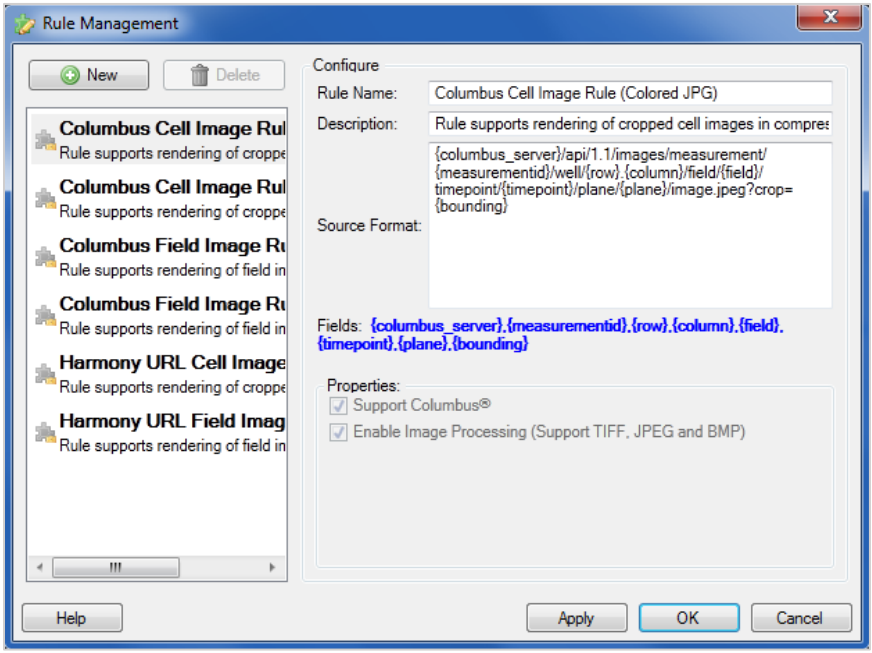
rules. Predefined rules are available to all users and cannot be modified by end users. Group rules are available to all the users of the group and user rules are available to the user who defined them.

9.16.2.1 Accessing Rule Management

1. Select **Tools > Image Discovery > Manage Rule.**



The **Rule Management** dialog is displayed, like the example shown below.



Note: The Rule Management dialog can also be accessed from the Insert Image Column and Edit Image Column dialogs via the **Manage Rules** button.

There are three types of rules: predefined rules, group rules and user rules. Predefined rules are provided, by default, with the software and available to all users and cannot be modified or deleted. User Group rules are only available by the user who created the rule, and Group rules are available to all users in a group.

New rules can be added by clicking on the **New** button. Existing rules can be deleted by selecting the rule from the list and clicking on the **Delete** button. You will be prompted to confirm the deletion. Pre-defined rules cannot be deleted. Refer to the following sections for additional details on adding new rules and deleting existing rules.

9.16.2.2 Pre-defined Rules

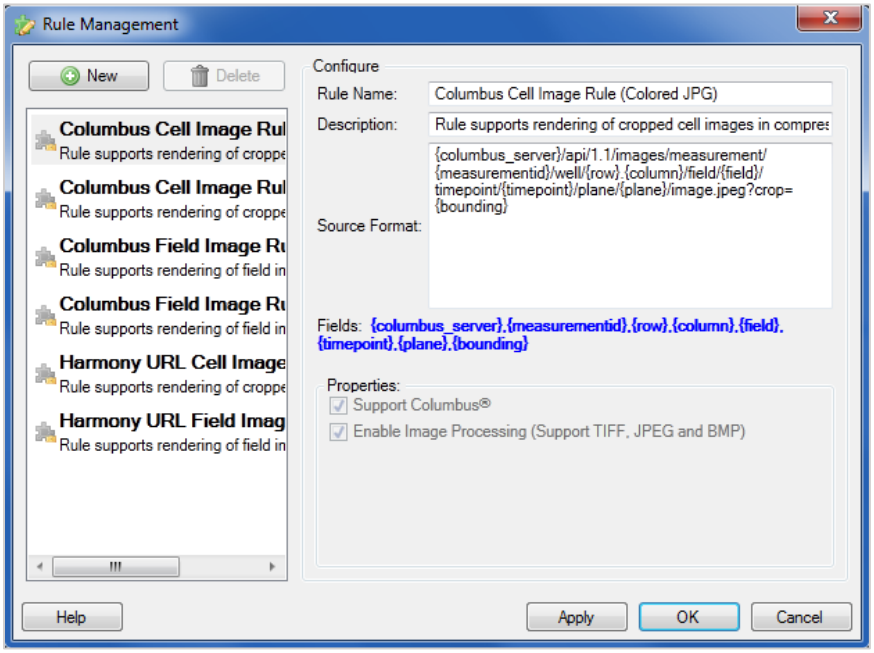
By default, the following pre-defined rules are included:

Rule Name	Description
Columbus Cell Image Rule (Colored JPG)	Rule supports rendering of cropped cell images in compressed, colored JPG format with cell level result files from Columbus 2.6 and higher.
Columbus Cell Image Rule (Raw TIFF)	Rule supports rendering of cropped cell images in TIFF format with cell level result files from Columbus 2.6 and higher.
Columbus Field Image Rule (Colored JPG)	Rule supports rendering of field images in compressed, colored JPG format with aggregate/well result files from Columbus 2.6 and higher.
Columbus Field Image Rule (Raw TIFF)	Rule supports rendering of field images in TIFF format with aggregate/well result files from Columbus 2.6 and higher.
Harmony URL Cell Image Rule	Rule supports rendering of cropped cell images in TIFF format with cell level result files from Harmony.
Harmony URL Field Image Rule	Rule supports rendering of field images in TIFF format with result files from Harmony.

Note: Pre-defined rules cannot be deleted.

9.16.2.3 Understanding the Source Format

The source format text points to the image source to be retrieved. Content contained within the curly brackets are fields. Each field in the source format will be replaced for each row with either a fixed value or a column related value.



For example, using the following Source Format:

{columbus_server}/api/1.1/images/measurement/{measurement}/well/{row}.{column}/field/{field}/timepoint/{timepoint}/plane/{plane}/image.jpeg?crop={bounding}

1. The rule format below will be mapped like the following:

http://165.88.124.88/api/1.1/images/measurement/4/well/5.16/field/1/timepoint/1/plane/1/image.jpeg?crop=[23,1,75,42]

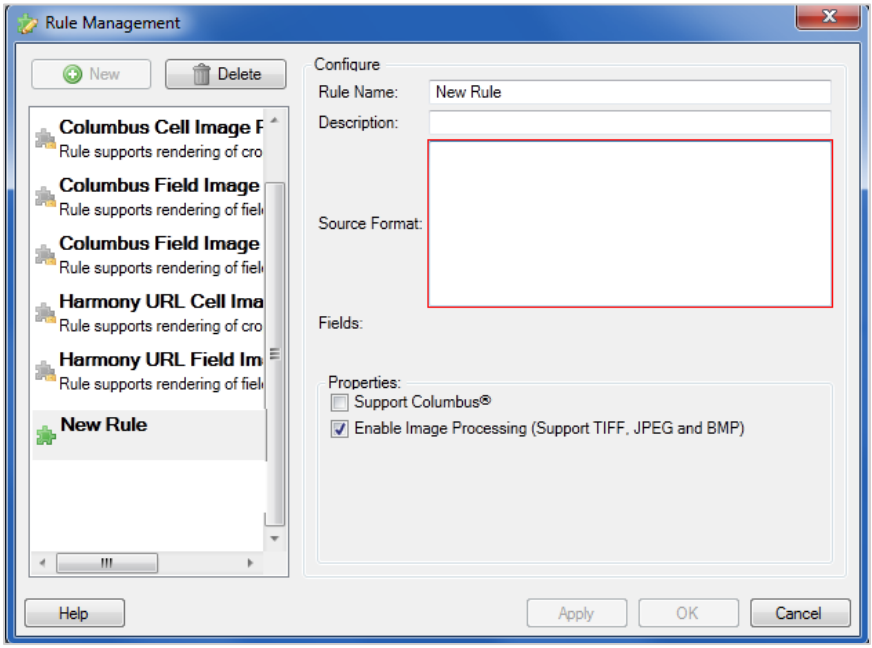
2. All the fields {in curly brackets} are replaced with concrete values. This mapping occurs for each data row and is responsible for the images populating the image column.

9.16.2.4 User Rules

User Rules can be added by the currently logged in user and are only available to that user.

To add a new user rule:

1. From the Rule Management dialog, click on the **New** button.

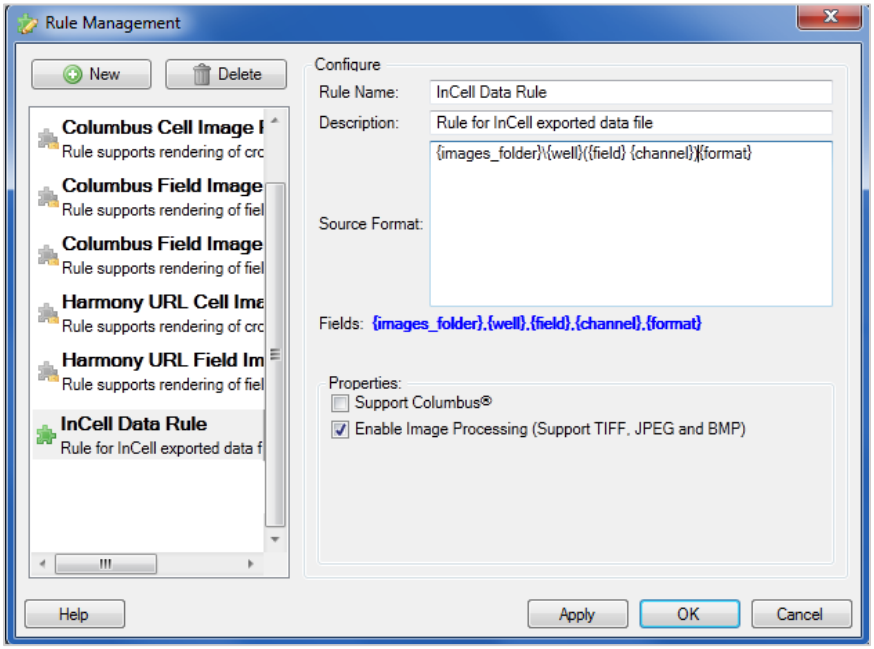


2. In the Configure group box, enter a unique Rule Name and optional description. By default, the Rule Name will be ‘New Rule’.
3. Enter the source format for the new rule. The fields will be extracted from the source and displayed below the Source format text box.

Note: The content is validated as you are typing. Incorrect syntax will be highlighted in red. The **Apply** and **OK** button will be disabled if the current rule is invalid.

4. Define the Properties for the new rule. Enable the Support Columbus® checkbox if the rule requires the extraction of images from Columbus®. Enable the Enable Image Processing (Support TIFF, JPEG and BMP) checkbox if the rule requires the images to be processed by Arcadia. The options are not mutually exclusive, and both options can be selected for a user rule.

Tip: The image displayed in the Images Settings preview panel may be different from the image which is displayed in the visualization because the rule to generate the images may not support “Image Processing” property. In this case, images in the visualization may display the original image without any processing. If settings are modified through the Image Settings panel, Image Discovery, by default, uses the “Image Processing” property to process the image and therefore the image in the preview panel may differ from that displayed in the visualization. The Thumbnails panel also, by default uses the “Image Processing” property to process the image, and therefore thumbnails may differ from the image displayed in the visualization.



5. Click **Apply** to apply the changes without closing the Rule Management dialog. Clicking **OK** will save the new rule and close the Rule Management dialog. Click **Cancel** to discard any changes and close the dialog. User rules are appended to the bottom of the Rules list.

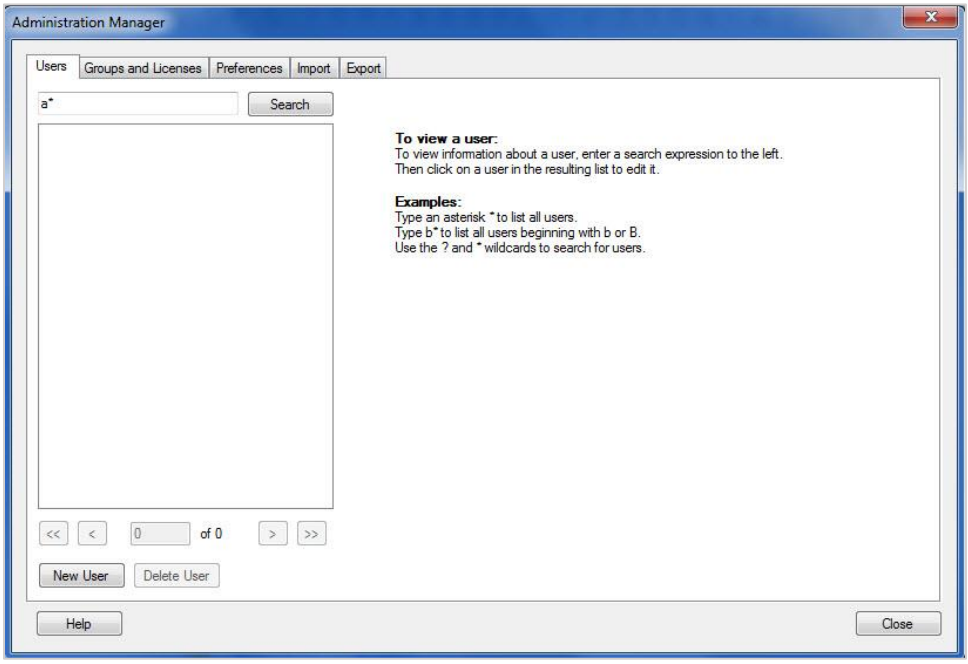
Note: User rules can be deleted by selecting the rule and clicking on the **Delete** button.

9.16.2.5 Group Rules

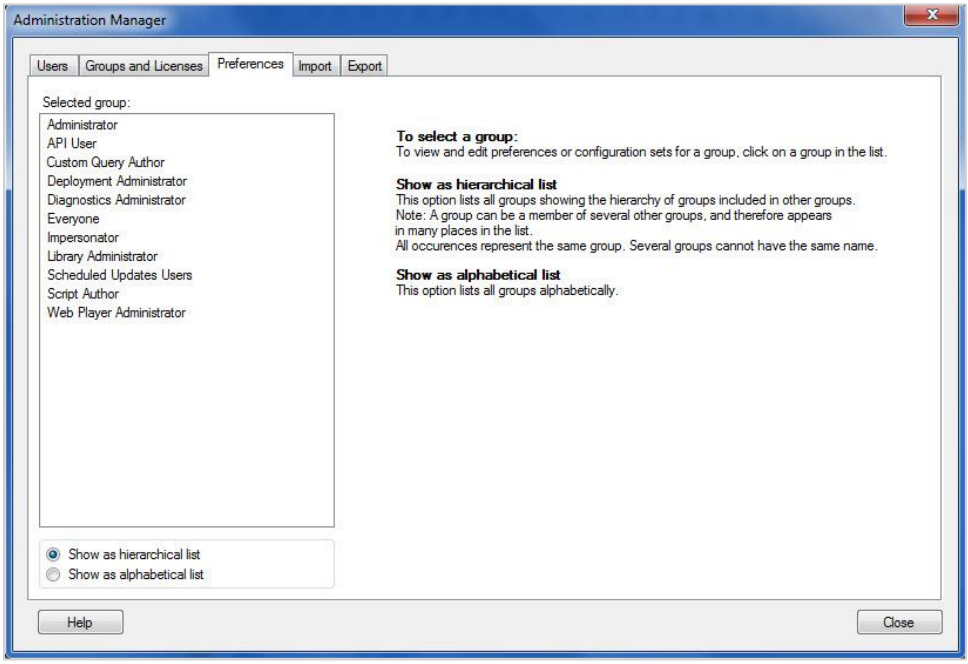
Rules can be defined for specific groups defined within Spotfire®. Group rules are defined and managed through the Administration Manager in Spotfire®. Only members belonging to the group can see the Group rule in the Rule Management dialog.

To define a group rule:

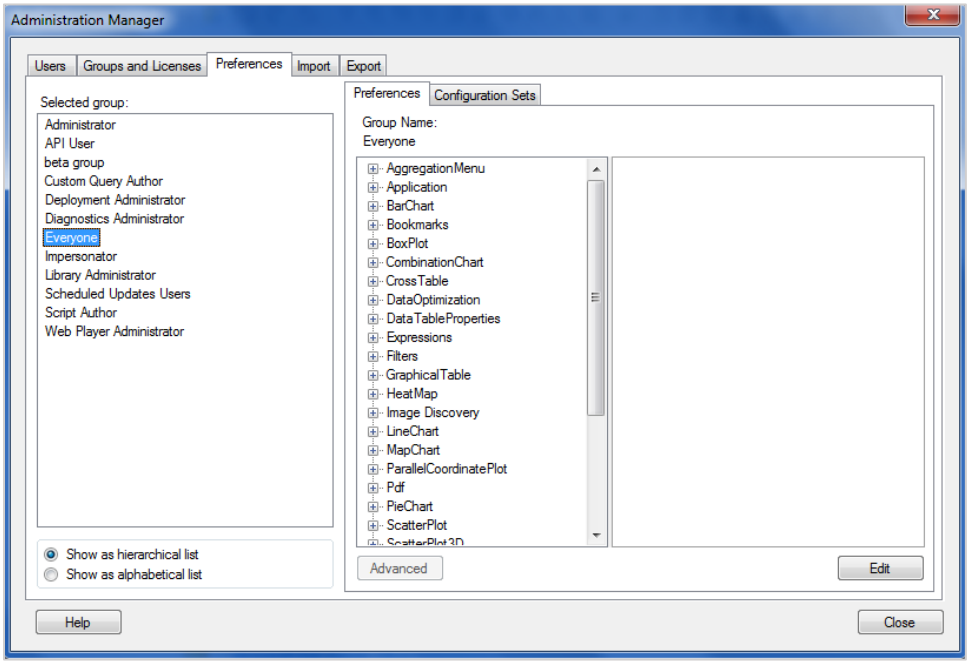
1. Open the Spotfire® client, and logon as a Spotfire® Administrator. From the **Tools** menu, select the **Administration manager…** sub-menu item. The **Administration Manager** window opens.



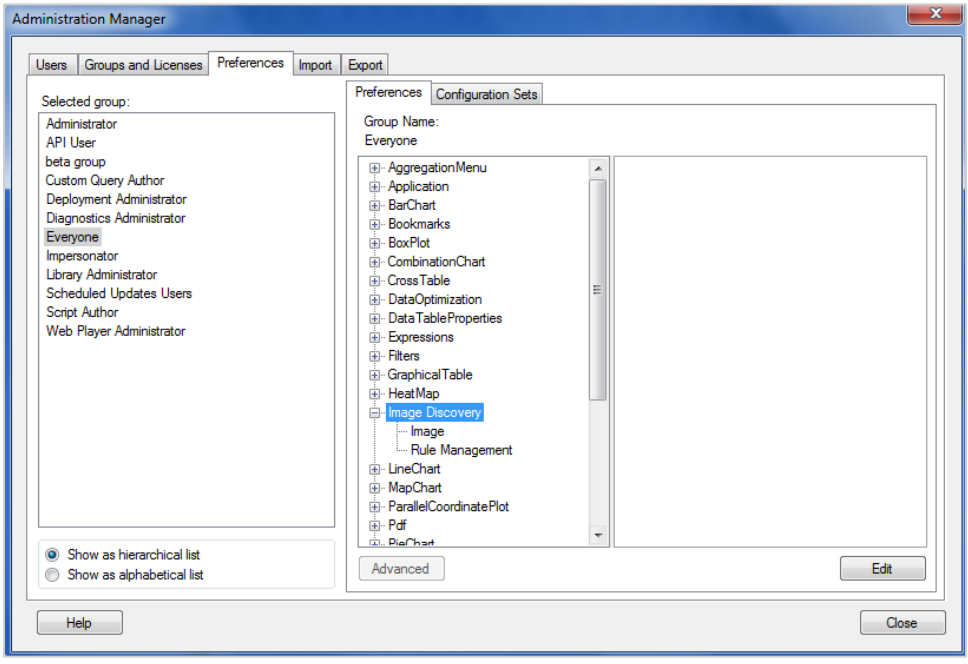
2. Click on the **Preferences** tab.



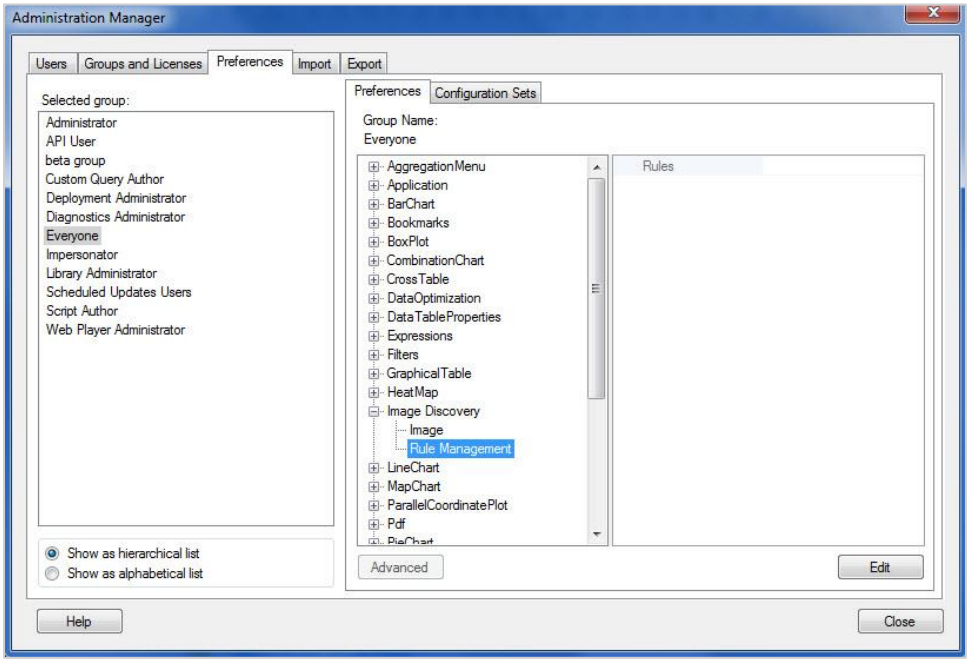
3. Select the **Group Name** to which the group rule should be applied.



4. Select the **Preferences** tab in the right-hand panel. Expand the **Image Discovery** category in the tree.

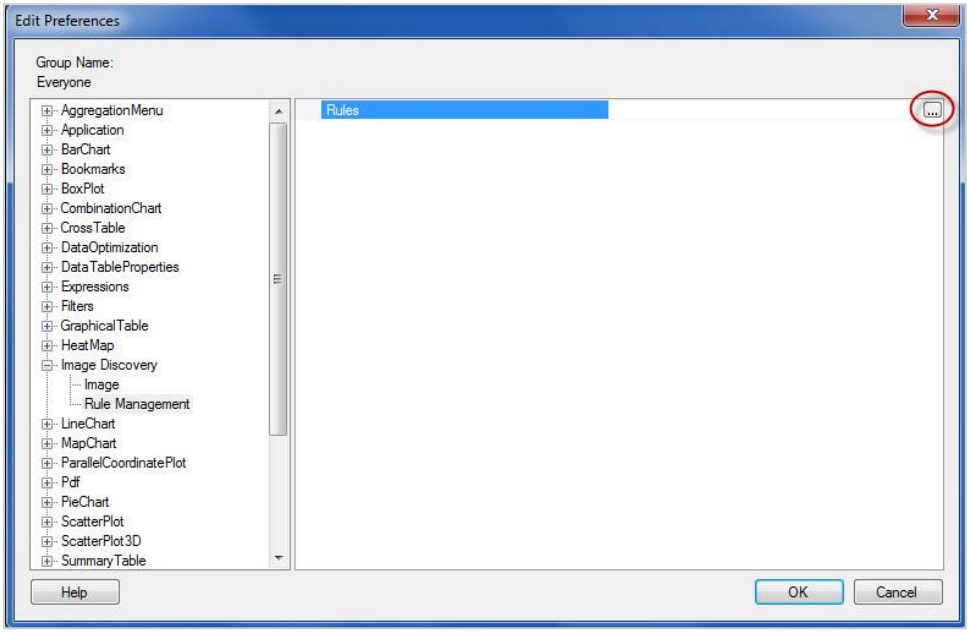


5. Select the **Rule Management** sub-category in the tree.

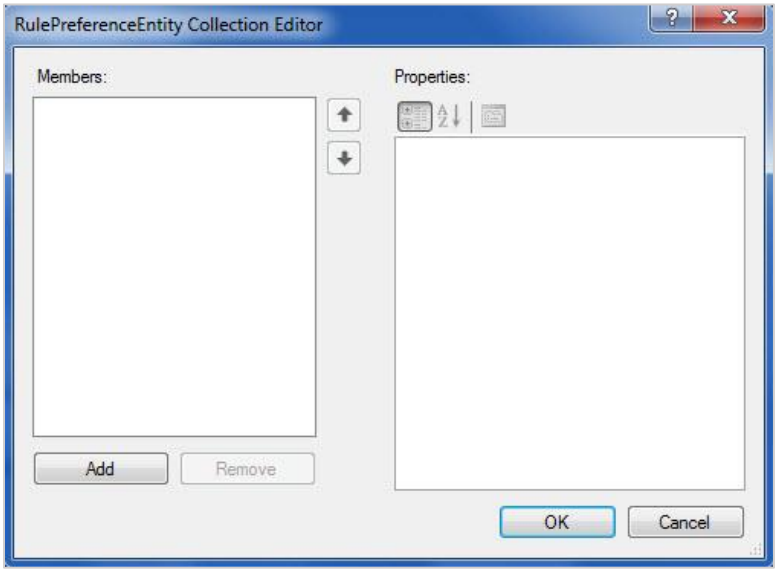


6. Click on the **Edit** button.

The **Edit Preferences** dialog opens.

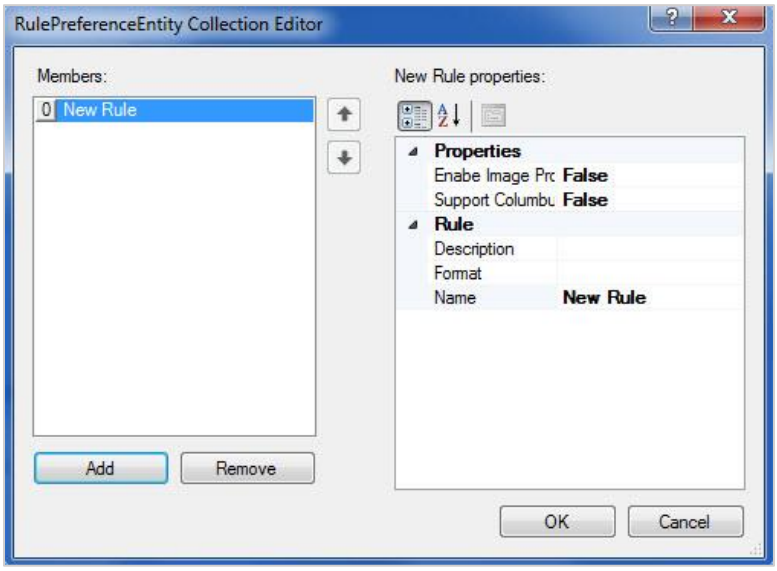


7. Select Rules cell in the grid in the right-hand panel, a … button is displayed. Click on the … button to open the **Rule Preference Entity Collection Editor** dialog, like the example shown below.



From here, you can define new group rules.

8. Click on the **Add** button to add a new group rule. By default, the rule is named ‘New Rule’.



9. Define the Rule Properties (Enable Image Processing, Support Columbus).
10. The default rule name can be modified by entering a unique new name, and an optional description can be entered.
11. Enter the rule format for the new group rule.

Note: It is recommended that you ensure that the group rule name is not duplicated, and that rule format is valid.

12. If applicable, click **Add** to enter another new group rule. Selecting a rule from the Members list and clicking on the **Remove** button will delete the group rule.

By default, the configurable settings of the rules are grouped by categories. The settings can be grouped alphabetically by clicking on the  button located above the grid. Clicking on the  button will re-group the settings by categories.

13. Click **OK** to close the Rule Preference Entity Collection Editor dialog or cancel to discard the changes.

Changes are not saved until you click **OK** in the Edit Preferences dialog. If the **Cancel** button is selected in the Edit Preferences dialog, all changes to the group rules are discarded. Changes to group rules will take effect after you restart Spotfire®.

Every group has its own rule preference settings. If a group is a child group of another group, then the child group's rule settings will take precedence. For example, if Group B is a child Group of Group A, then members of Group B will use the rule settings for Group B. If Group B does not have any rule settings, then the rules settings for Group A will take effect.

If a user belongs to multiple groups, then the rule settings of the Group with higher privileges takes precedence.

9.16.3 Token Management

To retrieve images from external image sources, (i.e. Columbus®), you will need to provide authentication credentials. After successful authentication, a token will be saved for future use.

Currently only Columbus authentication is supported. Other image source authentication may be supported in the future.

Columbus Authentication

When inserting an image column using a rule that supports Columbus, a Columbus Authentication Credentials dialog is displayed. A User Name and Password must be correctly entered for the specified Columbus Server to retrieve images.

You will be notified if the connection fails to complete, or if invalid credentials are entered.

Upon successful authentication, the Authentication Credentials dialog closes, the token is saved, and the image retrieval begins automatically.

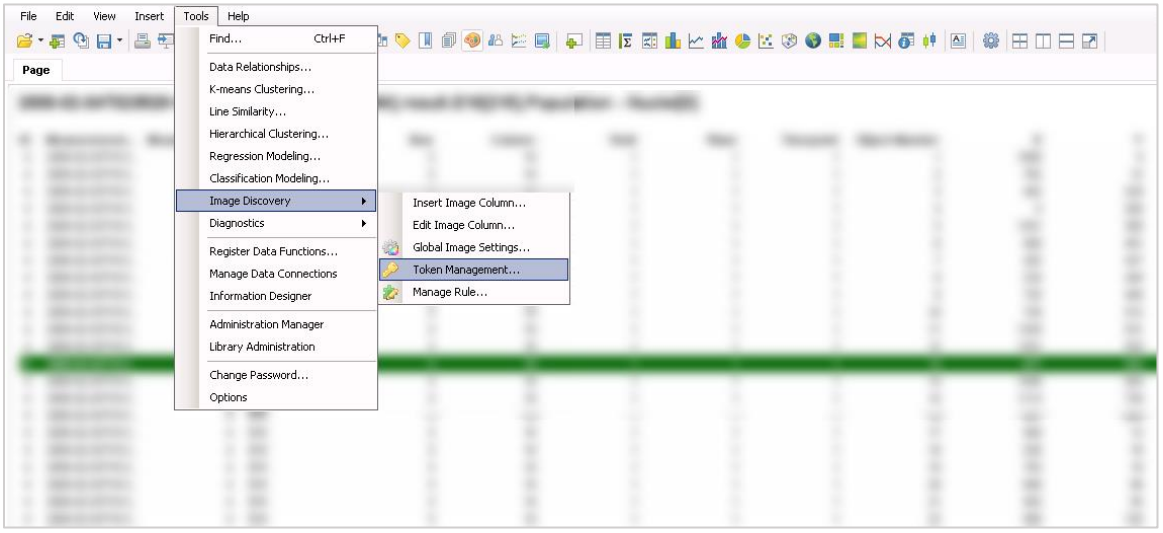
Alternatively, you can click the **Cancel** button to skip the authentication and can use the Token Manager dialog to re-authenticate again.

Token Management

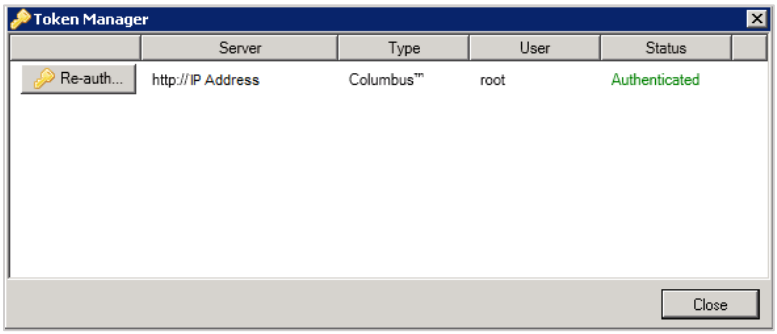
Via the Token Manager, you can switch credentials for an existing token or resume a cancelled authentication.

To access the Token Manager:

1. From the **Tools** menu, select the **Image Discovery>Token Manager** sub-menu item.



The **Token Manager** dialog is displayed.



The dialog lists all tokens for the current document, one line per server. Authenticated servers are listed as **"Authenticated"** in the Status column and highlighted in green. If a server's status is listed as "**Unauthenticated"**, then its status is highlighted in red. In this case, authentication against the server was cancelled by the user.

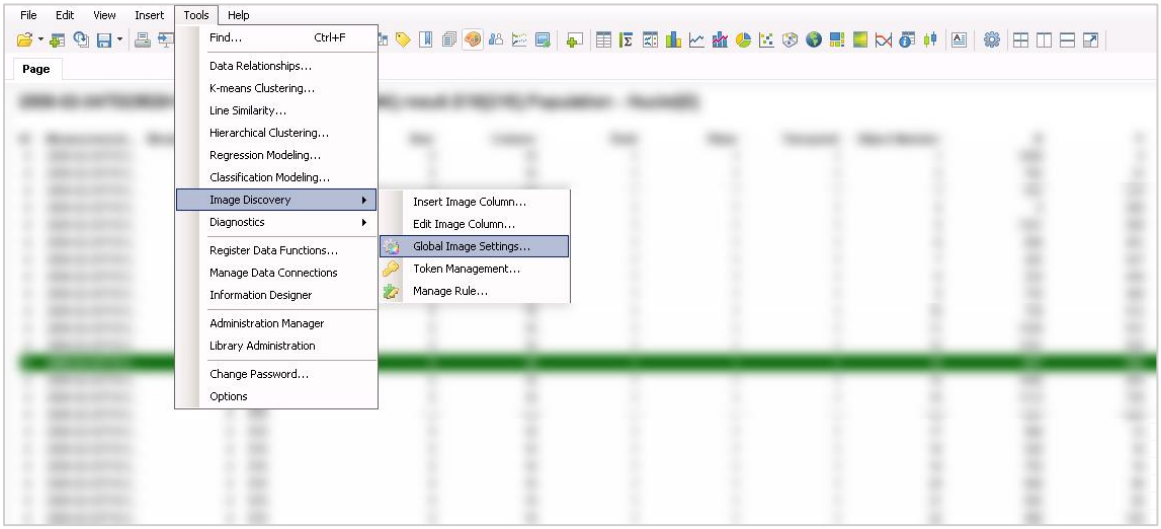
To replace the credentials for an authenticated server or resume a cancelled authentication, click on the **Re-auth…** button for the corresponding server. The Columbus Authentication Credentials dialog will be displayed. From here, you can enter the User Name and Password for the Columbus server.

9.16.4 Global Settings

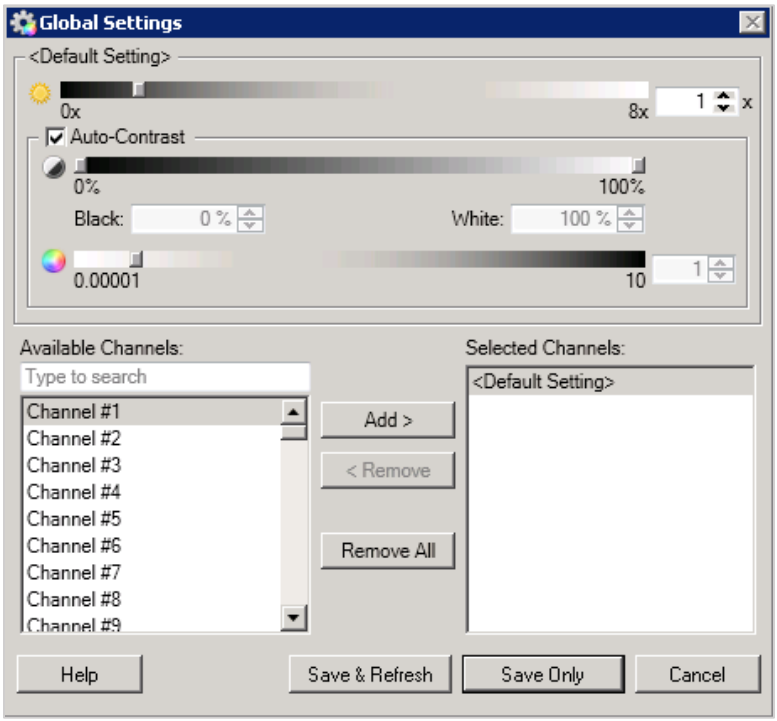
Through the Global Settings, you can define the settings for the image view in a global location and these settings can be used for all images in the current document. Global settings are user specific settings, and therefore are not saved with the document (*.dxp).

To access Global Settings:

1. From the **Tools** menu, select the **Image Discovery >Global Image Settings** sub-menu item.



The **Global Settings** dialog is displayed.



The Global Settings dialog can also be accessed by clicking on the  icon on the Image Settings panel.

From the Global Settings dialog, you can define default global image settings which are saved for the current user. These settings are then applied to all images in the current document.

Note: Customized setting defined through the Image Settings panel may affect the image column for the selected row. New image settings will override any previous settings.

Using the sliders, or the spin controls, the brightness can be set. By default, the brightness is set to 1. The brightness is limited to a maximum value of 8.

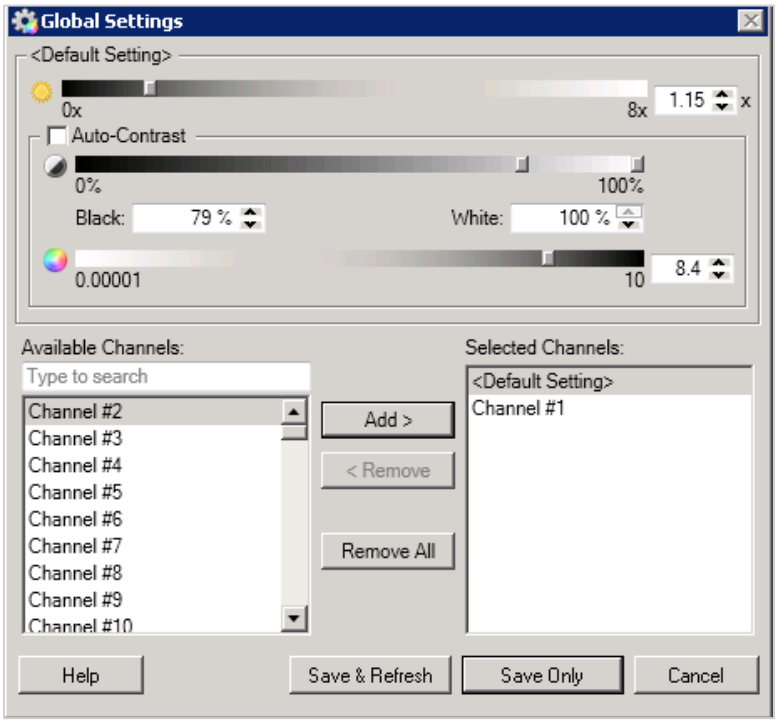
By default, the Auto-Contrast checkbox is enabled. Alternatively, you can disable the checkbox and use the sliders or spin controls to define the black/white contrast (percentage) and gamma setting manually.

By default, the settings are applied to all channels. However, you can select up to 99 custom channels and specify the global settings for each channel. These custom channels are named Channel #1 to Channel #99.

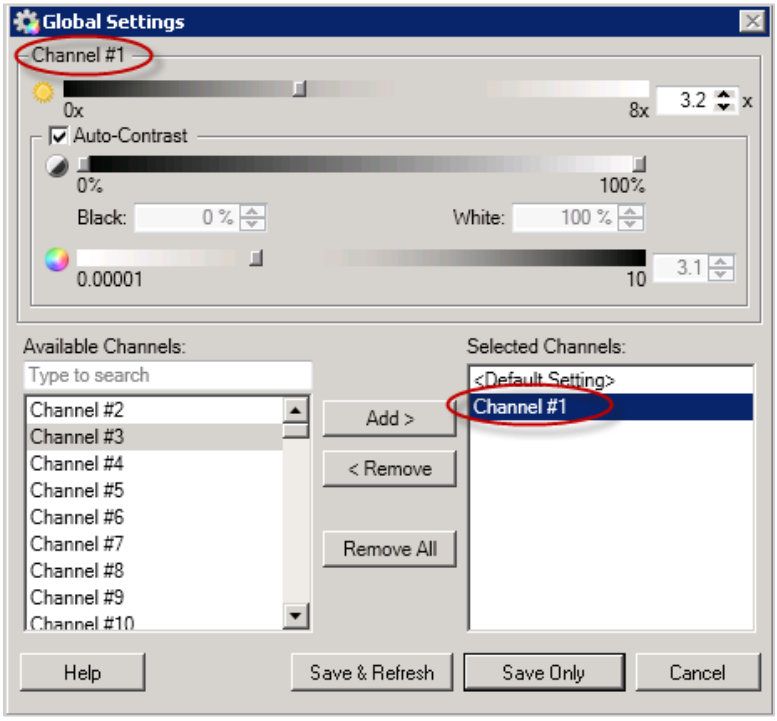
To define custom settings for a channel:

1. From the Available Channel list, select a channel and click on the **Add** button. Alternatively, you can double click on the selected channel. Multiple channels can be added, however only one channel can be added at a time.

The Channel Search field text box (located above the Available Channels list) allows you to locate the channel quickly without having to scroll through the channel list.



The selected channel is added to the Selected Channel list. It is no longer listed in the Available Channels list.



2. Highlight the channel in the Selected Channels list for which you want to adjust the global image settings. The select channel (e.g. Channel #1) is also displayed above the panel settings.
3. Adjust the brightness, and contrast/gamma settings for the selected channel.
4. Continue to add and define custom channels as required.

The **Remove** button removes selected channels from the Selected Channels list and sends them back to the Available Channels List. Except for the **<Default Settings>** channel, the **Remove All** button removes all channels from the Selected Channels list. The **<Default Settings>** cannot be removed. The Default Settings will be applied to all indexed channels that exist in the TIFF images that are not defined as custom channels through the Global Settings dialog.

Note: Although you can define the settings for up to 99 custom channels, not all images have that many channels. Normally each image has 3 (or 4) channels of image data. In this case, global settings for any additional channels are ignored.

Clicking the **Save & Refresh** button will save the settings and apply the global image settings to the current document. Clicking the **Save Only** button will save the settings, but it will not apply them to the current document. To abort any changes, click on the **Cancel** button. All changes will be lost.

9.16.5 Inserting Images Column

Image Discovery allows you to specify the source format of images relevant to the current data set in Spotfire®.

Once a data file is opened in Spotfire®, an image column can be inserted into the data table via the **Tools > Image Discovery > Insert Image Column** sub-menu item in Spotfire® to open the Insert Image Column dialog. From here, you can set a name for the new image column and select a pre-defined rule. You also have the option to set up a new rule from this dialog.

You can set the image source (Local Images Service or Remote Images Service) for the current image column.

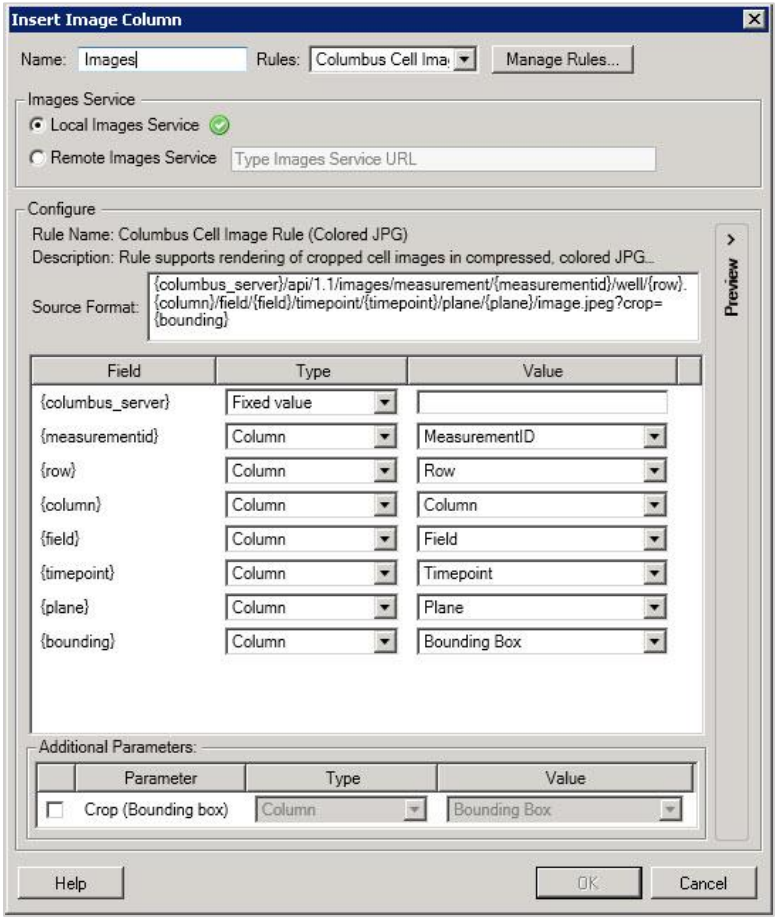
You can configure the selected rule by manually editing the Type/Value for the fields identified by the rule. Modification to the fields here does not impact the pre-defined rule, and it only applicable to the image column currently being inserted.

You can preview the image path from the Insert Image Column dialog by expanding the Preview panel.

Once complete, images for the associated data table can be retrieved via the images service and displayed in the newly inserted column.

To insert an image column:

1. Open the Spotfire® Client and open a document. From the **Tools** menu, select the **Image Discovery > Insert Image Column** sub-menu item to open the **Insert Image Column** dialog, like the example shown below.



Alternatively, you can right-click on any cell in the data table and select the **Image Discovery > Insert Image Column** sub-menu item to open **the Insert Image Column** dialog.

2. Enter a Name for the new column. By default, this column is named Images. Subsequent image columns added to the same table are named Images2, Images3, etc. by default.
3. From the Rules drop down list, select the rule that represents the image data being retrieved. The dropdown list will contain all pre-defined rules, all user rules created by the current user as well as group rules specific to the group to which the current user belongs.

Note: The **Manage Rules** button located to the right of the Rules drop down list will open the Rule Management dialog, allowing you to create new rules, modify and/or delete existing rules. Refer to the Rule Management section for more information.

4. From the Image Service group box, select the Images Service to use to retrieve and process images from the image source. To use the Local Images Service deployed with Spotfire®, enable the **Local Images Services** radio button. A green checkmark will appear once the connection to the image service is verified. If you want to use a remotely deployed image service, enable the **Remove Images Service** radio button. Enter the Images Service URL. The URL will be validated as you are typing. A green checkmark will appear once the connection to the image service is verified. If the connection is unsuccessful, a red stop sign is displayed.

Note: The URL must always be preceded with “http://”.

5. Review the source format for the selected rule and verify that it meets the requirements needed to retrieve images for the current data column from the image source.

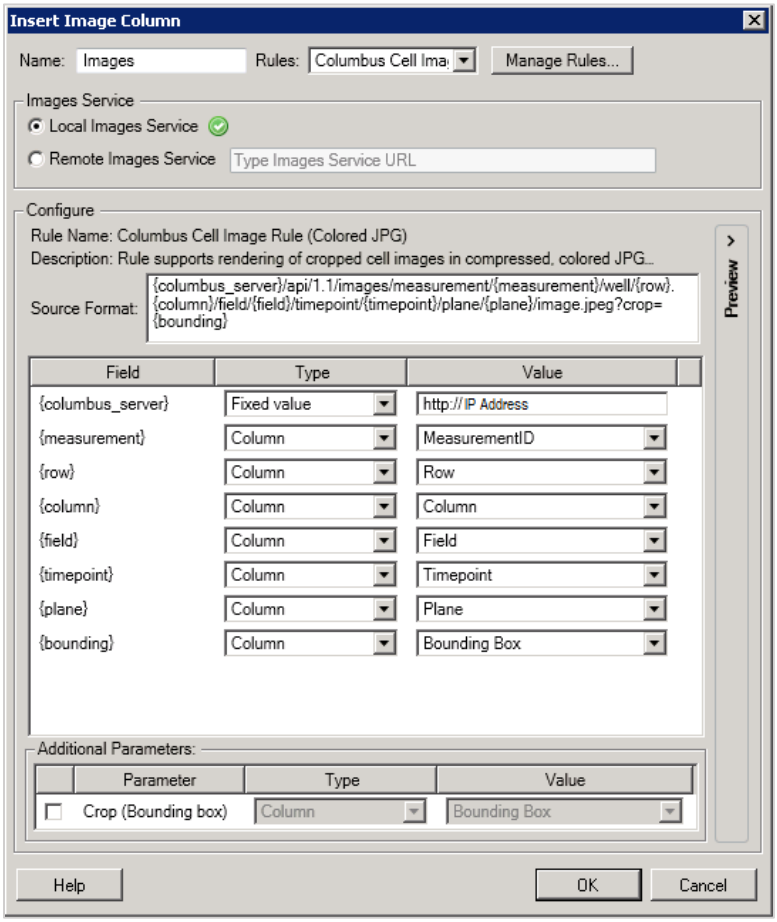
The source format text points to the image source to be retrieved. Content contained within the curly brackets are fields. Each field in the source format will be replaced for each row with either a fixed value or a column related value. Here you can provide each field with a fixed value or choose a column from the current data table for each field.

Fields can be defined to identify the name or location of images, as well as the column(s) in the data table can be set to fields to identify images.

Review the fields list and make any revisions to the field type and/or value. Fields can be added or removed if the rule does not meet the requirements. Modifications made here do not impact the selected rule; they only affect the image column being currently inserted.

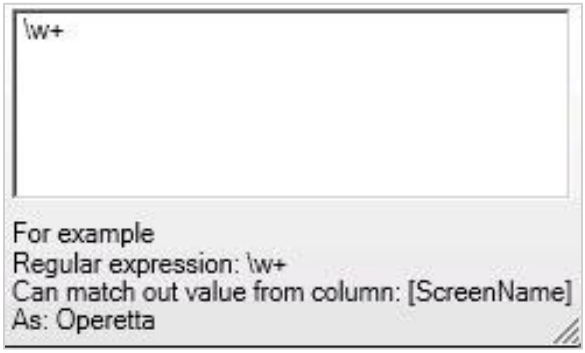
There are four field types available:

Fixed value – user can enter a fixed value for the type. In the example shown here, the Server URL for the Columbus Server is entered.

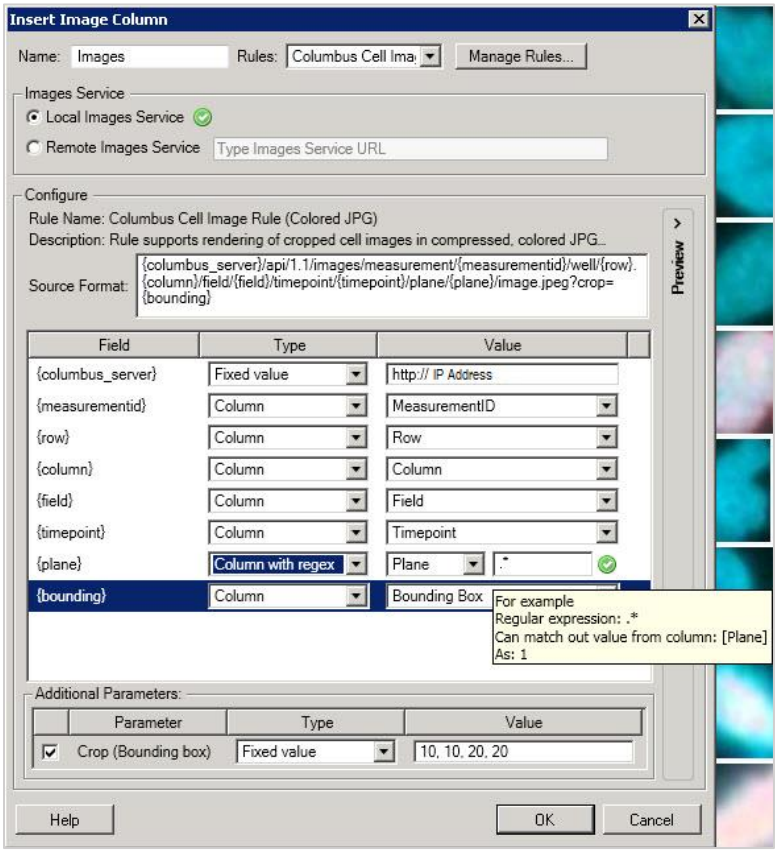


Column – the user can select a column from the data table and select the column value.

Column with regex – the user can select a column from the data table and use C# regex (regular expression) to extract value from the column value. A green checkmark is displayed if a valid regular expression is entered. You can preview the result in a popup dialog while entering the expression, like the example shown below.



In the example below, ".*" indicates to extract all text in the column value.

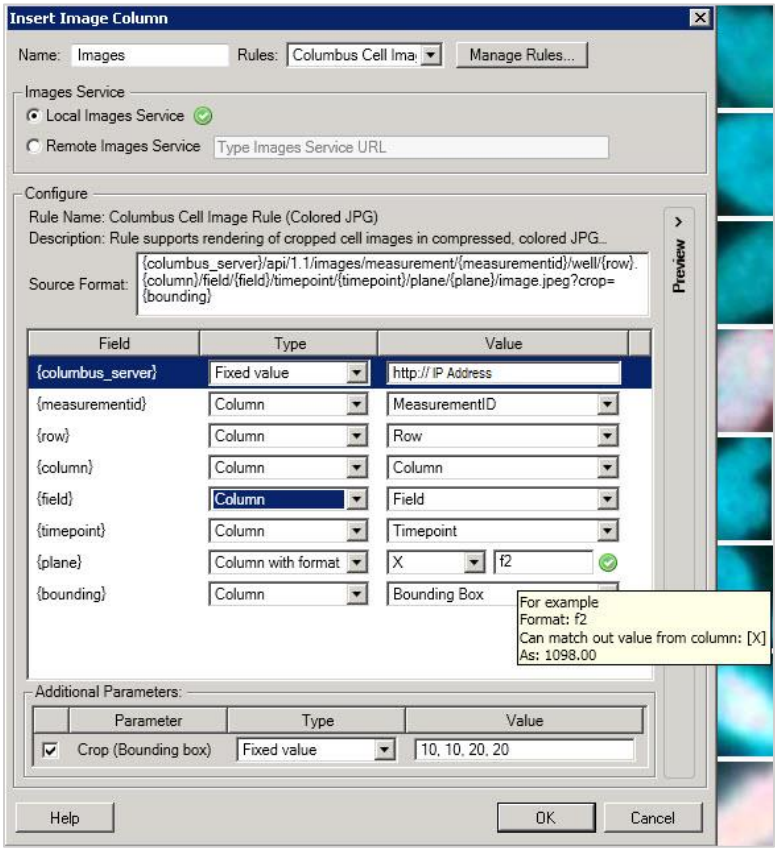


Regular Expression Examples:

Regular Expression	Description
\w	Matches any word character, from ‘A’ to ‘Z’, ‘a’ to ‘z’ and ‘_’.
\w+	Matches any word. A word means combination of at least one-word character.
\W	Matches any non-word character.
\s	Matches any white-space character.
\S	Matches any non-white-space character.
\d	Matches any decimal digit.
\d+	Matches any decimal string.
\D	Matches any character other than decimal digit.

For additional information on regex usage, refer to https://msdn.microsoft.com/en-us/library/az24scfc(v=vs.110).aspx .

Column with format - User can select a column from the data table and use C# formatting string to format the column value. In the example below, “f2” indicates to format the column value as a float value and keep 2-digit decimals.



Format String Examples:

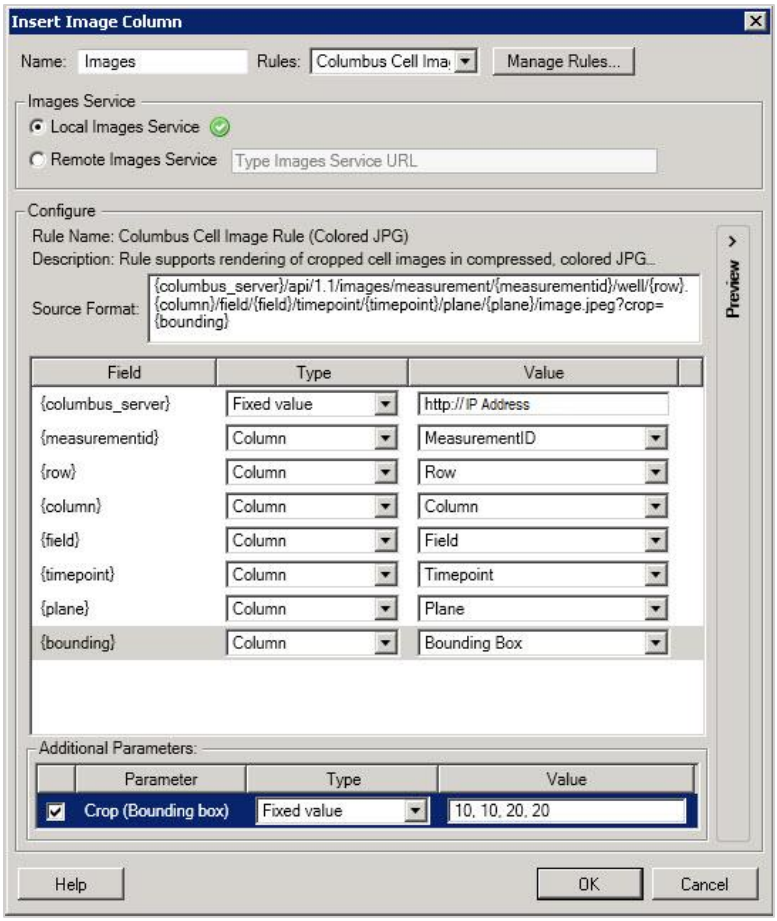
Format String	Description
AAA{0}BBB	Embeds the column value to AAA and BBB. You can replace AAA and BBB to any text applicable.
D6	Format a numerical value to 4 digit integer. (i.e. 1234 will be formatted to 001234 and -1234 will be formatted to -001234). 6 in this example, is just an exemplary value can be modified is required.
F2	Format a numerical value to float and keep 2 digit decimal point (i.e. 12.001 will be formatted to 12.00).

For more details related to format string, please refer to https://msdn.microsoft.com/en-us/library/dwhawy9k(v=vs.110).aspx and https://msdn.microsoft.com/en-us/library/az4se3k1(v=vs.110).aspx

6. Add Crop (bounding box) via the Additional Parameters group box. Enabling the Crop (Bounding box) checkbox allows you to specify the type and value. This option is only applicable if the selected rule has the 'Image Processing' property enabled (e.g. supports TIFF, JPEG and BMP).

Note: There are two types of image cropping available. If defined through the [Bounding Box] column in the rule's source format, the cropping parameter is sent to the Columbus Server, and is cropped on the Columbus Server. If the Crop (Bounding box) checkbox is enabled in the Additional Parameters group box, the cropping parameter is

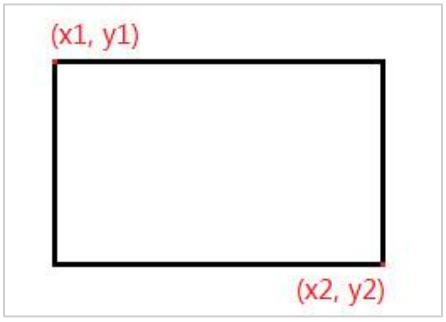
sent to the Image Service, and the cropping process occurs locally. Both cropping options can be defined, however in most cases the [Bounding box] column is used so the image is only cropped once.



The following format is required:

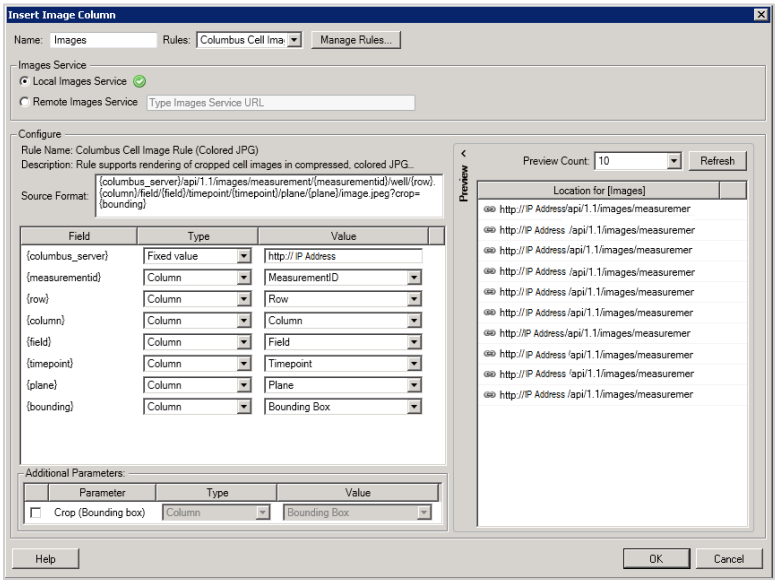
x1, y1, x2, y2. There parameters can be contained within the following: { }, [ ], or ( ).

(x1, y1) and (x2, y2) are the coordinates of the top-left, and bottom-right vertexes of the expected cropping rectangle.



Previewing the Image Paths

Expand the Preview side panel to display the image paths of the images being retrieved from the image service for the current data column.



Hover over the image path to view the path in its entirety. From the Preview Count drop down list, you can specify to view 10, 20 or 50 image paths Location for [Images] view. Click on the **Refresh** button to update the display after changing the Preview Count.

Note: When clicking **OK** to insert the image column, by default, two columns will be inserted; the image column, as well as the Location for [Images] column. By default, the Location for [Images] column is a hidden column. You can select to view the Location for [Images] column in the current data table.

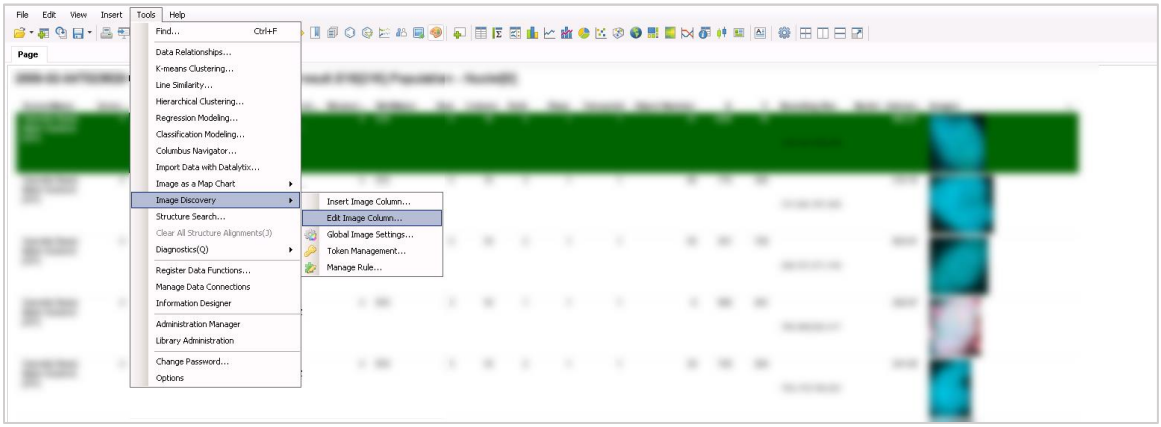
To view Location for [Images] column in the current data table:

1. Right-click on the table visualization and select **Properties**.
2. Select **Columns** in the left-hand panel. From the **Available columns** list, add the Location for [Images] column to the **Selectedcolumns** list
3. Click **Close**.

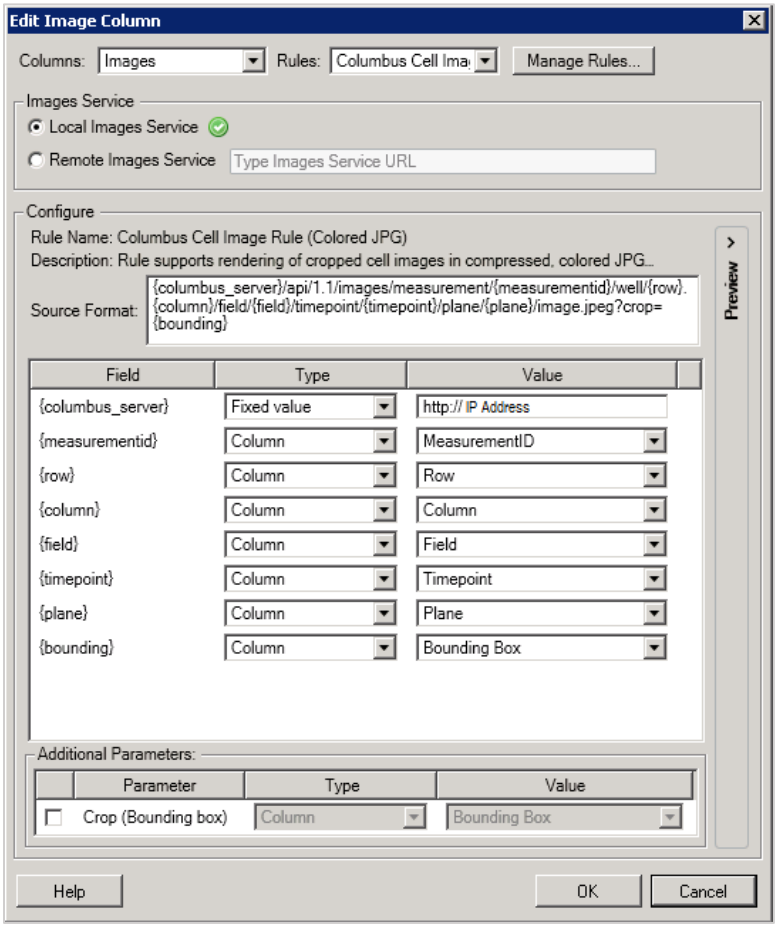
9.16.6 Editing Images Column

To edit an image column:

1. Open the Spotfire® Client and open a document with an Image column(s). From the **Tools** menu, select the **Image Discovery >Edit Image Column** sub-menu item.



The Edit Image Column dialog is displayed like the example shown below.



Alternatively, you can right-click on any cell in the data table and select the **Image Discovery >Edit Image Column** sub-menu item to open the **Edit Image Column** dialog.

To edit an image column:

1. From the Columns drop down list box, select the image column to edit.
2. Make the necessary changes to the Image Column configuration. Refer to Inserting Images Column.

Note: The rule being edited here is the rule saved in the images column when it was inserted and may differ from the original rule. Any changes made to the rule's source format via the Manage Rules (Rule Management) dialog will not affect a saved rule.

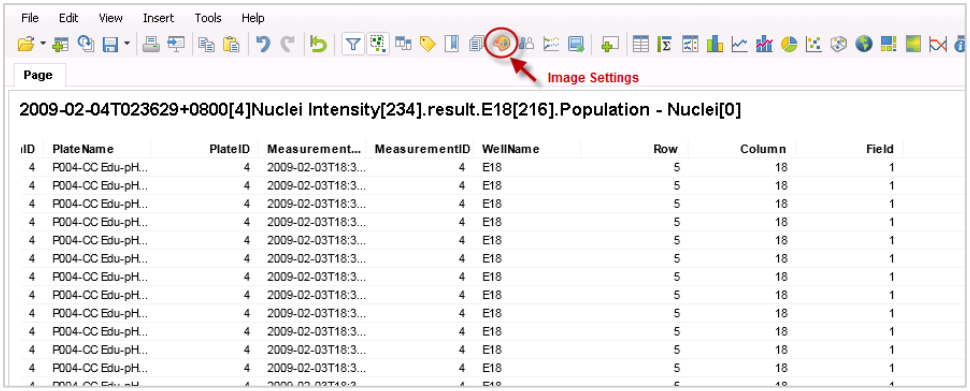
9.16.7 Image Settings

Via the Image Settings panel, you can define the image rendering settings for specific images within a specific image column in the data table. These settings can be defined on a per image basis, and do not have to apply to other images in the table and do not affect the Global Settings.

To navigate to the Image Settings panel:

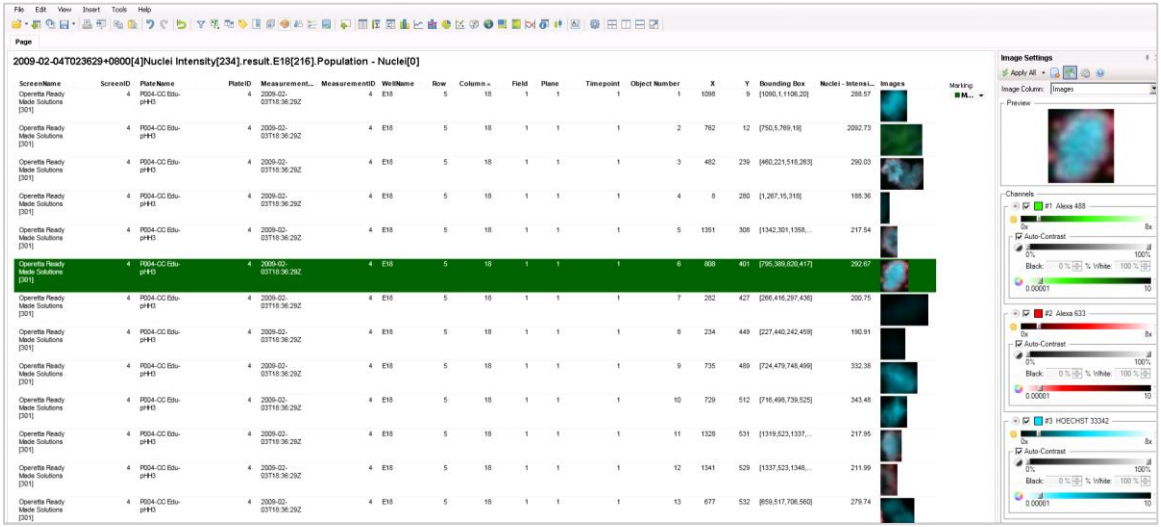
1. Open a document containing at least one image column.

2. In Spotfire®, click the **Image Settings** button  in the main toolbar.

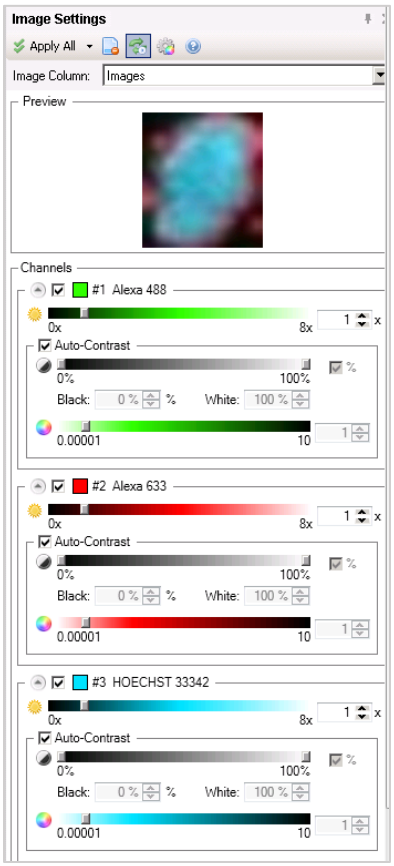


In Spotfire® 10.3 or later, open **View > Image Settings.**

The Image Settings panel opens, by default, to the right-hand side of the data table, like the example shown below. By default, the Image Settings panel opens above the Filters panel.



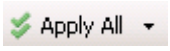
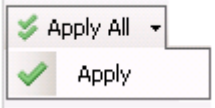
From here, you can define image settings such as channel selection, channel color, brightness, and contrast and gamma settings. When the Image Settings panel is opened for the first time, the image settings are reflective of those defined by the Global Settings.



About the Image Settings Toolbar

The image below shows the toolbar for the Image Settings panel.

The following table explains the icons found on the Image Settings toolbar, as shown in the image above.

Button / Control	Description
	**Apply All** - Applies the image settings to all images in the selected image column. Settings are persisted and the image is re-rendered when the user clicks on another image in the column.
	**Apply** – Applies the image settings to the selected images in the selected image column
	**Auto Apply Settings** – changes to the image settings will take effect immediately in the image column when row selection is changed.

	If the Auto Apply Settings is not triggered, changes to the image settings will not take effect until the **Apply** or **Apply All** button is clicked.
 	**Auto Refresh Preview** – by default, the image preview is refreshed automatically when the image settings are changed. If the user selected the Auto Apply Settings above, the Auto Refresh preview will change to  and a **Refresh Preview** button will be displayed above the Image Preview area. You will need to click on the **Refresh Preview** button to see the effects of the image settings in the Preview area. If **Auto Apply** button is selected again, the Auto Refresh Preview will be triggered again, and the **Refresh Preview** button will no longer be available for selection.
	Opens the Global Settings dialog
	Opens the Help window for the Image Settings panel.

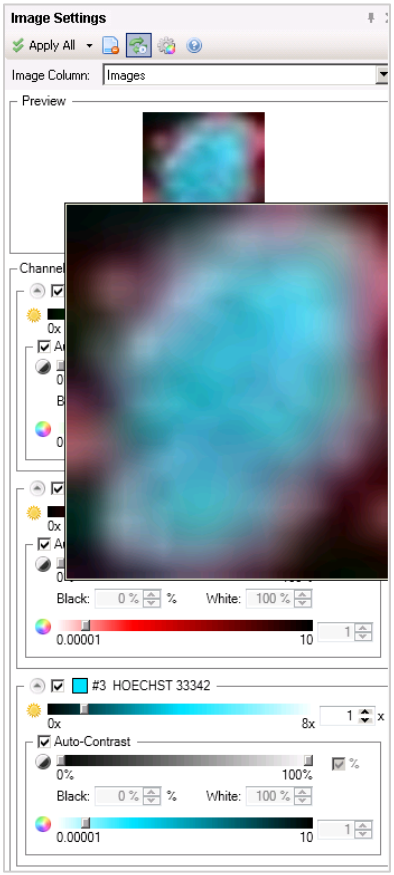
Image Column

Directly below the Image Settings toolbar, there is an **Image Column** drop down list containing all the image columns that have been added to the current data table. The image settings are only applicable to the selected column.

Previewing the Image

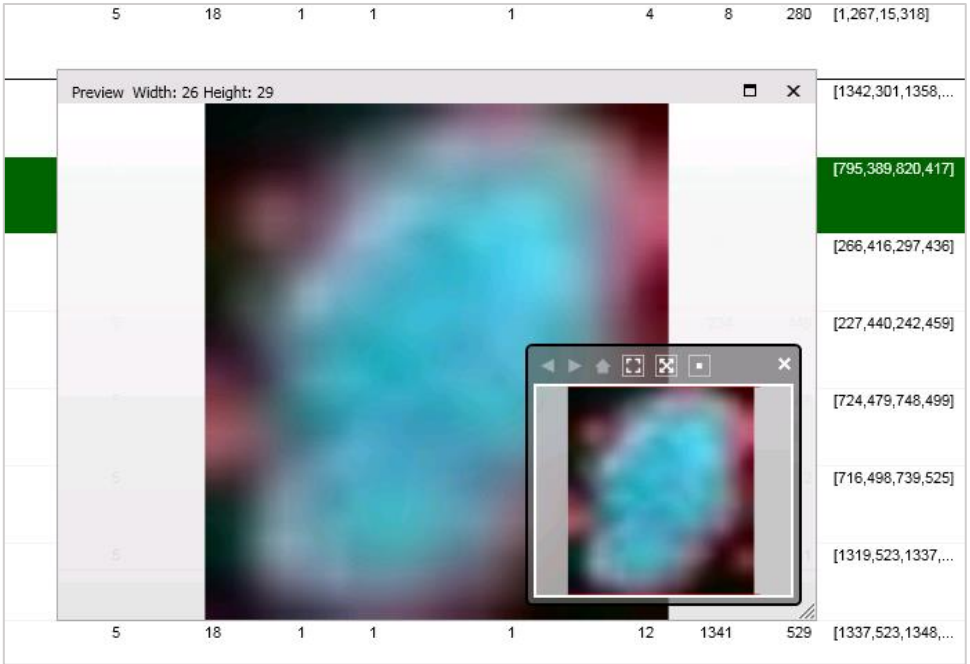
The selected image for the selected Image column is displayed in the Preview area. Based on the selections made in the toolbar (Auto Apply Settings, Auto Refresh Preview), the image will be updated accordingly.

Hovering over the image will zoom in on the image, like the example shown below.



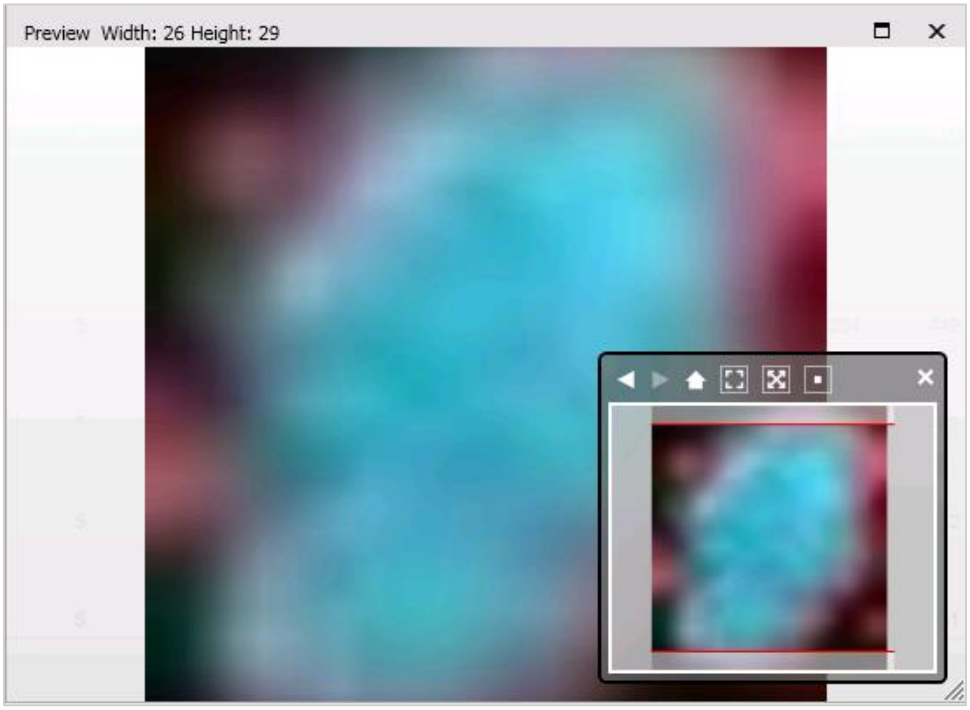
Using the Image Viewer

Clicking on image in the preview panel will open the **Image Viewer** for the selected image. You can re-size the image viewer window as required.



By default, there is a **View Finder** window in the lower-right-hand corner. From here, you can zoom in/out to see the image in greater detail.

By holding down the Shift key, and simultaneously scrolling the mouse wheel, you can zoom in /out of the image. If the image is too large to be displayed in the Image Viewer, a red box is displayed in the View Finder, identifying the area that is currently displayed in the Image Viewer. You can drag the red box to the area of focus.



About the View Finder Toolbar

The image below shows the toolbar for the View Finder.

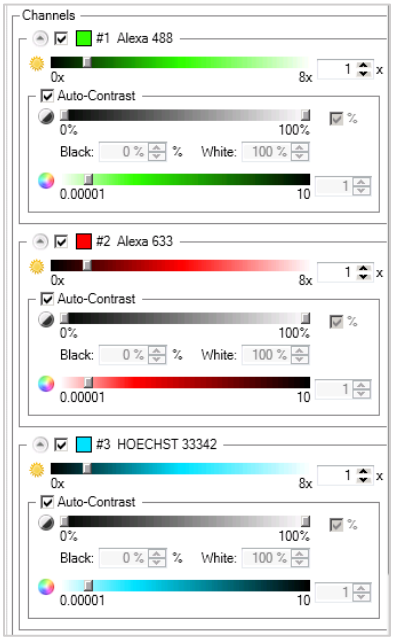
The following table explains the icons found on the View Finder toolbar, as shown in the image above.

Button / Control	Description
	Go Back – reverts the Image Viewer to its last current state. Go Forward – undoes the previous Go Back operation.
	Home- restores the image to its original state.
	Fit Content within Bounds – zooms into the image to fit the current Image Viewer window size.
	Fill Bounds with Content – zooms the image to fill the current Image Viewer window size.
	Center Content – the center of the image is displayed.

	Closes the Image Finder. To re-open the View Finder, hover the mouse over the lower-right-hand corner of the Image Viewer. The Show View Finder  icon will be displayed. Click on the icon to open the View Finder.

Channel Settings

In the Channels grouping of the panel, all the channels of the first selected image will be listed. You can use the expand ()/fold () buttons to expand/fold the detailed setting of channels.

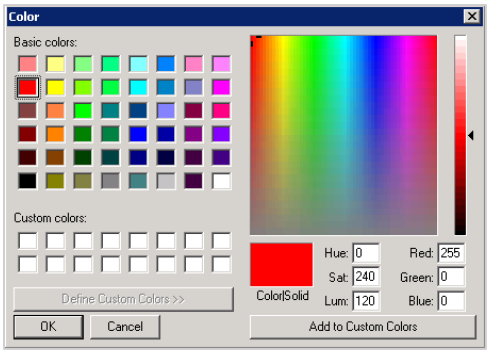


Next to the expand/fold button is the channel selection checkbox (). If the checkbox is enabled, the corresponding channel is enabled and will participate in rendering of the merged image nor shown in image thumb panel.

Note: If a channel is disabled, then all setting controls for that channel will be disabled. If all channels are deselected, then a black image will be displayed in the Preview area.

To the right-hand side of the channel selection checkbox, is the channel color selector (). By default, it displayed shows the false-color of the channel stored in the image.

You can change the channel color by clicking on the channel color selector to open a standard Windows color picker dialog.



Select the color and click **OK.** The selected color will be applied to the selected channel.

Using the sliders, or the spin controls, the brightness can be set. By default, the brightness is set to 1. The brightness is limited to a maximum value of 8.

By default, the Auto-Contrast checkbox is enabled. Alternatively, you can disable the checkbox and use the sliders or spin controls to define the black/white contrast (percentage) and gamma setting manually. You can adjust the black/white contrast using absolute value by disabling the Percentage checkbox and setting the absolute values using the spin controls.

Note: The Black (Minimum B/W value) is always 0. The White (Maximum value) is dependent on the image's format (bit depth of the image).

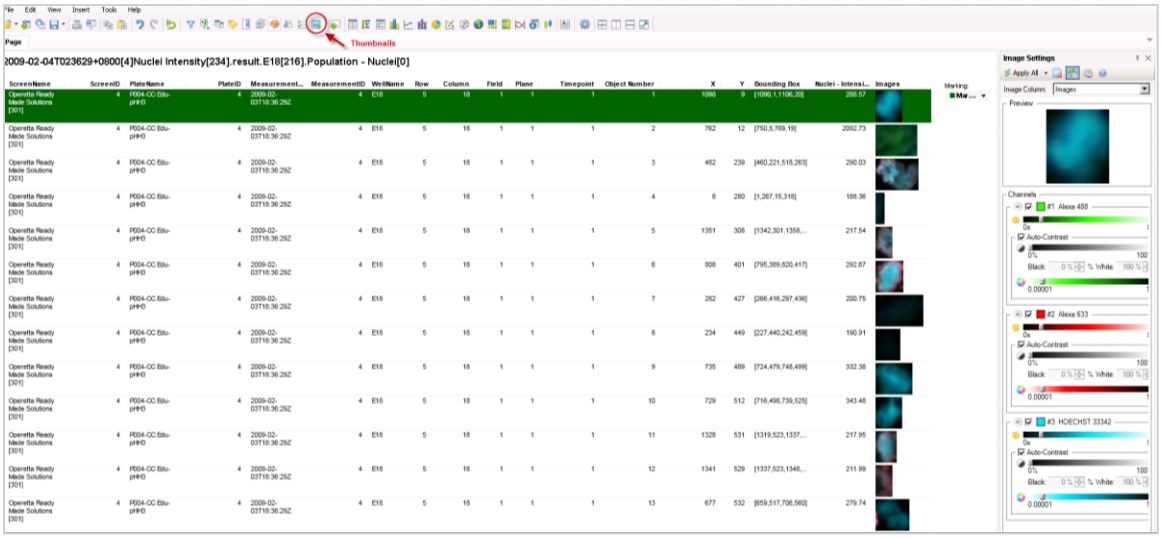
9.16.7.1 Image Thumbnails

Via the Thumbnails panel, you can render cell specific thumbnails for a selected row in the data table.

To navigate to the Thumbnails panel:

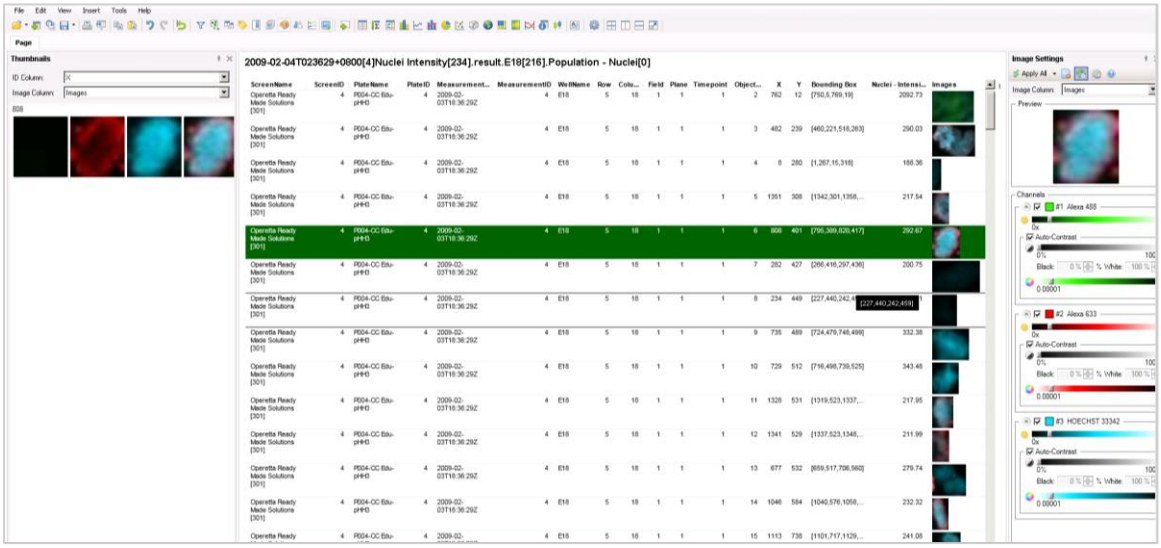
1. Open a document.

2. In Spotfire®, click the **Thumbnails** button  in the main toolbar.



In Spotfire® 10.3 or later, open **View > Thumbnails**.

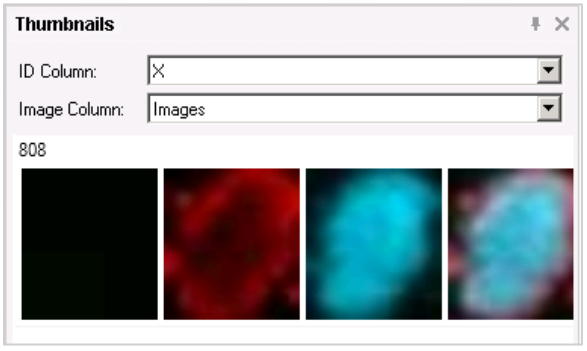
The Thumbnails panel opens, by default, to the left-hand side of the data table, like the example shown below.



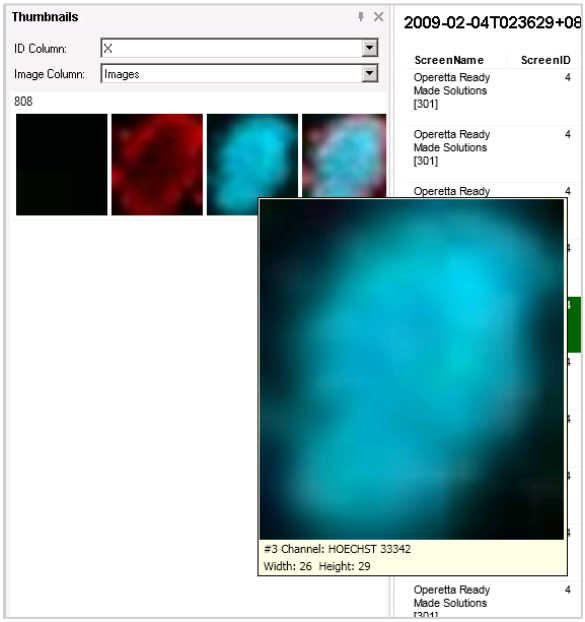
Viewing the Thumbnails

The Thumbnails panel automatically detects the ID Column and will display it in the ID Column drop down list by default. Alternatively, you can choose to change the ID column by selecting another column from the dropdown list.

If the current document has one or more image columns, the image columns are listed in the Image Column drop down list on the Thumbnails panels. You can select one or multiple image columns by enabling the respective checkbox for each image column.



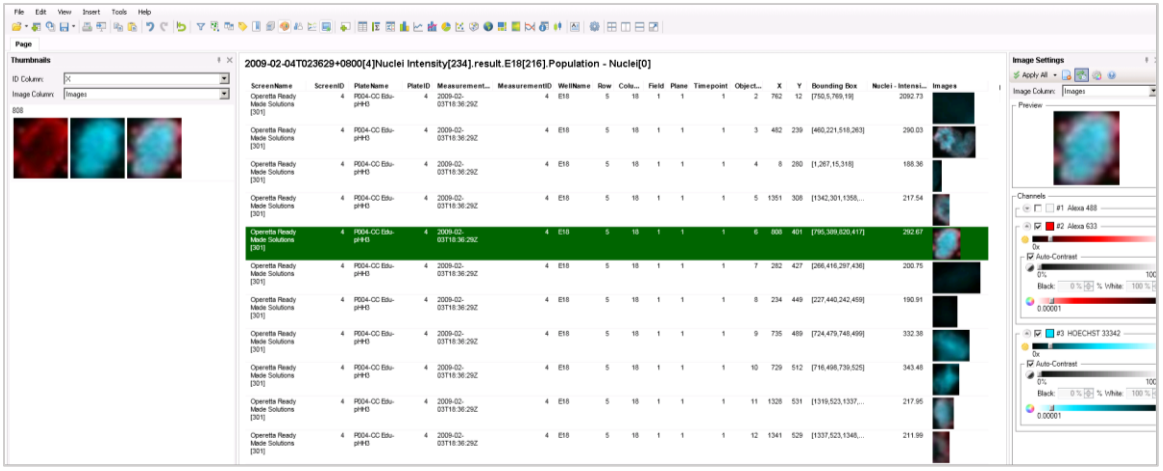
Hovering over a thumbnail will display a preview of the thumbnail along with the name of the channel and the dimensions (resolution) of the thumbnail, like the example shown below.



When one image column is selected, and the selected images in the column have multiple channels, the thumbnails of each channel will be displayed by Channel index, as well as a thumbnail representing a merged image of all the channels.

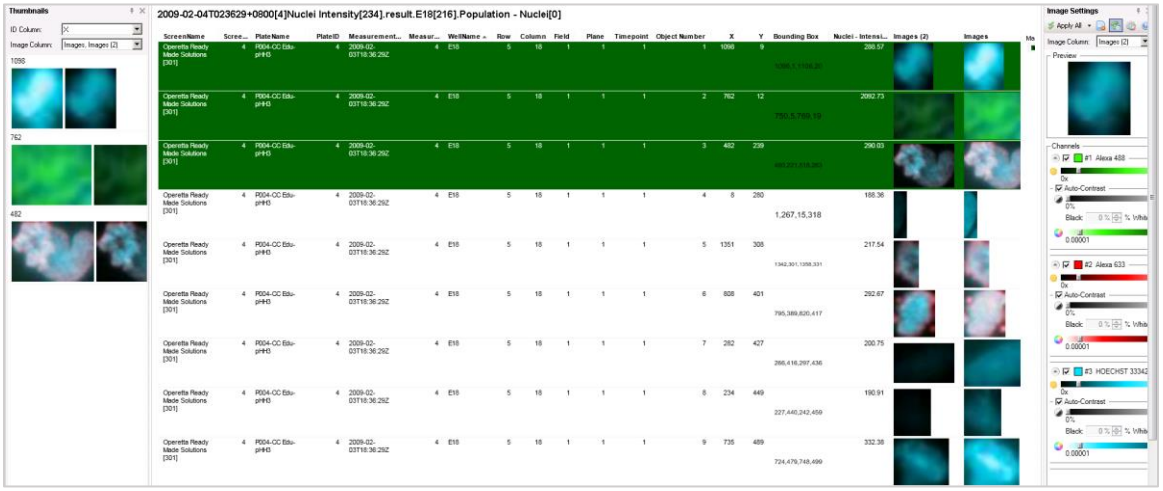
For example, a selected image column has three channels (red, green, blue). A thumbnail is displayed for each channel as well as a merged image of all three channels.

You can choose to disable one or more channels of the selected images through the Channel Settings in the Image Settings panel. In this case, the thumbnail for the disabled channel will not be displayed in the Thumbnail panel, like the example below.



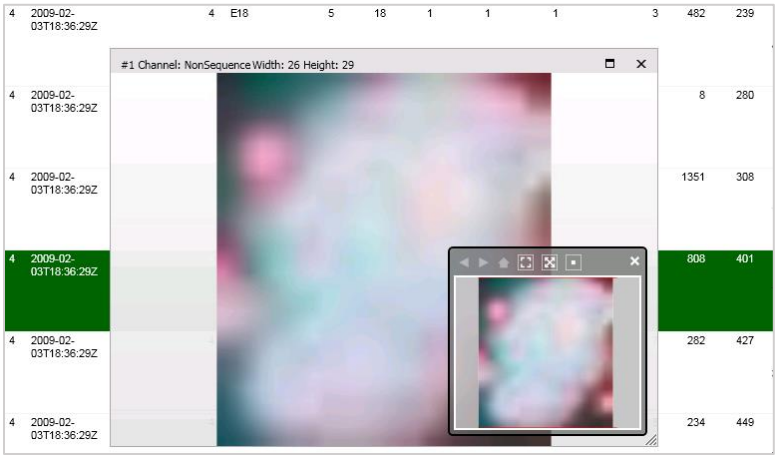
You can select multiple image columns in the Thumbnail panel to display the images in multiple columns. For example, if three image columns have been inserted into the data table which displays three different fields of an image, you can select all three image columns in the Thumbnails panel to display together in one row.

When multiple image columns are selected, and the selected image columns have multiple channels, only one thumbnail per column (merged image of all channels). In the following example, two image columns are selected, and two thumbnails per row are displayed.



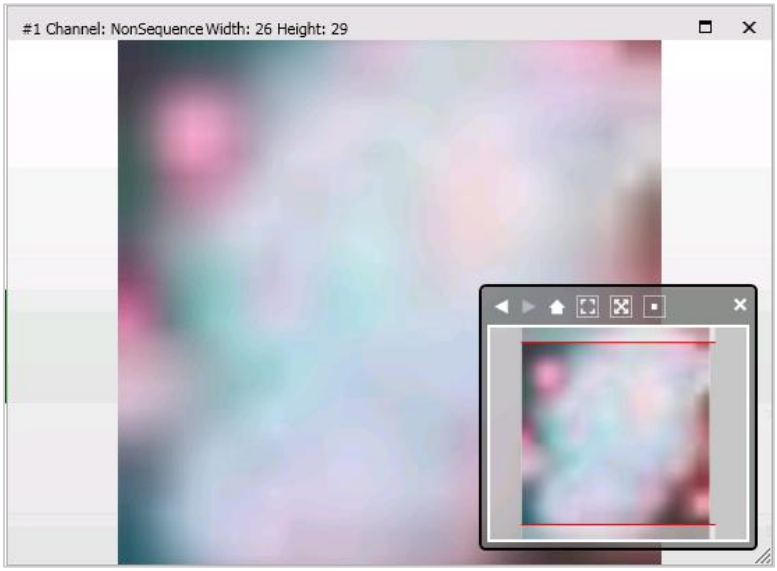
Using the Image Viewer

Clicking on one of the thumbnails in the Thumbnail panel will open the **Image Viewer** for the selected image. You can re-size the image viewer window as required.



By default, there is a View Finder window in the lower-right-hand corner. From here, you can zoom in/out to see the image in greater detail.

By holding down the Shift key, and simultaneously scrolling the mouse wheel, you can zoom in /out of the image. If the image is too large to be displayed in the Image Viewer, a red box is displayed in the View Finder, identifying the area that is currently displayed in the Image Viewer. You can drag the red box to the area of focus.



About the View Finder Toolbar

The image below shows the toolbar for the View Finder.

The following table explains the icons found on the View Finder toolbar, as shown in the image above.

Button / Control	Description
	Go Back – reverts the Image Viewer to its last current state. Go Forward – undoes the previous Go Back operation.
	Home- restores the image to its original state
	Fit Content within Bounds – zooms into the image to fit the current Image Viewer window size.
	Fill Bounds with Content – zooms the image to fill the current Image Viewer window size.
	Center Content – the center of the image is displayed.
	Closes the View Finder. To re-open the View Finder, hover the mouse over the lower-right-hand corner of the Image Viewer. The Show View Finder  icon will be displayed. Click on the icon to open the View Finder.

Displaying Thumbnails

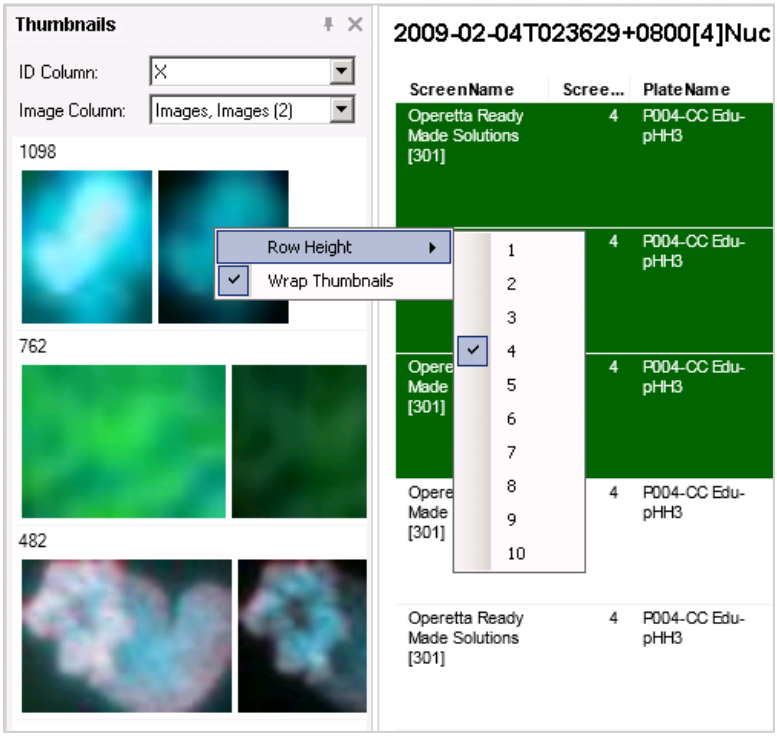
You can adjust the display of the thumbnails in the Thumbnail panel by adjusting the row height of the thumbnails and specifying if the thumbnail display should be wrapped.

Adjusting Row Height

When a Thumbnail panel is opened, the row height will automatically default to 4.

To adjust the row height:

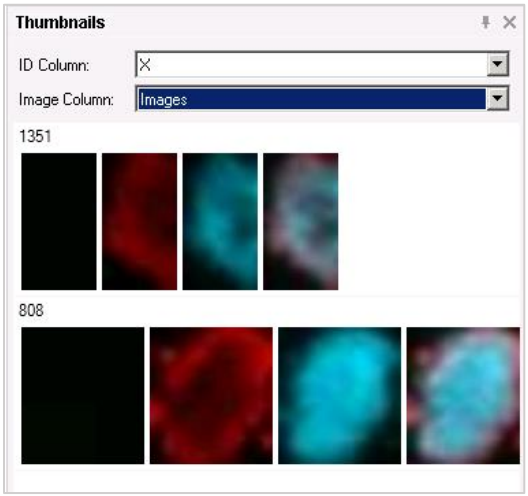
1. Right-click on a thumbnail.
2. Select the **Row Height** menu item and select the row height.



The row height will be adjusted for all thumbnails in the panel.

Wrap Thumbnails

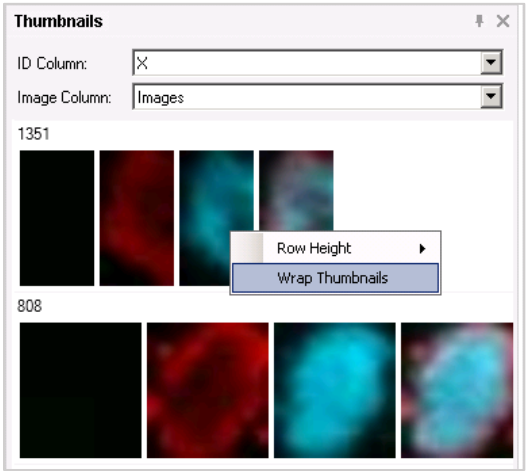
By default, thumbnails are not wrapped and are displayed in a singular row in the Thumbnail panel, like the example shown below.



Alternatively, you can choose to display the thumbnails with wrapping in the Thumbnails panel.

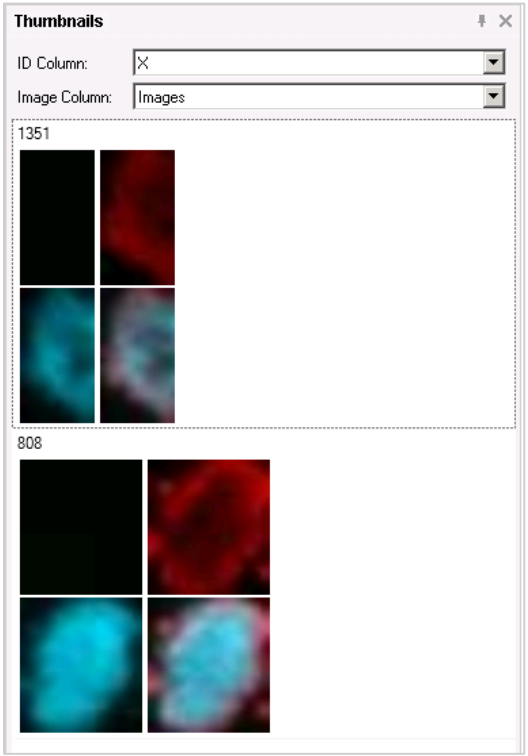
To display thumbnails with wrapping:

1. Right-click on a thumbnail.



2. Enable **Wrap Thumbnails** menu item.

The Thumbnails for each row in the data table are wrapped in the Thumbnail panel.



9.16.7.2 Image Renderer Settings

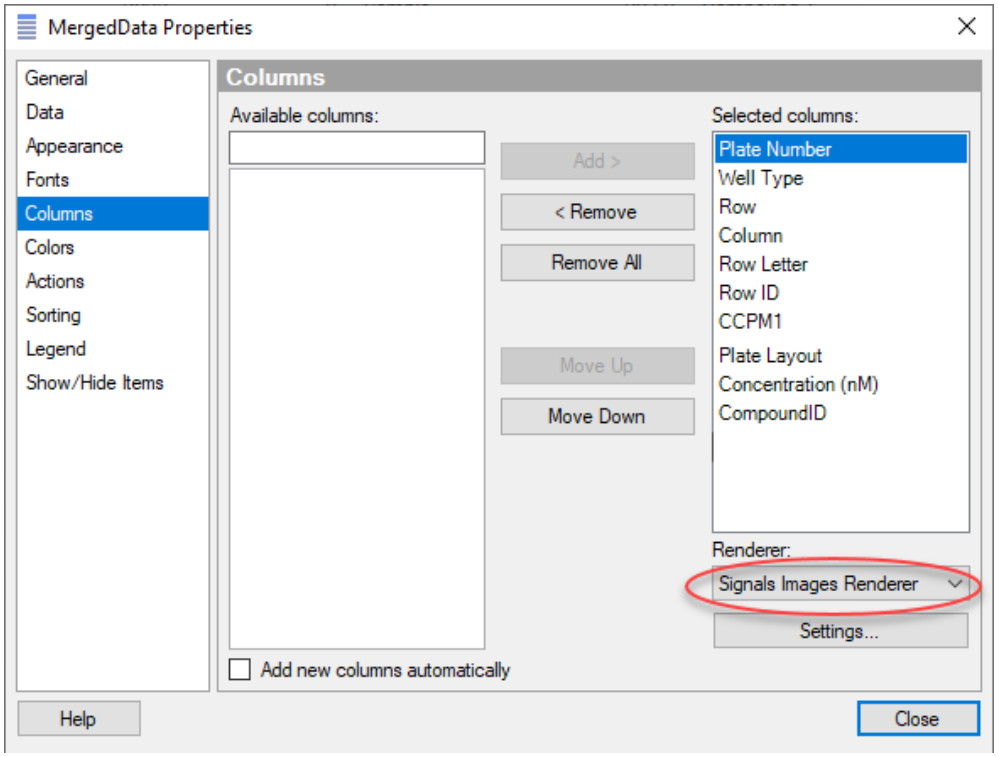
By default, Image Discovery uses the customized Spotfire® renderer to render images. Renderer settings may affect how the images are rendered in Spotfire®.

Renderer settings can be modified through the **Images Renderer Setting** dialog.

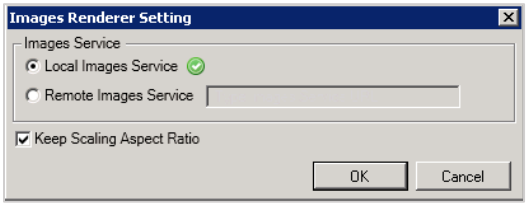
Note: Refer to the Spotfire® online help for general information about renderer selection and settings.

To access Images Renderer Setting dialog:

1. Right-click on a table visualization and select **Properties**.
2. Select **Columns.** From the Selected Columns list in the right-hand panel, select one Image Column. The renderer for the selected Image Column is displayed as '**Signals Images Renderer'**.



3. Click on the **Settings** button to open the **Images Renderer Setting** dialog.



From here, the Images Service can be modified.

The Keep Scaling Aspect Ratio checkbox is enabled by default. This enables the original ratio of the image to be maintained when the image is rendered. Disabling this checkbox will ignore the original image ratio and may result in the image being stretched when rendered.

9.16.8 Appendix A - Troubleshooting

Q1: Why is the Local Image Service unavailable?

A: The local Image Service is launched by the Image Discovery extension; it requires one network port to provide service for Image Discovery.

Windows Vista and higher operating systems, by default, enable User Access Control (UAC). This may block the Image Service from accessing a port. Image Discovery will automatically launch one script to enable ports for the Image Service. If the script execution fails, open Windows command line as an Administrator (right-click on command line icon, choose Run as administrator), and run the following command to enable a port for the Image Service:

netsh http add urlaclurl=http://127.0.0.1:<Port Number>/ user=Everyone

The port number should be 9251 to 9261; therefore, it is necessary to run this command ten times for port 9251 to 9261.

Note: When using a language setting different from English, the user “Everyone” may have a different name in that language

Q2: Why can’t I see the image from my shared network folder in Image Discovery?

A: The following checkpoints should be verified:

- Check the rule format and preview the image path result. Verify that the image can be accessed via the path.
- Check the image rule. If ‘Enable Image Processing’ option is enabled, it only supports the processing of TIFF, JPG, and BMP formats. It cannot process GIF and PNG format.
- Check if the network share folder requires authentication. Open the shared folder in Windows Explorer and enter your user name /password to login. Refresh the image column in Spotfire®.
- Check local system C: disk. Ensure that there is at least 12GB of free space.

How do I refresh / re-render the images in the Thumbnail panel or image column??

Q3: How do I refresh / re-render the images in the Thumbnail panel of the image column

A: Due to network timeout or a system configuration problem, the images in the image column may not render. It may be necessary to refresh the image data, and request Spotfire® to re-render the images.

There are several ways to refresh images in an image column:

1. Open **Image Discovery > Edit Image Column** dialog. Select the image column to refresh and click **OK** without making any changes to the dialog. The image column will be refreshed.
2. Changing the image column row height can re-render the images.
3. Re-loading data through Spotfire® native functionality can refresh and re-render the images in the document.

Note: Only images in unfrozen columns will be refreshed.

Q4: Why is it recommended to save data in the data table as embedded data?

A: Image columns in a document are generated by a calculation function; therefore, changes made to the images in the image column may be lost when data is re-loaded or re-calculated. If a user saves and re-opens the document, the calculation function will be re-calculated, and changes will be lost. Saving the data as embedded data can maintain the custom image rendering settings such as channel selection and brightness/contrast/gamma values.

Image Discovery will automatically detect and prompt a warning when saving data linked to original data source. It is recommended that you save the data as embedded data for images.

Q5: Why is the image display in Image Setting panel different from the image displayed in the visualization?

A: The image displayed in the Images Settings preview panel may be different from the image which is displayed in other visualizations because the rule to generate the images may not support the “Image Processing” property. In this case, images in the visualization may display the original image without any processing.

If settings are modified through the Image Settings panel, Image Discovery, by default, uses the “Image Processing” property to process the image and therefore the image in the preview panel may differ from that displayed in the visualization.

Q6: Why is the authentication status always set to “Authenticated” event even if the user changes the credentials in Columbus®?

A: Although Image Discovery can save the authentication status and token information in the document, re-opening a document in Spotfire® will not perform a re-authentication. In this case, the status in Token Management will remain as “Authenticated” even if user changes the credentials in Columbus®. The images in the image column may not display if the credentials are changed. You will need to re-authenticate via the Token Management dialog.

Q7: Why can’t I connect to remote Image Service? Why can’t I use remote Image Service to render images?

A: The following checkpoints should be verified:

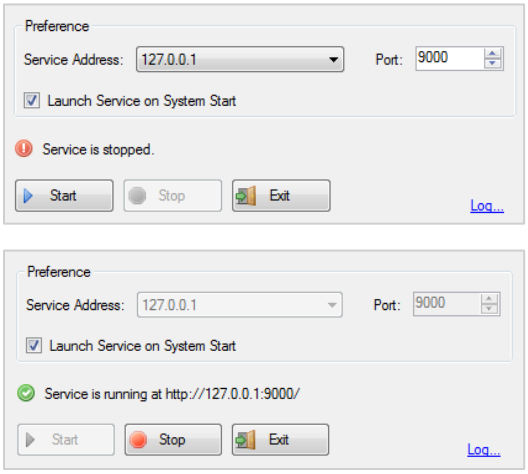
1. Check your network connections and make sure the connections between the Image Service and the client machines are successful. Be aware that IT policy in your organization may block the ports used by Image Service. In such cases, you will need to open the ports (9251 to 9260) for Image Service. You may need to contact your IT department for privileges to perform this operation.
2. Verify the image source URL. If possible, avoid using local paths like such as “C:\ImgFolder\Img.jpg”. This path, for example, points to Img.jpg in ImgFolder on the local C:\ drive on the server that Image Service is deployed. If you expect the remote Image Service to render images located on a client machine, you will have to either share the image folder or copy/move them to shared folders and ensure that Images Service can access these folders.
3. Verify that the image path or URL is available for the remote service by trying to view it with a Web Browser. The machine hosting the remote Image Service may have network connection or privilege issues to the image source.
4. Confirm that the remote Image Service has started successfully. Verify its IP address and port. If you expect the standalone Image Service to be accessed from other machines, you should not use IP address 127.0.0.1 as the service address. There may be multiple network adapters installed on other machines, especially other Servers. Each connects to a different network and has a different IP address. In this case, ensure that the Image Service uses the IP address to which the clients have access.

If the Image Service fails to start, replace the <IP Address> and <Port Number> with the IP address and port number the Image Service used in the following command:

netsh http add urlaclurl=http://<IP address>:<Port Number>/ user=Everyone

Run with an Administrative command shell and start the Image Service again.

Refer to Q1 for details. The service address and port can be configured on Signals Images Service console.



Q8: How do I enable Image Discovery diagnostics and logging? Where do I locate the Image Discovery log file?

A: By default, the Image Discovery extension enables diagnostics and logging. You can navigate to the log file from your local file location: C:\Users\[user name]\AppData\Local\Signals_ImagesService.

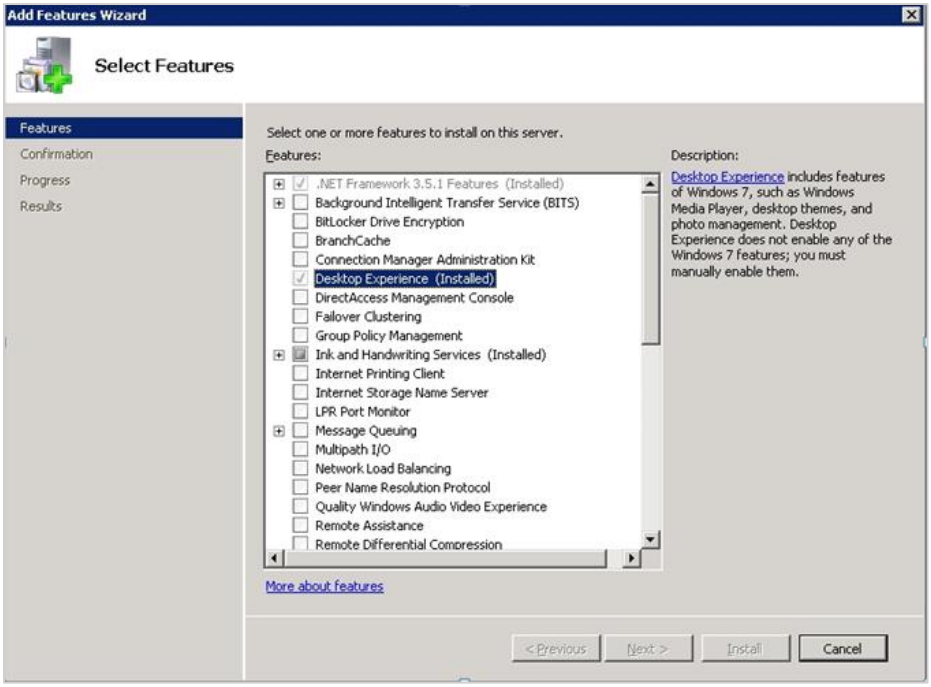
There are two log files:

- InternalImagesService.log: the local Image Service log
- ImagesDiscovery.log: Image Discovery extension log

Q9: How to view TIFF image in Windows Server system?

A: By default, the Windows Server system does not contain the Desktop Experience feature. Therefore, the TIFF image cannot always be rendered in Image Discovery.

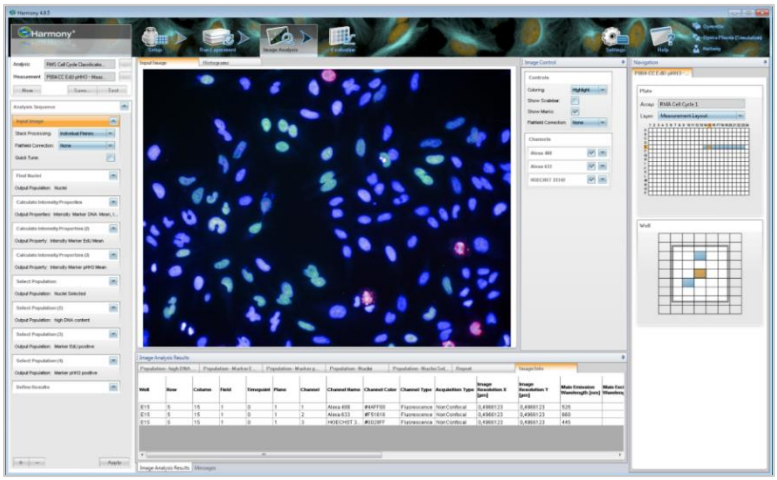
To install this feature, open **Server Manager > Features > Add Feature > Select Desktop Experience**.



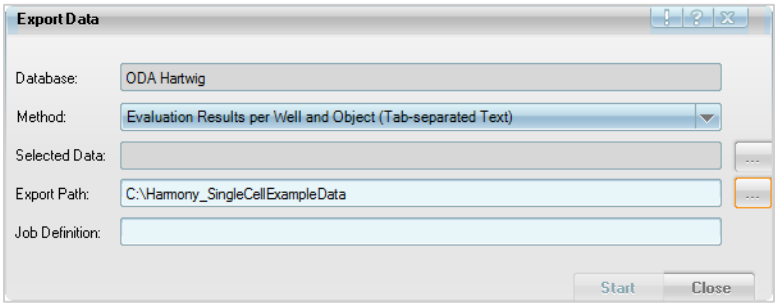
9.16.9 Appendix B - Single Cell Image with Harmony (Operetta and Opera Phoenix) Data

Data Preparation

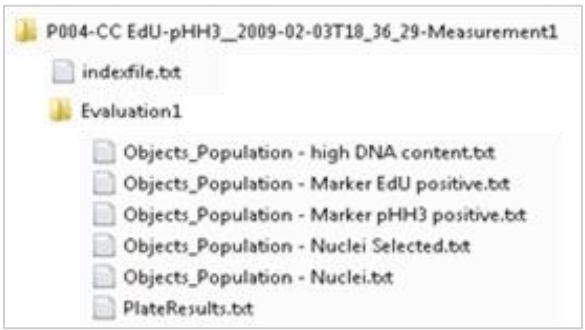
1. Open **Harmony**. Select one image analysis result in the “Image Analysis” page. Select the channels, well and fields in one well to verify the image analysis result.



2. Export data to one result folder per plate in the specified directory “Export Path”.



The exported folder structure resembles the following:

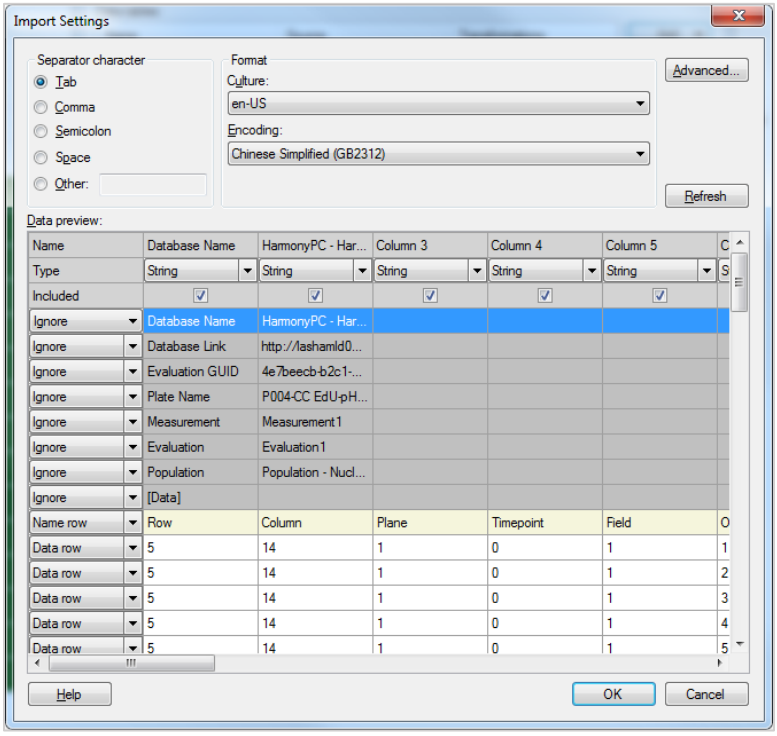


“indexfile.txt” is a tab separated file (UTF-8 encoding) containing the metadata for all images taken on the plate

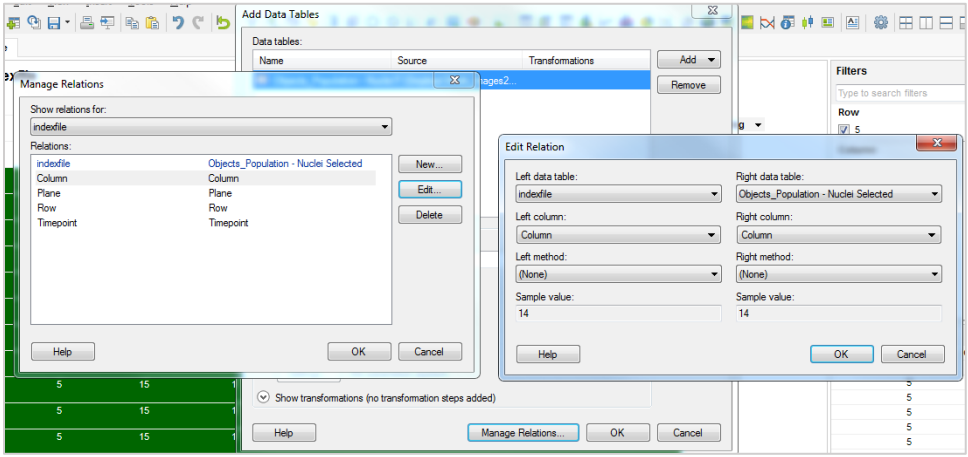
Import data into Spotfire®

1. Navigate to Spotfire® **File menu>Open** to open “indexfile.txt” file.

The index file contains metadata of all images which will be used to link cell data in different analysis result. For example, select **File>Add Data Tables** to add “Objects Population - Nuclei Selected.txt” file in Evaluation1 folder. Ignore the header information in text file as shown in the example below.



2. Click **Manage Relations** in **Add Data Tables** dialog.

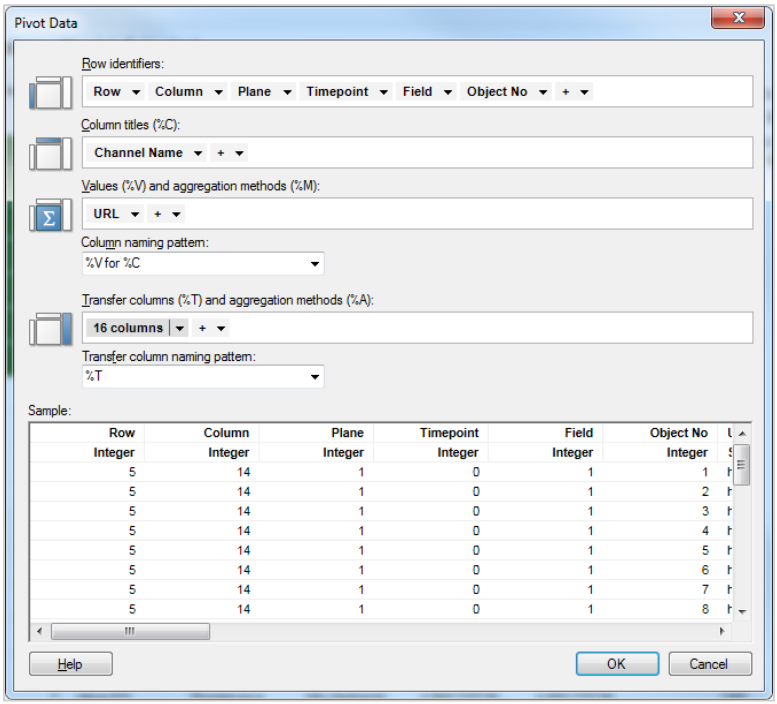


To add relations with current index file data:

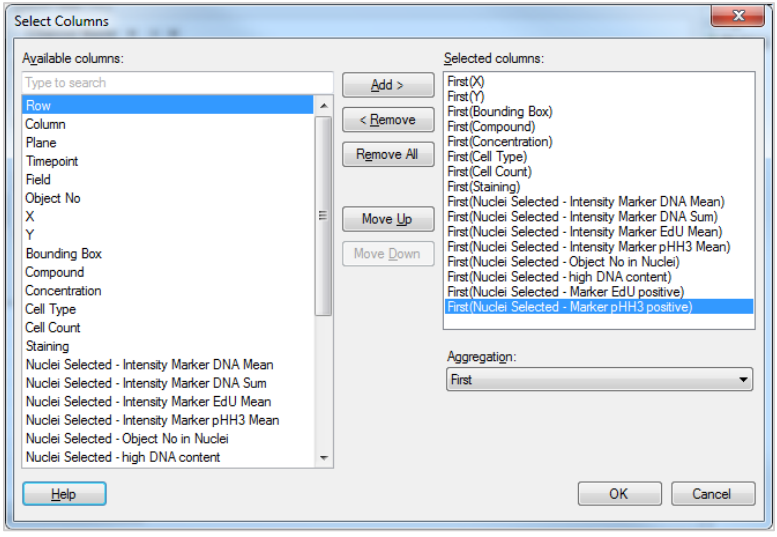
1. Select **Insert > Columns**.
2. Select **‘From Current Analysis’** and choose Source location from index file table, then match the columns:

Row => Row
Column => Column
Plane => Plane
Timepoint =>Timepoint
Field => Field

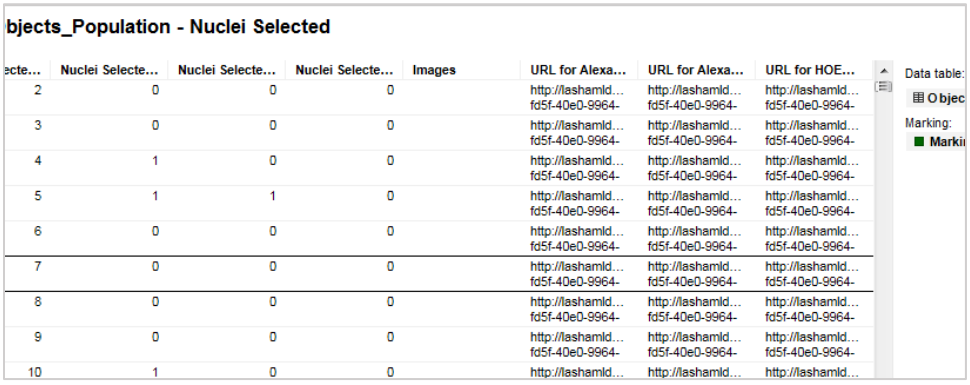
3. Select join method as LeftOuterJoin. Then select the columns you want to view in this analysis table.
4. Select **Insert>Transformations** and add the transformations **"Pivot".** Set pivot data parameters as shown below.



5. Set the transfer columns to use First() aggregation method for other columns.

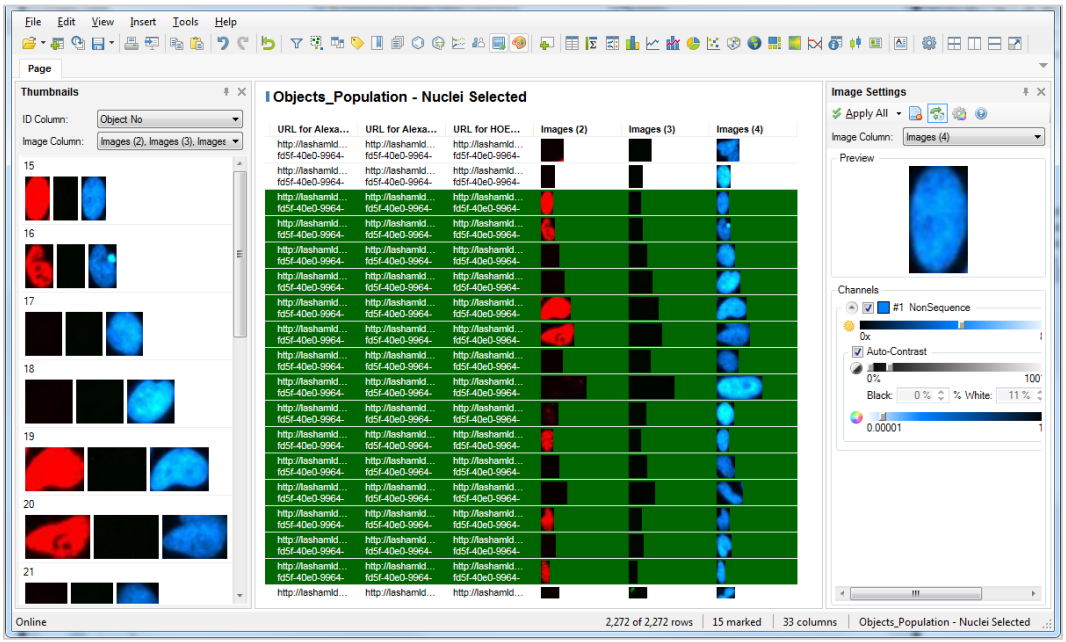


Now you can see three new URL columns for separated channels.



Insert image columns to view the single cell image

1. Navigate to **Tools > Image Discovery>Insert Image Column.** Choose a Harmony rule in drop down list. There are two Harmony rules available for selection:
   - **Harmony URL Cell Image Rule**– Rule supports rendering of cropped cell images in TIFF format with cell level result files from Harmony.
   - **Harmony URL Field Image Rule**– Rule supports rendering of field images in TIFF format with result files from Harmony.
2. Choose the value for {URL} field. For example, choose “URL for Alexa 488”to display the image which has channel “Alexa 488”.
3. Add three image columns for three channels, to view the single cell images in Spotfire® with Harmony result data.



10. Calculations Explorer

10.1 Introduction

The main goal of the **Calculations Explorer** is to provide the user with a simple way of storing and re-executing calculation templates that are frequently used. This tool allows the user to save an existing Workflow while extracting the most important factors required for re-executing the analysis using Calculations Explorer (CE) Templates. In this way, the user can recreate that Workflow with a different dataset by providing matching functionalities which will allow the analysis to be recomputed on the new dataset (provided that the data follows a similar structure to the required parameters previously used).

The **Calculations Explorer** also provides:

- **Built-in fitting models**: These can be used from within the Calculations Explorer Template framework, extending the available fitting capabilities provided by default in Spotfire® curve fitting models. The new fitting capabilities provided are accessible from the settings panel in any scatterplot within the CE Template. In the **Calculations Explorer Controls** panel, open the **Manage Fittings** tab and select **+ New** (Create New Fit). The fitting configuration UI will open and provide the option to fit different curves. The options available in the '**Fitting Type'** dropdown are:
    - Logistic Regression - 3 parameter
    - Logistic Regression - 4 parameter
    - Logistic Regression - 5 parameter
    - Custom Fitting
    - Schild analysis model (only available to those users that have the Signals VitroVivo Metastore configured)
- **Additional fitting models**: These curve fit equations are available from the Signals VitroVivo Metastore and are categorized below. For full details on these curve fit equations, please see *Revvity Signals VitroVivo Curve Fitting Descriptions.* Ten categories are available:
    - Logistic Regression Equations
    - Exponential Equations
    - Linear Equations
    - Biphasic Equations
    - Bell-shaped Equations
    - Site Binding Equations
    - Enzyme Kinetics Equations
    - Shift Equations
    - Polynomial Equations
    - Special Case Equations
- **A set of built-in expressions**: These are available to Calculations Explorer users when creating new calculated expressions. Four additional categories are available:
    - Signals normalization functions
    - Signals outlier functions
    - Signals statistical functions

    - Signals sub-setting functions

- **A curve interpolation tool:** This allows the users to interpolate values on a selected curve.
- **A curve fitting parameters table:** This is extracted automatically from any fitted curve and added to the CE Template and contains the initial fit settings. This table can be configured to include thumbnail figures of the fitted curves to allow a quick evaluation of the fits.

  Note: The order of the logistic regression columns can be changed, and the table can be configured to appear on the left or right-hand side of the scatterplot. Both preferences are remembered by the CE Template.
- **A method to parametrize calculated columns based on the values of another column:** This allows the user to define calculations that are based on a property and set the property to different values based on a second column.
- **An easy way of adjusting specific curve fit settings based on a parameter column:** This allows the user to define specific curve fit settings based on the values in a parameter column.

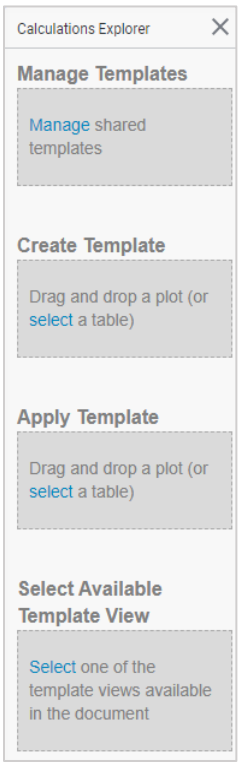
The **Calculations Explorer** requires some additional installation and configuration. For a detailed description of the required configuration, please refer to the *Revvity Signals VitroVivo Installation Guide*. The remaining part of this section describes the functionalities provided by the Calculations Explorer.

10.2 Launching the Calculations Explorer

To launch the **Calculations Explorer** from in Spotfire®, a table needs to be present in the Spotfire® document. Once the table is available select **Tools > Signals Calculations Explorer** from the menu. This will open the Calculations Explorer panel.

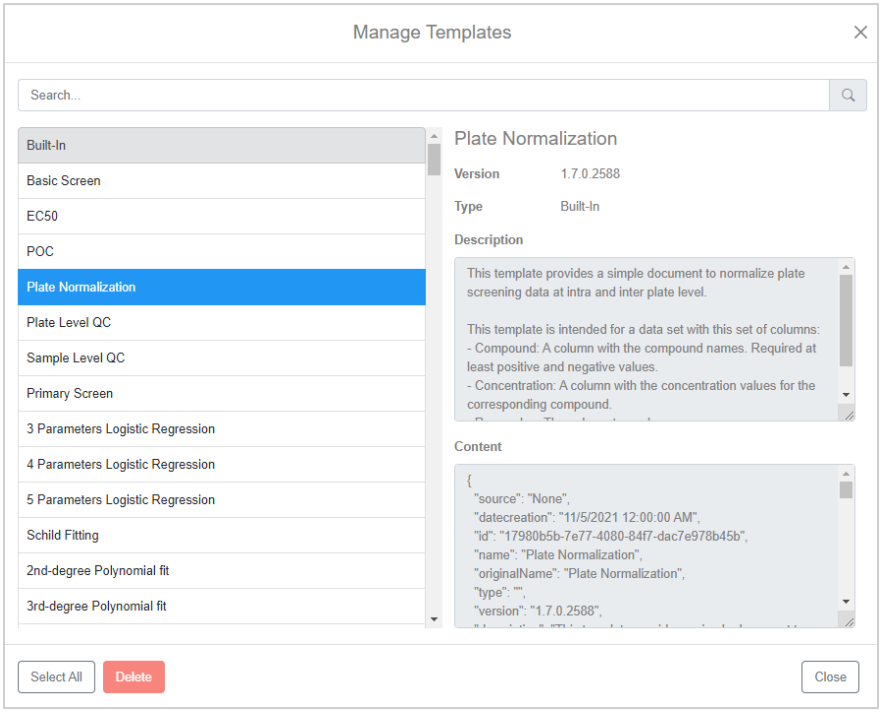
The **Calculations Explorer** panel will appear on the left-hand side of the Spotfire® interface. The Calculations Explorer section contains the following areas:

- **Manage Templates:** Used to manage CE Templates.
- **Create Template:** Used for creating a new CE Template from a table loaded in the document.
- **Apply Template:** Used for applying an existing CE Template using a table available in the document.
- **Select Available Template View:** Used for selecting one of the available CE Template views in the document (if applicable).



10.3 Managing Templates

To manage a CE Template, click on the **Manage** link in the **Manage Templates** section. This will show a list of the available CE Templates in the shared storage. Here the user can do the following:



- **Search:** Enter a search term to filter the CE Templates by.
- **Select All:** Selects all the CE Templates except the Built-in CE Templates. This facilitates bulk removal of CE Templates.
- **Select Templates:** To select a CE Template click on the CE Template name and a description of the CE Template together with its contents will be displayed in the right-hand side of the **Manage Templates** panel.
- **Delete Templates:** To delete a CE Template from the shared storage select the CE Template and select **Delete**. This option is not available for the Built-in CE Templates.
- **Close:** Close the window and exit.

10.4 Creating a New CE Template

To create a new CE Template with the **Calculations Explorer**, first create a document containing the analysis to perform. This analysis should also include any calculated columns required in the analysis, as well as any visualizations, fittings, and hierarchies needed. Add a description if desired to serve as the text displayed in the right-hand side of the **Manage Templates** panel if the CE Template is stored in the Signals VitroVivo Metastore.

Important: This document should only contain the analysis to be reproduced by the Calculations Explorer, it should not contain additional Apps or CE instances.

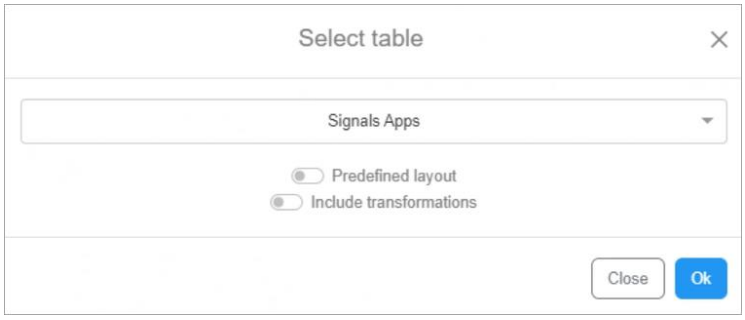
10.4.1 Creating a CE Template

To create a new CE Template, proceed in one of two ways:

1. Drag one of the existing visualizations to the **Create Template** area. This will create a CE Template based on the table underlying the visualization dragged upon the area.
2. Click on the **select** link in the **Create Template** area and select a table from the **Select table** dropdown menu, then **OK**. Selecting **Cancel** will leave the document unchanged.

   Note: Switching between page layout and predefined layout is only supported when using the **select** link in the **Create Template** area. When dragging the visualization to the **Create Template** area the default mode is always used.

   Note: Including transformations is only supported when using the **select** link in the **Create Template** area, as the toggle that allows the inclusion of transformations is not displayed when dragging the visualization to the **Create Template** area.



The CE Template is created in the following manner:

- If the **Predefined layout** toggle is on:
    1. All hierarchies are extracted from the document.
    2. The hierarchy levels are used to populate the CE Template views dropdown menu. If no hierarchy is present the default “All” view will be the only available option for the created CE Template.
    3. Any implicit hierarchy such as those defined in a trellis will be created as a hierarchy.
- If the **Predefined layout** toggle is of (default)
    1. The names of the pages in the document are used to populate the dropdown menu.
    2. All supported visualizations are extracted and made available in the corresponding views.

Note: Not all visualizations are supported by the **Calculations Explorer** Templates. There may also be specific configurations in some of the supported visualizations that are not kept and carried-on once the CE Template is created. It is strongly recommended to review the generated CE Template to ensure all the necessary information is available and in the desired format, before re-applying the CE Template for additional analysis on other datasets. The visualizations currently supported are:

- Scatterplot
- Barchart
- Piechart
- Boxplot
- Heatmap
- Linechart
- Summary table
- Cross table
- Graphical table

Within a specific visualization, customization of some features may not be supported within the **Calculations Explorer** Template. For additional information, please refer to Appendix A: Visualization Elements Supported in the CE Templates.

- All Calculated numeric columns are made available as columns in the aggregated tables over the different levels of the extracted hierarchies.

  Note: Columns are **not** removed from a document when removing a CE Template if they are shared by another CE Template.

- **Calculation groups**: The calculation groups are columns that were used to segment other columns within the Spotfire® analysis document. They are extracted from three sources:
    1. From each level of the hierarchical columns.
    2. From calculated columns. The group is the first column involved in an OVER clause. i.e: a calculated column with this expression, 'Avg([c1]) over ([c2])' will be assigned to the 'c2' group.
    3. From visualizations. The group is the column of the trellis or color axis. Additionally, for some visualizations the group can be obtained from other axes. In scatterplots it can be extracted from 'Shape by' axis, in line charts from the 'Line by' axis, or in heatmaps from 'X' or 'Y' axes.

All the CE Template elements (calculated columns and visualization) are assigned to the corresponding group. The default group is RowId.

- **Custom Expressions**: Custom expressions are supported in certain parts of the document and will be recreated using the new data as part of the CE Template application. The points at which custom expressions are supported are:
    1. Supported visualization axes
    2. Lines
    3. Data limiting expression
    4. Calculated columns
- **Transformations**: Some of the operations available in Spotfire® are also available as part of a CE Template, these include:
    1. Add Columns
    2. Add Rows
    3. Freeze columns
    4. Certain transformations:
        - Pivot
        - Unpivot: In this case, all the columns used in the unpivot must be matched
        - Calculate new column
        - Calculate and replace column
        - Change data types
        - Filter rows (except for dynamic categorical values which are not supported)
        - Exclude columns
        - Replace value

    Note: To include transformations the user must enable the 'Include transformations' toggle accessible when selecting the **Select** link in the **Create Template** area.

    Note: When including transformations in a CE Template, the CE will include all transformations affecting the selected table in the document. Because of this, when creating a CE Template with transformations you should ensure there are no transformations included that may have been added by Apps or other elements that should not be included in the CE Template, like IRLS or exclusion columns.

    Note: When adding transformations, no previous Calculations Explorer instance should be present in the document that uses the same table as input.

10.4.2 Normalization Templates

Calculations Explorer Templates can be used by the **Extensible Normalization** App to normalize data. To make a Calculations Explorer Template available as a shared normalization method usable by the Extensible Normalization App there are some special requirements it must meet.

- The CE Template must use as input only three columns with specific names.
    1. PLATE (string)
    2. CONTROLS (string).
    3. FEATURE (numeric)

- The control values from the CONTROL column defined should be named:
    1. POSITIVE
    2. NEGATIVE
- The CE Template produces as output a single **Normalized** column.
- The CE Template Type is set to **Normalization**.

If a CE Template meets these requirements once it is saved in the Signals VitroVivo Metastore it will be available as a shared normalization method in the **Extensible Normalization** App.

10.5 Exploring a CE Template

When a CE Template is executed in addition to replicating the analysis contained in the CE Template, the user will have several exploration tools that are integrated into the **Calculations Explorer** panel.

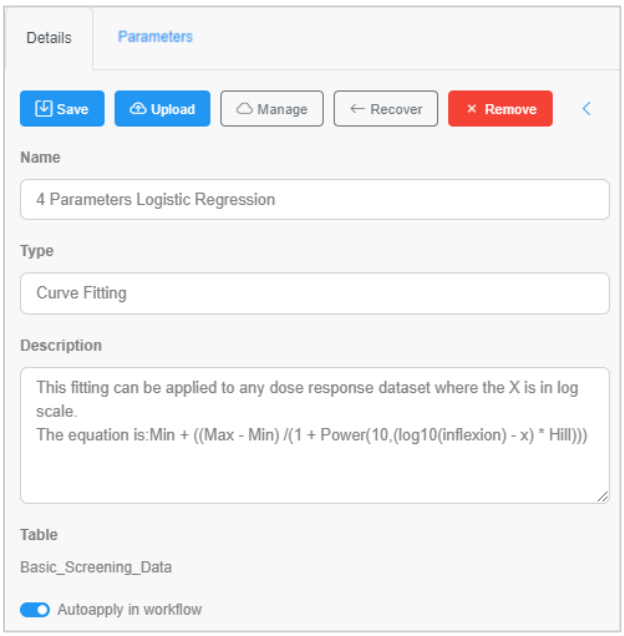
10.5.1 Saving and Configuring a CE Template

To access the CE Template configuration options, select the *Cog Wheel* icon available on the right-hand side of the Calculation CE Template name (left-hand panel).

Refresh: Selecting the 'Refresh template panel' icon will refresh the CE Template panel and update the CE Template. This will include saving to the CE Template any changes that have been made in supported visualizations.

Settings: Selecting the 'Template settings' *Cog Wheel* icon, a window containing CE Template details appears. This includes a **Details** tab and a **Parameters** tab.

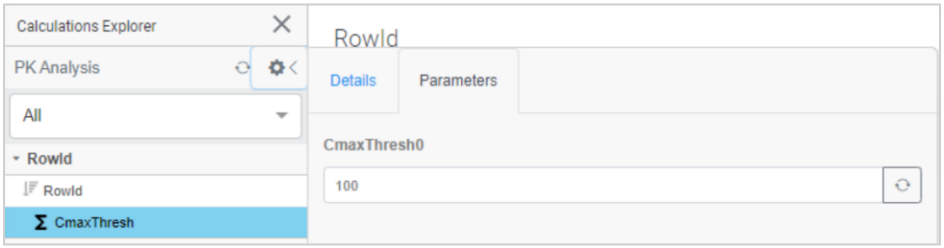
The **Details** tab includes the name and description of the CE Template as well as the configuration settings.



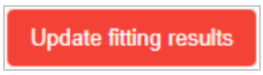
The different sections available in this tab are:

- **Save to local file (** Save **):** Saves the CE Template as a local file in *.json* format.
- **Save to shared repository (** Upload **):** Uploads the CE Template to shared storage in the cloud.
- **Manage Shared Templates (** Manage **):** Opens the Manage Templates UI.
- **Recover Template (** Recover **):** Returns to the last saved version of the CE Template.
- **Close Template (** Remove **):** Closes the CE Template and removes CE Template information from the document. A confirmation dialog will be displayed before this operation is performed.
- **Close Settings (** **):** Closes the settings extension panel.
- **Name:** To modify the name click on it and type the new name desired.
- **Type:** This is the type to which the CE Template belongs and can be edited directly.
- **Description:** To add a description, select the description area and enter the desired description.
- **Table:** This section is for information purposes only and non-editable. It displays the name of the table that is being used by the CE Template.
- **Autoapply in workflow:** When enabled, this toggle allows the CE Template to be automatically applied when added to a Workflow. Note that if the expected columns are missing, the user will be prompted to perform matching manually.

The **Parameters** tab contains a list of those document and table properties that are used within CE Template expressions and allows the modification of these properties from this tab:



Update fitting results: The **Update fitting results** button is displayed when there are some columns from the fitting table that are added to the main table and the fitting has changed. Clicking on this button will update the results in the main table and the button will disappear.



10.5.2 Changing a CE Template View

All views in the document will be available from the dropdown menu below the CE Template name. Here the user will have one entry for each of the hierarchies or pages available in the original document, and those created from the implicit hierarchies in the document, as well as an 'All' option. By default, the first available view will be selected. The 'All' option displays an overview of the top levels extracted from the initial document, as well as all the visualizations that were created from the initial document. It is important to note that when more than one hierarchy is present, updating the CE Template is only possible from the 'All' view.

10.5.2.1 Views Dropdown Menu

This dropdown menu contains one entry for each of the hierarchies available as well as an additional '**All'** category.

- **All:** In the '**All'** view all the levels of the imported hierarchies are displayed and the visualizations displayed are tables and visualizations created by the user for each one of these levels. The tables might contain one or more calculated columns with the aggregated value over the corresponding level. This last statement applies to each one of the calculated columns that the user has generated prior to creating the CE Template.
- **Hierarchy entries:** When a hierarchy entry is selected, the visualizations data displayed are tables and visualizations created by the user corresponding to the aggregations for each one of the hierarchy levels. In contrast to the 'All' view, in the hierarchy visualization mode, the generated table(s) and visualization(s) are designed to perform a drill-down analysis, where only the top level is populated, and lower levels of the hierarchies get populated as the user selects data from upper levels.

10.5.2.2 View Exploration

When a hierarchy is selected, an entry for each of the levels of the hierarchy is displayed in the configuration panel. The following controls are available for the hierarchy exploration mode:

- **Collapse/Expand level components (►):** Collapses or expands the calculations and/or visualizations corresponding to the level, allowing a detailed exploration of the hierarchy level.

When the components for a level are expanded some additional controls for the individual levels are available.

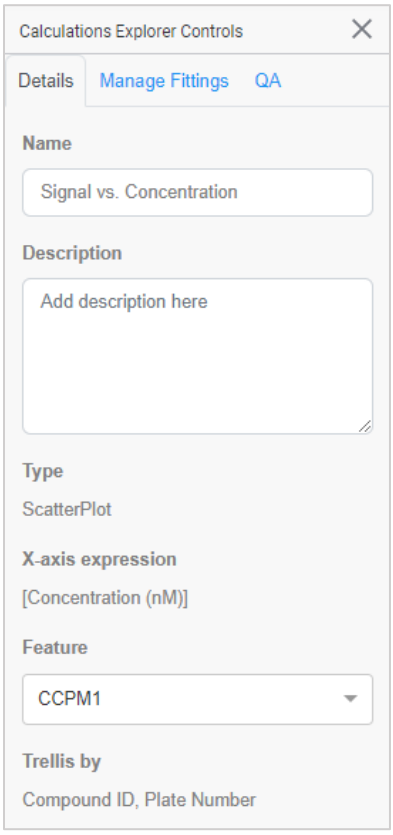
  - **View/Hide level component (👁):** Allows toggling '**on**' and '**off**' individual components (calculations or visualizations) depending on this level.
  - Selecting an entry launches the **Calculations Explorer Controls** panel.

- **Show/Hide filter:** Shows or hides the filter panel corresponding to this hierarchy.
- **Show/Hide Visualization (👁):** Allows toggling '**on**' or '**off**' all the components (calculations and visualizations) that depend on this level (if any).

10.5.3 Calculations Explorer Controls Panel

The **Calculations Explorer Controls** panel appears on the right-hand side of the Calculations Explorer and can be launched by selecting any visualization, or by selecting an entry in the hierarchy on the left-hand panel. To close this panel, select the **X** in the upper-right-hand corner. There are up to three available tabs depending on the entry type:

- **Details tab:** Shows the settings for the specific component of the level. The specific settings displayed will depend on the component type:

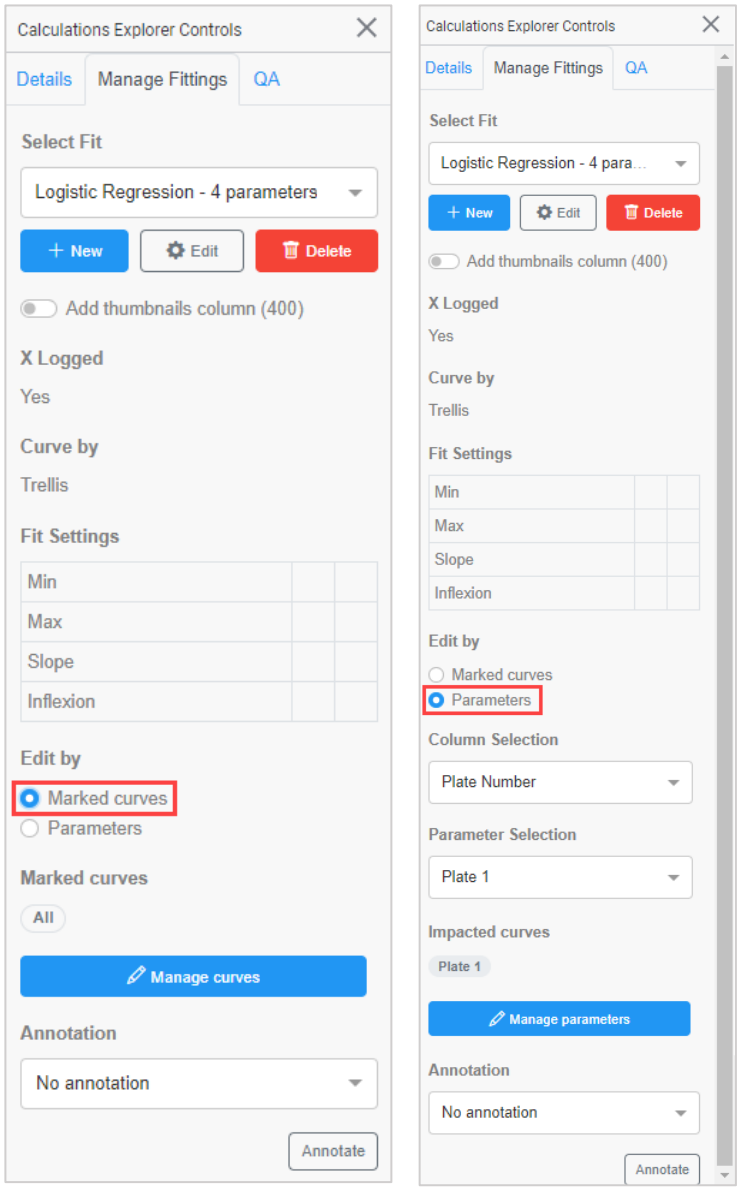


The available elements in this panel are:

  - **Name:** Editable by user.

- **Description:** Editable by user.
- **Type:** Type of visualization (visualization components).
- **X-axis expression:** When relevant to the plot (visualization components).
- **Y-Axis**: Allows the user to select the feature (Y-axis) to be used for the fitting.
- **Trellis by**: When relevant, displays how the current selection is trellised.

- **Manage Fittings tab:** This tab is present in those cases where a visualization contains a curve fit and allows the user to add or remove curve fits, as well as modify the parameters of the fit for those curves that are selected (or for all curves in the case when none are selected). If no fittings are present only the **Create New Fit** button will be present in the panel.



The available elements in this panel are:

- **Select Fit:** A dropdown that allows the user to select a fit from those present in the visualization.
- **Create New Fit:** Selecting the **+ New** button opens the Fit Settings UI.

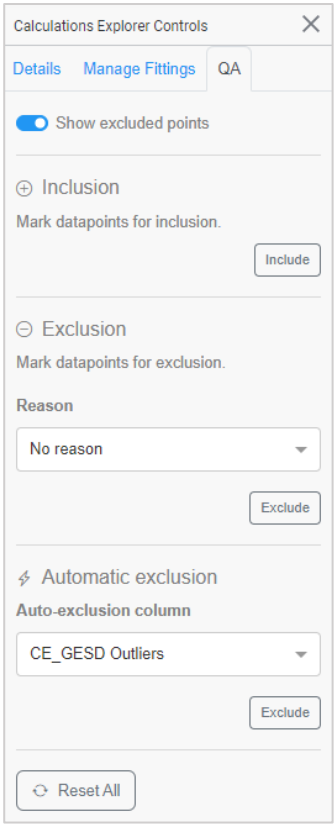
- **Edit Fit:** Selecting the **Edit** buttons allows the modification of the fit selected in the dropdown.
- **Delete Fit:** Selecting the **Delete** button deletes the fit selected in the dropdown.
- **Add thumbnails column (400):** Toggles on / off the option to display curve fit thumbnails.
- **X Logged:** Displays information regarding if the X value is treated as log scale (only present in Logistic regression fits).
- **Curve by:** The columns used to identify each curve, if any.
- **Fit Settings:** The curve fit settings together with the initial values used for the fitting, if any.
- **Edit by:** This radio button allows the user to choose the way in which the settings will be applied, when editing the settings for a specific subset of curves. When changing between the two radio button options, the templates should be refreshed. Note this feature is only available for the default template mode and is not available for the predefined layout templates.
    - **Marked curves:** Applies the changes to the marked curves (default)
    - **Parameter:** Applies the changes to those curves with a specific parameter value.

      Note: The list of points belonging to a single curve should all contain the same parameter value. If this is not the case the results may not be as expected.

      Note: The number of unique values in the parameter list is limited to 500.
- **Marked curves:** Is shown when **Edit by** is set to Marked curves. Displays the list of currently marked curves. If the number is larger than five it will show the first ones together with the total number of curves selected. These will be the curves on which any changes in the **Starting parameters** section below will be applied. If nothing is selected, ‘All’ is displayed.
- **Column selection:** Is shown when **Edit by** is set to Parameter. Displays a dropdown with the list of columns suitable for selection as parameter columns (columns with string values). When changing between different values in the dropdown, the templates should be refreshed.
- **Parameter selection:** Is shown when **Edit by** is set to Parameter. Displays a dropdown with the list of values in the column chosen in the column selection dropdown. The curves in the table containing this value will be the ones that appear in the Impacted curves list.
- **Impacted curves:** Is shown when **Edit by** is set to Parameter. Displays the list of curves on which any changes in the **Starting parameters** section below will be applied. If the number is larger than five it will show the first ones together with the total number of curves selected.
- **Edit:** Opens the **Marked Curves Settings UI** and contains the following elements:
    - **Fitting Type:** Indicates the fitting type.
    - **Fitting Name:** Shows the fitting name.
    - **Equation:** Displays the equation for the curve.
    - **Parameters:** Here the user can modify the parameters for the marked curves. Note the user cannot modify the settings for parameters that are based on columns as this setting is always global.

        - **Load:** Opens the **Apply fitting settings** menu, where the user can select from a dropdown list a desired fit. This feature allows the user to easily apply standard curve fit adjustments to perform QA/QC more quickly.
        - **Save:** Saves the fit settings under a new name to be accessible as a future standard curve fit adjustment.
        - **Close:** Exits the window with no changes.
        - **OK:** This button will apply the fit parameter changes configured to the marked curves. Any other curves will be unaffected. The Calculations Explorer will also perform a check to ensure the equation is valid (i.e. all parameters defined, and syntax is correct).
    - **Annotation:** Provides a dropdown menu to help users mark and track which curves have been reviewed, the default annotations are: No annotation, Approved, Rejected, Needs Review. Select the **Annotation** button to confirm the dropdown selection. Custom annotations can be added by the administrator using the instructions in section *Curve Annotation Reasons Configuration*.

- **QA tab:** This tab is present in those cases where a visualization contains a curve fit and allows the user to exclude or include data points in the fittings, show or hide those excluded points, as well as modify the fitting parameters for subsets of curves.



The available elements in this panel are:

- **Show excluded points:** This toggle will switch between showing and hiding those points that have been excluded from the visualization. When shown they will be marked as a red "**X**" icon.
- **Include:** Includes the marked datapoints in the fitting.

- **Exclude:** Excludes the marked datapoints from the fitting.
  - **Reason**: Allows the user to select from a pre-defined dropdown list the reason for point exclusion. Custom reasons can be added by the administrator using the instructions in section *Point Exclusion Reasons Configuration*.
- **Automatic Exclusion:** Uses the criteria provided by the Boolean column selected in the Auto-exclusion column dropdown to exclude data points.
- **Reset All:** Reset the inclusion status of all the datapoints to 'included'.
- **Update fitting results:** This button will only be displayed if some of the fitting columns are copied to the main table and the fitting has been updated. Clicking it will update the main table and hide the button (see description above).

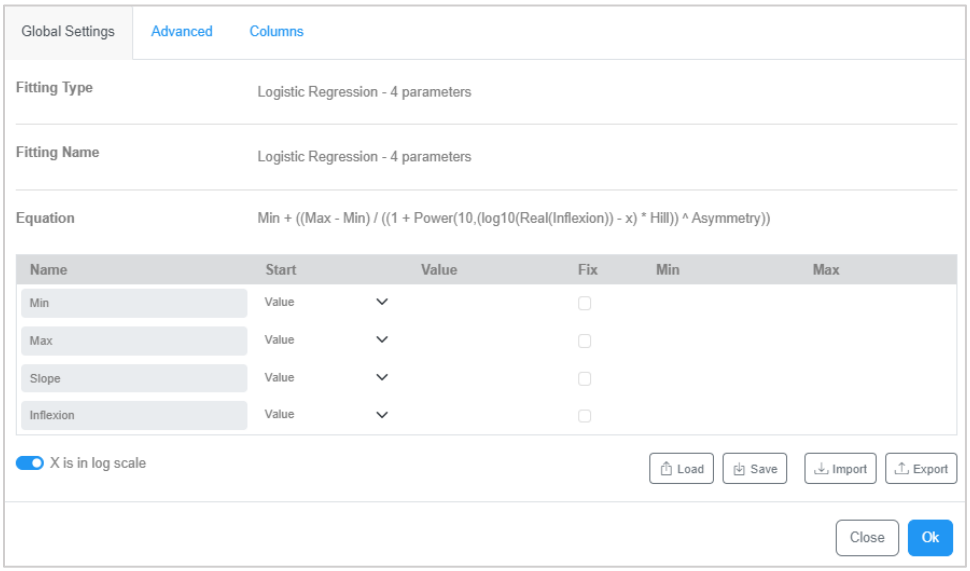
10.5.4 Creating a New Curve Fit

To create a new curve fit the user will open the **Calculations Explorer Controls** panel corresponding to a scatterplot, go to the **Manage Fittings** tab and select **+ New** button (Create New Fit).

This will open the fittings UI which has three different tabs:

- **Global Settings:** This will allow the user to select the type of fit and the initial parameters and constraints.
- **Advanced:** Allows the selection of the algorithm to be used as well as the weighting of the Y values, if any.
- **Columns:** Allows the user to configure which extra curve points to be added (in Logistic regression fits) as well as if the coefficient and parameter columns should be added as columns to the original table and if the thumbnails of the curve should be added to the fitting table.

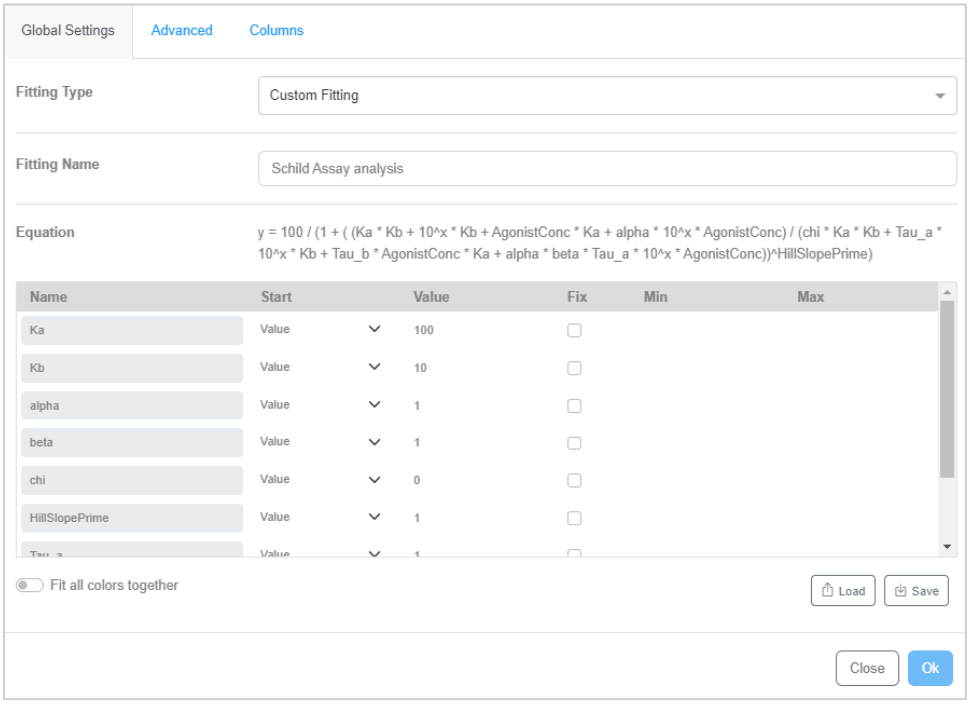
In the **Global Settings** tab, the user has the following options:



- **Fitting Type:** This provides a dropdown to choose between the 3,4 and 5 parameter logistic regression curve fits, a custom fit, or the Schild analysis model. Additional curve fits are available from the Signals VitroVivo Metastore and are discussed in detail in the *Revvity Signals VitroVivo Curve Fitting Descriptions* guide.

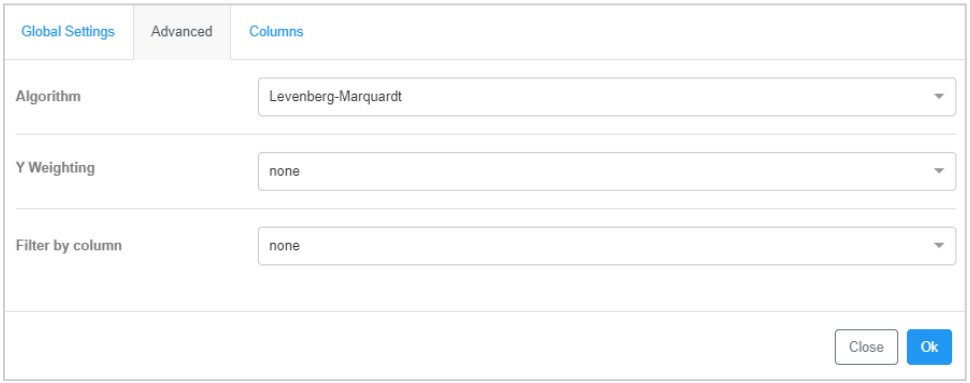
- **Fitting Name:** The name the fitting will have in the analysis.
- **Equation:** This shows the equation that will be used. For logistic regression fits, this equation is read only. For custom fits the user should provide the equation. The equation should follow the R syntax and start with y = and the rest will be the equation used for the fitting. The independent variable (x) should be written in lower case.
- **Parameters:** This section will contain the equation parameters, prepopulated in the case of the built-in fits with the required number of parameters. For custom curve fits the user should add to this section the parameters that have been used in the equation using the + and – controls. The user can set the starting value for each parameter in the fitting. This value can be:
    - A fixed value for all the curves “Value”.
    - A value calculated from the data for each of the curves. The available options for this are: Min, Mid and Max points for both the X and Y axes.
    - Column: In this case a column should be selected. The values of this column will be used as an additional independent variable during the calculation of the fitting.

  Additionally, the parameter can be fixed to the starting value by selecting the **Fix** checkbox or constrained to a value between those values entered in the **Min** and **Max** textboxes, if any. Column parameters are always shown as fixed.
- **X is in log scale:** This toggle only appears for applicable logistic regression fits, indicating the data is visualized in log scale and therefore takes this into account when performing and visualizing the fit.
- **Load:** Opens the **Apply fitting settings** menu, where the user can select from a dropdown list a desired fit. This feature allows the user to easily apply standard curve fit adjustments to perform QA/QC more quickly.
- **Save:** Saves the fit settings under a new name to be accessible as a future standard curve fit adjustment.
- **Import:** Opens the ‘Open File’ menu, allowing the user to select a CE Template and sets the fitting parameters based on an imported local .json file.
- **Export:** Opens the ‘Export fit settings’ menu, allowing the user to select the location to export the current CE Template as a .json.
- **Fit all colors together:** This toggle only appears in the Schild analysis model or in custom fits where one of the parameters is a column. When toggled ‘on’ (default for Schild Analysis) it will fit all data points that are colored according to the parameter column as a single curve. This allows the different subsets of the data (i.e. concentrations in the case of Schild) to be fit together, while still displaying separate colored curves for each of the subsets.



- **Close:** Exits the window with no changes.
- **OK:** This button will apply the fit parameters. The Calculations Explorer will also perform a check to ensure the equation is valid (i.e. all parameters defined, and syntax is correct).

In the **Advanced** settings there are three options available:

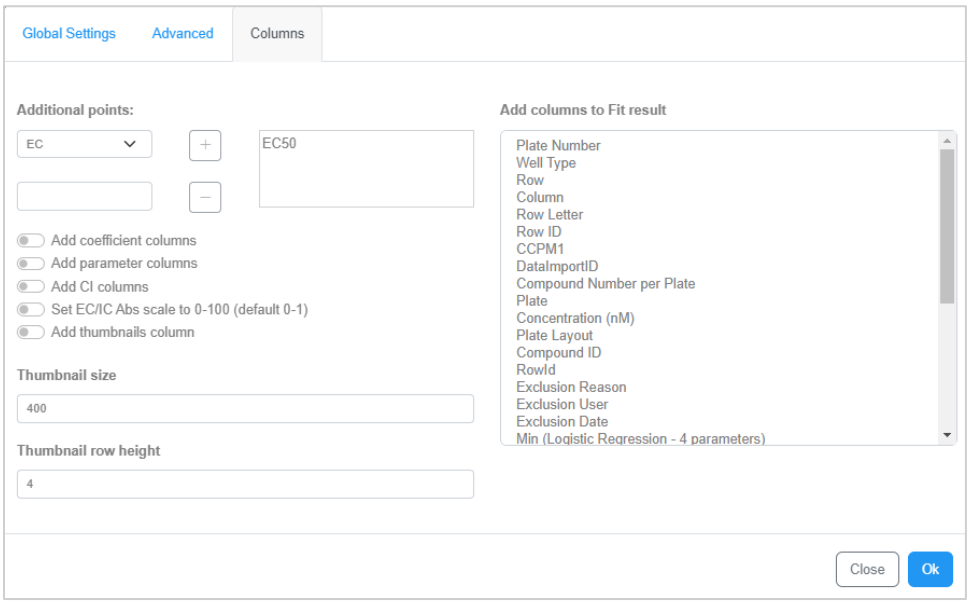


- **Algorithm:** This allows the selection of the fitting algorithm to be used.
  - **Levenberg-Marquardt:** (Default) uses the LM algorithm to perform the fitting.
  - **Levenberg-Marquardt Robust:** This uses an IRLS implementation of the LM to perform the fitting while down-weighting those points that are potentially outliers. The information on the weighting of the different points is returned in a separate table.
    - **B:** Allows the user to set the tuning constant.
    - **Sensitivity:** Defines a residual threshold as a percentage of the max-min value. If no point has residuals above the threshold. IRLS will not be run.

- **Excluded values limit:** Sets a threshold for the maximum number of points that can be excluded by IRLS, if this is exceeded the fit will fail.
- **IRLS weight display threshold:** Allows the user to specify the IRLS cut-off for defining the datapoint shape.

- **Y weighting:** This dropdown provides a selection of weighting options for the Y values. It is only available when Levenberg-Marquardt is selected. The options available are:
  - **None:** No weighting is performed
  - **Poisson (K = 1):** Weighting is done by 1/Y
  - **Relative (K = 2):** Weighting is done by $1/Y^2$
  - **SD:** Weighting is done by 1/SD
- **Filter by column:** This selector will display Boolean columns present in the document and allows the user to filter out those curves that are marked as 'false'. In this way fitting for those curves will not be attempted.

In the **Columns** tab the user can add additional columns to the resulting tables from the fitting table or add existing columns from the data table to the fitting table.



- **Additional points:** This is only available for the built-in logistic regression fits. It allows the user to add more IC or EC points to the curve in addition to the EC50.
- **Add coefficient columns:** This will add the coefficient columns to the initial table.
- **Add parameter columns:** This will add the parameter columns to the initial table.
- **Add CI columns:** This will add columns containing the Confidence Intervals for some of the parameters.
- **Set EC/IC Abs scale to 0-100 (default 0-1):** This allows the user to set the desired EC/IC50 Abs scale.
- **Add thumbnails column:** This will add a thumbnails column to the fitting table. This toggle can also be accessed directly from the **Manage Fitting** panel for convenience.

- **Thumbnail size:** This allows the configuration of the size in pixels of the thumbnail that will be added to the fitting table. Defaults to 400.
- **Thumbnail row height:** This allows the user to configure the row height of the fitting table when the Thumbnail is added. Defaults to 4.
- **Add columns to Fit result:** Allows the user to add the values from columns in the original data to the fitting table.

  Note: This option assumes the value will be unique for each of the fit results and will take the first value for each of the fits.

10.5.5 Updating a CE Template

The aim of the update functionality is to provide the user a way in which to customize an existing CE Template so that those custom changes are saved within the CE Template itself. This allows the user to change those visualizations shown in the CE Template or add new calculation columns to the CE Template. After making these changes and selecting the update button, the CE Template will be updated. In this manner, a user applying this saved CE Template later will generate a document that includes the newly added changes.

Note: The update action will save all the CE Template visualizations that are displayed when selecting the update button, this means that if the user is showing a view that hides some of the visualizations, these will be removed from the new CE Template. Because of this and to avoid inadvertently removing visualizations by mistake, the update is only available from the “All” view in those cases where the CE Template to be modified contains more than one hierarchy (view).

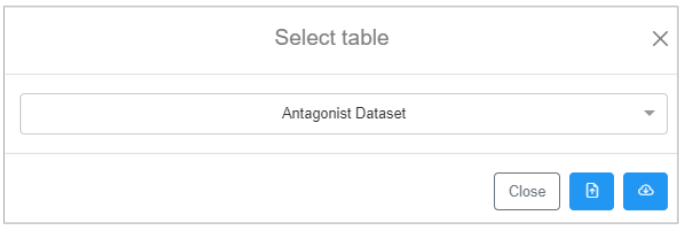
Note: The update action will not update changes done in the CE Template tree by renaming elements of the tree using the Calculations Explorer panel. If these need to be modified, they should be modified by changing the name at the corresponding document level.

10.5.6 Applying a CE Template

The aim of applying an existing CE Template to a new dataset is to generate the CE Template elements (calculations, visualizations, and hierarchies) based on the new table. The CE Templates are applied within the page context, therefore, when applying a CE Template all the visualizations in the current page are deleted and the page is populated with the visualizations of the CE Template.

To apply a CE Template to a new dataset using the Calculations Explorer, you need to first create a document containing the new data where the CE Template should be applied. Once complete, open the **Calculations Explorer** and either select the **‘*select*’** link in the **Apply Template** section, or drag the desired table into the **Apply Template** area.

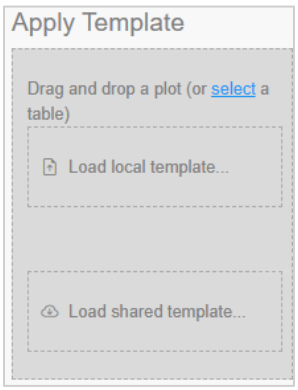
When selecting the ‘*Select*’ link in the **Apply Template** area, a ‘*Select table*’ window is displayed.



This window contains the following options:

- **Dropdown table selector:** Allows the user to choose the table where the CE Template should be applied.
- **Close:** Closes the window with no change in the document.
- **Load local template ( ):** Allows the selection of a local *.json* file.
- **Load shared template ( ):** Allows the selection of a CE Template from the remote server.

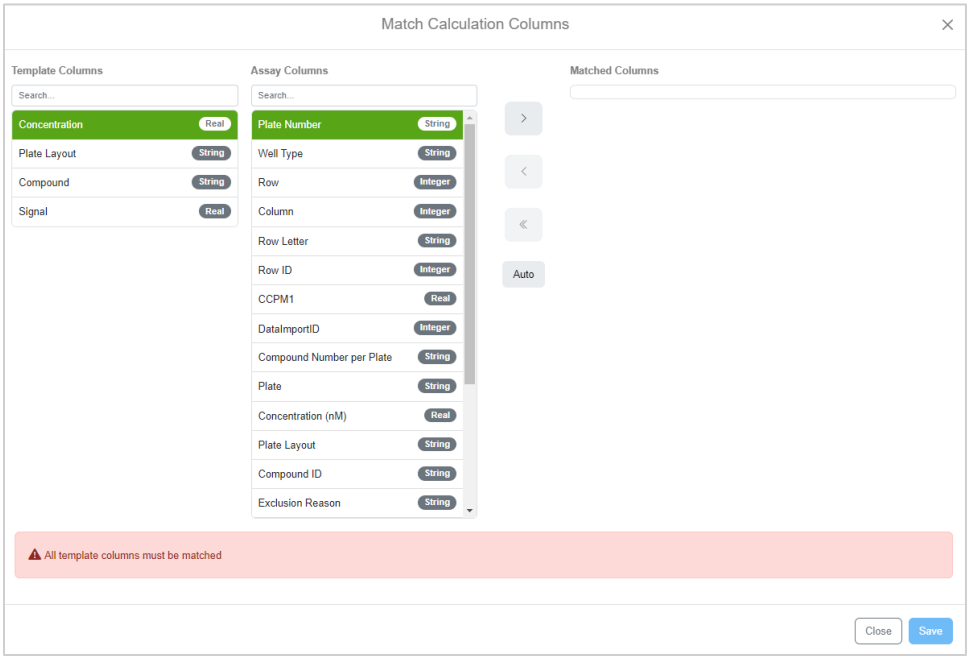
Dragging the table onto the **Apply Template** will expand that area and show two sections.



- Drag the table to the '*Load local template…*' to load a CE Template from a local *.json* file.
- Drag the table to the '*Load shared template…*' to load a CE Template from a shared repository.

10.5.7 Matching the Data to the CE Template

When a CE Template is loaded to be applied to a new dataset, the columns from the dataset need to be matched to the ones required by the original CE Template. To do this the **Match Calculation Columns** interface is used.



In this window, select the columns from the CE Template to be matched to the ones of the new dataset. Execution of the CE Template will only take place if all columns from the CE Template are matched correctly.

The following options are available:

- **Pair (>):** Matches the currently selected columns and moves them to the matched columns area.
- **Unpair (<):** Unmatches the currently selected column and moves it back to the unmatched columns area.
- **Unpair All (<<):** Unmatches all columns currently matched and moves them back to the unmatched columns area.
- **Auto:** This button will attempt to automatically match the columns. This matching is based on the name and annotation type of the columns and may not be able to match all available columns. However, it can be useful to get an initial set of columns matched.

Note: To use the annotation type in the CE Template matching the data in the table needs to contain annotation types, which can be added by the **Grid Plate Editor** App. In addition, the CE Template to be applied must also be created from data that contains these annotation types.

Note: Calculated columns cannot be used for matching.

In some cases, the values of the columns also need to be matched. In such cases, the **Match Column Values** dialog will be displayed after the matching of the column to allow the user to match the values correctly:

The **Match Column Values** interface provides analogous controls to the ones in the **Match Calculation Columns**. In the case of the calculated columns, the column values are the unique values of the calculated column that contain literal values (quoted strings) in the expression.

In case the user matched the columns using the **Auto** functionality, a red information icon (ⓘ) will be displayed when the columns require value matching. Selecting the red icon will open the **Match Column Values** interface.

A red alert icon (⚠) will also be displayed together with an error message if there is any problem in the matching of the columns (such as different data types).

Once all columns (and any values, if needed) are matched, the **Save** button will become enabled. Select **Save** and the CE Template will be applied and executed on the new dataset.

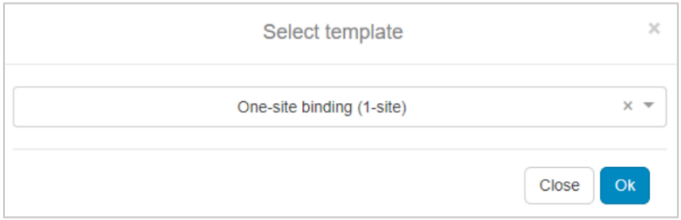
Once the CE Template has been applied, the results can be explored as described above.

Note: In the case of the matching of values, although all the values available will be presented, only those involved in the calculations need to be matched for the correct functioning of the CE Template.

10.6 Viewing an Existing CE Template

If a CE Template is already present in the document and a new tab is opened an additional option will be available in the **Calculations Explorer** pane, **Select Available Template View**:

When the user clicks the **Select** link, the **Select template** UI is displayed allowing the selection of one of the tabs that contain CE Templates present in the document:



Here the user has the following options available:

- Dropdown to select one of the CE Templates containing tabs in the document.
- **Close:** Close the **Select template** UI with no changes.
- **OK:** Add the CE Template from the selected tab to the current tab.

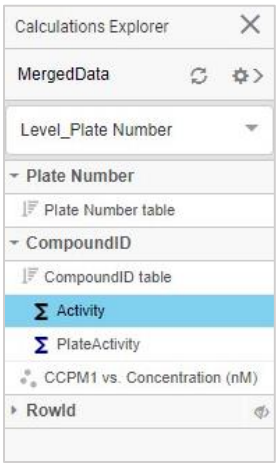
This allows the user to add the same CE Template multiple times and have different views of it in different tabs, or to add several different CE Templates in the same document in different tabs.

10.7 Column Parametrization

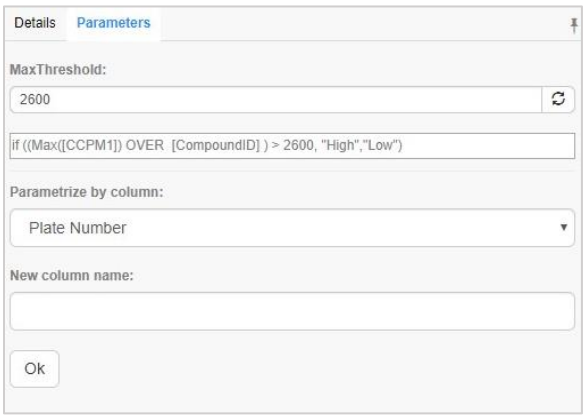
The **Calculations Explorer** provides a way to set the property value used in certain calculations to different values based on a second column.

To do this, perform the following steps:

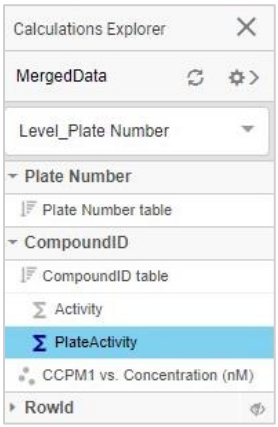
- Create a column with a calculation that includes a property value as part of a numerical comparison. I.e. `if ([column1] < ${Threshold}, "Low","High")`
- The property to be parametrized must be part of the first term of an “if” statement.
- Create a **Calculations Explorer** CE Template and go to the column entry in the Calculations Explorer hierarchy in the left-hand side panel.



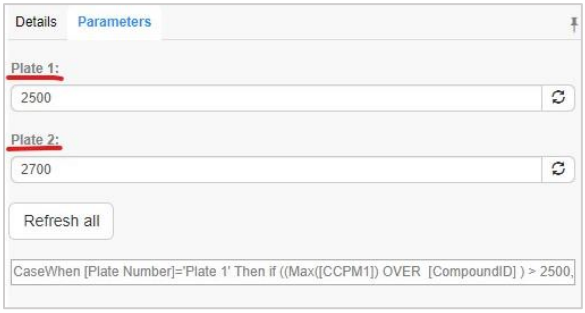
- Mark the created column to expand the right-hand side panel and select the **Parameters** tab.
- Select the column by which you would like to parametrize and click **OK.**



- Update the **Calculations Explorer** Template to add the new parametrized column to the hierarchy.



- Mark the newly created parametrized column, expand the right- hand side panel and select the **Parameters** tab.
- You will now see a separate property has been created for each value in the column selected to parametrize the calculation.



- The value for each of these properties can be set independently from the Parameters tab in the right-hand panel when marking the parametrized column or from the main parameters tab in the document.

Note: Parametrization is supported only for a maximum of 25 unique values in the column that is chosen for parametrizing. If more values are selected a warning is issued and the first 25 are used and the rest are ignored.

10.8 Curve Fitting

In addition to creating and applying CE Templates, the **Calculations Explorer** provides the user with additional fitting models that may be used in any Spotfire® document including a CE Template.

These additional fitting models are:

- 3 Parameter Logistic Regression
- 4 Parameter Logistic Regression
- 5 Parameter Logistic Regression
- Custom Fitting
- Schild analysis model

Note: Additional fitting models are available from the Signals VitroVivo Metastore. For full details on these curve fit equations, please see Revvity *Signals VitroVivo Curve Fitting Descriptions.*

The following section describes the general use of these fitting models.

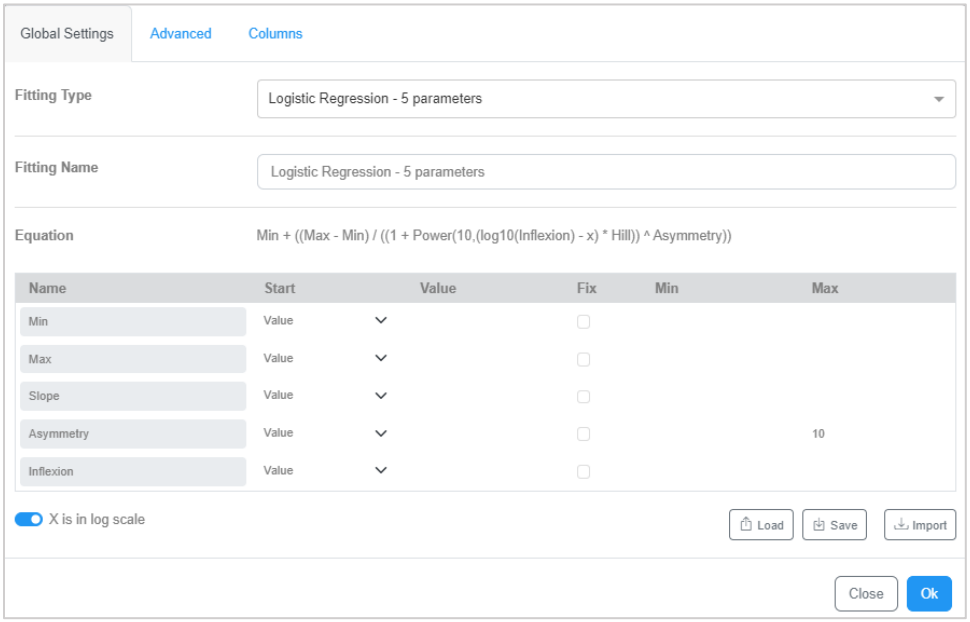
10.8.1 Fitting Model Selection

To use any of these fitting models, the user needs to follow these steps:

- Create a visualization that supports the addition of the curve model (scatterplot).
- Create a Calculations Explorer Template.
- Open the **Calculations Explorer Controls** panel for the scatterplot component within Calculations Explorer.
- In the **Manage Fittings** tab click on the **+ New** button (Create New Fit).

10.8.2 Logistic Regression Fitting Configuration

When **+ New** is selected, the Fitting UI is opened in the **Global Settings** tab. Select one of the Logistic regression curve fits, a configuration window is displayed to configure the curve fitting parameters:



The fitting parameters interface allows the user to set the initial values used by the algorithm for any of the curve parameters. Note that the fitting parameters may also be set from a .json file using the **Import** button.

- **Fitting Parameters:** This section allows the user to provide the initial values used for each of the curve parameters during the fitting process. The parameters available are dependent on the curve fitting model selected:
    - **Min:** This is the lower asymptote of the curve and can be set in all cases.
    - **Max:** This is the upper asymptote of the curve and can be set in all cases.
    - **Slope:** This is the slope of the curve and can be set in the 4- and 5-parameter fit. In the 3-parameter fit it will always be 1 (or -1 if the curve decreases).
    - **Asymmetry:** This is the asymmetry of the curve and can be set in the 5-parameter fit. In the 3- and 4-parameter fit it will always be 1.
    - **Inflexion:** This is the inflexion point of the curve and can be set in all cases. However, as this would correspond to the IC50 (or EC50) in the symmetrical curves, it is usually one of the parameters that the user aims to obtain, as opposed to the Max and the Min, that are frequently fixed.

  The values that can be used are:
    - Specific values provided by selecting “Value” from the dropdown menu and entering the required value. In this case the value will be the same for all the curves if more than one curve is being fitted.
    - A value calculated from the data. Using the dropdown menu, the user can set the initial value for the parameter to the following data dependent values by selecting the one of interest from the dropdown menu:
        - Min X or Y: The minimum value for the X or Y axes values respectively.
        - Max X or Y: The maximum value for the X or Y axes values respectively.
        - Mid X or Y: The mid value for the X or Y axes values respectively.

The “Fix” checkbox allows this value to be fixed to the value entered, as opposed to being used as an initial estimation of the parameter that will be modified by the fitting algorithm.

The Max and Min columns allow the resulting fit value to be constrained to a specific range.

- **X is in log scale:** This toggle will indicate if the plot has an X-axis in log scale. Frequently, dose-response data is represented with the X in log scale to reveal the characteristic ‘S-shape’ that these curves display. To visualize the data with the ‘***X in log scale***’, right-click on the **X-axis** and ensure the **Log scale** option is marked.

In addition to the settings provided in the **Global Settings** tab there are optional settings that may be changed by the user in the **Advanced** and **Columns** tabs as mentioned above. In the **Advanced** tab the following settings are available:

- **Algorithm:** Here it is possible to select the algorithm used for the fitting. The choices are Levenberg-Marquardt or Levenberg-Marquardt Robust, which is an implementation of the LM algorithm that uses IRLS to down-weight the outliers during the fitting process. This improves the quality of the fits.

  When using Levenberg-Marquardt Robust there are a few additional options that can be configured:

  - **B:** This allows the user to modify the default value of the B parameter in the IRLS that is used for the Tukey biweight function
  - **Sensitivity:** IRLS sensitivity is defined as a residual value between 0 and 100 where 100 is the min-max interval. If no point has a residual larger than the set threshold. IRLS is not run.
  - **Excluded values Limit:** The user can set a percentage value in this box. If a higher percentage of the points in the curve are weighted as 0 by the IRLS the IRLS fit will fail
  - **IRLS weight display threshold:** Here the user can set the threshold weight above which the point in the scatterplot is displayed as weighted.

- **Y Weighting:** Here the user has a choice of weighting methods that can be used to weight the Y values the methods available are:

  - **None**
  - **Poisson (K = 1):** Here the error if each point is divided by the Y.
  - **Relative (K = 2):** The error of each point is divided by Y^2.
  - **SD:** The error of each point is divided by the SD of the Y at that position. This will only work if there are multiple replicates of each concentration.

    This option is not available when using the Robust method of the LM, as both methods are designed to cope with deviations in a different manner and combining them could produce unexpected results.

- **Filter by column:** This dropdown allows the user to use a Boolean column to determine which curves should be fit, avoiding in this way unnecessary fits where the result is known not to be useful.

In the **Columns** tab the user can add additional columns:

- **Additional points:** This allows the user to optionally specify additional points in the curve to be displayed in the visualization and added to the curve fitting parameters table (i.e. the user can add the IC90 by selecting IC in the dropdown menu and entering 90 in the text box and clicking on the “+” button, or the

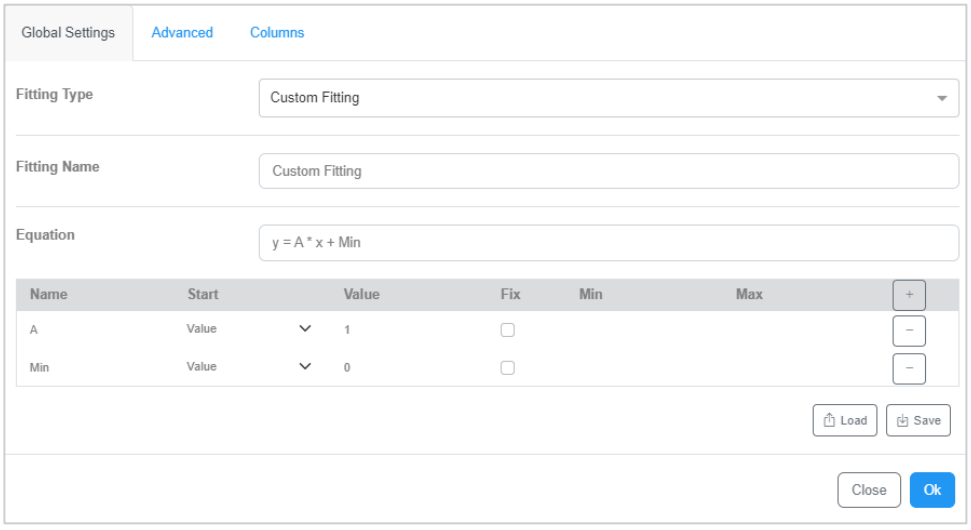
EC40 by selecting EC in the dropdown and entering 40 in the textbox.). Points can also be removed by selecting them in the list box and clicking on the “-” button.

- **Add coefficient columns:** When the toggle is active the rSquared value and IC/ECX from the calculated curve will be added to the table. Default is toggled ‘off’.
- **Add parameter columns:** When the toggle is active all the curve parameters from the calculated curve will be added to the table. Default is toggled ‘off’.
- **Add CI columns:** This will add columns containing the Confidence Intervals for some of the parameters.
- **Set EC/IC Abs scale to 0-100 (default 0-1):** This allows the user to set the desired EC/IC50 Abs scale.
- **Add thumbnails column:** This toggle will add to the fitting table a column containing the thumbnails of every curve. Default is toggled ‘off’.
- **Thumbnail size:** This is the size in pixels of the thumbnail. Default is ‘400’.

  Note: If the row size is small these thumbnails will be scaled accordingly, and will look smaller, this is the pixel size they would have when displayed at 100% size.
- **Thumbnail row height:** This will set the size of the row. Default is ‘4’.
- **Add columns to fit result:** Add columns from the data table to the fit result table.

10.8.3 Custom Curve Fitting Configuration

When the **Custom Fitting** curve fit is selected, a configuration window is displayed to configure the curve parameters. Many of the controls are the same as in the built-in Logistic regression curves. Highlighted below are the key differences.



The Fitting Parameters interface allows the user to configure the following:

- **Fitting Name:** The name to be used for the curve
- **Equation:** The formula to be used for the fit in the form: y = a + b * x

The syntax should be R compatible and each parameter identified by a letter in the equation should be assigned an initial value or value range as described below.

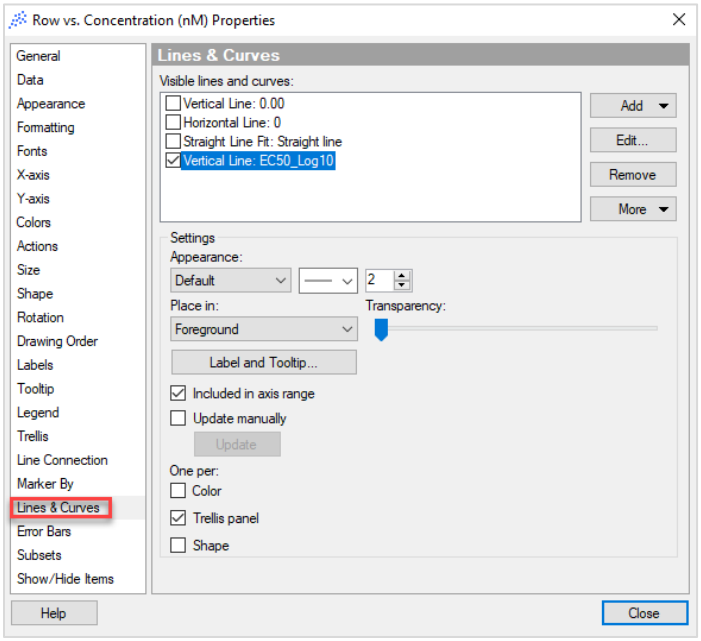
- **Curve Parameters Table:** This section allows the user to provide the initial values used for each of the curve parameters during the fitting process. The initial value for each of the parameters can be set under the **'Start'** column as:
    - **Value:** The initial value used for estimating the fitting can be set in the **'Value'** column.
    - **A value calculated from the data**. Using the dropdown menu under **'Start'**, the user can set the initial value for the parameter to the following data dependent values by selecting the one of interest from the dropdown menu:
        - Min X or Y: The minimum value for the X or Y axes values respectively.
        - Max X or Y: The maximum value for the X or Y axes values respectively.
        - Mid X or Y: The mid value for the X or Y axes values respectively.

      The "Fix" checkbox allows this value to be fixed to the value entered, as opposed to being used as an initial estimation of the parameter that will be modified by the fitting algorithm.

      The Max and Min columns allow the resulting fit value to be constrained to a specific range.
    - **Column:** an existing column from the input data table. The **Fix** column will automatically be enabled when **Column** is selected.
    - **Fit all colors together:** This allows the user to fit separately or together those curves with different colors that are based on a column parameter. Therefore, this toggle only appears when one of the parameters is a column.
- **Load:** Opens the **Apply fitting settings** menu, where the user can select from a dropdown list a desired fit. This feature allows the user to easily apply standard curve fit adjustments to perform QA/QC more quickly.
- **Save:** Saves the fit settings under a new name to be accessible as a future standard curve fit adjustment.
- **Close:** Selecting Close will close the dialog without modifying the document.
- **OK:** Selecting OK will perform the fitting.

10.8.4 Fitting Results

The result of the fitting models is analogous to the result of those fitting models provided by Spotfire®. The curve will be plotted over the data, and the precise parameters of the fit can be seen by hovering over the resulting curve. The visibility status of the IC/EC50 lines (if applicable) are remembered by the CE Template and can be configured by right-clicking on a plot **> Properties > Lines & Curves**, then toggling on/off the desired line(s).



The table containing the fitting parameters for the curve is added automatically to the document, but it is not displayed by default. The fitting parameters table contains for each curve the estimated value for each parameter, the initial fit settings, as well as the rSquared value and the 95% confidence intervals.

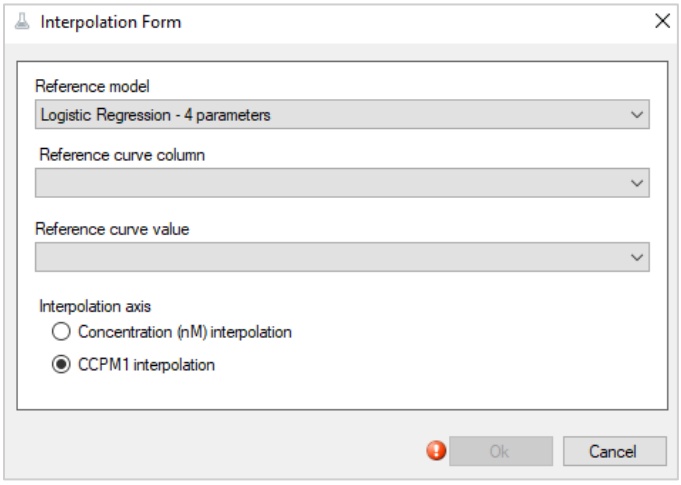
If the LM Robust fitting is used, a fit data table is also generated in the document. This table is not displayed by default. The *Fit* data table contains for each data point the X and Y value as well as the Y predicted by the fit for that point and the IRLS weights for each of the datapoints.

10.9 Curve Reference Interpolation

This functionality allows the user to create an interpolation column based on one of the existing curve fits in the document.

10.9.1 Curve Reference Interpolation Usage

To use the Curve Reference Interpolation tool right-click on the visualization containing the curve fit that you wish to interpolate and select “Curve Reference Interpolation …”. This will open the **Interpolation Form** UI:



- **Reference model:** Select the reference curve fit from the dropdown. This dropdown will display those fitting models available in the selected visualization.
- **Reference curve column:** The column that contains the reference value.
- **Reference curve value:** Select the reference partition category from the partitions available in this dropdown. This menu will display a list of the categories into which the data is partitioned in this visualization by the reference curve column.
- **Interpolation axis:** Select the axis on which the interpolation should be performed.

Once the settings are configured the tool will generate a calculated column that will provide the value for the corresponding axis of the selected reference curve in each of the rows in the data table based on the other axis value.

In those cases where the user has multiple partitions (i.e. compound and plate), if the user selects as partitions for the trellis both compound and plate and the reference column selected is compound, a separate interpolation will be performed for each of the values in the plate column.

10.10 Custom Calculations

The **Calculations Explorer** provides access to a set of calculations available within the Spotfire® **Add calculated column** interface available from the main toolbar by navigating to **Data > Add calculated columns**. These calculations are divided into the following categories:

- **Signals normalization functions:** Contains functions used for data normalization.
- **Signals statistical functions:** Contains statistical functions that are not available out of the box in Spotfire®.
- **Signals outlier functions:** Contains functions commonly used for outlier detection.
- **Signals subsetting functions:** Contains sub-setting functions to simplify sub-setting operations that could else require longer or more convoluted Spotfire® expressions.

10.10.1 Signals Normalization Functions

Within this category the following normalization functions are available:

- **ControlPercentage:** Calculates the Percentage of control using the formula: (value / Avg(c)) * 100, where c is the control (positive or negative) value.
    - The first argument is the value column.
    - The second is the type column which should contain a label for each of the samples.
    - The third is the name of the control as indicated in the labels from the chosen type column.
- **Median Polish:** This implements an Intra-plate normalization method that normalizes the values using the Median Polish algorithm (Tukey, J. W. (1977). Exploratory Data Analysis, Reading Massachusetts: Addison-Wesley) to remove row and column biases individually for each plate.
    - The first argument is the value column.
    - The second is the plate column.
    - The third is the name of the row column.
    - The fourth is the column.
    - If the controls should be excluded from the normalization 3 additional optional arguments are needed.
        - The fifth argument is the controls column which should contain a label for each sample.
        - The sixth is the label of the Negative control in the chosen controls column.
        - The seventh is the label of the Positive control in the chosen controls column.
- **NPA:** Calculates the Normalized Percent Activation using the formula (value - mean(n) / mean(p) - mean(n)) * 100, where p and n are the values for positive and negative controls.
    - The first argument is the value column.
    - The second is the type column.
    - The third is the label of the negative control in the chosen type column.
    - The fourth is the label of the positive control in the chosen type column.
    - If a fifth column with plate data is passed as an argument, the statistics will be determined separately for each plate.
- **NPI:** Calculates the Normalized Percentage Inhibition using the formula (mean(p) - value / mean(p) - mean(n)) * 100, where p and n are the values for positive and negative controls.
    - The first argument is the value column.
    - The second is the type column.
    - The third is the name of the negative control.
    - The fourth is the name of the positive control.

- **RobustControlPercentage**: Calculates a robust version of the Percentage of Control by using the Median in place of the average. The formula used is: (value / Median(c)) * 100, where c is the control (positive or negative) value.
    - The first argument is the value column.
    - The second is the type column.
    - The third is the name of the control.
    - The fourth argument is the plate column and is optional.
- **ZScore:** Calculates the Zscore.
- **RobustZScore:** Calculates a robust version of the ZScore by using the Median in place of the average, and MAD in place of the StdDev.

10.10.2 Signals Statistical Functions

Within this category the following statistical functions are available:

- **CV:** Coefficient of Variation
- **POC:** Percentage of Positive Control
- **ZPrimeFactor:** Calculates the Z'Factor using the formula 1 - (3 * (StdDev(p) + StdDev(n)) / abs(Avg(p) - Avg(n))), where p and n are the values for positive and negative controls.
    - The first argument is the value column.
    - The second is the type column.
    - The third is the name of the negative control.
    - The fourth is the name of the positive control.
- **RobustZPrimeFactor**: A robust version of the ZPrimeFactor that uses the MAD and Median in pace of the StdDev and average.
- **SignalToBackGround:** Provides a measure of the signal available.
- **SignalToNoise:** Provides a measure of the signal available.
- **SSMD:** Calculates the strictly standardized mean difference using the formula (Avg(p) - Avg(n)) / sqrt((StdDev(p) ^ 2) + (StdDev(n) ^ 2), where p and n are the values for positive and negative controls.
    - The first argument is the value column.
    - The second is the type column.
    - The third is the name of the negative control.
    - The fourth is the name of the positive control.
- **SSMDSampleHits**: Calculates the strictly standardized mean difference (SSMD) for each compound using the formula (Avg(v) - Avg(n)) / (sqrt((2/k) * (n-1)) + StdDev(n)), where v is the compound values, n is the values for negative controls and k is the number of negative controls minus 2.48.

- The first argument is the value column.
- The second is the type column to group by.
- The third is the name of the negative control.

10.10.3 Signals Outlier Functions

Within this category the following outlier functions are available:

- **GESD:** The GESD Method (Generalized Extreme Studentized Deviate Test) performs a Rosner's GESD method to detect up-to maximum of n likely outliers in an approximately normal distribution.
  - The first argument is the value column to determine the outliers.
  - The next n arguments (optionals) are the group by columns. A set of columns to segmentate the outlier detection.
  - The last argument, max_n (optative) is the maximum percentage of outliers permitted to be identified. If is not given then the max_n is 20%.
- **GESDwithSpecificLimits:** The GESD with specific limits Method performs a Rosner's generalized extreme studentized deviate method to detect up - to a maximum of n likely outliers in an approximately normal distribution.
  - The first argument is the value column to determine the outliers.
  - The second argument n_max is the number of outliers to detect.
  - The third argument indicates if all outliers should be ignored if more than n_max outliers are detected. Values can be 'Yes' or 'No'.
  - The next n arguments (optional) are the columns to group by. Up to 4 columns can be added here to segment the outlier detection.
  - The next argument, n_end (optional) is the number of points at the end of the order columns that should not be considered outliers.
  - The last argument, order column (optional) is a column that will be used to sort the value column and determine which are the first and last elements if n_end is > 0.

  The main differences with the previous are that the number of outliers provided are the absolute numbers, not a fraction, and the possibility of ignoring the results if more than the specified outliers are detected, as well as ignoring the result for the first and last elements from a second column. The aim of this last option is to ignore the first and last concentration points in dose response curves, for compounds that may have a very steep change in activity at certain concentration ranges.
- **OutliersByColumn**: Provides an indication of the number of outliers present in a column as determined by using the interquartile range.
  - The first argument is the value column.
  - The second is the aggregation (over) column.
  - The third (optional) is IQRThreshold. If is not provided then the IQRThreshold defaults to 1.

10.10.4 Signals Subsetting Functions

Within this category the following subsetting functions are available:

- **IsInsideInterval:** Returns TRUE if the column value is within a certain interval.
- **ValueIfEqualTo:** Returns the value of a column if the column value is equal to certain condition. If a second column is provided the value returned is that of the second column.
- **ValueIfGreaterThan:** Returns TRUE if the column value is greater than a certain value. If a second column is provided the value returned is that of the second column.
- **ValueIfInsideInterval**: Returns the value of a column if the column value is within the specified interval. If a second column is provided the value returned is that of the second column
- **ValueIfLowerThan:** Returns TRUE if the column value is smaller than a certain value. If a second column is provided the value returned is that of the second column.
- **ValueIfNotOutlier:** Returns the value if the value is not an outlier over a given column else it returns NULL. Outlier condition is determined using the formula (Xi > (Q3 + (f * IQR)) or (Xi < Q1 - (f * IQR)), where Xi are the values and f the numeric IQRThreshold.
- **ValueIfOutsideInterval:** Returns the value of a column if the column value is outside the specified interval. If a second column is provided the value returned is that of the second column.

10.11 Additional Configuration

10.11.1 Signals VitroVivo Metastore Service Configuration

To store the CE Templates in a shared repository accessible to different users, the Metadata Store Service needs to be configured.

To configure the shared repository so it is accessible from the Calculations Explorer, the following steps should be followed:

- Log into Spotfire® as an Administrator.
- Open **Tools > Administration Manager**.
- Select the **preferences** tab.
- Select the user group for which the shared preference should be configured.
- In the list of preferences edit **Signals Screening Metastore > Service URL**.
- Add the URL to the server where the Metadata Store Service is running.

10.11.2 Curve Annotation Reasons Configuration

In case the user wants to add a custom annotation reason to the marked curves annotations this preference needs to be configured.

To configure the annotation reason so it is accessible from the **Calculations Explorer**, the following steps should be followed:

- Log into Spotfire® as an Administrator.
- Open **Tools > Administration Manager**.
- Select the **preferences** tab.
- Select the user group for which the shared preference should be configured.
- In the list of preferences edit **Calculations Explorer > Annotation > Reasons**
- Add the desired annotation name

Note: To add multiple annotations, separate each by a semicolon. The new annotations will be added alongside the default annotations which cannot be removed.

10.11.3 Point Exclusion Reasons Configuration

In case the user wants to add a custom annotation reason to the excluded point annotations this preference needs to be configured.

To configure the annotation reason so it is accessible from the **Calculations Explorer**, the following steps should be followed:

- Log into Spotfire® as an Administrator.
- Open **Tools > Administration Manager**.
- Select the **preferences** tab.
- Select the user group for which the shared preference should be configured.
- In the list of preferences edit **Calculations Explorer > Exclusion > Reasons**
- Add the desired exclusion name

Note: To add multiple exclusion reasons, separate each by a semicolon. The new exclusion reason will be added alongside the default exclusion reason which cannot be removed.

10.11.4 Curve Fit Table Layout

In case the user wants to modify the default position of the *Curve Fit* table from the right of the scatter plot to the left of the scatter plot, this preference needs to be configured. By default, this preference is empty which defaults as ‘False’ and leaves the table on the right-hand side of the scatter plot.

To modify the *Curve Fit* table position to the left, the following steps should be followed:

- Log into Spotfire® as an Administrator.
- Open **Tools > Administration Manager**.
- Select the **preferences** tab.
- Select the user group for which the shared preference should be configured.
- In the list of preferences edit **Calculations Explorer > Layout > Show fit table on the left > True**

10.12 Troubleshooting

10.12.1 Unable to Create a CE Template from Table

If an error message stating the CE Template cannot be created because the table selected contains no calculation columns or supported visualizations associated, ensure the correct table is selected and that it contains at least one calculated column, or alternatively that the selected table is used as the underlying data for one of the supported visualizations.

10.13 Built-in CE Templates

The **Calculations Explorer** also provides the user with a set of built-in CE Templates that can be used out-of-the-box for analyses and as a basis for the development of custom CE Templates by the user.

Note: Additional fitting models beyond what is listed below are available from the Signals VitroVivo Metastore. For full details on these curve fit equations, please see *Revvity Signals VitroVivo Curve Fitting Descriptions.*

The built-in CE Templates provided are the following:

- Basic Screen
- EC50
- POC
- Plate Normalization
- Plate Level QC
- Sample Level QC
- Primary Screen

A more detailed description of each of these can be found below.

10.13.1 Basic Screen

The aim of this CE Template is to provide the user with a basic screening analysis. It contains a set of calculations to cover different screening aspects, such as:

- Plate QC using Z'Factor, outlier and missing value detection.
- Sample level QC.
- Normalization of valid values.
- POC/IC50 analysis on normalized values where the outliers have been removed.

This CE Template is intended for a plate-based dataset with the following set of columns:

- **Compound:** A column with the compound names. From the values in this column at least one positive and one negative control value are required.
- **Concentration:** A column with the concentration values for the corresponding compound.
- **Raw_value**: The column to analyze.
- **Layout columns:** Plate and well columns.

10.13.2 EC50

This CE Template provides a simple document to explore the data and obtain the EC50 from dose-response data. This is a common procedure in screening assays.

This CE Template is intended for a dataset with the following set of columns:

- **Compound:** A column with the compound names.
- **Concentration:** A column with the concentration values for the corresponding compound.
- **Raw_value:** The column to analyze.

10.13.3 Percentage of Positive Control (POC)

This CE Template provides a simple exploration of the data using the Percent of Positive Control which is one of the percentages of control (POC) measures used in screening assays.

This CE Template is intended for a dataset with the following set of columns:

- **Compound:** A column with the compound names.
- **Concentration:** A column with the concentration values for the corresponding compound.
- **Raw_value:** The column to analyze.

10.13.4 Plate Normalization

This CE Template provides a simple document to normalize plate screening data between plates.

This CE Template is intended for a plate-based dataset with the following set of columns:

- **Compound:** A column with the compound names. From the values in this column at least one positive and one negative value are required.
- **Concentration:** A column with the concentration values for the corresponding compound.
- **Raw_value:** The column to analyze.
- **Layout columns:** Plate, column, row and well columns.

10.13.5 Plate Level QC

This CE Template provides a set of calculations to evaluate the quality of the experiment at the plate level as well as identify outliers, missing values, and any bias in the distribution these may have over the plates.

This CE Template is intended for a plate-based dataset with the following set of columns:

- **Compound:** A column with the compound names. From the values in this column at least one positive and one negative value are required.
- **Raw_value:** The column to analyze.
- **Layout columns:** Plate and well columns.

10.13.6 Sample Level QC

The aim of this CE Template is to visualize the distribution of several features over the samples to assess the quality of the data and detect outliers at the sample level easily.

This CE Template is intended for a plate-based dataset with the following set of columns:

- **Compound:** A column with the compound names.
- **Concentration:** A column with the concentration values for the corresponding compound.
- **Raw_value:** The column to analyze.
- **Plate column:** A column with the plate identifier.

10.13.7 Primary Screen

The aim of this CE Template is to provide the user with a quick visual overview of the plate quality using the information from the positive and negative controls in each plate together with an overview of the signal distribution for each of the compounds over the different plates.

This CE Template is intended for a plate-based dataset with the following set of columns:

- **Compound:** A column with the compound names.
- **Raw_value:** The column to analyze.
- **Type:** A column with the plate layout information. From the values in this column at least one positive and one negative value are required.
- **Plate column:** A column with the plate identifier.

10.14 Calculations Explorer Quick Start Guide

To demonstrate the use of **Calculations Explorer**, an example dataset (DemoDataset.txt) and CE Templates are provided along with the software. The data in this example (DemoDataset.txt) is a screening dataset containing multiple compounds screened over a set of plates using multiple concentrations.

10.14.1 Loading the Data

Load the data into Spotfire® with the default settings (**Data > Add Data**) and browse to the DemoDataset.txt in your local files and select **Open**. A default data view is displayed, select **OK** to import the data, and **OK** again in the resulting **Add data to analysis** menu. Select **Start from visualizations > Table**. A table visualization containing the following information appears:

- **Plate layout information:** PlateID, Well, Row, Column
- **Measurement values for 4 different features:** Feature1 – 4.
- **Compound information:** Generic compound name, Sample type, Compound
- **Concentration information:** Displayed for each well

10.14.2 Applying a CE Template

To apply the POC example CE Template follow these steps:

- Open the **Calculations Explorer** (**Tools > Signals Calculations Explorer**) and drag the data table by selecting its name to the **Apply Template > Load shared protocol** option in the configuration panel.
- Select the POC CE Template from the available Built-In CE Templates and select **OK**.
- The **Match Calculation Columns** dialog will open with a list of the columns from the CE Template that require matching and the columns available in the document.
- Click on **Auto** to perform an automatic matching of the columns and values where possible. Compound is matched automatically as well as the Positive and Negative values of the compound column (see info icon to the right of *Compound* = *Compound* in the **Matched Columns** area).
- Match the RAW_value column with the feature from the demo dataset that you would like to analyze, in the case of this example, Feature 4, by selecting each, then selecting the (>) **Pair** icon. Once this is done the **Save** button is enabled.
- Click on **Save**. The CE Template is executed, and the different analysis views are now available.

10.14.3 Exploring the CE Template

By default, a table listing each compound and its corresponding POC will be shown in the visualization area. A bar chart of the *POC per Compound* is also displayed. This visual allows you to easily select the compounds with the POC of most interest.

From the Calculations Explorer panel on the left-hand side, expand the ‘Compound’ and ‘RowId’ levels to view a list of each available visualization. As an example, you can view the original dataset by selecting the Show/Hide icon next to the RowID table.

10.15 Appendix A: Visualization Elements Supported in the CE Templates

10.15.1 Supported for All Visualizations

Note: Only those elements specifically listed below are supported in the CE Templates. Spotfire® elements and visualizations not explicitly indicated are not supported and will not be kept when updating the CE Template.

- Title
- Data
  - WhereClauseExpression
  - Limit data using expression
- Description
  - ShowDescription
- UseSeparateColorForMarkedItems
- Legend

    - Visibility
    - Width
- Subsets based on custom expressions

10.15.2 Elements from All Supported Visualizations with X and Y Axes

- Axis
    - Expression
- XAxis
    - TransformType
    - Reversed
    - AxisMode
    - Expression
    - ShowAxisSelector
    - ShowAxisSelector
- YAxis
    - TransformType
    - Reversed
    - AxisMode
- ColorAxis
    - Expression
    - AxisMode
- Trellis
    - RowAxis.Expression
    - ColumnAxis.Expression
    - PageAxis.Expression
    - PanelAxis.Expression
    - TrellisMode
    - ManualLayout
    - ManualColumnCount
    - ManualRowCount
- MarkerByAxis
    - Expression
- Data
    - LimitingMarkingsEmptyBehavior

10.15.3 Specific Visualization Supported Elements

10.15.3.1 Scatterplots

- SizeAxis
    - Expression
    - ScaleType
    - AxisMode
- MarkerSize
- MarkerClass
- ShapeAxis
    - Expression

    - AxisMode
- Transparency
- YAxis
    - IndividualScaling
    - IndividualScalingMode
- XJitter
- YJitter
- Curve Label settings
- LabelImageSize
- LabelVisibility
- LabelColumn.Expression
- LineConnection
    - Width
    - UseMarkerColor
    - ShowArrows
    - IsBackground
    - ConnectionAxis.Expression
    - ConnectionAxis.AxisMode
- Lines & Curves
    - ReferenceLineFittingModel
    - StraightLineFittingModel
    - PolynomialFittingModel
    - LogisticRegressionFittingModel
    - ReferenceCurveFromTableFittingModel
    - PowerFittingModel
    - ExponentialFittingModel
- Details
    - DisplayMode
    - Marker.Visible
    - Shape.Visible
    - Size.Visiblec
    - X.Visible
    - Y.Visible

10.15.3.2 Line Chart

- LineByAxis
    - Expression
    - AxisMode
- BreakOnEmpty
- CompensateForMissingTimeSeriesValues
- LabelVisibility
- LineWidth
- MarkerSize
- ShowMarkers
- SteppedLines
- ShowIndividualScaling
- ShowMarkerLabels
- ShowLineLabels

- Transparency
- YAxis
    - IndividualScaling
    - IndividualScalingMode

10.15.3.3 Bar Chart

- BarWidth
- LabelSegments
- LabelVisibility
- HundredPercentBars
- LabelCompleteBar
- LabelOrientation
- LabelPercentageDecimalDigits
- MaxNumberOfLabels
- SortSegmentsBySize
- ReverseSegmentOrder
- SegmentLabelInformationType
- Orientation
- SortedBars
- Transparency
- StackMode
- YAxis
    - IndividualScaling
    - IndividualScalingMode
- Lines & Curves

10.15.3.4 Box Plot

- Show95PercentConfidenceInterval
- ShowDistribution
- UseRelativeScale
- BoxWidth"
- MarkerSize"
- ComparisonCircles.Visible
- ComparisonCircles.Span
- ComparisonCircles.AlphaLevel

10.15.3.5 Heat Map

- MeasureAxis.Expression
- SortOrder
- ColumnDendrogram
    - Visible
    - Span
    - PruningLevel
    - ShowPruningLine
    - UseDataTable

- RowDendrogram
  - Visible
  - Span
  - PruningLevel
  - ShowPruningLine
  - UseDataTable

10.15.3.6 Summary Table

- CategoryAxis.Expression
- ShowGridlines
- AutoAddNewColumns

10.15.3.7 Pie Chart

- Trellis
  - RowAxis.Expression
  - ColumnAxis.Expression
  - PageAxis.Expression
  - PanelAxis.Expression
  - TrellisMode
- ColorAxis
  - Expression
  - AxisMode
- SizeAxis
  - Expression
  - AxisMode
- SectorSizeAxis
  - Expression
  - AxisMode
- SizeAxis.ScaleType.ContinuousScaleType
- Data.LimitingMarkingsEmptyBehavior

10.15.3.8 Cross Table

- ColumnAxis
  - Expression
  - AxisMode
- MeasureAxis
  - Expression
  - AxisMode
- RowAxis
  - Expression
  - AxisMode
- ShowColumnGrandTotal
- ShowAxisSelectors
- ShowContinuousColor
- ShowRowGrandTotal

- ShowTopNColumns
- ShowTopNRows
- Color Schemes

10.15.3.9 Table plots

- Color Schemas
- Columns
- Column sorting
- Row values sorting
- RowHeight
- HeaderHeight
- WrapCellText
- WrapHeaderText

10.15.3.10 Graphical Tables

- RowAxis (only for calculated new value)
  - Expression
  - AxisMode
  - ShowAllCategories
- Column sorting
- Row values sorting
- ShowCellBorders
- ShowHeaderColumnHeader
- ShowHeaderRow
- ShowRowAxisSelector
- ShowTopNRows
- TopNRowCount

11. SPR Domain Apps

11.1 Introduction

This section will describe the Signals Apps provided to analyze data in the **Surface Plasmon Resonance (SPR)** domain.

11.2 Launching Apps

To launch an App from Spotfire® Analyst, from the main toolbar, navigate to **Tools** > **Signals Apps.** This will open the Signals Apps tab. Alternatively, if this tab is already present in the document, simply navigate to it using the appropriately labelled tab at the bottom of the screen.

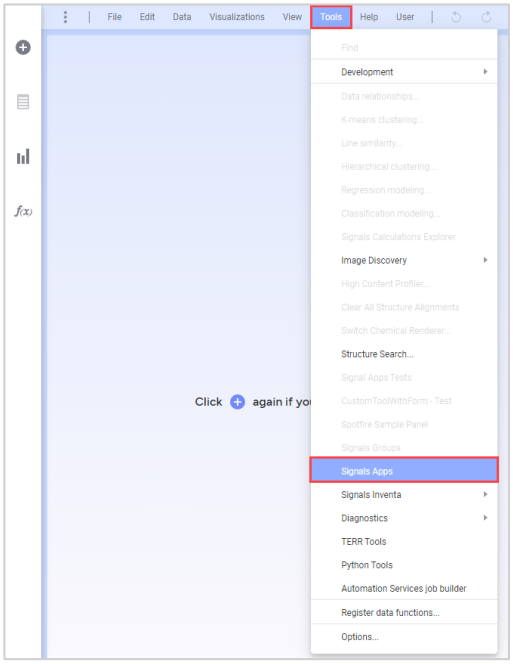


Figure 11-1: Launching Signals Apps

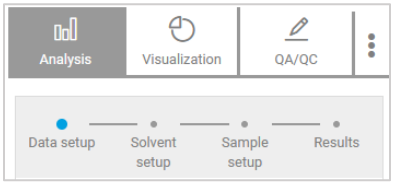
The Signals Apps tab will display the available Signals Apps. Navigate to the **Surface Plasmon Resonance (SPR) Apps** domain. Select the desired App card to launch the App.

Figure 11-2: SPR Apps on the Signals Apps Page

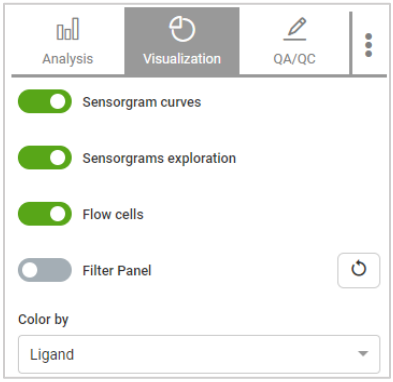
11.3 App Tab Overview

Most SPR Apps contain a user interface area with three tabs (or sections), **“Analysis”, “Visualization”, “QA/QC”** and an additional panel **“…”,** displayed in the top-left-hand corner. Apps differing from the general structure described below are described in detail in their corresponding App section.

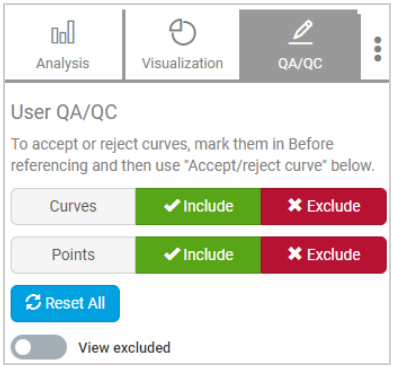
- **Analysis:** Contains the main controls for using the App and may contain one or more steps, as shown in the example below. The current step is highlighted in blue.



- **Visualization:** Provides App-specific toggles to control which visualizations are displayed and options for customizing those visualizations, as shown in the example below.



- **QA/QC:** Contains data exclusion controls that allow the user to select specific data points from the *Sensorgram(s) details* / *Sensorgrams Before and After* visualizations and exclude them from downstream analysis. It also allows the user to either show or hide those excluded data points from the visualizations.

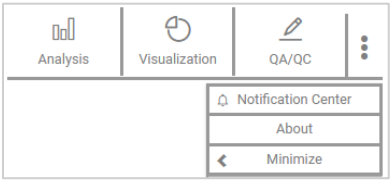


The controls available in this menu are:

- **Curves**:
  - **Include** all data points belonging to those curves for which any data point is selected
  - **Exclude** all data points belonging to those curves for which any data point is selected
- **Points**:
  - **Include** all selected data points
  - **Exclude** all selected data points
- **Reset All**: Resets the tags in all data points so they are all included in downstream analyses.

View excluded: Allows the user to toggle on or off the display of those data points that have been excluded. When set to "On" those data points that are excluded from the fitting are visible in the plot. However, they are represented as a red ***'X'*** to distinguish them from those data points that will be included in the downstream analysis, which are colored gray.

- **'…':** Contains additional information related to the App. Here users can find the "Notification Center", "About", and the ability to "Minimize" the User Interface.



- **Notification Center:** Permits access to detailed information, warnings, and error messages produced while performing an analysis. Selecting a title will filter the messages by type.
- **About:** Contains general information about the App, including simplified App name and version:
- **Minimize:** Collapses the UI to provide visualizations more space in the document.

11.4 SPR Data Import App

The **SPR Data Import** App provides an automated solution for importing SPR data into Spotfire® as well as a set of interactive visualizations allowing manual exploration and analysis of the imported curves. This App can import data in a variety of formats, depending on the instrument source.

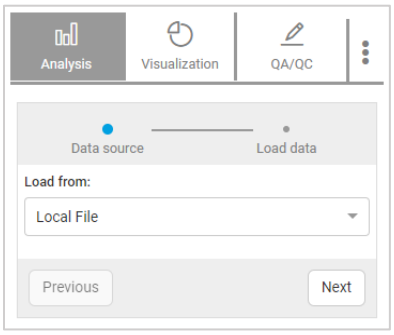
From the Signals Apps page, select the **SPR Data Import** App card. A new tab containing the App will be added to the document and the App will be launched.

Figure 11-3: SPR Data Import App Card

11.4.1 Configuring the SPR Data Import App

11.4.1.1 Analysis Tab

The **Analysis** tab guides the user in preparing and loading the data with the following two steps: *Data source* and *Load data*.



The user has the following controls available from the *Data source* step of the **Analysis** tab**:**

1. From the '**Load From:'** dropdown, select if the data should be loaded from:
    - An external file via *'Local File'.*
    - An existing table from Signals via '*Signals Experiment'.*
2. Select **Next** to continue to the *Load data* step and proceed with one of the following options depending on if the data is loaded from an external file, or from Signals.
3. *In the case of an external file via 'Local File':*

    From the '**Select import format:'** dropdown, select the source instrument from which the data is generated. The available options are:

    - **Biacore T200 / S200:** The App supports importing of .blr files from T200 and S200 instruments.
    - **IBIS MX96:** The App supports importing of .ibmx files exported from the instrument.
    - **BIAevaluation B4000:** The App supports importing curve files exported in text format from the BIAevaluation software. The exported file should contain the following columns: "Cycle", "Curve", "Sample", "Conc (µg/ml)", "Capture", "X", "Y", "Y, ref sub".
    - **Octet RED384:** The App supports importing .frd files exported from the instrument.
    - **Carterra:** The App supports importing of .sprdata files output by the Carterra instrument.
    - **Biacore Insight Export:** The App supports importing .json files exported from the Biacore Insight software. These files should preferably contain raw sensorgram data.

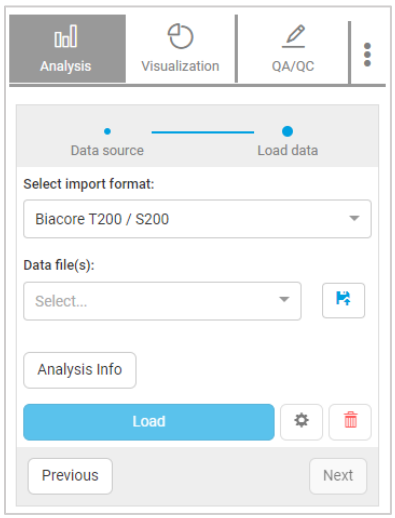
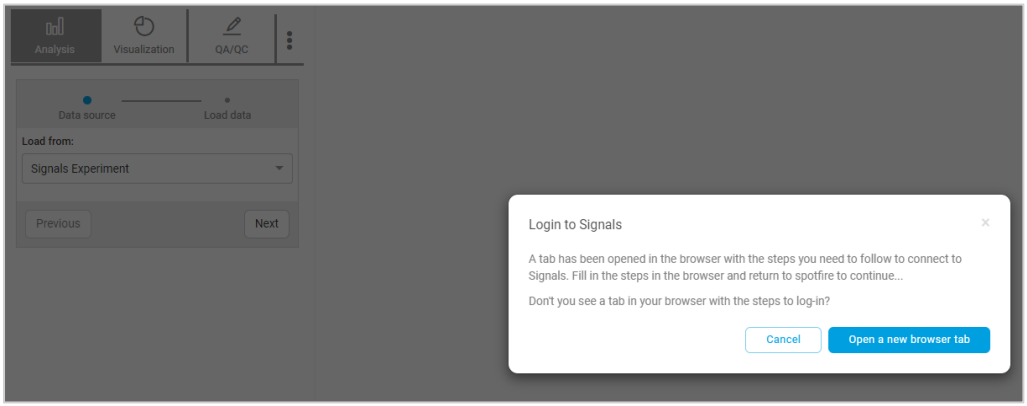
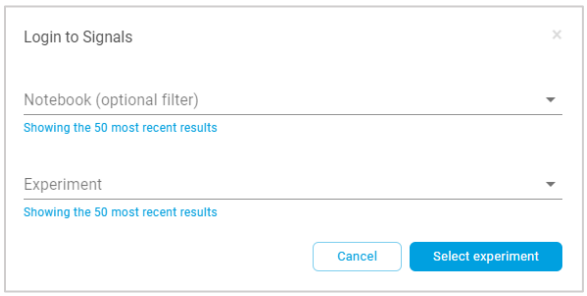


Figure 11-4: Data Import Settings – Select Import Format

In the case of an external file via 'Signals Experiment':

- A "Login to Signals" dialogue box displays in Spotfire® and redirects to the Signals Login page in the browser.





Note: Logging into Signals will allow access to Signlas Notebook(s) configured by the System Administrator.

Because a Signals Experiment has been selected previously:

- A warning message is raised indicating that there are no valid files if there are no valid files (any file in the experiment matches with any of the instrument types).
- The '**Select import format:'** is set to the default value: Biacore T200 / S200 if there are files belonging to different instrument types.
- The '**Select import format:'** is automatically set to the corresponding instrument type if there are valid SPR files belonging to any of the instrument types.

4. For loading the files:

In the case of an external file via 'Local File':

Add the desired data file(s) by selecting the **Upload file** icon and navigating to the file(s) locally. Select **Open** to import the files and note the blue progress bar under the icon, indicating the file(s) are successfully importing. If necessary, files can be removed by selecting the ***'X'*** icon next to the file name.

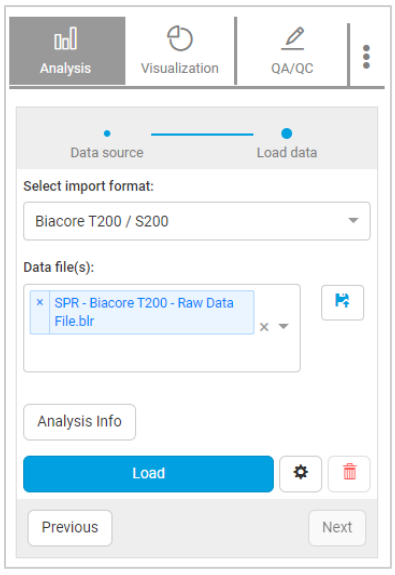
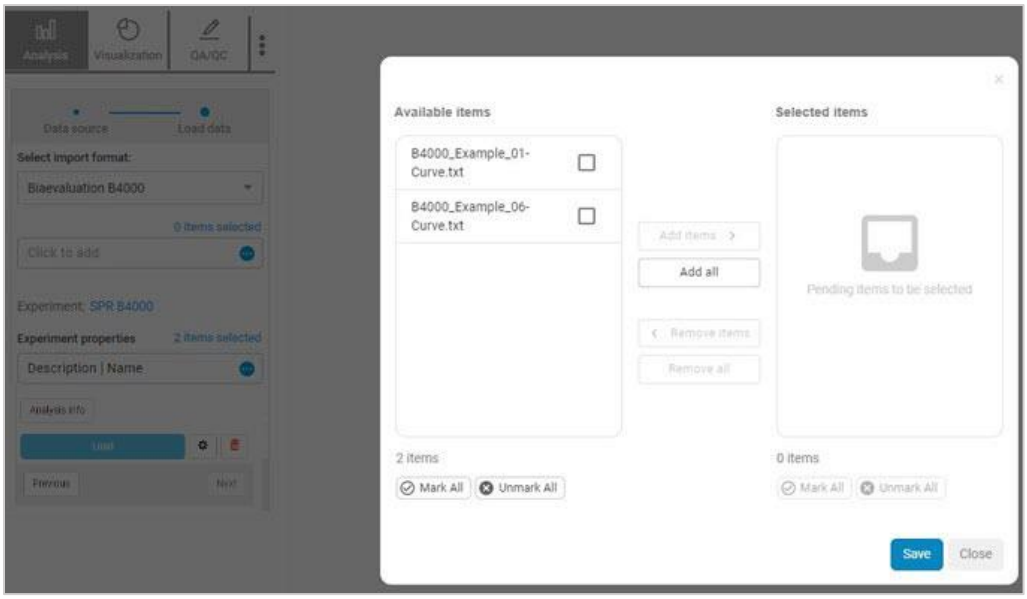


Figure 11-5: Data Import Settings – Uploaded Data File

In the case of an external file via 'Signals Experiment':

- Add the desired data file(s) by clicking on the three dots next to the **Upload file** selector. When the popup is open, select the desired file from the list of available files**,** and click on **'Add items'** to add it to the **'Selected items'** list.
- Confirm the selection by clicking on the **Save** button, the popup is closed and the 'Upload file' selector is set to the file you have chosen.

Note: For IBIS and B4000, multiple files are not supported.

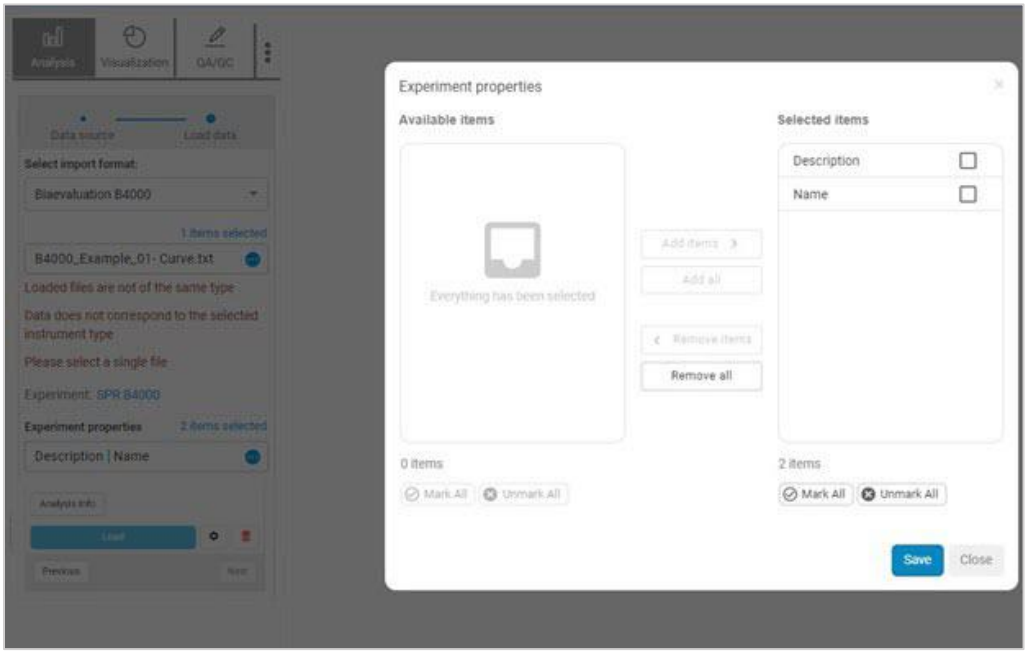


If necessary, before clicking on **Load**, files can be removed by selecting the three dot icon next to the file name. In the resulting menu, select the file to be removed from **Selected items** and click on **Remove all** or **Remove items**.

5. Information about the experiment appears next to the Experiment label. Select the Experiment name to be redirected to the browser to see the experiment.

6. The experiment properties can be configured by selecting the three dots to the right of **Experiment properties**. A popup will appear with all the properties marked as **Selected items**. Select a subset and click on **Remove items** to remove the selection from the list of properties or click on **Remove all** to discard them. Select **Save** to close the menu and the **Experiment properties** component is now updated to show all the selected properties.

   Note: The Experiment properties are the same as defined in the Signals Notebook template used. These properties are configured by the administrator. This Experiment properties information is available in the Results Table and in the Analysis Info Table.



7. The **Settings** icon will be enabled for specific instrument types once an appropriate file is uploaded, allowing the user to optionally select the columns corresponding to the ligand, analyte, and concentrations. Select **Close** to return to the main **Analysis** tab. See below for an example of the menu when a Biacore T200/S200 instrument file is imported. Select **Load** to load the selected datafiles into the App and for storing the experiment properties in the *Analysis Info* table and in the *Results Table*. Once loaded, the remove option is disabled for the loaded files.

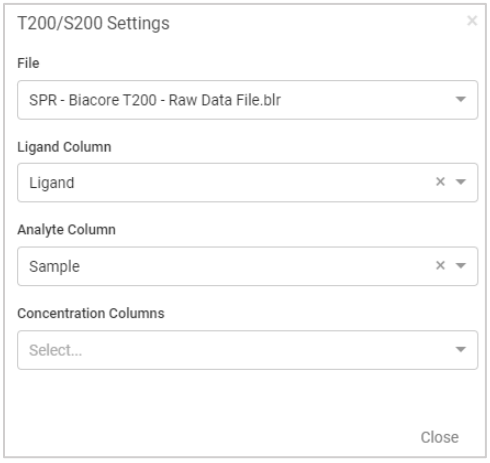


Figure 11-6: T200/S200 Settings UI

8. The **Trash** icon is enabled after selecting **Load**. Selecting the **Trash** button will remove all loaded files and reset the App.
9. When loading Local files, the **Analysis Info** button launches the **Analysis Settings** menu, used to enter optional experiment level details. In the case of a Signals Experiments, this information is loaded from the Notebook properties. Select **Save** to include the entered details as SPR metadata.

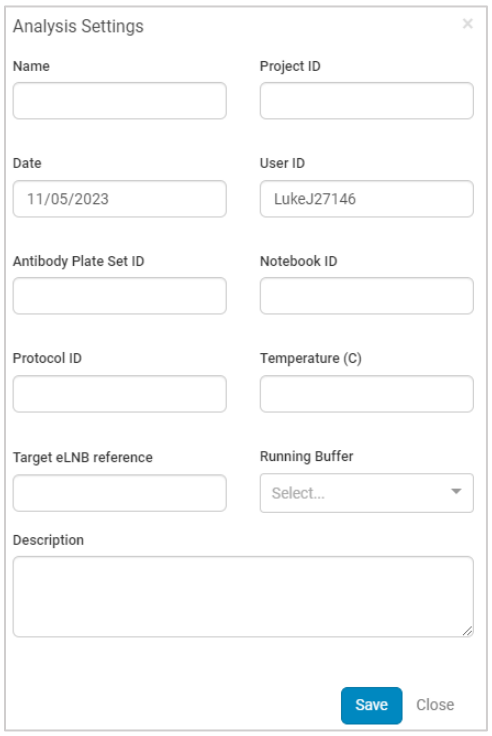


Figure 11-7: Optional Analysis Settings Menu

11.4.1.2 Visualization Tab

The **Visualization** tab provides finer control over what is displayed in the visualizations pane using the following controls:

- **Sensorgram curves:** This toggle will hide or show the *Sensorgram(s) details* visualization.
- **Overview:** This toggle will hide or show the *Curve exploration* table visualization.
- **Sensorgram curves lines:** This toggle will hide or show the *Sensorgrams* plot visualization.
- **Show raw data:** This toggle will hide or show raw data in the *Sensorgrams* plot and the *Sensorgram(s) details* visualization.
- **Show report points:** This toggle will hide or show the report points on the *Sensorgram(s) details* visualization.
- **Show detailed view**: This toggle will hide or show additional details in the *Curve exploration* table visualization.
- **Trellis by Experiment:** This toggle is enabled by default and will trellis the Sensorgram(s) details visualization by experiment.

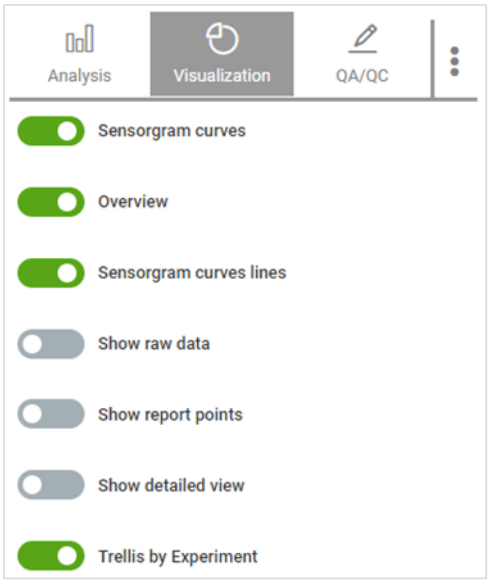


Figure 11-8: Visualization Tab

11.4.1.3 QA/QC Tab

The **QA/QC** tab contains data exclusion controls that allow the user to select specific data points from the *Sensorgram(s) details* plot and exclude them from downstream analysis. It also allows the user to either show or hide those excluded data points from the visualizations. See the section *App Tab Overview* for additional details.

11.4.2 SPR Data Import App Results

Once the data is loaded, the right-hand side of the screen will populate with the **SPR Data Import** App visualizations.

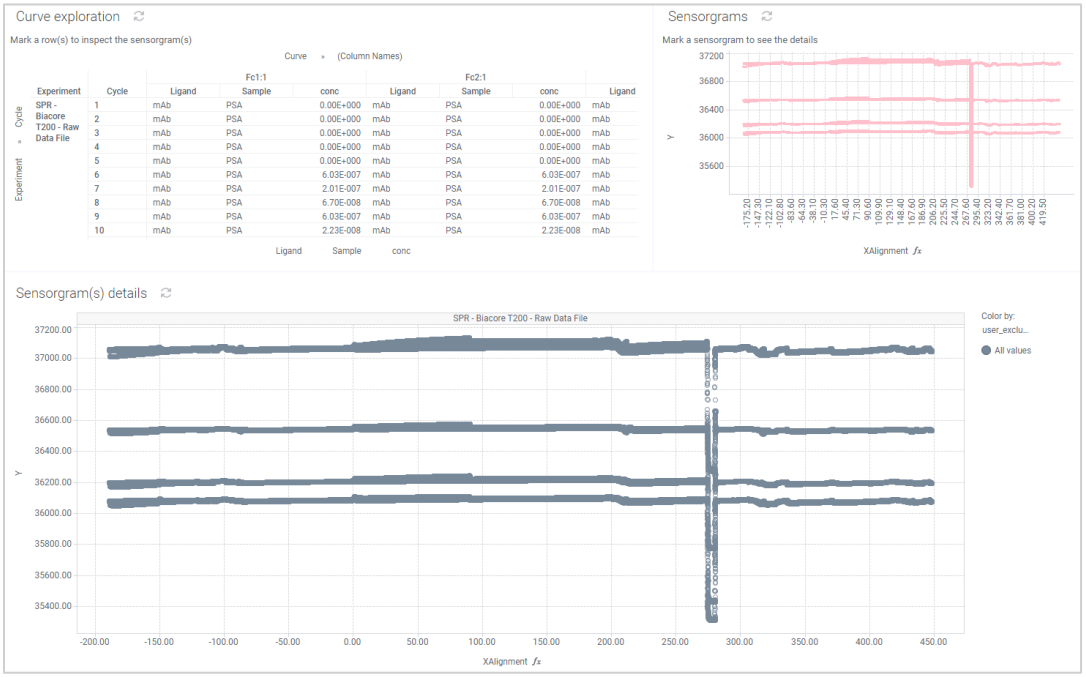


Figure 11-9: SPR Data Import App Visualizations

The visualizations generated by the SPR Data Import App are the following:

- *Curve exploration*
- *Sensorgrams*
- *Sensorgram(s) Details*

11.4.2.1 Curve Exploration Table

The *Curve exploration* table contains the imported sensorgrams in list form and by selecting curves of interest, it can be used to control the sensorgrams that are displayed in the *Sensorgram(s) details* visualization.

Multiple sensorgrams can be selected by highlighting and selecting the desired rows using the *Control* while selecting with the mouse. All rows can be selected by clicking on the top row and then navigating to the last row while holding the *Shift* key. Alternatively, to select all the sensorgrams after a subset of curves has been selected, right-click on the *Curve exploration* table **> Marked Rows > Unmark**. This deselects the highlighted sensorgrams. Right-click again in the *Curve exploration* table **> Marked Rows > Invert**.

Column	Description
Experiment	Shows the name of the file to which the data corresponds
Cycle	Cycle number of the curve
Ligand	Name of the ligand used in the experiment (molecule bound to the surface of the flow cell)

Sample	Name of the sample used in the experiment (molecule that is injected over the flow cell)
conc	Molar concentration of the sample
Curve	Curve ID of the curve

11.4.2.2 Sensorgrams Visualization

Sensorgrams: the pink sensorgrams in the upper right-hand corner control the sensorgrams displayed in *Sensorgram(s) details* visualizations below. To magnify regions of interest in the *Sensorgram(s) details* visualizations, right-click and drag the desired region in the pink sensorgrams. Double-click on the pink sensorgrams plot to restore all the data in the *Sensorgram(s) details* visualization.

11.4.2.3 Sensorgram(s) Details Visualization

Here, the imported sensorgrams can be viewed where each of the individual measurements is represented as a separate data point.

- Information on exclusion state of the data points is displayed, with the excluded data represented as red x points and the data included for downstream analysis as grey circles. (See section *App Tab Overview* for additional details on excluding data).
- Report points are displayed as vertical lines in the visualization if present and the **Show report points** toggle is on.
- By default, the aligned data is displayed, which facilitates the viewing of report points.
- If multiple experiments are loaded, by default the data is trellised by experiment, it is possible to see all the data together with no trellis by changing the toggle in the Visualization tab.

11.5 Preprocessing Apps

The preprocessing Apps provide the tools to prepare raw SPR data for kinetic and steady state analyses. They include the following Apps:

- Zeroing
- Alignment
- Cropping
- Referencing
- Blank Subtraction
- Solvent Correction

The section below describes the general layout and visualizations within the preprocessing Apps. The sections following describe in detail the use and functionality of each preprocessing App.

11.5.1 Visualizations within Preprocessing Apps

The basic structure of the preprocessing Apps is shown in the following figure. It is divided into controls specific to each preprocessing step on the left-hand side within a box labeled **Data Configuration**. The body of each preprocessing App contains a *Curve exploration* table, a *Sensorgrams* plot (pink sensorgrams) and the *Before* and *After* visualizations.

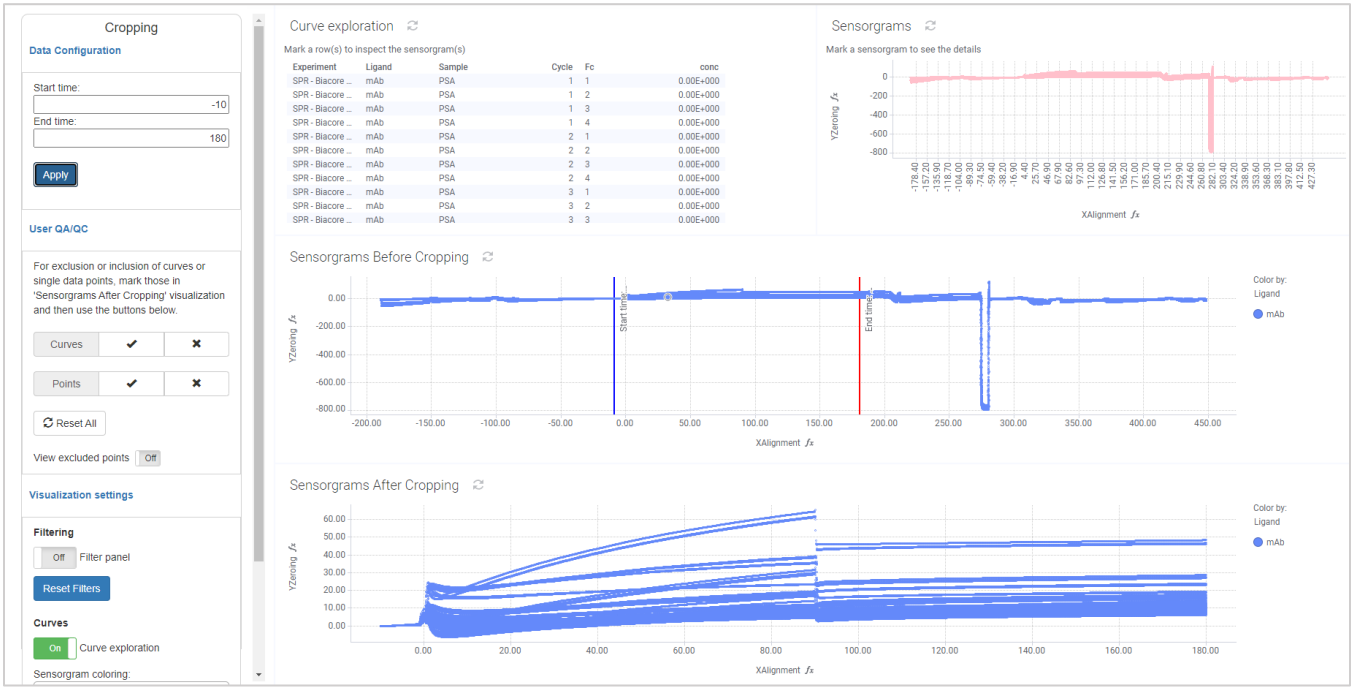


Figure 11-10: Basic Structure of the Preprocessing Apps

11.5.1.1 Curve Exploration Table

The *Curve exploration* table contains the imported sensorgrams in list form and by selecting curves of interest, it can be used to control the sensorgrams that are displayed in the *Before* and *After sensorgrams* visualizations.

Multiple sensorgrams can be selected by highlighting and selecting the desired rows using the *Control* while selecting with the mouse. All rows can be selected by clicking on the top row and then navigating to the last row while holding the *Shift* key. Alternatively, to select all the sensorgrams after a subset of curves has been selected, right-click on the *Curve exploration* table **> Marked Rows > Unmark**. This deselects the highlighted sensorgrams. Right-click again in the *Curve exploration* table **> Marked Rows > Invert**.

Column	Description
Experiment	Shows the name of the file to which the data corresponds
Cycle	Cycle number of the curve
Ligand	Name of the ligand used in the experiment (compound bound to the surface of the flow cell)

Sample	Name of the sample used in the experiment (compound that is injected over the flow cell)
Conc	Molar concentration of the sample.
Fc	Flow cell ID for the curve

11.5.1.2 Sensorgrams Visualization

Sensorgrams: the pink sensorgrams in the upper right-hand corner control the sensorgrams displayed in the *Before* and *After sensorgrams* visualizations. To magnify regions of interest in the *Before* and *After sensorgrams* visualizations, right-click and drag the region in the pink sensorgrams. Double-click on the pink sensorgrams plot to restore all the data in the *Before and After sensorgrams* visualizations.

11.5.1.3 Sensorgrams Before and After Visualizations

The information displayed in these sensorgrams depends on the preprocessing App. In all cases what is displayed in the *"Sensorgrams Before …"* visualization is the sensorgrams before applying the preprocessing step and the *"Sensorgrams After …"* visualization is the sensorgrams after applying the preprocessing step.

11.5.1.4 Visualization Settings

Each preprocessing App contains a **Visualization settings** section/tab that provides finer control of the visualization displays. The visualization controls available in the different Apps vary depending on the visualizations available in the results section.

Controls available in this menu are:

- **Sensorgrams curves**: This toggle will hide ("Off") or display ("On", default) the *Before and After* sensorgram visualizations.
  - **Sensorgrams exploration**: This toggle will hide ("Off") or display ("On", default) the *Curve exploration* table and the pink *Sensorgrams* visualization.
  - **Filter Panel**: This toggle will provide access to the filter panel which will open to the right-hand side of the visualizations and allows filtering using criteria specified by the user.
  - **Reset Filters**: This button will reset the filtering, so all the data that was filtered out is again visible.
  - **Color by:** This dropdown allows the user to change how the *Before* and *After* visualizations are colored (Ligand by default). Note that in some pre-processing Apps this dropdown is named "Sensorgram coloring:".

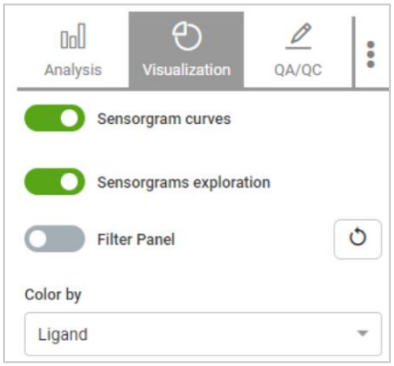


Figure 11-11: Visualization Settings

11.5.2 Zeroing App

The **Zeroing** App aligns the sensorgrams on the Y-axis.

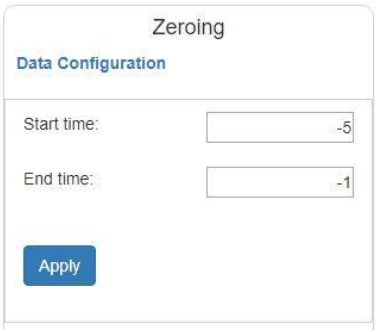
Figure 11-12: Zeroing App Card

11.5.2.1 Configuring the Zeroing App

Controls required to zero the sensorgrams can be found in the box in the left-hand panel, labeled “Data Configuration”.

To zero the sensorgrams:

1. Set the **Start time** for the start for zeroing. The location of this point on the data is marked by a blue vertical line in the *Sensorgrams Before Zeroing* visualization.
2. Set the **End time** for end of zeroing. The location of this point on the data is marked by a red vertical line in the *Sensorgrams Before Zeroing* visualization.
3. Click **Apply** to perform the zeroing operation. This operation averages the Y-axis values between the input start and end times and sets this value equal to zero on the Y-axis. The entire sensorgram is adjusted on the Y-axis based on these zero values.



The start and end times can also be set within the *Sensorgrams Before Zeroing* plot by clicking on the X-axis and dragging the mouse to select the required region, or simply selecting an area of the Sensorgram plot. When set in this manner the **Start time** value will be set to the lowest marked value and the **End time** value to the highest marked value on the X-axis.

11.5.2.2 Zeroing App Results

Once the curves are zeroed, the right-hand side of the screen will display both the *Sensorgrams Before Zeroing* and *Sensorgrams After Zeroing* visualizations, and the controls in the user interface will be available to further refine the analysis.

11.5.3 Alignment App

Alignment is performed automatically for T200 *.blr* files, Carterra files, and Insight exported files during **SPR Data Import**. In the case of other instruments, curve alignment is supported for IBIS MX96 data, where if needed, should be performed using the **Alignment** App.

Figure 11-13: Alignment App Card

11.5.3.1 Configuring the Alignment App

Once the **Alignment** App is open, the controls needed to align the sensorgrams are visible in the box on the left-hand side of the screen.

To align the data from flow cells in series to a common time point on the X-axis, select **Apply** in the **Automatic Alignment** section. This is currently the only option provided by the **Alignment** App.

Curve alignment will subtract from each point the value of the injection start as recorded in the corresponding files. This will align the injection start of each of the sensorgrams to zero seconds on the X-axis.

11.5.3.2 Alignment App Results

Once the curves are aligned, the right-hand side of the screen will display both the *Sensorgrams Before Alignment* and *Sensorgrams After Alignment* visualizations, and the controls in the user interface will be available to further refine the analysis.

11.5.4 Cropping App

The **Cropping** App crops the sensorgrams to the region of interest and removes extraneous data.

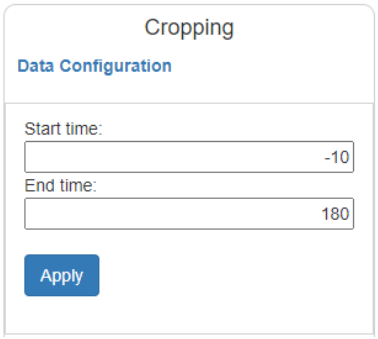
Figure 11-14: Cropping App Card

11.5.4.1 Configuring the Cropping App

Cropping is used to magnify the region of interest, typically the association and dissociation phases of the sensorgrams. Controls required to crop the sensorgrams are in the box on the left-hand side of the screen labeled **Data Configuration**.

To crop the data:

1. Set the **Start time** for cropping. The location of this point is marked by a blue vertical line in the *Before Cropping Sensorgrams* visualization.
2. Set the **End time** for cropping. The location of this point is marked by a red vertical line in the *Before Cropping Sensorgrams* visualization.
3. Click **Apply** to crop the data.



The **Start** and **End** times can also be set within the *Sensorgrams Before Cropping* plot by clicking on the X-axis and dragging the mouse to select the required region, or simply selecting an area of the sensorgram plot. When set in this manner, the **Start time** value will be set to the lowest marked value and the **End time** value to the highest marked value on the X-axis.

11.5.4.2 Cropping App Results

Once the curves are cropped, the right-hand side of the screen will display both the *Sensorgrams Before Cropping* and *Sensorgrams After Cropping* visualizations, and the controls in the user interface will be available to further refine the analysis.

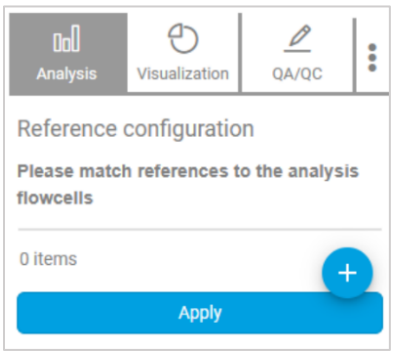
11.5.5 Referencing App

The **Referencing** App provides a solution for subtracting the sensorgram data corresponding to a reference flowcell from the sensorgram data to be analyzed. This App also allows the manual exploration and analysis of the final referenced curves. A set of interactive visualizations are provided to easily explore the results.

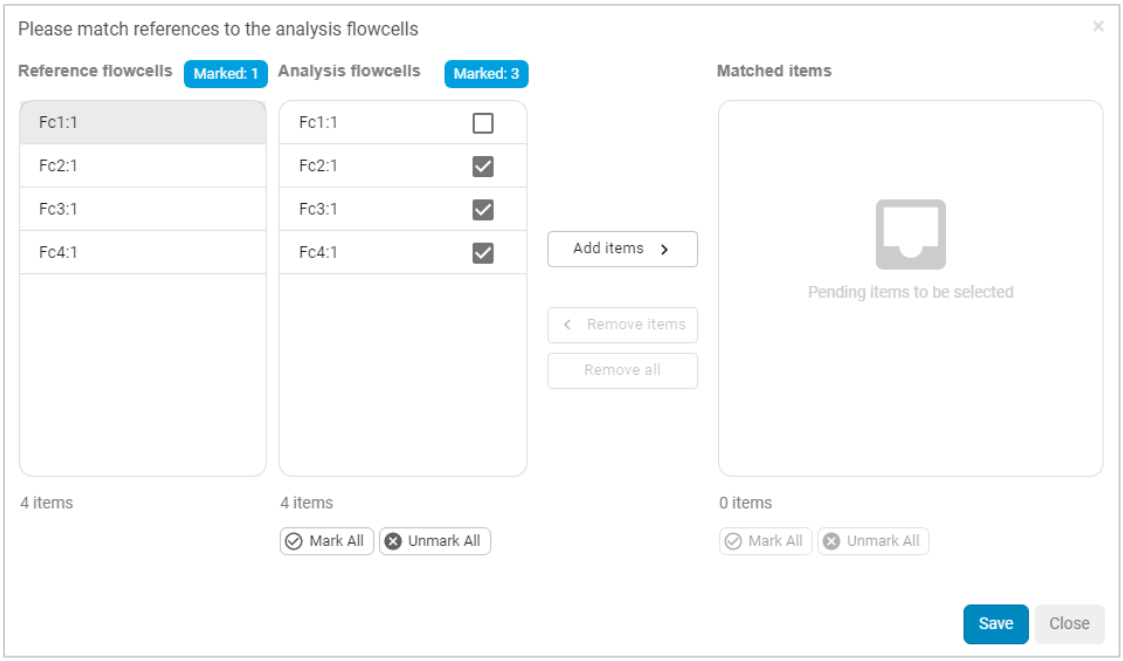
Figure 11-15: Referencing App Card

11.5.5.1 Configuring the Referencing App

From the **Analysis** tab, the **Reference configuration** section allows the user to launch the matching interface by selecting the + icon.



In the matching interface, the user can choose the reference cell and all those cells that will be referenced to it. A single reference cell can be chosen and any number of cells to be referenced to it by using the checkbox in the 'Analysis flowcells' list. Additionally, the search box can be used to limit the displayed flowcells in any of the lists and simplify the selection.



Once the correct matching has been chosen select **Add items** to move the selection to the 'Selected items' box.

Remove items can be used to remove all marked matches from the 'Selected items' box and **Remove all** will reset all matches.

Once the matching is satisfactory select **Save** to store the matches and display them in the **Referencing** App left-hand-side panel.

Selecting **Close** will close the matching interface without any changes to the document.

Once the reference flowcells have been matched as needed and before selecting **Apply**, the sensorgrams of the reference surface and the ligand surface before referencing are visible.

> **Note:** Circular referencing is not allowed where a cell is used as a reference while being itself referenced to another. In this case an error will be raised when attempting to save the matching.
>
> **Note:** Flowcells that are referenced to themselves will result in 0 for all referenced values.

Select **Apply** and the *Sensorgrams After Referencing* visualization is displayed.

Figure 11-16: Referencing App Visualizations

11.5.5.2 Referencing App Results

Once the curves are referenced, the right-hand side of the screen will display both the *Sensorgrams Before Referencing* and *Sensorgrams After Referencing* visualizations, and the controls in the **Visualization** and **QA/QC** tabs will be available to further refine the analysis.

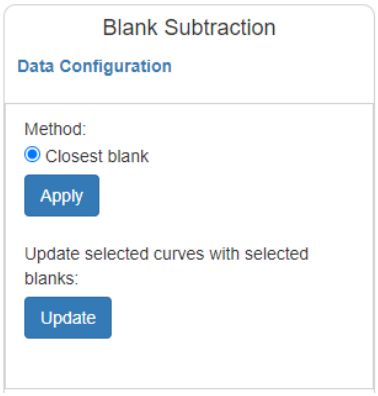
11.5.6 Blank Subtraction App

The **Blank Subtraction** App allows the user to select and subtract a blank from the analyte sensorgrams.

Figure 11-17: Blank Subtraction App Card

11.5.6.1 Configuring the Blank Subtraction App

Once the Blank Subtraction App is open, the *Blanks Table, Blanks Sensorgrams, Curve exploration, Sensorgrams* and the *Sensorgrams Before Blank Subtraction* plot are visible. Controls to perform the blank subtraction of the sensorgrams are shown in the **Data Configuration** box on the left-hand side of the screen.



To blank subtract the data:

- Set the **Method** for blank subtraction. In the current implementation the only method to perform blank subtraction is 'Closest blank'. This method uses the closest blank injection preceding the analyte injection of interest to perform blank subtraction. If no blank is available before the analyte injection, the closest blank after the injection is used.
- Select **Apply** and the data from closest blanks will be subtracted from the analyte sensorgrams.
- **Update**: The update button allows the sensorgrams used for blanking to be changed for selected curves. To do this:
    a. Select the blank to be used from the *Blanks Table*. If more than one blank is selected the first one will be used.
    b. Select the sensorgrams to be updated from the *Curve exploration* table.
    c. Select the **Update** button and the blank subtracted from the selected sensorgrams will be updated to the one selected in the Blanks Table.

Figure 11-18: Blank Subtraction Visualizations After Applying

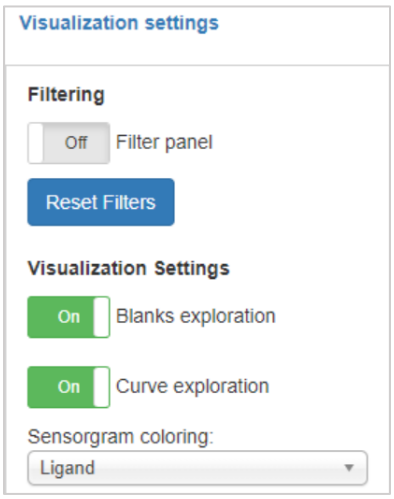
11.5.6.2 Blank Subtraction App Results

When blanks are subtracted, the right-hand side of the screen will show both the *Sensorgrams Before* and *After Blank Subtraction* and the controls for the Visualizations within Preprocessing Apps in the user interface will be available to further refine the analysis.

After executing blank subtraction, two additional columns will be added to the *Curve exploration* table that provide information on the blank used in the blank subtraction operation for each of the sensorgrams.

Column	Description
ref_curve	Shows the blank sensorgram used for the blank subtraction of this curve
ref_cycle	Shows the cycle used for the blank subtraction of this curve

The **Visualization Settings** box in the lower left-hand corner in **Blank Subtraction** App has additional functionality beyond what is provided in each preprocessing App. It contains buttons to turn off the visualizations used for the custom blank subtraction.



- **Blanks exploration**: The Blanks exploration toggle will hide (“Off”) or display (“On”, default) the *Blanks Table* as well as the pink Blanks *Sensorgram* plot.

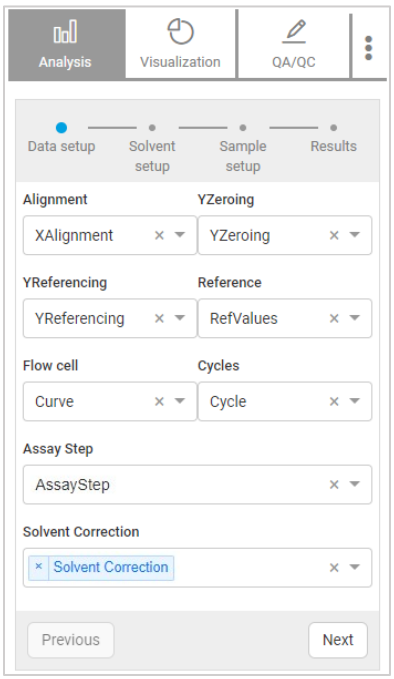
11.5.7 Solvent Correction App

The **Solvent Correction** App provides functionality to correct any signal distortions caused by the solvent used in small molecule signal measurements. The App allows the selection of solvent correction regions in the solvent cycle, as well as the position on the reference to be used for the correction. Using the provided data, the solvent correction curves are fitted and subsequently the correction to be used for each sample is estimated.

Figure 11-19: Solvent Correction App Card

11.5.7.1 Configuring the Solvent Correction App

The **Analysis** tab contains a wizard to navigate through the different steps in the configuration of the required parameters.



As seen in the figure above, the UI provides a 3-step wizard, before the results, that simplifies the configuration of the analysis:

- **Data setup:** The solvent setup step provides the following controls:
    - **Alignment**: Allows the selection of the alignment column if different from the one automatically selected.
    - **YZeroing**: Allows the selection of the Y zeroed column if different from the one automatically selected.
    - **YReferencing**: Allows the selection of the Y referenced column if different from the one automatically selected.
    - **Reference**: Allows the selection of the reference values if different from the one automatically selected.
    - **Flow cell**: Allows the selection if the flow cell column if different from the one automatically selected.
    - **Cycles**: Allows the selection of the cycle column if different from the one automatically selected.
    - **Assay Step**: Here the user can select the column that contains the information on the type of cycle.
    - **Solvent Correction**: This control allows the selection of the specific values for the **Assay Step** column selected in the previous control that indicates a Solvent cycle.

  The visualization area will display the *Solvent cycles*, the *Solvent cycles referenced*, the *Sample cycles* and the *Sample cycles referenced*.

  Once configured, click **Next** to proceed to the next step.
- **Solvent setup:** This grid allows the definition of the regions where deltas for the solvent correction curves will be determined. A default value of 10 seconds before injection start and end is set. Upon any change of the default value(s) and applying it, the *Solvent cycles* visualization will be updated to reflect it. The *'Solvent*

setup' step provides the information for each of the sample cycles that is read automatically from the file and needed to calculate the deltas. It allows the user to modify any of the information as needed. The available columns are:

- **Experiment:** Experiment name
- **Phase:** The solvent correction phase
- **Window range**: The range over which the results will be calculated
- **Time before Start of injection:** Seconds before the event corresponding to the phase start
- **Time before Stop of injection:** Seconds before the event corresponding to the phase end
- **Event Alignment Start of injection:** Time at which the phase start of injection is set
- **Event Alignment end of injection:** Time at which the phase end of injection is set

Once the events are correctly setup, select **Next** to proceed to the next step.

- **Sample setup:** For the *Sample setup* step, the configuration is very similar to the previous step, except that instead of defining regions as for solvent injections, the sample injections are configured.

  The configuration is very similar, as are the visualizations. These consist of a data grid where the correction region can be edited and the Sensorgrams, showing the sample cycles zeroed, are plotted. The vertical line shows the position of the correction region.

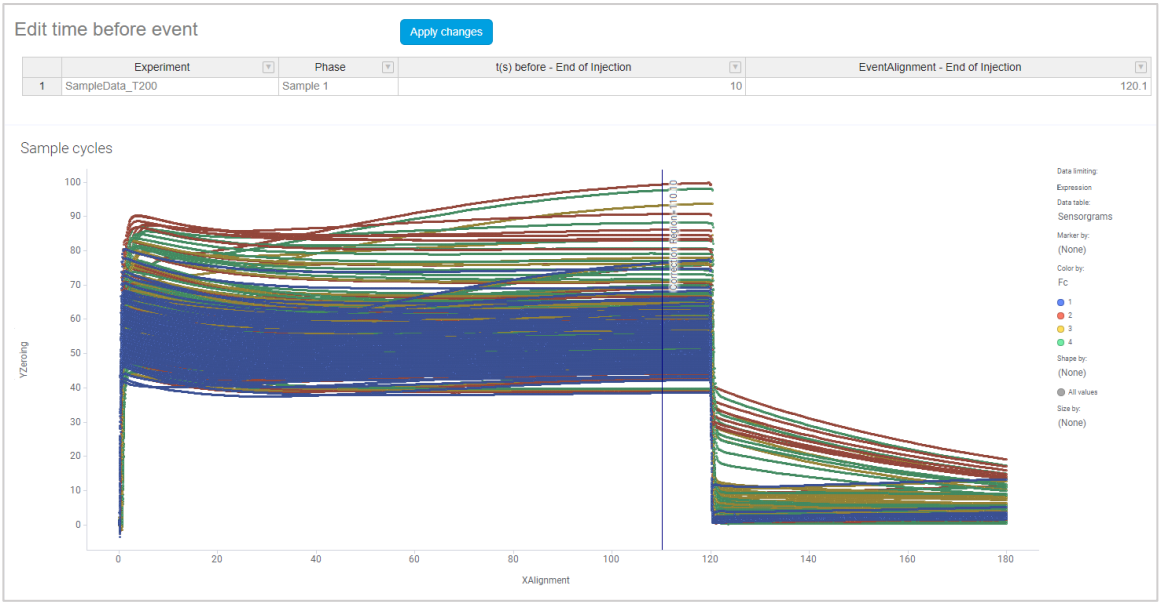
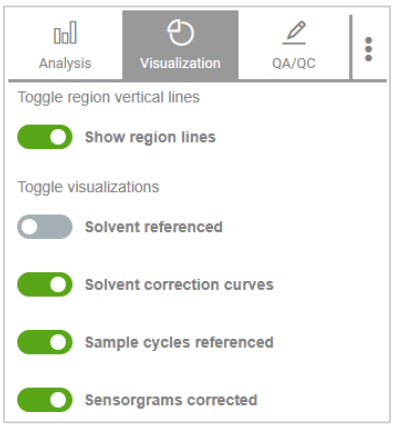


Figure 11-20: Sample Setup Visualizations in the Solvent Correction App

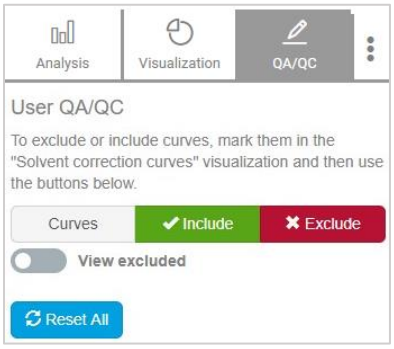
In the **Visualization** tab, there are several toggles available to control the elements displayed:



The available options are:

- **Toggle region vertical lines:** This will show or hide the vertical lines in the solvent cycles referenced visualization (hidden by default).
- **Toggle visualizations:** Here there are four toggles corresponding to the four visualizations available in the page that can be either turned on or off. By default they are all on except for *Solvent cycles referenced*.

The **QA/QC** tab contains the necessary controls to exclude or include data points and to toggle on or off the visualizations of those excluded data points.

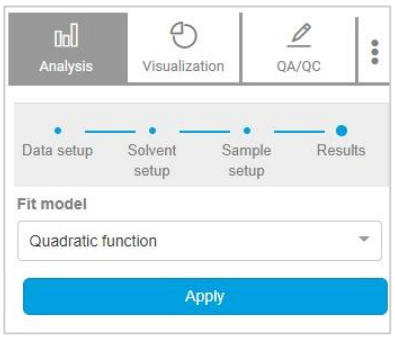


11.5.7.2 Solvent Correction App Results

Once the data configuration is complete, select **Next** to display the results page. In this page there are three visualizations:

- *Solvent Correction curves*
- *Sample cycles referenced*
- *Samples corrected*

In the UI there is a set of controls to configure the solvent correction curve fit.



Here the fitting model can be selected. The available models are:

- Quadratic
- Linear
- None (No correction)

Once the required model is selected, select **Apply** and the analysis will complete, displaying the corrected results in the *Sample cycle corrected* visualization.

11.6 QA-QC App

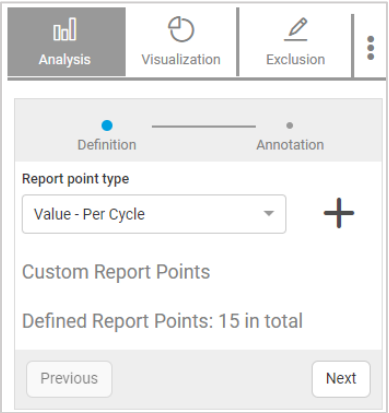
The **QA-QC** App supports the addition of custom report points, displaying a set of visualizations that provide for each of the sensorgrams in the analysis, information on the response at specific points of the curve marked by the report points.

Figure 11-21: QA-QC App Card

11.6.1 Configuring the QA-QC App

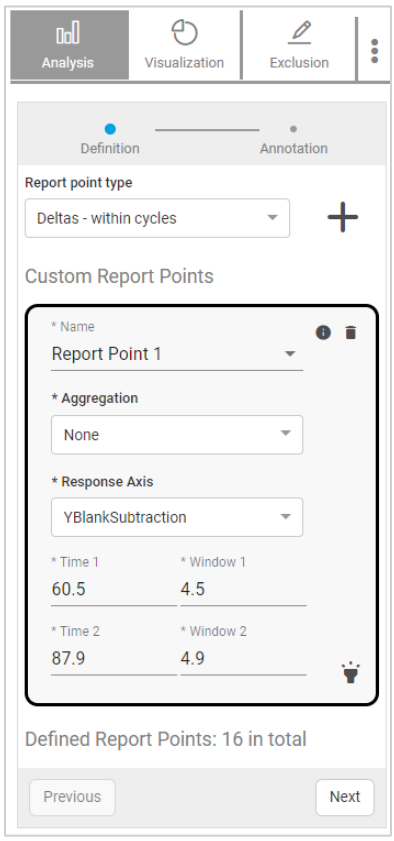
11.6.1.1 Analysis Tab

The **Analysis** tab contains two steps, the *Definition* step, and the *Annotation* step. The *Definition* step allows the user to add custom report points and displays the current total number of the defined report points in the bottom-left hand corner.

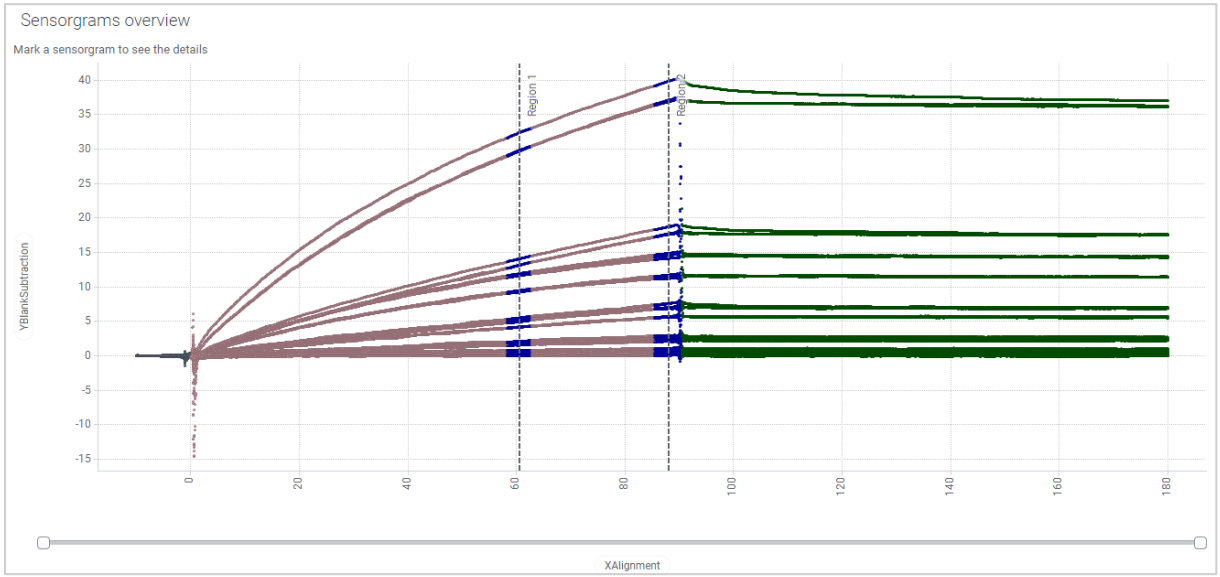


To add a custom report point:

1. Select the desired **Report point type** from the corresponding dropdown (*Value per cycle* or *Deltas within cycles*).
2. Selecting the **+** button to reveal the **Custom Report Points** menu.
3. Enter the require information:
    - Provide a **Name** for the report point or use the default naming convention.
    - Select the desired **Aggregation** from the dropdown menu.
    - Select the desired **Response Axis** from the dropdown menu.
    - Enter a **Time 1** and **Window 1** value, or select and drag to create a window in the *Sensorgrams overview* visualization. Note the region will be highlighted in orange. Select **Set Region** to populate the values in the left-hand panel and the region will turn blue.
    - Enter a **Time 2** and **Window 2** value (for *Deltas within cycles* only) using the same method as described above for the second region.

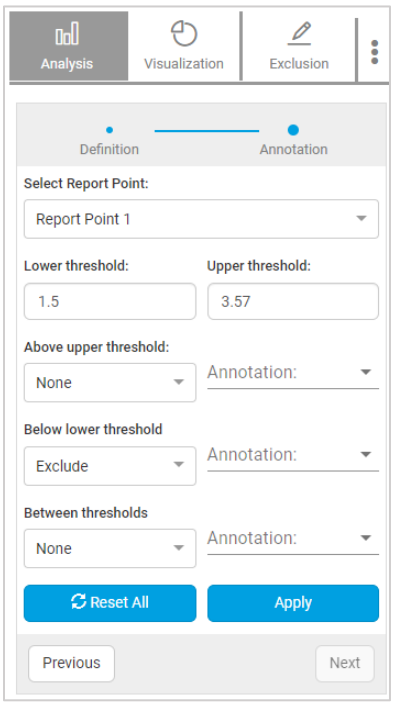


To highlight a specific report point in the *Sensorgrams overview* visualization, select the '*Select as Active*' icon and note the dashed vertical line(s) added to the graph and the associated blue region(s).



For additional information on a report point, hover over the icon. To delete a report point, select the icon.

Continue to add report points as desired using the **+** button, then select **Next** to advance to the *Annotation* step.

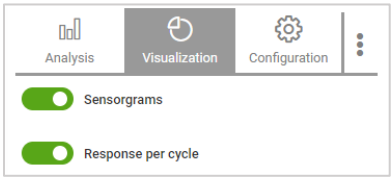


The *Annotation* step contains controls for setting report point specific limits for the inclusion and exclusion of points. The following menu options are available:

- **Report Point:** This dropdown allows the user to select the desired report point as defined in the previous *Definition* step.
- **Lower threshold:** A user defined threshold, set to 0 by default.
- **Upper threshold:** A user defined threshold.
- **Above upper threshold:** Default None, can be set to exclude or include points above the upper threshold.
- **Below lower threshold:** Default None, can be set to exclude or include points below the lower threshold.
- **Between thresholds:** Default None, can be set to exclude or include points between the set thresholds.
- **Annotation:** A dropdown with controlled vocabulary set by the administrator to assign details regarding a threshold.
- **Reset All:** Resets all user exclusions, even those defined in other Apps.
- **Apply:** Applies the defined thresholds to the *Sensorgrams overview* and *Report Point* visualization.

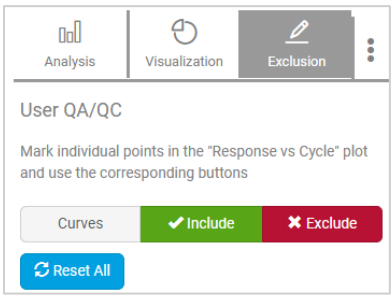
11.6.1.2 Visualization Tab

This tab contains toggles to turn on/off the *Sensorgrams overview* visualization via the **Sensorgrams** toggle and the *Report Point* visualization via the **Response per cycle** toggle.



11.6.1.3 Exclusion Tab

This tab contains data exclusion controls that allow the user to select specific data points from the R*eport Point (Response vs Cycle)* visualization and exclude them from downstream analysis. It also allows the user to either show or hide those excluded data points from the visualizations.



11.6.2 QA-QC App Results

The selected report point in the *Annotation* step dropdown is displayed as a scatterplot in the lower half of the screen, containing the response at the selected report point for each of the cycles and trellised by flow cell. A maximum of four flow cells can be displayed simultaneously. Excluded points are marked in red on both the *Sensorgrams overview* and *Report Point* visualizations.

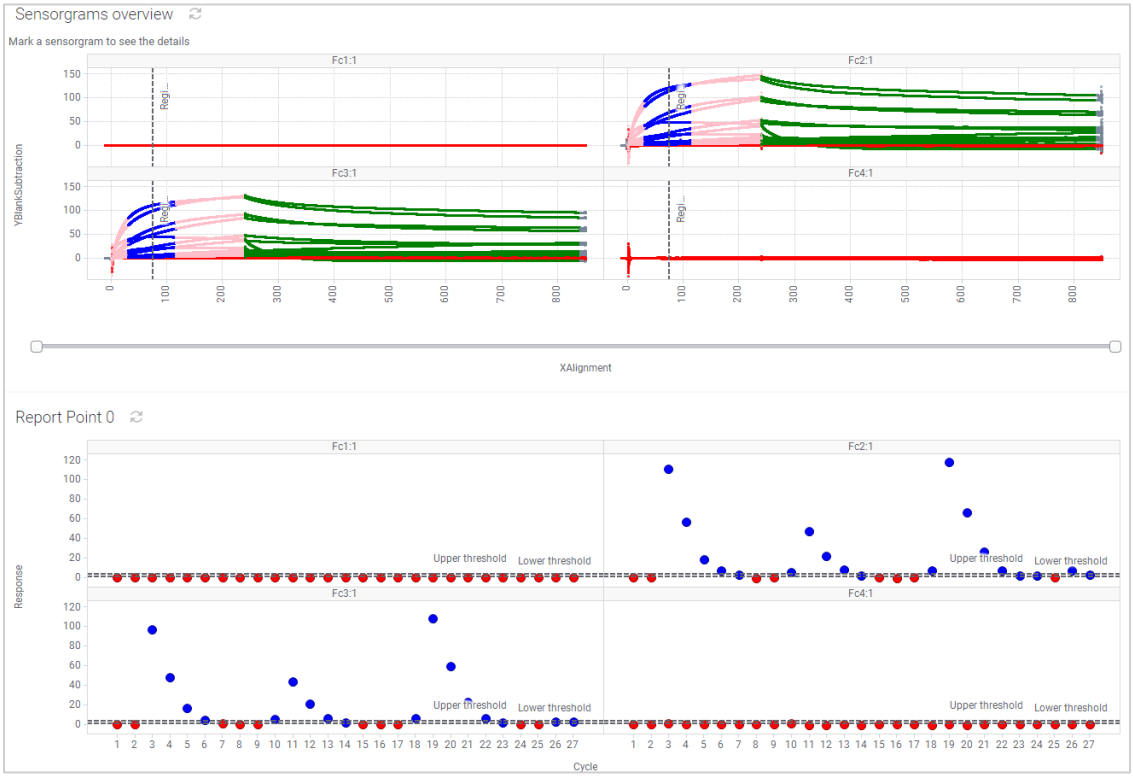


Figure 11-22: Scatterplot of Response at the Selected Report Point for each Cycle, Trellised by Flow Cell

Note: In the current implementation the report point values in the **QA/QC** App are calculated dynamically using the sensorgram data. This allows the user to interactively alter the number of datapoints used for the calculation of the report point value (i.e if the user wants the value to be calculated using a different number of datapoints instead of the default number used by the instrument they can set this number in the textbox and the values will be calculated using the specified number of points around the location of the report point). Because of this, the values may in some cases differ from those that have been read into the Report *Points* table.

11.7 Relative Active Concentrations App

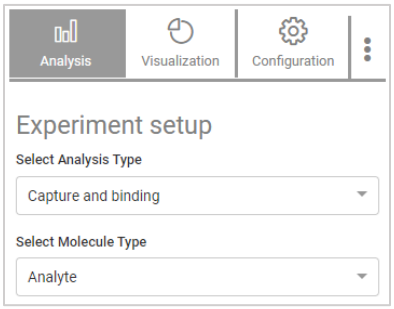
The **Relative Active Concentrations** (RAC) App provides the user with a set of analyses and visualizations to determine relative responses of different curves compared to a reference.

Figure 11-23: Relative Active Concentrations App Card

11.7.1 Configuring the Relative Active Concentrations App

11.7.1.1 Analysis Tab

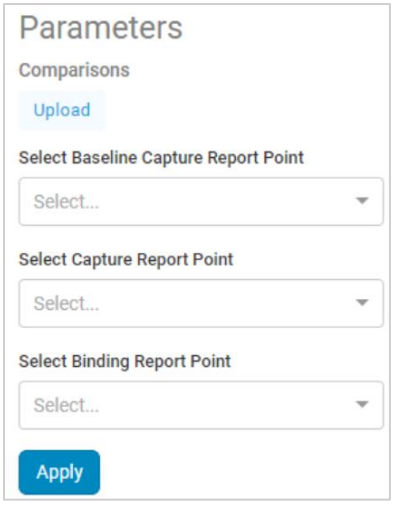
The **Analysis** tab contains controls that allow the selection of the (1) experiment setup, and (2) parameters.



To complete the **Experiment setup** procedure:

- In the **Select Analysis Type** dropdown, select the type of analysis to be used:
  - **Capture and binding:** In this case a capture phase is followed by a binding phase
  - **Direct binding:** Only the binding phase is present
- In the **Select Molecule Type** dropdown, the user should choose the variables to compare:
  - **Analyte:** Will perform the analysis over the analytes
  - **Ligand:** Will perform the analysis over the ligands

Once the type of experiment has been set up, the parameters can be configured in the **Parameters** section.



In this section the following options are available:

- **Comparisons:** The **Upload** button allows the user to upload the file that contains the information on the different data groups to compare. The format required is a columnar tab separated file with the columns:

- **Group** (optional): This is the group containing several samples to be compared to the same reference.
- **Cycle** (optional): This contains the cycle to which the sample corresponds.
- **Analyte** (mandatory): This is the name of the column used as analyte in the analysis, it could be an analyte or a ligand depending on the **Experiment setup** used.
- **Analyte Type** (mandatory): This indicates if the sample is a Reference or a Target. Each block should contain a single reference, if more than one is present only one will be used.

- **Select Baseline Capture Report Point** (Capture and binding only)**:** This dropdown will contain a list of the available report points and allows the selection of the report point to be used as the capture baseline reference.
- **Select Capturing Surface (if any)** (Direct binding only)**:** This dropdown will contain a list of the available flow cells in the analysis which allows the selection of one (or none) as the reference capture surface.
- **Select Capture Report Point:** This dropdown will contain a list of the available report points and allows the selection of the report point to be used as the capture reference.
- **Select Binding Report Point:** This dropdown will contain a list of the available report points and allows the selection of the report point to be used as the binding reference.
- **Apply:** This will perform the analysis.

11.7.1.2 Visualization Tab

This tab is currently not used by the App.

11.7.1.3 Configuration Tab

This tab contains a toggle to turn on/off **Guided Mode**.

11.7.2 Relative Active Concentrations App Results

When the App is initially opened, only the *Sensorgram* visualization is displayed, showing a representation of the sensorgrams available in the document.

When the comparisons file is uploaded, a *Comparisons* table is displayed allowing the user to explore the comparisons that will be performed before applying the analysis.

After selecting **Apply**, an *RAC per Standard* histogram will be displayed containing a visual representation of the relative active concentration of the different samples as well as a measure of the error in case replicates are present. This histogram is trellised by the ligand if it is a sample analysis and the different analytes are grouped together according to the reference against which they were compared.

In addition to the visualizations created, the *RACTable* containing detailed information on the analysis is created in the document. This table contains the following columns:

Column	Description
Experiment	This shows the name of the file
Standard	The name of the reference used
Analyte	The name of the target
Ligand	The name of the ligand
Group	The block to which the comparison belongs
Curve	The flow cell corresponding to the curve
Cycle	The cycle number
BRAnalyte	Binding Ratio for the analyte
BaseCapRPAnalyte	Baseline Capture value for the analyte
CapRPAnalyte	Capture value for the analyte
BindRPAnalyte	Binding value for the analyte
BRStandard	Binding Ratio for the standard
BaseCapRPStandard	Baseline Capture value for the standard
CapRPStandard	Capture value for the standard
BindRPStandard	Binding value for the standard
RAC	The calculated value for the RAC
Significance	A Boolean indicating if the difference is significant
Repeat	This contains the repeat information
Analyte Type	This reflects if the Analyte is a reference or not
RACMethod	The analysis type used

11.8 Multi-Cycle Kinetics App

The **Multi-Cycle Kinetics** App allows the fitting of the sensorgram data to obtain association and dissociation rate constants, k_a and k_d, as well as Rmax, bulk refractive index (RI) or mass transport coefficient (k_t) using the analytical solution derived in Sigmundsson, K. et al. as described in the Appendix 13.1.2.

Figure 11-24: Multi-Cycle Kinetics App Card

Once the App is open, the visualizations before kinetic analysis can be seen.

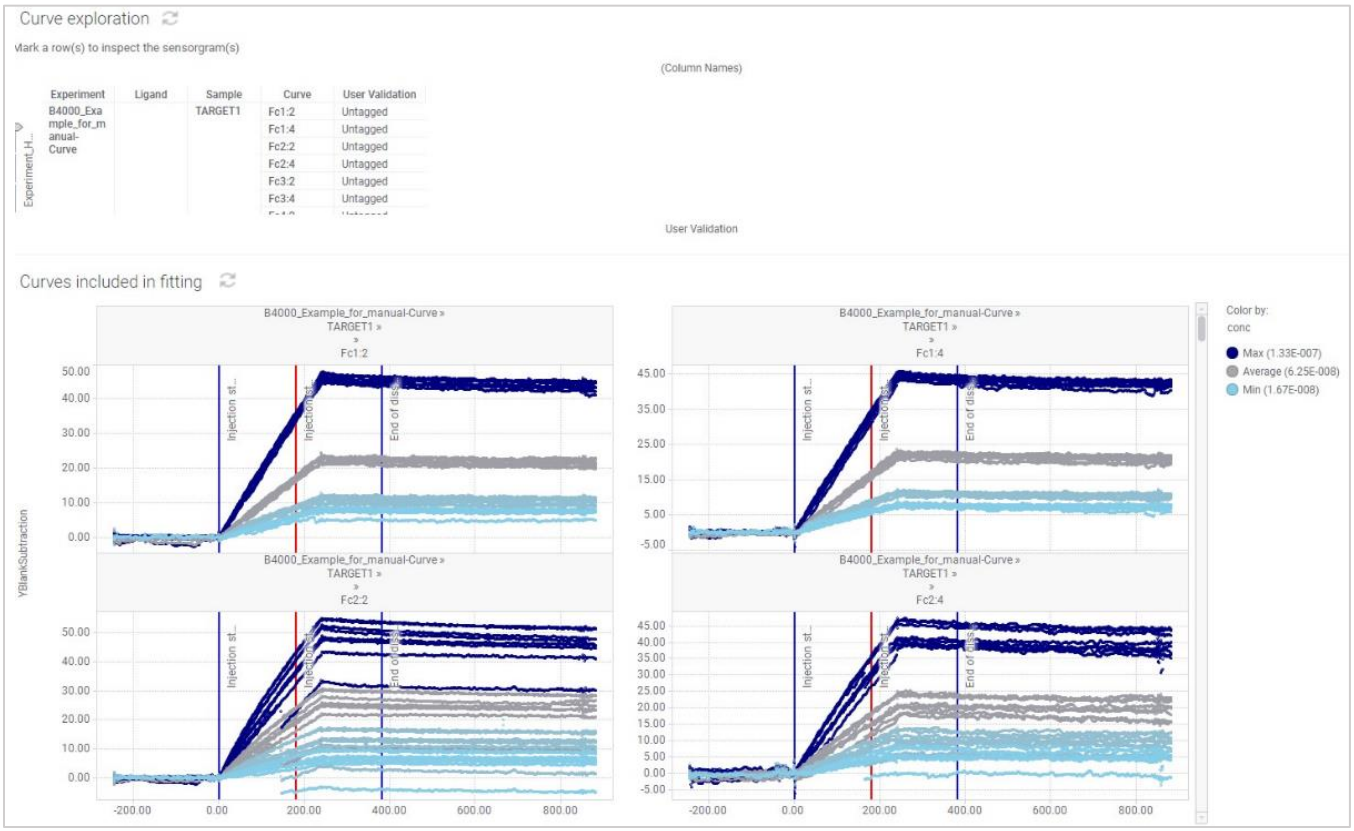
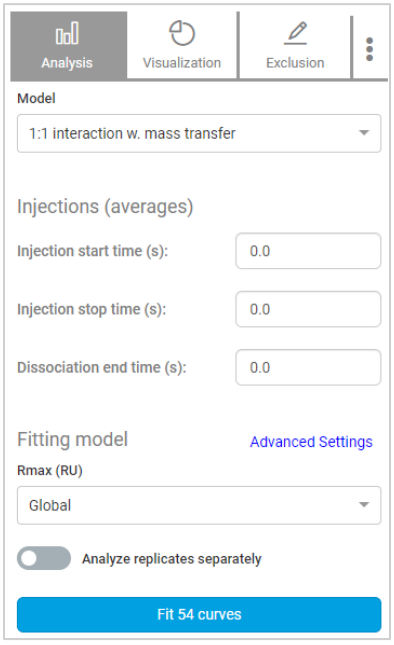


Figure 11-25: Multi-Cycle Kinetics App Visualizations before Kinetic Analysis

11.8.1 Configuring the Multi-Cycle Kinetics App

11.8.1.1 Analysis Tab

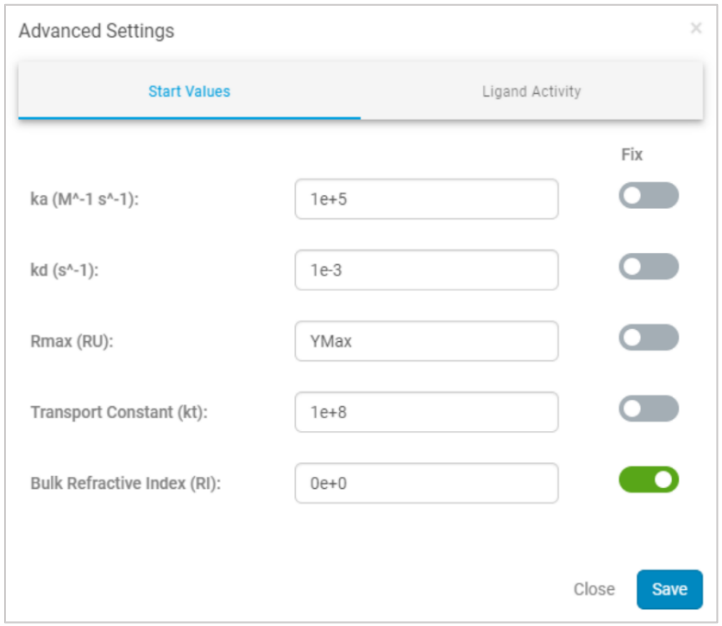
The **Analysis** tab contains controls that allow the configuration of the settings required for kinetic fitting analysis. The curves that will be fitted are those marked in the *Curve exploration* table, or all curves if nothing is marked in the table. Marking the set of curves to be fit can be done by selecting the different curves while holding the *Shift* or *Control* keys. Clicking on the names in the different hierarchy levels of the table will mark all those curves belonging to that level. The slider on the hierarchy axis controls how fine-grained the hierarchy represented in the table is.



The number of curves currently marked can be seen in the **Fit** button at the bottom of the left-hand **Analysis** panel. These are also the curves that will be shown in the *Curves included in fitting* visualization. Once the curves to be analyzed are marked the analysis settings can be configured in the left-hand panel. Here the user can configure the following:

- **Model:** This dropdown allows the selection of the fitting model. The available options are:
    - **1:1 interaction w. mass transfer:** This model will fit the complete sensorgram beginning at the start of injection and finishing at the end of dissociation.
    - **Dissociation only:** This model will fit the dissociation section of the sensorgram beginning at the end of injection and finishing at the end of dissociation.
- **Injection start time (s)**: Set the time of the injection start. This time will default to the one extracted from the data if present and is represented in the sensorgram visualization as a blue line. If more than one value is present in the file, the single most common value will be displayed. This option is not displayed when using the ‘Dissociation only’ model.
- **Injection stop time(s)**: Set the time of the injection stop. This time will default to the one extracted from the data if present and is represented in the sensorgram visualization as a red line. If more than one value is present in the file, the single most common value will be displayed.

- **Dissociation end time(s)**: Set the end time of dissociation to include in the kinetic analysis. This time will default to the one extracted from the data if available and is represented in the sensorgram visualization as a blue line. If more than one value is present in the file, the single most common value will be displayed.
- **Advanced Settings:** This dialog can be used to modify the starting parameters used for the kinetic fitting, as well as to configure the required information for the ligand activity analysis, if needed. The initial values used in the fitting can be set from the **Start Values** tab:

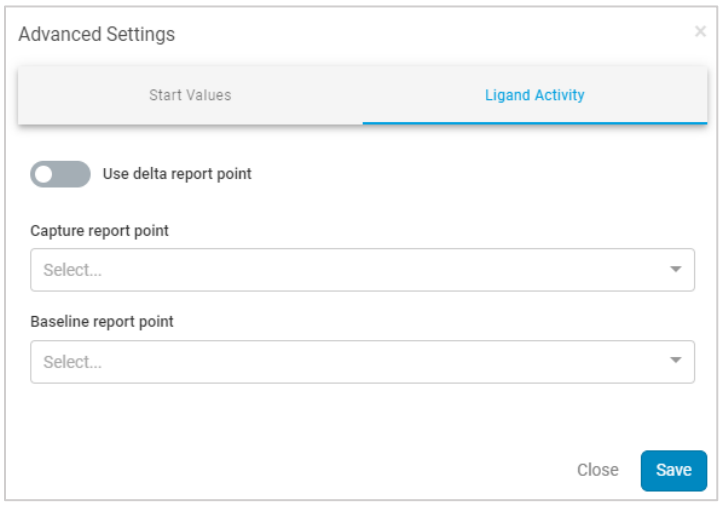


  - The **ka** textbox contains the initial estimation for the ka to be used in the fitting procedure. If desired, another starting point can be entered.
  - The **kd** textbox is the initial estimation for the kd to be used in the fitting. If desired, another starting point can be entered.
  - The **Rmax (RU)** textbox contains the initial estimation for the Rmax to be used in the fitting procedure. This box can hold a number in RU or the word “YMax”, which will allow the Rmax value to be found during kinetic fitting.
  - The **Transport Constant (kt)** textbox contains the initial starting point for kt to be used in the fitting procedure. If another starting point is desired, it can be entered.

    Note: kt is the mass transport coefficient in units RU/(M*s).
  - The **Bulk Refractive Index (RI)** textbox contains the initial estimation for the RI to be used in the fitting procedure.

    Note: Each of the values above can be fixed to the defined value by enabling the corresponding toggle.

The **Ligand Activity** tab allows the user to select the capture and baseline report points for those analyses that require the calculation of the ligand activity. Bear in mind to do this, the Molecular weight should be loaded into the **SPR Data Import** App.



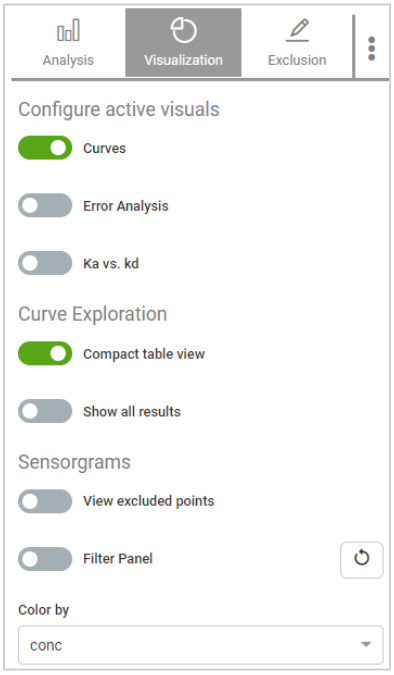
- **Use delta report point**: This toggle allows the user to switch between the use of a single delta report point and two separate report points defined by the **Capture** and **Baseline** report points.
- **Capture Report Point**: If a report point is selected, this will be the reference capture point to be used in the calculation of the ligand activity. Visible when **Use delta report point** toggle is off.
- **Baseline Report Point**: If a report point is selected, this will be the baseline report point to be used in the calculation of the ligand activity. Visible when **Use delta report point** toggle is off.
- **Delta report point**: The report point selected will be used for the calculation of the ligand activity. Only available when the **Use delta report point** toggle is on.

When using the 'Dissociation only' model the only available option is the kd.

- **Fitting model**: The user can configure the fitting by using the controls in this section. These controls are only available for the 1:1 interaction w. mass transfer model. The procedure to configure them is the following:
  - The **Rmax (RU)** dropdown menu can be set to:
    - **Global** (Default)**:** The Rmax is calculated for each group of sensorgrams within an analyte series, which is the set of sensorgrams that share the same Sample, Ligand and Flow cell/Spot.
    - **Local:** The Rmax is calculated for each individual curve.
- **Analyze replicates separately:** This toggle will set the analysis of the replicates separately. By default, the replicates will be analyzed together.
- **Fit:** Select the **Fit curves** button to start the analysis. The number of curves that will be fit is displayed on the button to help ensure the desired analysis is included.

11.8.1.2 Visualization Tab

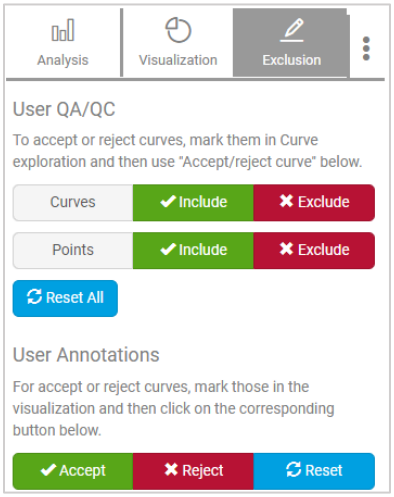
The **Visualization** tab in the **Multi-Cycle Kinetics** App has additional functionality to customize the displayed analysis elements:



- **Curves:** This toggle button will show or hide the *Curve exploration* table and the *Sensorgrams* visualization.
- **Error Analysis:** Only visible after executing the analysis. Shows or hides the residuals visualizations. These are the “Error over time” and “Error distribution” plots. This visualization is not available for the ‘Dissociation Only’ model.
- **Ka vs. Kd:** Only visible after executing the analysis. Shows or hides the Ka vs. Kd isoaffinity plot. This visualization is not available for the ‘Dissociation Only’ model.
- **Compact table view**: This toggle will switch between a compact version of the *Curve exploration* table containing only the basic columns (“On”) and an extended version of the *Curve exploration* table that contains additional SE information without being the full table (“Off”).
- **Show all results**: This toggle will show in the results table those additional columns that calculated by the analysis but are not displayed by default, such as the confidence intervals and the starting parameters used.
- **View excluded points**: This toggle will allow those excluded points to be visualized as red X symbols on the plot.
- **Filter Panel**: This toggle will show or hide the filter panel, the reset button located to the right-hand side of it will reset the filtering scheme.
- **Color by:** This dropdown allows the user to change how the *Before* and *After* visualizations are colored (Ligand by default).

11.8.1.3 Exclusion Tab

The **Exclusion** tab allows the user to override the automatic classification of the fitted curves into ‘Accepted’ or ‘Rejected’ curves. It provides the following options:



User QA/QC section:

Here the user may exclude or include curves or curve points as described in previous sections.

User Annotation section:

Here the user can override the automatics acceptance or rejection of the different curves.

- **Accept:** Set the selected curve(s) as accepted regardless of the automatic classification.
- **Reject:** Set the selected curve(s) as rejected regardless of the automatic classification.
- **Reset:** Reset the classification to the default provided by the automatic classification.

11.8.2 Supporting Visualizations

11.8.2.1 Residual Visualizations

There are two visualizations in this group which are hidden by default and can be viewed by setting the '**Error analysis'** toggle to "On". "Residuals" are the distance between the data and the fitted line.

- **Error over time**: a plot of the residuals over time.
- **Error distribution**: A histogram representation of the residuals.

Figure 11-26: Error Analysis Visualizations

11.8.2.2 ka vs kd Visualization

This visualization shows a scatterplot of ka versus kd for the samples. It is hidden by default, and can be viewed by toggling the **Ka vs. kd** toggle in the **Visualization** tab. The dotted lines in the plot represent the location of equilibrium dissociation constants (KD) calculated from the ratio of kd/ka. This visualization is also available with more interactive nature in the **Hit Selection** App Results sections.

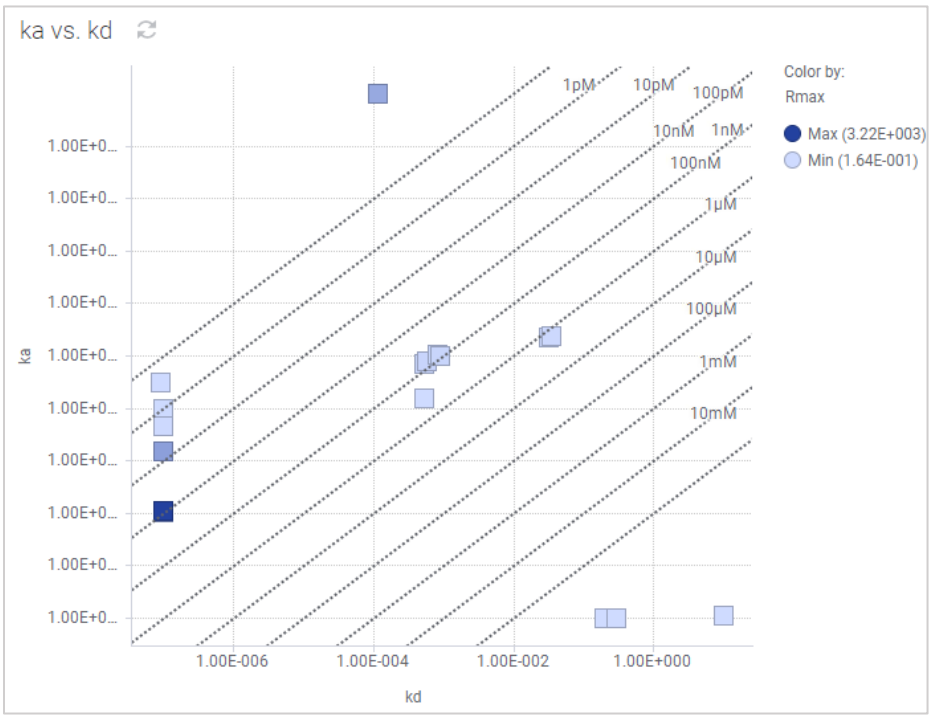


Figure 11-27: ka vs. kd Visualization

11.8.3 Multi-Cycle Kinetic Analysis App Results

Once the sensorgrams are fitted, the fit is overlaid on the sensorgram data and the controls in the user interface will be available to further explore and refine the analysis.

To change the sensorgram display there are multiple options:

- Multiple sensorgrams can be selected in the *Curve exploration* table by highlighting and marking the desired rows using the Control and/or Shift keys while marking with the mouse. All rows can be marked by clicking on the top row and then navigating to the last row while holding the Shift key. Alternatively, to mark all the sensorgrams after a subset of curves has been marked, right-click on the *Curve exploration* table, choose “Marked Rows” and “Unmark”. This unmarks the highlighted sensorgrams. Then right-click again on the *Curve exploration* table, choose “Marked Rows” and “Invert”.
- The hierarchy levels can be used for more fine-grained marking of the curves. The curves marked in the *Curve exploration* table are the ones that will be displayed in all other visualizations.
- By default, the number of sensorgrams in the visualization is set to a 2x2 trellis but can be adjusted by using the Spotfire® Trellis function. This is done by right-clicking on the visualization area of the sensorgrams and selecting **Properties**. Choose **Trellis** and, in the resulting window, select **Panels and Manual Layout** as shown below. Select the number of rows and columns that will be visible within the sensorgram visualization. The sensorgrams can be scrolled through using page down or the scroll bar.

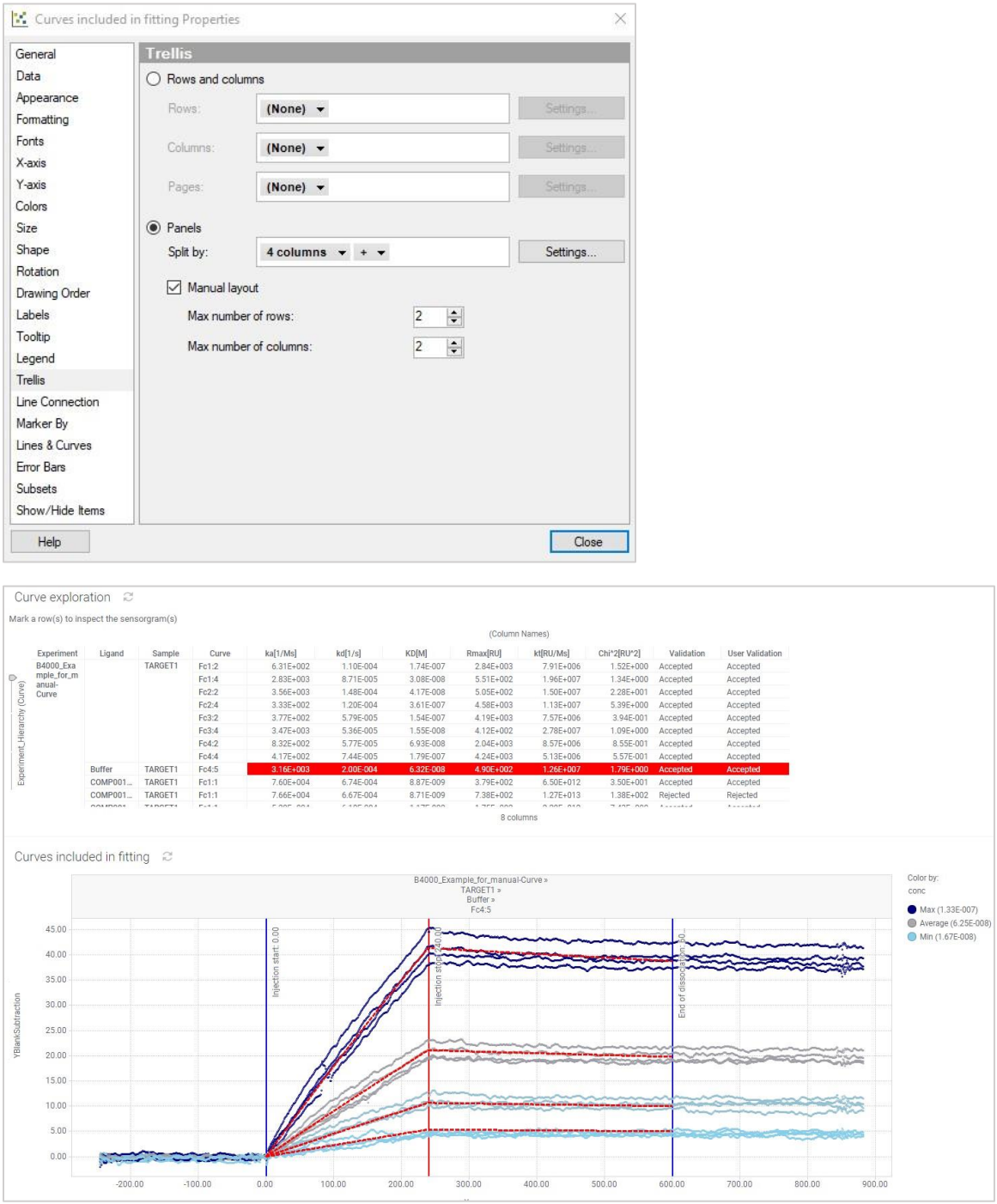


Figure 11-28: Multi-Cycle Kinetics App Visualizations

- **Curve exploration table**: This visualization will have an additional set of columns populated with the values obtained from fitting the sensorgrams or set of sensorgrams within an analyte series, which is the set of sensorgrams that share the same Sample, Ligand and Flow cell/Spot.

Column	Description
Conc [nM]	The concentration of the sample used
Ka [1/Ms]	The estimated association rate constant, ka, obtained by fitting the sensorgram(s)
CI-Lower (ka)	The lower bound of the Confidence interval for the ka
CI-Upper (ka)	The upper bound of the Confidence interval for the ka
SE (ka)	The estimated SE of the ka
kd [1/s]	The estimated dissociation rate constant, kd, obtained by fitting the sensorgram(s)
CI-Lower (kd)	The lower bound of the Confidence interval for the kd
CI-Upper (kd)	The upper bound of the Confidence interval for the kd
SE(kd)	The estimated SE of the kd
KD [M]	The estimated KD for the sensorgram(s) from the ratio of the kinetic rate constants kd/ka
SE (KD)	The estimated SE of the KD
Rmax [RU]	The estimated Rmax obtained by fitting the sensorgram(s)
CI-Lower (Rmax)	The lower bound of the Confidence interval for the Rmax
CI-Upper (Rmax)	The upper bound of the Confidence interval for the Rmax
SE (Rmax)	The estimated SE of the Rmax
kt [RU/Ms]	The estimated mass transport coefficient kt, in units RU/(M*s), obtained by fitting the data
CI-Lower (kt)	The lower bound of the Confidence interval for the kt
CI-Upper (kt)	The upper bound of the Confidence interval for the kt
SE (kt)	The estimated SE of the kt
RI [RU]	The estimated Bulk Refractive Index, RI, obtained by fitting the sensorgrams
Chi^2 [RU^2]	The χ^2 calculated for the provided fit

MAE	Mean Absolute Error of the fit
Validation	This is the automatic Acceptance or Rejection of the curve as estimated by the app. Individual curves are rejected when CHIsquare value is >10% R_{max}
User_validation	This will have the same value as the previous column unless the user overrides the automatic acceptance criteria
Ligand_Activity	The Ligand activity. This is calculated as the ratio between the Rmax measured for the curve and the Theoretical Rmax calculated as (RU Ligand * MW Analyte * No Binding sites Ligand) / MW Ligand. This would reflect the actual amount of analyte that has bound to the ligand vs. the amount of Ligand that should bind if all the Ligand bound in the capture phase is able to bind the maximum amount of ligand.
CI-Lower (Ligand_Activity)	The lower bound of the Confidence interval for the Ligand activity
CI-Upper (Ligand_Activity)	The upper bound of the Confidence interval for the Ligand activity

In addition to these columns, there are some additional columns that provide the information on the starting parameters used for the fitting to ensure the analysis is easily reproducible. By default, not all these columns are displayed in the *Curve exploration* table. To show them all, set the '**Show all results'** toggle to on in the **Visualization** tab. If '**Show all results'** toggle is off a subset of these columns is shown with only the main parameters and the validation results displayed when the '**Compact table view'** toggle in the **Visualization** tab is set to on.

Note: When using the 'Dissociation only' model some of these columns will not be provided.

Once the kinetics analysis is performed the curves are automatically classified by the software as accepted or rejected (**Validation** column on the table). The rejection criteria used in this automatic validation is: Chi^2 > Rmax * 0.1.

11.9 Steady State Analysis App

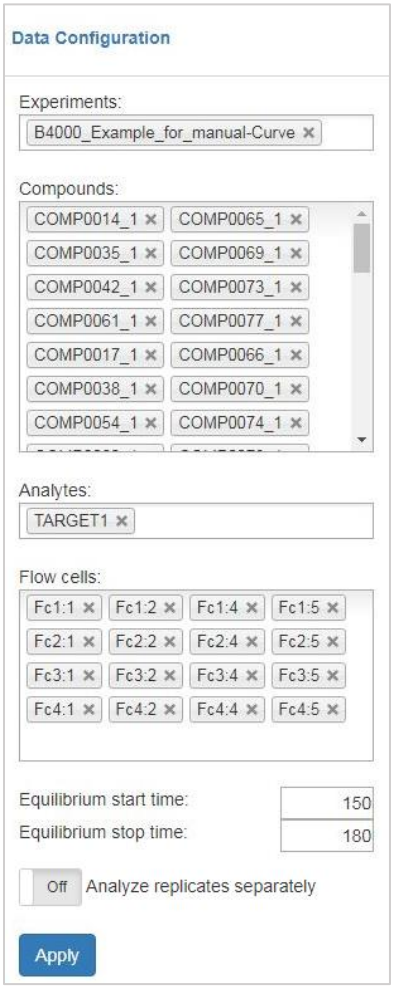
The **Steady State Analysis** App calculates the equilibrium dissociation constant, KD, based on concentration vs. response units (RU) when the interaction has reached a steady state (plateau).

Figure 11-29: Steady State Analysis App Card

Once the App is open, the sensorgrams before the steady state analysis are visible.

11.9.1 Configuring the Steady State Analysis App

The **Data Configuration** panel contains controls that allow the selection of the data points to be included in the fitting. These controls are the following:



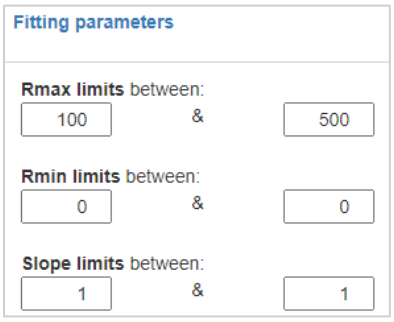
To select the data to be fitted the procedure is the following:

- In the **Experiments** box, select the experiment or experiments to be used.
- In the **Compounds** box, select the ligands to be included in the fitting.
- In the **Analytes** box, select the samples to be included in the fitting.
- In the **Flow cells** box, select the flow cells to be included in the fitting.
- Equilibrium region:
    - In the **Equilibrium start time** box, set the lower bound of the equilibrium region on the X-axis, and it will appear represented in the *Sensorgram* visualization as a blue line.

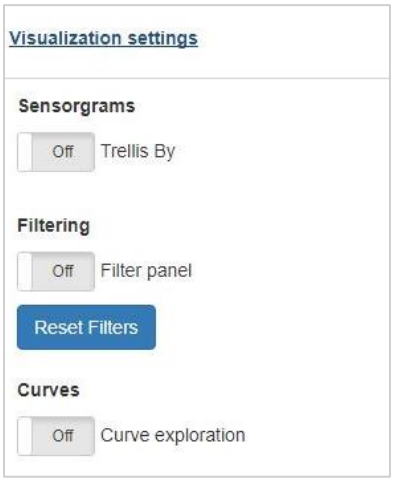
- In the **Equilibrium stop time** box set the upper bound of the equilibrium region on the X-axis, and it will appear represented in the *Sensorgram* visualization as a red line.
- Alternatively, the equilibrium region can also be set by selecting a range on the X-axis by clicking on the X-axis and dragging the mouse to select the required region, or simply selecting an area of the *Sensorgram* plot. When set in this manner, the **Start time** value will be set to the lowest marked value and the **End time** value to the highest marked value on the X-axis.

- **Analyze replicates separately:** Set this toggle to 'On' if you wish to analyze the replicates separately.
- Select **Apply** to execute the analysis. If you would like to set boundaries to the parameters do this before applying as described below.

The **Fitting parameters** section allows the user to set limits for the **Rmax**, **Rmin** and **Slope** of the calculated logistic regression curve fit. For each of the Rmax, Rmin and Slope the user can set lower and upper bounds using the available boxes.



The **Visualization settings** section in the **Steady State Analysis** App in addition to the functionality provided in the Preprocessing Apps section, also contains the trellis by toggle that allows the user to trellis the sensorgrams visualization into a 2x2 grid in order to facilitate the visual exploration.



11.9.2 Steady State Analysis App Results

When the steady state analysis is complete, the *Dose Response* plot and the *Curve exploration* table will be available for further exploration.

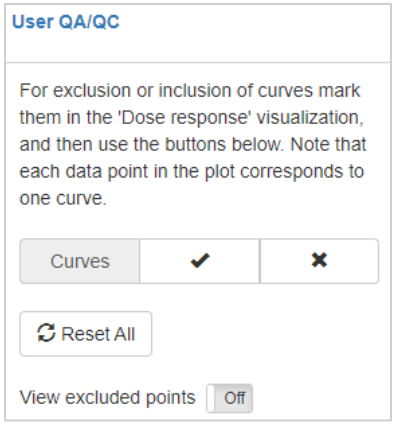
- **Curve exploration table**: This will have an additional set of columns populated with the steady state values calculated for the curves.

Column	Description
KD	The equilibrium dissociation constant, KD (M), obtained by Steady State analysis fit of the data. This is calculated from the following equation: Rmin + ((Rmax - Rmin) / (1 + (((10 ^ -PKD) / (10 ^ -x)) ^Slope)))
pKD	The negative $\log_{10}$ of the equilibrium dissociation constant KD obtained by steady state analysis fit of the data
PKD CI Lower	The lower limit for the confidence interval for the negative $\log_{10}$ KD obtained by steady state analysis
PKD CI Upper	The upper limit for the confidence interval for the negative $\log_{10}$ KD obtained by steady state analysis
ASYMMAX	The maximum asymptote
ASYMMAX CI Lower	The lower limit for the confidence interval for the estimated ASYMMAX
ASYMMAX CI Upper	The upper limit for the confidence interval for the estimated ASYMMAX

- **Dose Response:** This plots the Log10 concentration versus the Req (average of the RU region selected). The dose response fits are overlaid on the data.

After executing the steady state analysis, the user can work with the **User QA/QC** controls to exclude specific concentrations from the analysis, by selecting the points to be excluded from the *Dose response* visualization.

The controls available in this menu are limited to the controls for excluding full curves, as in steady state, the visualizations represent data from complete curves.



11.10 Non-Regenerative Kinetics App

The **Non-Regenerative Kinetics** App provides a solution for analysis of non-regenerative kinetics (single-cycle) in Spotfire®. A set of interactive visualizations are provided to easily explore the results.

Figure 11-30: Non-Regenerative Kinetics App Card

11.10.1 Configuring the Non-Regenerative Kinetics App

In the case of Non-Regenerative Kinetics data, in the .blr file there are usually several concentration columns that will correspond to different concentration injections during the cycle. These are selected in the **SPR Data Import** App from the ⚙ icon to correctly assign the respective concentrations to the different parts of the cycle.

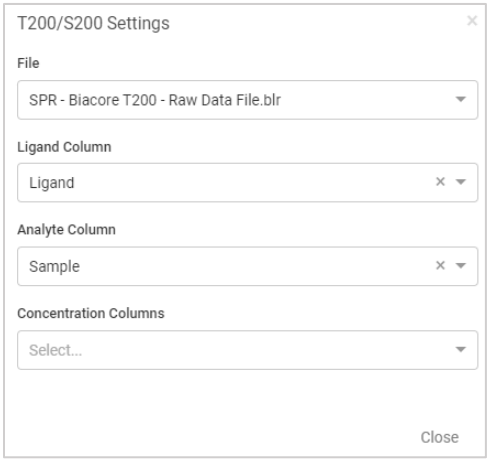
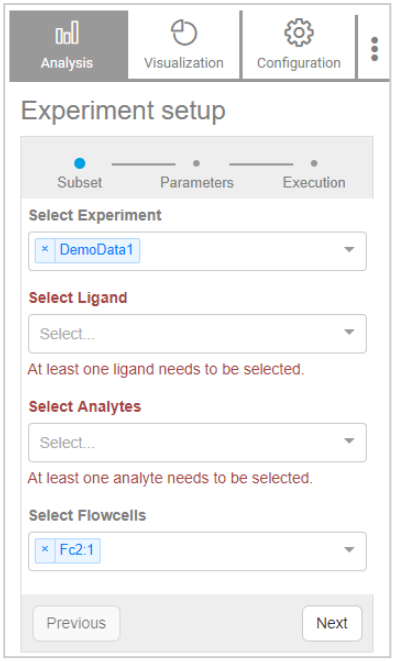


Figure 11-31: Settings UI with Concentration Columns Selector

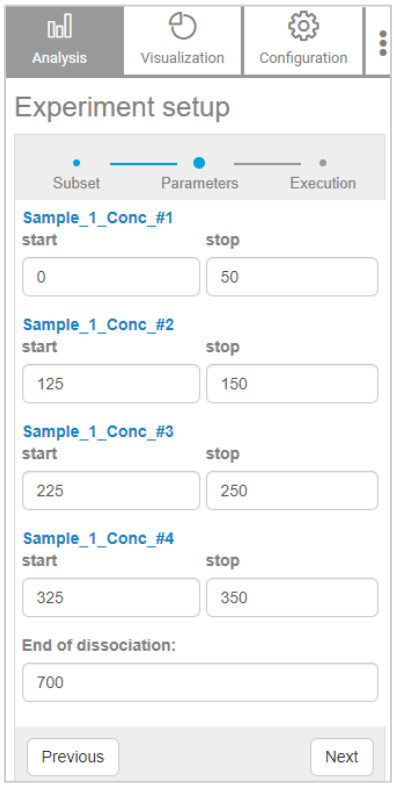
11.10.1.1 Analysis Tab

After performing any required preprocessing and loading the **Non-Regenerative Kinetics** App, the user will be presented with a configuration wizard. This experiment setup wizard has three steps in the **Analysis** tab:

- In the *Subset* step, the selection of the experiment, ligand(s), analyte(s) and flow cells on which to perform the analysis are completed.



- In the *Parameters* step, the user inputs the start and stop times of injection for each of the analyte concentrations as well as the end of the dissociation position.



- In the *Execution* step, the user can execute the analysis.

11.10.1.2 Visualization Tab

This tab is currently not used by the App.

11.10.1.3 Configuration Tab

This tab contains a toggle to turn on/off **Guided Mode**.

11.10.2 Non-Regenerative Kinetics App Results

Once the analysis is performed the visualizations shown by the App are the following:

Curve exploration table: This contains the values calculated from each of the fitted curves as well as information on the settings used in order to perform the fitting. The columns available are:

Column	Description
Experiment	The name of the experiment to which the row corresponds
Cycle	The number of the cycle
Curve	This shows the Curve ID
Ligand	The Ligand name
Sample	The sample name
Replicate	The replicate number
Conc	This contains the maximum concentration that was used in the curve
Kd	The estimated dissociation rate constant, kd, obtained by fitting the sensorgram(s) for the analyte series
kd_SE	Standard error of the kd
kd_CI_Lower	Lower limit of the kd confidence interval
kd_CI_Upper	Upper limit of the kd confidence interval
Ka	The estimated association rate constant, ka, obtained by fitting the sensorgram(s) for the analyte series
ka_SE	Standard error of the ka

ka_CI_Lower	Lower limit of the ka confidence interval
ka_CI_Upper	Upper limit of the ka confidence interval
KD_kinetics	The estimated KD obtained from the ratio of the kinetic rate constants kd/ka
Rmax	The estimated Rmax obtained by fitting the sensorgram(s)
Kt	The estimated mass transport coefficient kt, in units RU/(M*s), obtained by fitting the data
CHIsquare	The CHIsquare calculated for the provided fit

In addition, the table will contain information on the starting parameters that were used for the fitting.

Sensorgram: This panel shows a trellis containing the curves that have been analyzed, the visualization can be filtered to show only certain curves by selecting them in the *Curve exploration* table.

11.11 Hit Selection App

The **Hit Selection** App provides a solution for the selection and classification of SPR analysis results data in Spotfire® and the manual exploration and analysis of the results using a set of interactive visualizations.

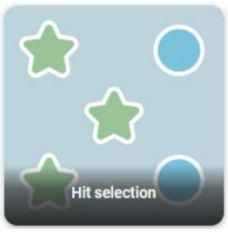
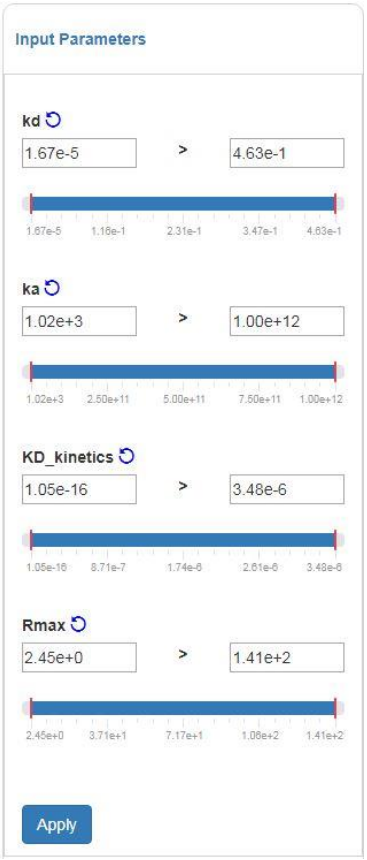


Figure 11-32: Hit Selection App Card

11.11.1 Configuring the Hit Selection App

The **Input Parameters** settings provide the ability to control what is classified as a hit.



Controls available in this menu are:

- **kd**: This control will allow the user to classify hits that are within the kd ranges set, using the slider or text boxes.
- **ka**: This control will allow the user to classify hits that are within the ka ranges set, using the slider or text boxes.
- **KD_kinetics**: This control will allow the user to classify hits that are within the KD_Kinetics ranges set, using the slider or text boxes.
- **Rmax**: This control will allow the user to classify hits that are within the Rmax ranges set, using the slider or text boxes.
- **Apply**: When clicked on, this button will update the visualizations with the automatic classification according to the current settings of the sliders within the **Input Parameters** section.

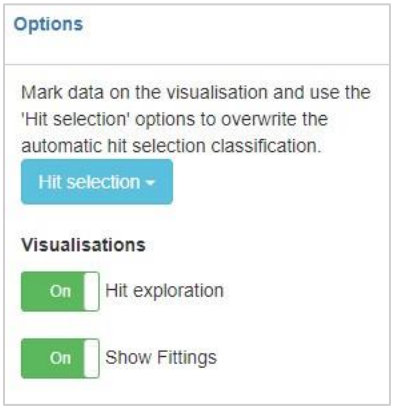
The **Options** section allows the user to override the automatic hit classification based on the **Input Parameter** settings. To override a subset of curves, select the curve(s) on the *Hit exploration* table or the *kon - koff* plot and use the options in the **Hit selection** dropdown:

- **As hit:** The classification is set to hit.
- **As a non hit:** The classification is set to non-hit.
- **Reset tags:** The classification is set to the value initially provided by the algorithm.

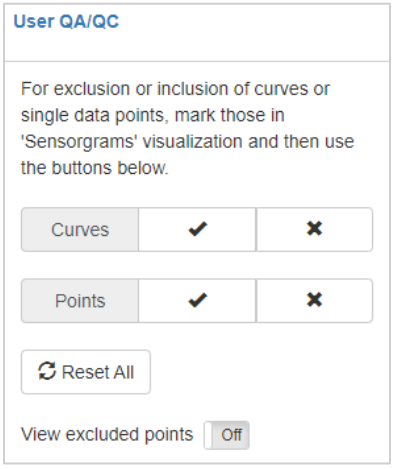
Under '**Visualizations'**, the user can change what is displayed in the visualization area of the App:

- **Hit exploration:** This toggle shows or hides the hit exploration table.

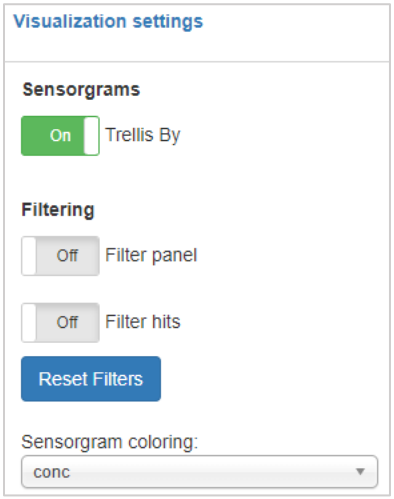
- **Show Fittings:** This toggle shows or hides the curve that is overlaid on the *Sensorgrams* visualization.



After executing the **Hit selection** App, the user can use the **User QA/QC** controls to exclude specific data points or curves from the analysis, by selecting the points to exclude from the *Sensorgrams* visualization.



The **Visualization settings** section in addition to the functionality described previously in the Visualizations within Preprocessing Apps section, also contains the **'Trellis By'** toggle that allows the user to trellis the *Sensorgrams* visualization into a 2x2 grid in order to facilitate visual exploration.



11.11.2 Hit Selection App Results

The **Hit Selection** App contains a *Hit exploration* table, *kon - koff* isoaffinity plot and *Sensorgrams* visualizations for all the data for which the kinetic analysis provided an ‘Accepted’ fitting. These visualizations are all interactive. Sensorgrams to be displayed can be selected by highlighting curves in the *Hit exploration* table or selecting points within the *kon - koff* isoaffinity plot.

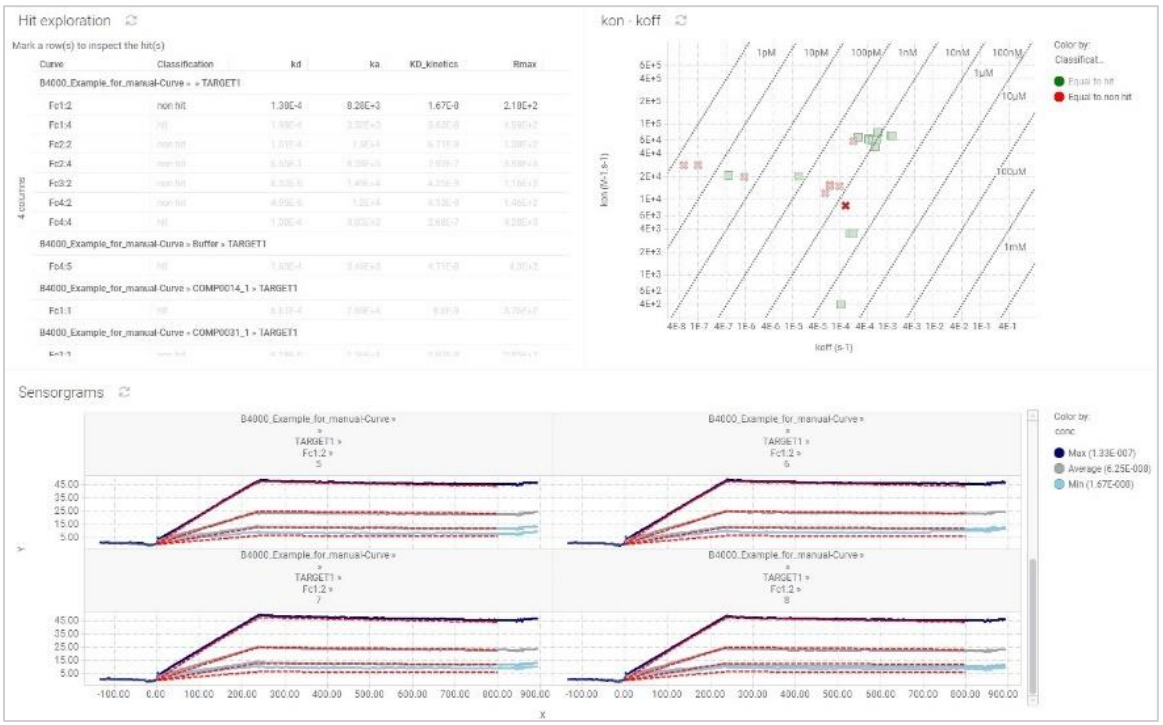


Figure 11-33: Hit Selection App Visualizations

The visualizations generated by the **Hit Selection** App are the following:

- **Hit exploration table**: This contains the values for the analyte series fitted in the **Multi-Cycle Kinetics** App that provided an ‘Accepted’ fitting (an analyte series is the set of sensorgrams that share the same Sample, Ligand and Flow cell/spot). The rows in the table can be used to control the sensorgrams that are displayed in the *Sensorgrams* visualization.

Column	Description
Curve	This shows the Curve ID
Classification	This shows the classification of the set of curves as “hit” or “not hit”.
Kd	The estimated dissociation rate constant, kd, obtained by fitting the sensorgram(s) for the analyte series in the Multi-Cycle Kinetics App.
Ka	The estimated association rate constant, ka, obtained by fitting the sensorgram(s) for the analyte series in the Multi-Cycle Kinetics App.

KD_kinetics	The estimated KD obtained from the ratio of the kinetic rate constants kd/ka.
Rmax	The estimated Rmax obtained by fitting the sensorgram(s).
CHIsquare	The CHIsquare calculated for the provided fit.

- **kon - koff**: This visualization shows a representation of the kon (Y-axis) versus the koff (X-axis) for all the sets of curves. It can be used to identify hits based on the desired kinetics parameters.
- **Sensorgrams**: This panel shows those curves selected in the *Hit exploration* and/or *kon – koff* visualizations. On the sensorgrams the fitting generated by the kinetics fit is overlaid.

Once the **Hit Selection** App has imported the data, the results can be explored using the *Hit exploration* and *kon - koff* visualizations as mentioned above.

11.12 Relative Potency App

The **Relative Potency** App allows the user to analyze 4-parameter logistic regression curves created from concentration-response data to identify differences in activity between the analyzed compounds.

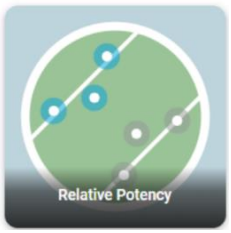


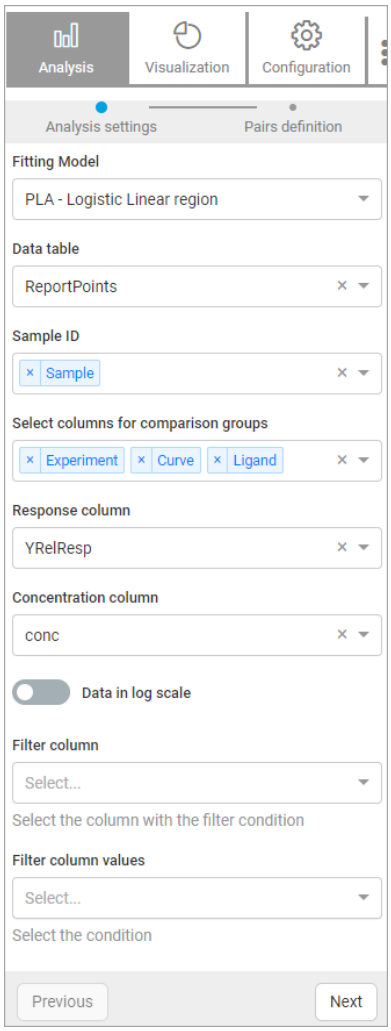
Figure 11-34: Relative Potency App Card

11.12.1 Configuring the Relative Potency App

11.12.1.1 Analysis Tab

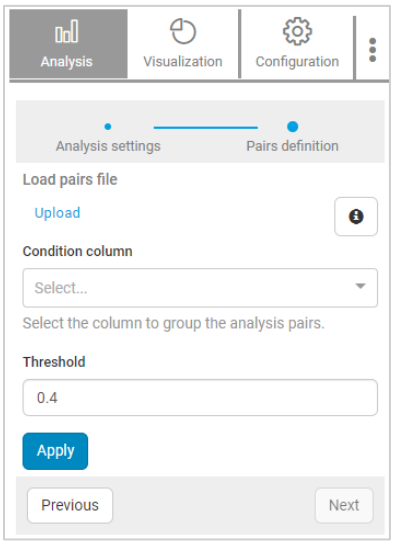
The **Analysis** tab guides the user in configuring the data settings with the following two steps: *Analysis settings* and *Pairs definition*.

1. In the *Analysis settings* step the user may add the underlying data from which the 4 parameter logistic regression used in the first step was calculated. This step is optional and can be left empty if there is no data available. It contains the following options:



- **Fitting Model:** This dropdown allows the user to select the fitting model.
- **Data table:** This is the table containing the underlying data for the fitting. It may be the same as the one containing the curve fits.
- **Sample ID:** The selected column identifies the sample ID column.
- **Select columns to create data groups:** The selected columns uniquely identify the data rows.
- **Response column:** This column contains the response value for the point.
- **Concentration column:** This column contains the concentration value for the column.
- **Data in log scale:** This toggle allows the user to indicate if the data in the table is in log scale.
- **Filter column:** This column indicates the column to be used for additional filtering in order to consider only a subset of the data for the analysis. This step is optional and can be left empty if there is no filtering required but may be useful if the dataset contains multiple entries of the same compound in different plates, as it will allow the user to perform an analysis per plate.
- **Filter column value:** This is the value of the column that will be used to select the subset of data to perform the analysis. This step is optional and can be left empty.

2. In the *Pairs definition* step, the user may add a file containing the list of pairs that should be compared, specifying if they should be grouped together using any condition column and the similarity threshold for the slope. It contains the following options:

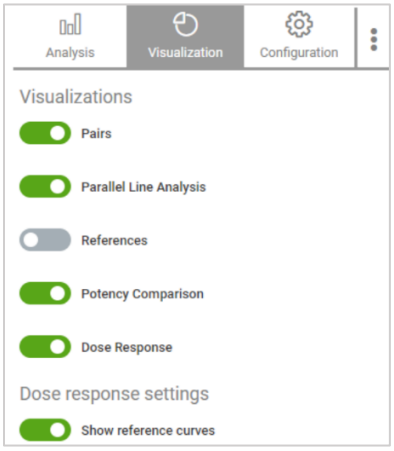


- **Upload:** This allows the user to upload a pairs file. This file contains a set of sample reference pairs to be compared in the analysis. It may also contain additional columns that will be used to group the comparisons. This allows the user to easily define triplets or groups of comparisons that can be analyzed together. The format of the pairs file is a tab separated text file with the following columns:
    - **Sample ID:** This is the data that needs to be evaluated.
    - **Control:** This is the reference data to which the same should be compared.
    - **Condition:** The column selected as condition will be used to group the analysis pairs and calculate the confidence information.

      Note: Alternatively, the user may complete the empty table in the right-hand panel for pairs configuration.
- **Condition column:** The column selected in this dropdown will be used to group the samples in the *Parallel Line Analysis* and *Potency Comparison* visualizations.
- **Threshold:** This value is the threshold used to determine if the slope of the two lines is similar enough to perform a meaningful parallel line analysis.
- **Apply:** Selecting this button will perform the relative potency analysis.

11.12.1.2 Visualization Tab

The **Visualization** tab provides a toggle to display several visualization options. Note that these do not appear until after the *Analysis settings* and *Pairs definition* steps are completed.



11.12.2 Relative Potency App Results

When the relative potency analysis has been performed the following visualizations are available:

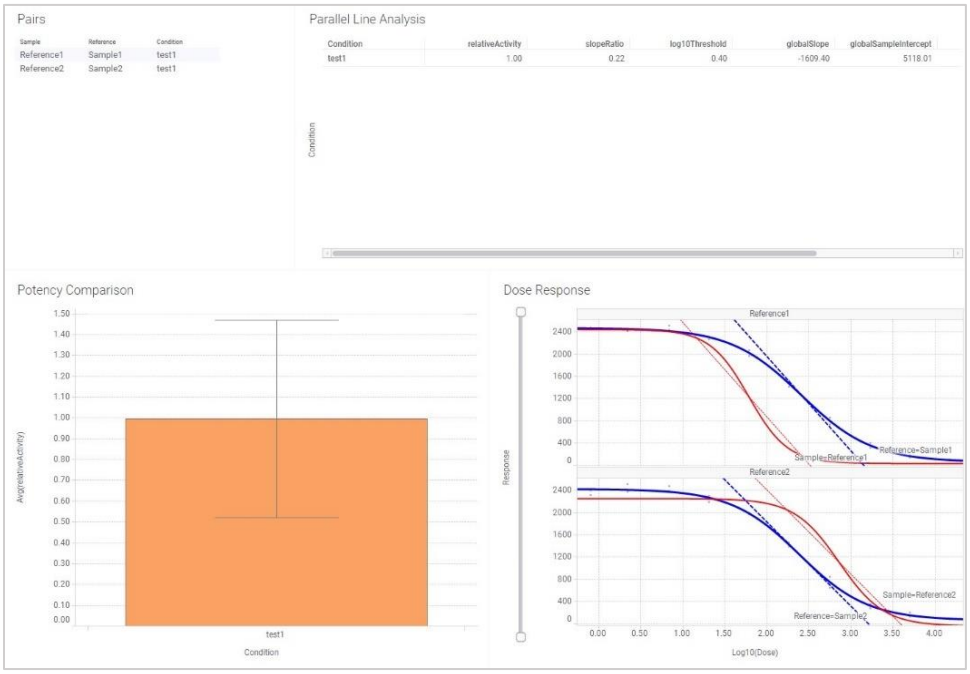


Figure 11-35: Relative Potency App Visualizations

- **Pairs**: This section provides the information on the pairings that the user has used and displays any additional condition columns that have been loaded and can be used for aggregating the data.
- **Parallel Line Analysis**: This table contains the result of the analysis. The following columns will be present:

Column	Description
Condition	The condition that has been selected in the ‘Condition column’ dropdown to aggregate the data for the analysis.

RelativeActivity	The relative activity of the sample compared to the reference.
slopeRatio	This is the ratio of the slopes.
Log10 Threshold	This is the slope thresholds used for the cutoff to decide if the curves are similar enough to perform the analysis.
globalSlope	The slope of the global fit.
globalSampleIntercept	The intercept of the sample when using the global slope.
sampleSlope	The slope of the Sample fit when fit individually.
sampleIntercept	The intercept of the Sample fit when fit individually.

- **Potency Comparison:** This bar chart will display the relative activity for each of the groups in the selected condition.
- **Dose Response:** These scatterplots will display the reference and sample curves as well as the global fits for each of them. In addition, the sample datapoints can be seen in order to identify any point that may be problematic.

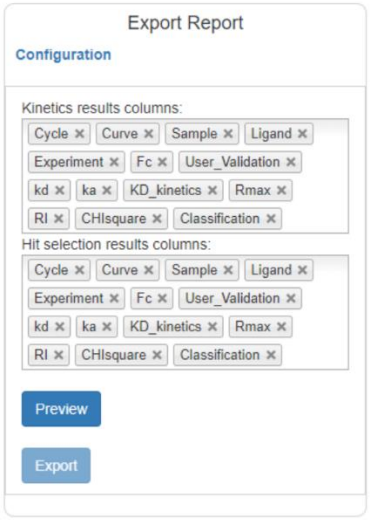
11.13 Export Report App

The **Export Report** App allows the selection and exporting of SPR analysis results in Spotfire®. A set of interactive visualizations are provided to easily configure the previewing and exporting of results.

Figure 11-36: Export Report App Card

11.13.1 Configuring the Export Report App

Once the App is open, the **Configuration** section is visible in the interface on the left-hand side of the screen.



To export the report:

1. Select the **Kinetics results columns** to be included in the report.
2. Select the **Hit selection results columns** to be included in the report.
3. Click **Preview** to visualize the report in the right-hand side of the screen.
4. Modify any elements as desired.
5. Click on **Export** to export the report as an .html file.

11.13.2 Export Report App Results

When the **Export Report** App preview has been generated, the right-hand side of the screen will allow the exploration of the report before exporting the final version.

The sections generated by the **Export Report** App in the report are the following:

- **Report experiment info**: This section provides the information on the experiment that the user has entered in the **SPR Data Import** App.
- **Report data summary**: This section provides the information on the results of the kinetic analysis and hit selection.
    - **Kinetic results**: This table provides the kinetic information selected in the configuration section for the kinetic results columns, together with a visualization of the sensorgrams with the fitted curves.
    - **Hits results**: This table provides the hit selection information selected in the configuration section for the hit selection results columns, together with a visualization of the Kon – Koff plot.

11.14 Export to TraceDrawer

The **Export to TraceDrawer** App allows the users to export the information in the loaded SPR files to a format accepted by TraceDrawer.

Figure 11-37: Export to TraceDrawer App Card

11.14.1 Configuring the Export to TraceDrawer App

Once the App is open, the **Configuration** panel is visible on the left-hand side containing three tabs:

- **Analysis:** This tab contains a single button to **Export** and will generate the required TraceDrawer files from the data that is visible in the *Curve exploration* table.
- **Visualization:** This tab contains two toggles:
  - **Filter panel:** This will toggle the view of the filter panel on or off (default on).
  - **Show injection start:** This toggle will turn on the visualization of the injection start positions on the *Sensorgrams to export* visualization (default off).
- **Configuration:**
  - **Guided Mode:** This toggle is currently not needed in the **Export to TraceDrawer** App and may be removed in the future.
  - **TraceDrawer license key:** Here the user can enter their Tracedrawer license key.

11.14.2 Export to TraceDrawer App Results

Once the user has used the filter panel on the right-hand side to limit the curves needed for the TraceDrawer analysis, select **Export** to generate a file for each of the pairs in the table that can be submitted to the TraceDrawer server for analysis.

Note: To submit the files to the TraceDrawer server for analysis it is necessary to have a valid TraceDrawer license which is not part of Signals VitroVivo and should be obtained separately from the TraceDrawer vendor.

12. In Vivo Domain Apps

12.1 Introduction

This section will describe the Signals Apps provided to analyze data in the **In Vivo** domain. To launch an App from Spotfire® Analyst, from the main toolbar, navigate to **Tools** > **Signals Apps.** This will open the Signals Apps tab. Alternatively, if this tab is already present in the document, simply navigate to it using the appropriately labelled tab at the bottom of the screen.

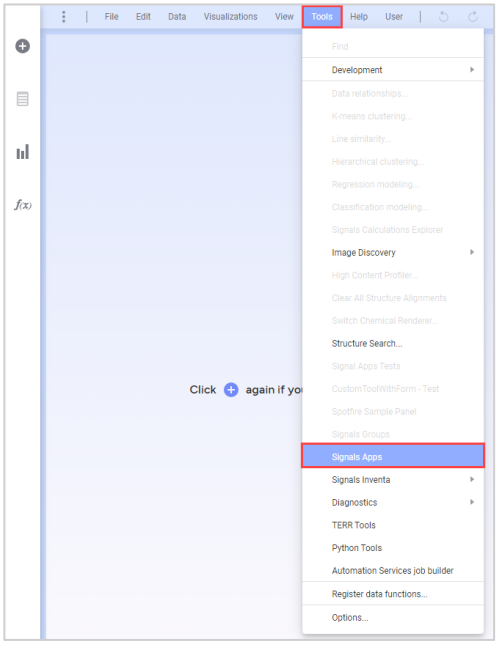


Figure 12-1: Launching Signals Apps

By default, the **Workflows tab** will open upon launch. Navigate to the **Apps tab** to view the available Signals Apps and scroll down to the **In Vivo** domain Apps area.

Figure 12-2: In Vivo Apps on the Signals Apps Page

12.2 PK Parameters App

The **PK Parameters** App allows the user to review and QA the Pharmacokinetics (PK) curves and calculate the final PK parameters and contains a user interface area containing three tabs, **“Analysis”, “Visualization”, “Configuration”** and an additional panel **“…”,** displayed in the top-left-hand corner.

Figure 12-3: PK Parameters App Card

12.2.1 Configuring the PK Parameters App

12.2.1.1 Analysis Tab

To load the required data in the App, use the **Mass spectrometer data** and **Dosing data** buttons. Selecting these buttons will open their respective data selection dialogs with two options: select the appropriate data table from the Spotfire® document or upload a new data table by selecting **Load new data**. To upload new data, browse to the appropriate .txt file(s) and select **Open**. Note that multiple files of the same format can be selected.

Once the data tables have been selected or uploaded note the dropdown menus for each column header. If the dropdown menus are not automatically populated, choose the matching column header from the loaded data table. After making the selections, click **Accept**.

The loaded Mass Spectrometer and Dosing data will be displayed in the following visualizations: *Concentration versus Actual Time*; *Mass Spectrometer value by Animal*; *Mass Spectrometer average by Cohort*; and *Dosing Information*.

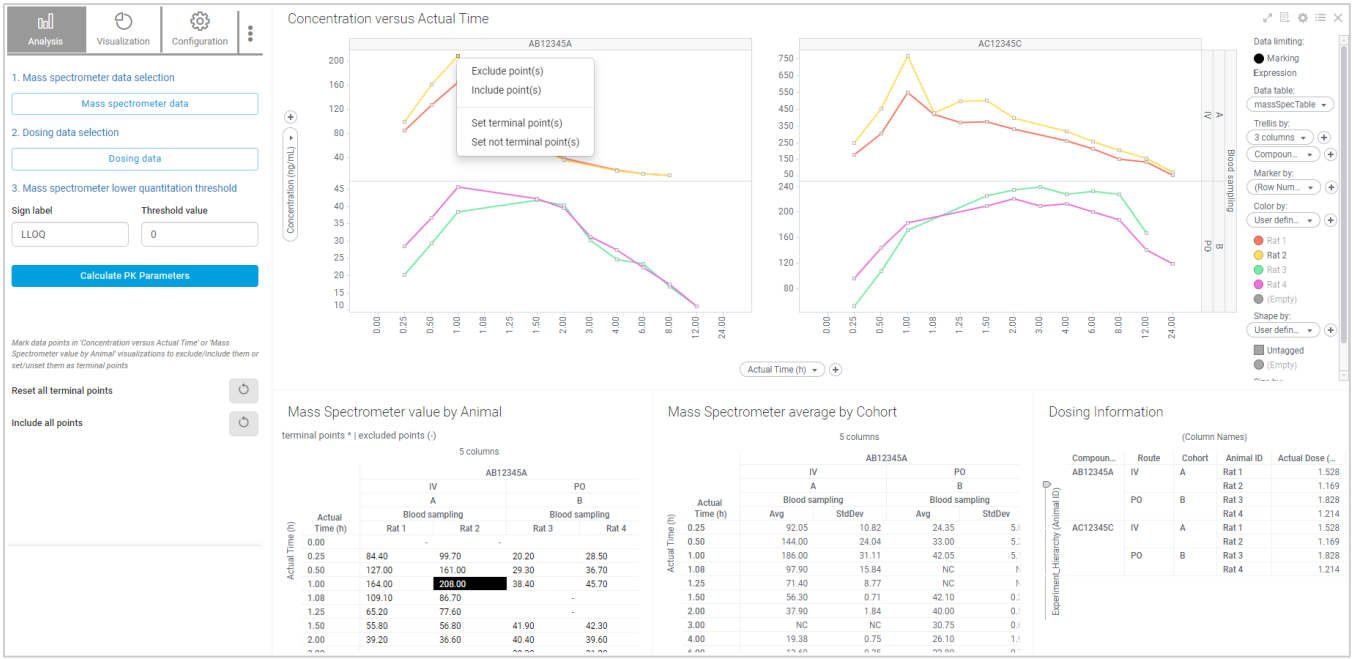


Figure 12-4: Example of the Visualizations Generated with Loaded Mass Spec and Dosing Data

The remaining controls in the left-hand panel of the **Analysis** tab consist of:

- **Mass spectrometer lower quantitation threshold:**
  - **Sign Label** (default LLOQ): An editable sign label.
  - **Threshold value** (default 0): Flags all values below the set mass spectrometer lower quantitation threshold. Corresponding data points will be filtered out of the analysis.
- **Calculate PK Parameters:** Calculates the pharmacokinetics (PK) parameters and generates the *Pharmacokinetics results* table in the bottom right-hand panel.

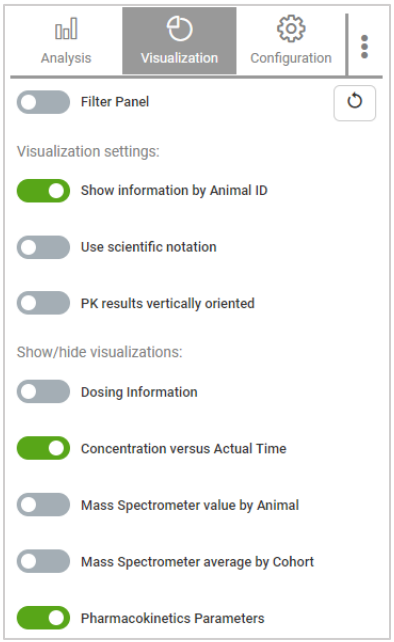
- **Reset all terminal points:** Any prior terminal point configuration will be removed so no mass spectrometer data points will be considered as terminal points.
- **Include all points:** All mass spectrometer data points previously excluded will be included back in the analysis.

If desired, points may be manually excluded by selecting them in the *Concentration versus Actual Time* or *Mass Spectrometer value by Animal* visualizations and selecting **Exclude point(s)** in the resulting pop-up menu. Excluded points are shown as a red X on the plot. Points may be included in the same manner.

Additionally, if desired, points may be marked as terminal points to estimate elimination half-life by selecting them in the *Concentration versus Actual Time* or *Mass Spectrometer value by Animal* visualizations and selecting **Set terminal point(s)** in the resulting pop-up menu. Terminal points may be unmarked in the same manner. Note that if no manual selection is made, the algorithm will automatically select the terminal points by computing the best fit line for all points after Tmax (requires at least 3 points after Tmax). Automatic selections should be reviewed or refined where changes are applied by recalculating the PK parameters.

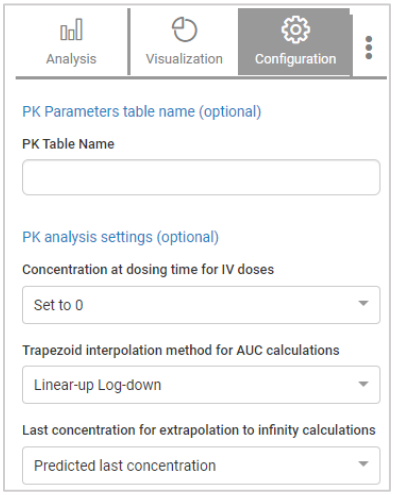
12.2.1.2 Visualization Tab

In addition to the **Filter Panel**, the following toggles are available to show / hide the available visualizations and offer additional visualization settings. Options and default states are dependent on the stage of the analysis (i.e. before or after PK parameters are calculated).



12.2.1.3 Configuration Tab

The **Configuration** tab contains the following settings:



- **PK Parameters table name (optional):** the user can provide an optional table name for the *PK Parameters* table in the **PK Table Name** textbox.
- **PK analysis settings (optional):** the user can adjust the PK analysis settings to ensure enough flexibility to meet their analysis requirements. Note that as it is most likely that the same setup is applied to all the studies of a specific type (e.g. to all the studies with bolus IV dose), when a Workflow is created the selected settings will be saved as Workflow parameters. Available options are:
    - **Concentration at dosing time for IV doses:** controls the concentration value at dosing time (*C0*) used/assumed for IV routes. Available options are:
        - **Concentration at time 0 provided**: the concentration at the dosing time should be included in the Mass Spec data provided. If missing, it will be assumed to be 0.
        - **Back extrapolation**: the semi-log line between the first two measured times is computed and that line is used to extrapolate backwards to dosing time.
        - **Same as c1**: the concentration at dosing time is assumed to be the same as at the first timepoint after dosing time.
        - **Set to 0**: the concentration at dosing time is set to 0.
    - **Trapezoid interpolation method for AUC calculations:** controls the interpolation method used for the AUC calculation. Available options are: **Linear-up Log-down** and **Linear**.
    - **Last concentration for extrapolation to infinity calculations:** controls which last concentration (*Clast*) is used for extrapolation to infinity calculations such as *AUC_0-inf*. Available options are:
        - **Predicted last concentration**: an estimation using a semi-log regression of the last concentration is used.
        - **Observed last concentration**: the actual last concentration observed is used.

12.2.2 PK Parameters App Results

Pharmacokinetics Parameters

(Column Names)

6 columns

Compound ID	Route	Cohort	Event Type	Anim... ID	Thumbnai... by Animal ID	Actual Dose (mg/kg)	Cmax (ng/mL)	Tmax (h)	Half-life (h)	Lambda -z	MRT (h)	CL (mL/min/kg)	Vdss (L/kg)	AUC 0-t (ug.h/mL)	AUC 0-Inf (ug.h/mL)	AUC extrapolated (%)	DNAUC 0-t (h.kg/L)	DNAUC 0-Inf (h.kg/L)	Bioavailability (%)	Terminal Points
5163638A	IV	A	Blood sampling	Rat 1		1.53	164.000	1.00	4.80	0.144	4.89	70.132	20.564	0.288	0.363	20.599	0.189	0.238		8, 6, 4 points automatically ...
				Rat 2		1.17	208.000	1.00	5.58	0.124	5.26	48.216	15.210	0.313	0.404	22.418	0.268	0.346		8, 6, 4 points automatically ...
	PO	B	Blood sampling	Rat 3		1.83	41.900	1.50	5.72	0.121	8.69	84.230	43.898	0.272	0.362	24.879	0.149	0.198	67.844	12, 8, 6 points automatically ...
				Rat 4		1.21	45.700	1.00	6.07	0.114	8.82	52.919	28.012	0.285	0.382	25.492	0.235	0.315	107.985	12, 8, 6, 4, 3 points automatically ...
5218131A	IV	A	Blood sampling	Rat 1		1.53	548.000	1.00	9.90	0.070	13.62	5.626	4.598	3.734	4.527	17.507	2.444	2.963		24, 12, 8, 6 points automatically ...
				Rat 2		1.17	769.000	1.00	10.47	0.066	14.04	3.398	2.862	4.662	5.732	18.667	3.990	4.905		24, 12, 8 points automatically ...
	PO	B	Blood sampling	Rat 3		1.83	239.000	3.00	11.93	0.058	18.60	5.624	6.278	2.483	5.416	54.148	1.359	2.964	75.333	12, 8, 6 points automatically ...
				Rat 4		1.21	221.000	2.00	23.53	0.029	34.66	2.654	5.520	3.766	7.626	50.621	3.101	6.280	159.633	24, 12, 8, 6, 4, 3 points automati...

Figure 12-5: Example of the PK Parameters Table

Note: For improved compatibility and mapping with Signals Inventa's Target Engagement Profile App, three additional columns have been added to the *PK Parameters* table; *Study ID, Species, & Formulation*. Note these are hidden by default but are exported with the *PKAnalysis* table.

13. Appendices

13.1 Appendix Kinetic Fitting Model

This section describes the fitting process implemented within the SPR Kinetic package, starting with a general overview of the fitting process (a standard nonlinear least squares optimization), then outlining several hacks introduced to either improve the quality of the fits or to accelerate the optimization process.

This section also discusses the method employed to obtain the estimates of the error for critical parameters.

13.1.1 Overview

At a very high level, the fitting process is a standard case of nonlinear least squares (NLS) optimization. Deviations from the standard NLS case are discussed in the sections below.

The mathematical model (K. Sigmundsson, 2002) governing the kinetic behavior of the underlying chemical process over time provide a function, $f_\Theta(t)$ depending on a vector of parameters of interest, Θ, where experimental data provide a set of noisy measures (t_i, y_i).

The fitting process looks for the vector of parameters Θ minimizing the expression.

$$\sum_i \left(f_\Theta(t_i) - y_i\right)^2$$

The average value of the errors $\left(f_\Theta(t_i) - y_i\right)^2$ is the reported χ^2 goodness of fit measure. It is very important to note that this indicator measures the fitting error and that this error has two different components:

- Measurement error, or the noise generated in the measurement process, which is mostly random and unstructured.
- Model or model specification error, which is systematic (e.g., the fits underestimate the curves in some periods and overestimate them at other periods).

Usually, an experiment consists not just on a single curve, but a set of curves built on different experimental conditions, i.e., at different concentrations. In such cases, the initial data consists of points (t_{ij}, y_{ij}) where j indexes the different curves and the fitting process minimizes.

$$\sum_i \sum_j \left(f_\Theta\left(t_{ij}\right) - y_{ij}\right)^2$$

In such cases, both a global χ^2 and individual χ^2 (one for each curve) are provided.

Also, when the experiment consists in more than one curve, it is possible to specify global parameters within the vector Θ together with local parameters (affecting a single curve). These options will be discussed below.

The minimization problem is using a general purpose optimization software that implements the L-BFGS-B algorithm (Several Authors, 2018) that will be discussed below.

13.1.2 The Fitting Function

This section describes the function $f_\Theta(t)$ introduced before and its implementation. It is the sum of two different parts, a nonlinear function described in (K. Sigmundsson, 2002) and a step function related to the refractive index.

13.1.2.1 The Sigmundsson Function

The Sigmundsson function is described in (K. Sigmundsson, 2002) and it is the solution to an ordinary differential equation. It does not consider the refractive index (RI) effect, which will be discussed below. At this point, it suffices to note that this effect is estimated outside the machinery employed to fit the Sigmundsson function.

This differential equation has two distinct parts. One of them applies to the injection period, when the analyte is being injected into the system at a given concentration. This injection period is followed by the decay phase, when injection stops and the analyte gets progressively removed.

The Sigmundsson function, i.e., f_Θ above, depends on the following parameters:

- k_a, the association parameter, that measures the intensity of the affinity between the ligand and the analyte (or the binding speed). This parameter is always calculated globally.
- k_d, the dissociation parameter, that measures how fast the dissociation occurs. It is also calculated globally.
- k_t, or the transfer parameter. It is related to the viscosity of the analyte and, therefore, how fast it can percolate from the dissolution bulk to the interaction surface and the other way around. It is usually estimated globally, although it is possible to set it to a fixed value.
- R_{max}, which is an upper bound on the amount of available ligand on the interaction surface, forcing the saturation of the interaction by depletion. It can be set as a global value, a local (one by curve) value, or fixed (user provided) value.

In a single experiment several curves are fitted at the same time. They all share the same analyte and ligand pair (therefore, parameters k_a, k_d and k_t are common among them), but these curves may have a different R_{max} parameter. The user can specify either a common R_{max} parameter for all the curves or require a specific one for each curve. In any case, these options do not change the implementation of the f_Θ function.

There are two different approaches to implement the function f_Θ above (Sigmundsson function):

- An implementation based on the numerical solution to these equations, i.e., integrating numerically the ODEs.
- An implementation based on the closed formula for these equations provided in Sigmundsson paper.

Both approaches have been implemented, tested, and compared for efficiency and accuracy. In summary, naive implementations of the closed formula have the worst performance. Sigmundsson closed formula involves the calculation of Lambert's W function, a rare transcendental function for some extreme cases where numerical instability issues occur. Finding the values of such function requires solving a nonlinear equation by iterative methods.

Numerical integration of the ODEs has a much better performance and accuracy. This integration was performed using C code and automatic ODE integrators and it is orders of magnitude more efficient than the naive implementation of Sigmundsson's equations.

However, the final version of the fitting process relies on a heavily tailored implementation of Sigmundsson's closed formulas using C and a series of numerical tricks to speed up the calculations. For instance, the solution to the nonlinear equations via iterative methods mentioned above can be found much faster using the fact that the function values that need to be calculated are sequential and close to each other; therefore, intermediate values used in the calculation of a curve point provide reasonable starting points for the iterative process required for the calculation of the next.

13.1.2.2 The Refractive Index

A vertical shift affecting the measurements might be observed in the injection phase in some cases. This vertical shift is a deviation of the given measurements with respect to the *ideal* curves provided on theoretical grounds by (K. Sigmundsson 2002).

This shift, a refraction index effect, is related to the underlying physics behind the surface plasmon resonance technology and needs to be corrected to get good shifts and, therefore, credible estimations for the curve parameters.

To account for this effect, the actual function that gets optimized,

$$f_\Theta(t)$$

is decomposed as

$$f_\Theta(t) = s_\alpha(t) + r_{ri}(t)$$

where s_α is the Sigmundsson function discussed above, depending on the vector of parameters α enumerated in the previous section, and $r_{ri}(t)$ is a step function with values:

- ri_1 in the injection phase (Value of the signal at the end of the Injection Phase)
- ri_2 in the decay phase (Value of the signal at the beginning of the Decay Phase)

Users can set whether these values are zero (for all curves), fixed to a given value (for all curves) or whether to find the optimal values, either locally or globally.

If no jump occurs $ri_1 = ri_2$ and thus ri=0

If this is not the case the final reported ri parameter is the difference between the ri_1 and the ri_2, which measures the vertical shift often observed at the end of the injection phase.

13.1.3 Accuracy Improvements

The relatively simple optimization setup described in the previous sections could potentially suffice to accurately and robustly fit the curves. From a conceptual, high level approach, there is little more to add to the description of the fitting process: it is simply a nonlinear least squares problem whose underlying nonlinear function is given by certain ordinary differential equation.

In practice, most of the complexity of the implementation of the fitting process is related to a series of hacks introduced to prevent optimization mishaps. That the process gets trapped in local minima providing very bad fits.

There are two related reasons for bad fits (or local minima):

- Excess of parameters (more parameters, particularly when many parameters are estimated locally, not only implies a slower fitting process, but also a higher risk of local minima).
- Far off starting values for the main parameters. Optimization methods generally require the user to provide starting values; the L-BFGS-B method we use, does. The optimization process is highly dependent (both in speed and in accuracy) of the quality of the provided initial values. But users cannot be always trusted to provide good starting values for the different parameters (say, no more than one order of magnitude away from the actual values), and the further away initial parameters are from the optimal ones, the higher the risk of the optimization process to get trapped in local minima.

In any case, no matter how good fits are, there are cases where some parameters will never be able to be accurately measured. For instance, a very high k_t value means that analyte moves almost freely between the solution bulk and the interaction surface. Beyond some threshold, it does not matter how high the k_d parameter is: it is large enough so as not to have any significant effect in the process. Therefore, it can never be precisely estimated. A

similar effect can be found on the estimation of the R_{max} parameters when curves seem far away from the saturation point. Also, some effects can mask each other; for instance, relatively high k_a and k_d parameters can cancel each other (particularly in contexts of very slow transfer).

Issues like those discussed in the previous paragraph can be traced and help explain some local minima arising in naive implementations of the fitting process that the hacks described below helped to solve. Most of these hacks are based on the study of specific features of the Sigmundsson function.

13.1.3.1 Parameter Bounds

The L-BFGS-B allows for the specification of bounds for variables. This feature has been used to reduce the search space and to rule out *forbidden* regions for some parameters.

In particular:

- The ri parameter has been restricted so that its (absolute) value cannot exceed 10% of the maximum curve value.
- The k_d parameter has been bounded to constrain it within 2 orders of magnitude of its starting value (to be described below).
- The k_t parameter has been restricted to a region within which it is operative (in the sense, discussed above, that beyond certain threshold, it does not add any constraint to the chemical process).

13.1.3.2 Initial Values for the Dissociation Parameter

The parameter k_d is critical and it is easy to see the optimization parameter get trapped in local minima unless a reasonable starting value is provided. Because of this, a reasonable default has been put in place.

This default is constructed based on the observation that the decay equations can be approximated by an exponential decay. Therefore, an exponential decay model (a much simpler nonlinear model that can even be linearized taking logarithms) is fitted on the decay and the resulting coefficient is used as starting value:

$$y_t = Kexp\big(-k_d(t - t_0)\big) + \epsilon_t$$

We have observed that this estimation often falls within one order of magnitude of the final fitted value. This approximation, though, is affected by two potential features:

- A vertical shift on the decay phase, that has been observed in many examples. For instance, when there is a fast decay that stops at some level 0. It should be noted that both exponential decay and Sigmundsson equations preclude decays to values different from 0.
- Very low k_t values. With very high k_t values, Sigmundsson equations are effectively exponential decay. However, deviations occur when k_t values are much smaller.

13.1.3.3 Model Refitting

When local parameter (R_{max} or ri) estimation are requested, the number of parameters increases, and the problem of bad initial values becomes more acute. The strategy that was adopted to get good starting parameters for these local values is to perform a global fit first. The output of this previous global fit is used as starting values for a second, local fit.

This affects performance, but in a limited way because of two reasons:

- The global fit is much faster than the local fit; so, the processing time will never double. There can be, at most, a 20% or 30% execution time increase.
- Having better starting values in the local fit accelerates the optimization process.

13.1.4 Performance Improvements

Several modifications have also been introduced to increase the performance. This section discusses the most important ones beyond the already mentioned implementation of the numerical solution of the Sigmundsson ordinary differential equation in the much faster C language.

13.1.4.1 Calculation of the RI Parameter

The optimization running time is superlinear on the number of parameters: doubling the number of parameters more than doubles the time required for the fit. The problem becomes particularly acute when local fits are requested.

It is possible though to exclude the calculation of the ri parameter from the L-BFGS-B algorithm using the following trick. Remember that the optimization process searches for the minimum of the expression:

$$\sum_i (f_\Theta(t_i) - y_i)^2$$

which, according to the decomposition mentioned above, can be rewritten as:

$$\sum_i (s_\alpha(t_i) + r_{ri}(t_i) - y_i)^2$$

or, finally, as the sum:

$$\sum_i (s_\alpha(t_i) + ri_1 - y_i)^2 + \sum_i (s_\alpha(t_i) + ri_2 - y_i)^2$$

corresponding to the association and dissociation phases respectively. There, the values ri_j are fixed numbers, and the minimum (for each α) is obtained at the point:

$$ri_j = \frac{1}{N}\sum_i \left(y_i - s_\alpha(t_i)\right)$$

13.1.4.2 Fitting Error Estimation

For most fitted parameters, in particular, k_a, k_d, k_t, and R_{max} parameters, an estimation of the fitting error via the standard deviation can be calculated. For example, in a given fit, the estimated value of the k_a parameter could be $2.75e5 \pm 2.4e4$, implying a standard deviation of $1.2e4$ or about 4%.

There are some caveats with this information that needs to be taken in to account. The first one is that the usage of the standard deviation (SD) is adequate for normal errors, but it could be misleading in other cases. In the context of Sigmundsson's equations, errors for k_a, k_d, and R_{max} are usually symmetric and the SD interpretation is quite unproblematic. However, uncertainty about the k_t parameter is often heavily skewed. It is important to note that k_t measures medium viscosity and that beyond certain threshold, viscosity plays no role: it either impedes reactions totally if it is very high or does not affect them at all if it is very low. Variations of viscosity beyond these thresholds are impossible to measure and these induce very high SD values reflecting the fact that, e.g., once the medium is too viscose, it does not matter if it is two, ten or a hundred times as viscose.

The second caveat is related to the correlation between parameters. SDs for each value are calculated conditional on the values of the remaining parameters. But it happens sometimes that the association (k_a) and dissociation parameters (k_d) are highly correlated. It is so because they have opposite effects, so they mask each other. Therefore, for a fixed value of k_d, it may seem that k_a has a low uncertainty, but the situation could change dramatically if we allowed k_a and k_d to vary (conditional on the value of the other parameters).

The conditional SDs are calculated as indicated below. It should first be noted that a Bayesian implementation of the fitting process along the lines of (Feng, 2015) would have automatically provided error estimations for all parameters. This approach is unfeasible because of efficiency considerations: the Bayesian fit of a set of curves may take minutes if not hours as opposed to very few seconds with our approach. However, it suggests a strategy for the efficient calculation of SDs.

It suggests that the shape of the density of the posterior distribution (the distribution providing the uncertainty of the parameters) is proportional to the expression:

$$\sum_i \sum_j \left(f_\Theta(t_{ij}) - y_{ij}\right)^2$$

considered as a function of the parameters Θ. If we take one of them, say θ (which can be, e.g., k_a) and leave the others fixed, we know that:

$$f(\theta) = \sum_i \sum_j \left(f_{\theta,\Theta'}(t_{ij}) - y_{ij}\right)^2$$

is proportional to the density function of the posterior distribution. If we knew $f(\theta)$, then we could use rejection sampling (Rejection Sampling, n.d.) to get a sample of potential parameter values according to the actual posterior (conditional) distribution and calculate their SD. But is possible to get an estimate of $f(\theta)$ by evaluating this expression on a grid of values around the optimum and using interpolation.

Because of regularity considerations (the density function is unimodal and smooth) it is only necessary to evaluate that expression about 20 times for each parameter, which does not add much overhead to the full estimation process.

13.1.5 References

[1] Feng, Thomas B., Feng AND Kepler. 2015. “Bayesian Estimation of the Active Concentration and Affinity Constants Using Surface Plasmon Resonance Technology.” *PLOS ONE* 10 (6). Public Library of Science: 1–17. DOI:10.1371/journal.pone.0130812.

[2] Sigmundsson, K., et al. 2002. “Determination of Active Concentrations and Association and Dissociation Rate Constants of Interacting Biomolecules: An Analytical Solution to the Theory for Kinetic and Mass Transport Limitations in Biosensor Technology and Its Experimental Verification.” *Biochemistry*. DOI: 10.1021/bi020099h.

[3] Rejection Sampling. n.d. “Rejection Sampling — Wikipedia, the Free Encyclopedia.” https://en.wikipedia.org/wiki/Rejection_sampling.

[4] Several Authors. 2018. *L-Bfgs*. https://es.wikipedia.org/wiki/L-BFGS.

13.2 Appendix Screening Data Model Information

To simplify the use of the Screening domain Apps, even when used independently of a preconfigured Workflow, Apps rely on a set of document properties that are used during the analysis to help define automatic App configuration. This organization is what we termed the **Screening data model** throughout this document.

When a dataset is loaded through the **Data Import** App, the App will mark the imported table as the *screening table* by setting a table property in the document. The purpose is to ensure downstream Apps select this table by default as the input table for analysis.

In addition to marking the table to be used, the different columns that are added will be marked as features or annotations which may be of different types. This will help provide limited subsets of columns as possible choices during App configuration. In this way, the user can have a set of more relevant selections when configuring the Apps as opposed to having a list with all existing columns in the document to choose from whenever configuring a parameter for an App.

13.3 Appendix Curve Fitting in the Calculations Explorer and New Curve Fitting App

The curve fitting performed in the **New Curve Fitting** App uses the same underlying fitting implementation as the curve fitting performed from the **Calculations Explorer**, however it is configured in a manner that facilitates usage within a VitroVivo Workflow.

This section provides some additional details on the implementation used in curve fitting, provided by the **Calculations Explorer**.

13.3.1 Fitting Algorithms

There are two options available for fitting in the **Calculations Explorer** when configuring a curve fit:

- Levenberg-Marquardt (default)
- Levenberg-Marquardt Robust

In the Levenberg Marquardt approach, the Levenberg-Marquardt algorithm is used to solve the non-linear fitting equation and provides as a solution those parameter values that minimize the fitting error. The way this fitting error is computed can be altered by using the Y weighting dropdown.

In the Levenberg-Marquardt Robust approach, the Levenberg-Marquardt algorithm is applied within an Iterative Reweighted Least Squares (IRLS) method. After each iteration of the fitting, the points with the largest residuals are reweighted using Tukey's Biweight method to reduce the influence of potential outliers in the fitting result. The number of outliers excluded, and the fitting algorithm sensitivity can be tuned using additional parameters provided.

13.3.2 Fit Status Results

If the fitting has been performed successfully, the fitting status returned is "OK".

If fitting is not achieved, the status will indicate this. If the fitting did not converge, this will be indicated in the status and *t* will be noted if the initial parameters are used as the fit results in case no valid number was provided.

When using the robust fitting method there are two additional status messages that may be displayed:

- **IRLS exclusion limit validation failed**: If more datapoints are excluded by IRLS than the threshold set by the user.
- **IRLS sensitivity validation failed**: If the sensitivity threshold is reached.

In these cases, no fitting will be returned.

In addition to the fit status, information will be provided if the confidence intervals could not be calculated.

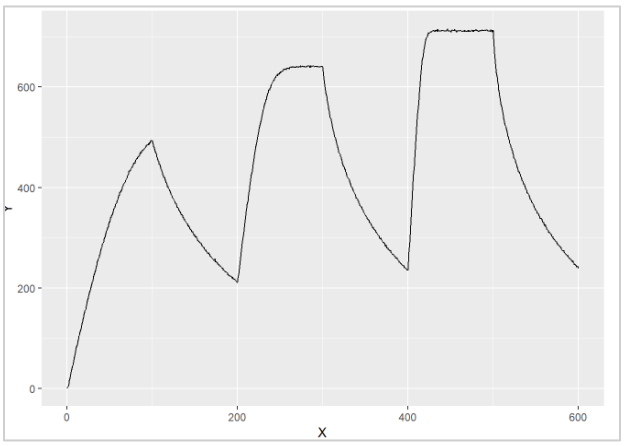
13.3.3 Curve Fitting Tables

Whenever a curve fitting is performed in the **Calculations Explorer** or the **New Curve Fitting** App, a new table is created that contains the fitting parameters and coefficients. This table will be updated whenever the curve fit changes and is the one that should be used as the source of any information based on the curve fitting results.

13.4 Appendix Non-Regenerative Kinetic Fitting Model

13.4.1 Introduction

This section describes the NRK (Non-Regenerative Kinetic) curve modeling approach. NRK curves are similar to multi-cycle curves where a regeneration step is performed between consecutive cycles, with the main difference being that several injection and decay cycles are run sequentially using different concentrations with no intervening regeneration cycles. An example of typical data for an NRK experiment is below:



Typical Aspect of an NRK Curve

The curve shown above, created synthetically according to the theoretic underlying model, has three injection and decay segments. In each segment, the analyte concentration has been doubled. In this type of setting it is not expected that during the decay segments the curve would return to the baseline that corresponds to no analyte bound to the ligand. Therefore, as opposed to what happens in multi-cycle kinetics, curves from consecutive concentration cycles would not start from zero.

13.4.2 Fitting Process

The fitting process used in the Signals VitroVivo Apps for these curves is a standard function minimization that looks at a set of parameters: θ (including ka, kd, kt, and Rmax) that minimize the expression:

$$\text{error}(\theta) = \sum_i (f_\theta(x_i) - y_i)^2$$

where the pairs (x_i,y_i) are the given data (or curve) and f_θ is a candidate curve built from the parameters θ.

The function error(θ) is minimized using standard optimization routines using R.

Note: Parameters ka, kd, kt, and Rmax have ranges that span over very different scales, producing numerical instability in optimization routines. To address this, and given that all these parameters are positive, as a normalization strategy they are transformed to log scale before optimization. They are optimized in log scale and are later transformed to the natural scale in the reports.

Note: The reported goodness of fit measure, the ChiA value, is precisely the minimum value of error(θ) divided by the number of observations.

The only non-standard component in the fitting process is the construction of function f_θ, which is based on the procedure used for the multi-cycle curves. Multi-cycle curves have two segments, (1) an injection segment originating from (0,0) (zero concentration at time 0), and (2) a decay segment originating at the endpoint of the previous segment. The values of f_θ are calculated using the solution to Sigmundsson equations in each segment.

In the case of NRK curves, several injection and decay segments are chained and the same solution to Sigmundsson equations is used sequentially, by taking the endpoints of each segment as the starting value for the next.

Although closed numerical approximations to Sigmundsson ODEs were used for the multi-cycle curves due to performance considerations, in the fitting of NRK curves the numerical ODE integration approach simplifies this *curve chaining* approach.

13.5 Appendix: Editable Data Grid App for SPR Domain

The **Editable Data Grid** App can be used to add or modify annotation data in the *ResultsTable* imported via the **SPR Data Import** App with the following functionalities:

- Ability to add and modify plate map metadata.
- Ability to add and modify concentration metadata.
- Ability to add and modify report points metadata.

Note that the user may insert data from the clipboard that is present in a grid-like structure, such as data copied from Excel, or insert data directly from a file and edit it before adding it to the results table.

13.5.1 Launching the Editable Data Grid

From the Signals Apps page, select the **Editable Data Grid** App card from the Screening domain. A new tab containing the App will be added to the document and the App will be launched.

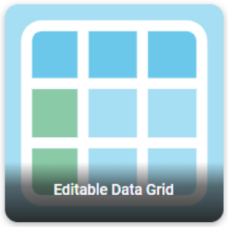
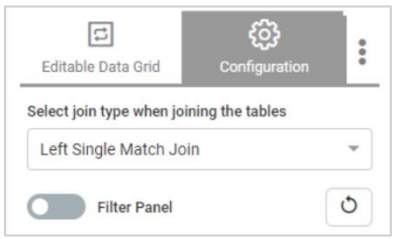


Figure 13-1: Editable Data Grid App Card

13.5.2 Configuring the Editable Data Grid

The user is presented with an empty data grid and a configuration UI containing two tabs, **Editable Data Grid**, **Configuration** and an additional panel "…", is displayed in the top-left hand corner.

- **Editable Data Grid:** This section contains the main controls that allow the user to load and save the table(s).
- **Configuration**: This section contains a dropdown menu to select join type between tables (i.e. 'Left Single Match Join' and 'Left Outer Join'). It also contains a toggle to display the **Filters Panel**.



The empty data grid can be used to paste data from the clipboard and will automatically resize to adapt to the size of the pasted data. The data grid always displays an empty row at the end to facilitate the addition of data.

13.5.2.1 Adding Plate Map Metadata

Add a **plate map file,** if desired. The plate map file allows the user to associate metadata to the ligand. The file must contain the names of the ligands in one column. The other columns can contain a notebook reference to the batch information for the ligand or another alias for the ligand name.

To add a plate map:

1. Ensure **(Create a new empty data grid)** is selected in the '**Populate editable Data Grid from:**' dropdown.
2. Either copy the desired data from an external source such as Excel and paste the data into the empty grid or select the '**Load file…**:' icon and browse to the file on your local file system.
3. In the '**Save data as**' section select '**New Table**' and give the new table a name in the provided textbox.
4. Select **'Apply'** to save the new data table to the Spotfire® document.

13.5.2.2 Adding Concentration Data

If the concentration data is provided in a separate file or needs to be updated in the results table due to an error in the input file, a text file containing the concentrations can be loaded to replace the existing values.

To replace concentration data:

1. Ensure **(Create a new empty data grid)** is selected in the '**Populate editable Data Grid from:**' dropdown.
2. Either copy the desired data from an external source such as Excel and paste the data into the empty grid or select the '**Load file…**:' icon and browse to the file on your local file system.
3. In the '**Save data as**' section select '**Replace column values**' and select '**ResultsTable'** from the '**Table to replace column values:**' dropdown. This step adds data from the data grid to an existing table replacing existing column values in the target table.
4. The user must now configure the matching between the data in the grid and the existing table by selecting the + icon. See section 9.8.1.1 for details on configuring matches.

13.5.2.3 Adding Report Points

New report points can be added to the *ReportPoints* table in the Spotfire® document for downstream analysis by selecting the *ReportPoints* table from the '**Populate editable Data Grid from:**' dropdown. To add a custom report point, fill in the empty row at the bottom of the data grid with the following information:

- **Report point**: Name of the report point.
- **Time (s):** Time at which the report point is defined relative to one of the regions in the curve.
- **Position:** The way in which the position is defined relative to the region in the curve.
- **Relative To:** The region relative to which the position is defined. These regions are the ones provided in the .blr file and are defined in the T200 instrument.
- **Window (s):** The window size in seconds that is used to calculate the report point.
- **Baseline:** Indicates if the report point is a baseline report point or not. This is important when evaluating the signal at each report point, as this will be calculated relative to the previous report point.

A new empty row will be added automatically to facilitate the addition of report points. Select **Apply** to add the report points to the *ReportPoints* table. Note that any report points can be edited directly from the grid as desired.

13.5.2.4 Data Grid Visualization

In addition to the controls available in the left-hand panel, the data grid section contains some additional controls that the user can access directly from it:

- **Reset grid:** This button will clear the data grid.
- **Column header controls:** Currently this control provides a single option to clear the contents of a column.
- **Right-click controls:** Currently right-clicking on the data contains the copy option that allows the user to copy the values of the selected cells to the clipboard.

Note: Spotfire® filtering is not integrated into the data grid. However, once the table is saved, if it is loaded in the Editable Data Grid the standard Spotfire® filtering options are available for use and the grid contents will be filtered accordingly.