

Chapter 15

Bruker Daltonics

Soeren-Oliver Deininger

Abstract In this chapter, a brief overview of running MALDI imaging experiments on Bruker flex instruments is given. General guidelines on the instrument setup and a discussion of the laser diameter settings in general are included. The basic operation of the flexImaging software is described, including the acquisition of the primary optical image, the teaching of the sample carrier, the selection of measurement areas, and the acquisition. The different ways to evaluate the results are also covered: The selection of peaks on a sum spectrum and the intensity distribution thereof, the comparison of selected regions on the tissue to find specific masses, and statistical evaluation, including principal component analysis, hierarchical clustering, and supervised classification, are also covered.

15.1 Preparing Instrument Parameters for MALDI Imaging Experiments

Before the acquisition of the matrix-assisted laser desorption/ionization (MALDI) image, the instrument parameters need to be appropriately set. Most instrument parameters need to be set only once for a given type of experiment; they include mass range and reflectron or linear mode, among others. Here, the sampling rate deserves more attention than in other applications, as it defines how many data points are recorded per second. The sampling rate is directly proportional to the size of the spectrum. A spectrum recorded with a sampling rate of 1 S/s will be 10 times larger than a spectrum with 0.1 GS/s. The chosen sampling rate should be as small as possible without compromising the peak shape in the spectrum; this is because for an imaging dataset with many thousands of spectra, the size of the spectra accumulates quickly. For a linear measurement of proteins, a sampling rate of 0.1 GS/s is usually sufficient.

S.-O. Deininger (✉)
Bruker Daltonik GmbH, Fahrenheitstrasse 428359, Bremen, Germany
e-mail: Soeren-Oliver.Deininger@bdal.de

15.1.1 Laser Diameter Settings

The smartbeam laser of flex instruments has an adjustable focus size that can be adjusted in several steps (e.g., ultralarge, large, medium, small, and minimum). The question of what exactly the diameter of the laser focus should be for each setting is often asked; unfortunately, that question is not easily answered. The diameter of a Gaussian-shaped laser beam profile is typically reported as full-width at half-maximum. Although the smartbeam laser has no Gaussian beam profile, the energy profile averaged over many shots has a Gaussian envelope. What matters in terms of the imaging experiment is the “effective” laser diameter that affects the sample; this effective diameter is strongly dependent on the laser energy. If, for example, a high laser energy level is needed, the effective diameter will be larger than with a lower laser energy level. The effective diameter is therefore also difficult to estimate; it can really only be estimated by acquiring spectra at a given position, moving the laser a defined distance, measuring again, and then looking at the distance at which the spectra still have comparable intensities. The recently introduced smartbeam II laser has a flat-top envelope and is therefore more robust in MALDI imaging.

In terms of MALDI imaging, the salient question is which lateral resolution can be achieved under practical conditions with real tissue. Here, we have observed that resolutions down to the 70- μm level can be obtained with the “large” laser setting and a sinapinic acid matrix, which is best suited for proteins. With a “medium” laser setting, ~50- μm resolution is possible with sinapinic acid. With a cyano-4-hydroxy cinnamic acid matrix, the “small” laser setting allows for an imaging resolution of approximately 25 μm . These values should be considered only guidelines, because they depend upon the laser energy, and the necessary laser energy in turn depends upon the matrix, the preparation, and the amount of salt and lipids in the tissue.

15.2 Data Acquisition and Processing with the Bruker Daltonics MALDI Imaging Solution

The core of Bruker Daltonics’s software for MALDI imaging is the flexImaging program. It is used both as an editor to set up the imaging run and to visualize and evaluate results.

flexImaging requires an optical image of the sample, which is referred to as the “primary image.”

15.2.1 The Primary Image

This optical image will later be used to define the measurement areas, and the results will be overlaid with this primary image. In setting up a new imaging sequence, the user is asked at one point to register this image in the instrument coordinates; this is

done by aiming with the crosshair of the instrument camera onto a specific feature and defining the corresponding feature in the optical image. Whether this primary image needs to be taken before or after the application of the matrix depends on how the matrix was prepared. If the matrix was prepared as a homogeneous layer – for example, by the ImagePrep station or a spray preparation – the primary image should be taken before matrix preparation. In this case, it is also necessary to put some teach marks onto the sample that can be clearly recognized under the matrix layer. Liquid Paper (“white-out”) has been shown to work nicely here.

If the sample has been prepared by spotting a grid of matrix droplets through the use of a robot, then the primary image needs to be taken with the matrix spots already on the sample, because the sample will be “tached” onto the corner positions of the grid.

For the acquisition of the optical image, a photographic slide scanner with an adaptor for microscopy slides is most convenient. Office scanners could also be used; however, the quality of the optical image is not as good as from a slide scanner. Microscopes can also be used, but the field of view is often too small to obtain an image of the entire sample.

15.2.2 Creating a New Imaging Sequence

If a new imaging experiment is being set up, first it is necessary to introduce the sample into the instrument. In flexImaging, the option to set up a new sequence is selected. The initial setup of the sequence is now controlled by the new sequence wizard; this wizard guides the user interactively through the necessary steps. First, a name for the sequence must be defined. On the next page, the wizard asks for the imaging mode; the user can choose from three modes of operation that are defined by the way the matrix is applied to the sample. In the case of a homogeneous matrix layer, the user must also select the option “uniformly distributed coating.” In this workflow, the user can arbitrarily choose the lateral resolution of the image in the measurement region. In the case of a robotically spotted matrix grid, the user chooses the “spot microarray” option; in this workflow, the lateral resolution of the imaging experiment is defined by the center-to-center distance between matrix spots, and so the user must “tell” the software at this point how many rows and columns are present or what are the center-to-center distances in the x - and y -directions. Alternatively, it is also possible to name a file with spot coordinates from the robot. The third workflow, “tissue profiling,” is used only if specific spots at arbitrary positions are being measured.

Next, the wizard asks for the primary optical image. After this is selected, the process of teaching begins. In this teaching, the optical image is correlated with the sample on the sample carrier. For three reference spots, the corresponding features are first selected with the crosshair of the MALDI camera optic and on the primary optical image. After this teaching is finished, the new sequence wizard closes and the primary optical image is displayed in flexImaging. Because of the teaching, the software can

now correlate each pixel on the primary image with the respective position of the sample carrier. flexImaging can now be used to remote control the sample carrier, to move to a specific point and take a spectrum, that is, check the quality of the preparation or estimate the appropriate laser energy level for the measurement (Fig. 15.1).

As a next step, the measurement region must be defined, either as a rectangular or as a polygon-shaped region. It is possible to define more than one region on the same sample. In the region's pane in flexImaging, the lateral resolution for the region can be defined, as well as the parameters for autoXecute, which defines the automatic measurement and processing. Setting up for automatic measurements is simple, because the laser is operated at a fixed laser energy level with a fixed number of shots to be accumulated. Once the measurement regions are selected, the measurement can start. The measurement can be triggered either for one sample directly from flexImaging or, if several samples are located on the same slide, by a batch-acquisition queue.

There are special considerations for mass spectrometry (MS)/MS imaging. Targeted analysis – such as small-molecule drug imaging – can be performed on time-of-flight (TOF)/TOF instruments; in this case, the only difference between TOF/TOF and an MS-only image is the fact that a parent mass must be selected as sequence property and an MS/MS autoXecute method must be chosen for the acquisition. It is recommended that this acquisition be undertaken in “fragments only” mode. The evaluation will be carried out later, on a characteristic fragment ion.

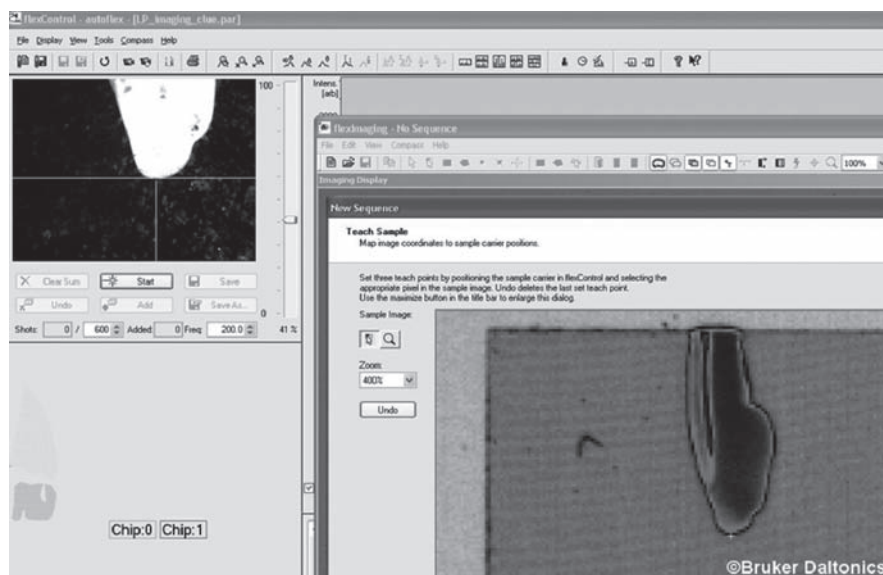


Fig. 15.1 Teaching during the setup of a run. The camera picture of the flex instrument is seen on the *left side* and the new sequence wizard of flexImaging with the optical primary image of the sample on the *right side*. The teaching is done by positioning the crosshair of the camera onto a specific point and then selecting the same point on the optical image. In this case, the reference point has been painted with Liquid Paper so that it is visible under the matrix layer

15.3 Evaluation of the Results

15.3.1 Distribution of Mass Signals

Once the measurement is finished, flexImaging automatically loads the imaging data. Also, during the acquisition, it is possible to load the already acquired data into flexImaging. flexImaging offers the option to read the raw spectra, the processed spectra, or peak lists. To load the processed spectra, it is recommended that baseline subtraction and smoothing be performed during acquisition.

By default, flexImaging displays an overall average spectrum for the dataset in the spectrum display. A mass marker can be used to select peaks in the overall spectra; at the same time, the intensity distribution of this signal will be shown in the image display as a color intensity (Fig. 15.2).

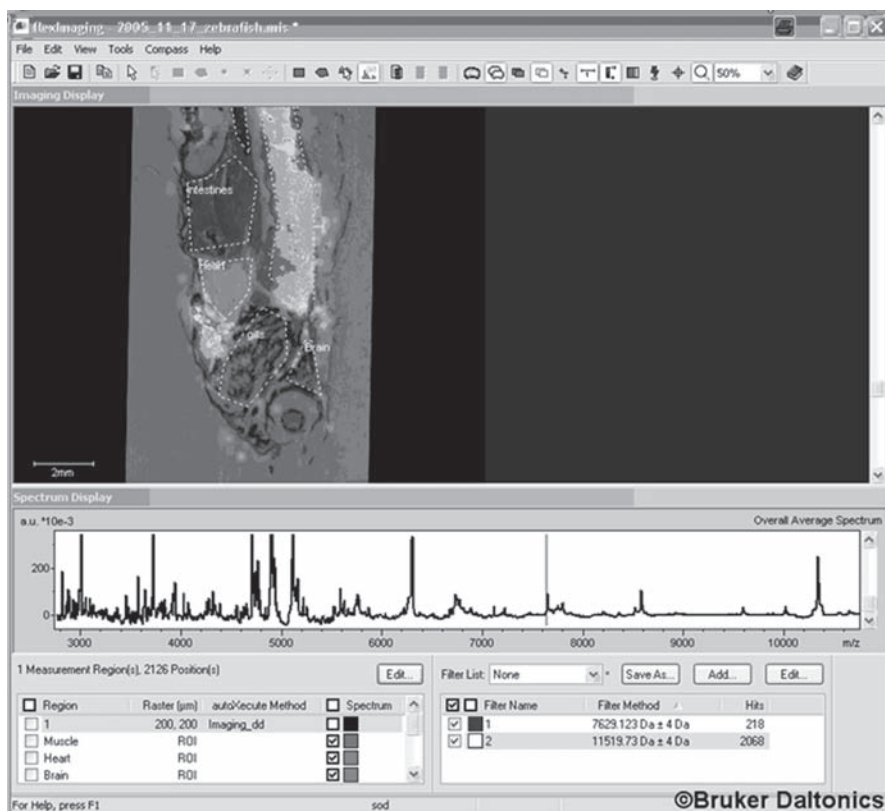


Fig. 15.2 Overview of flexImaging. In this case, an intensity map of the selected mass is displayed as an overlay over a co-registered, histologically stained whole-body section of a zebrafish. Different regions of interest have been defined on the zebrafish. The masses can be selected by clicking on the respective signal in the overall average spectrum as well as average spectra from different regions

The following information can be shown in the image display. The primary image can be switched on or off, a co-registered secondary image can be switched on or off, and the measurement region can be displayed; it is also possible to display regions of interest, the teach position, a scale bar, and the raster positions. In some cases, the primary image will show sufficient contrast to make the recognition of specific regions possible. For example, with brain sections, white or gray matter is easily recognized on an unstained section. In many cases, however, a detailed analysis of histology will require histological staining. For this reason, it is possible to co-register an image. This image could be a histological stain of a consecutive section, but it is also possible to remove the matrix layer after measurement with 70% ethanol and perform histological staining. In either case, this image can be co-registered in the “co-register image” dialog box (Fig. 15.3).

Here, the primary image is shown on the left side, with the image to be co-registered on the right side. The registration is done by selecting three corresponding spots on each image. It is now possible to overlay the histological image with the MS image. To get a clearer idea of the localization of the results, it is also possible to cross-fade between the two images. Usually, an image is not just analyzed by looking at the distribution of every single peak, because one wants to see characteristic differences between defined regions of the sample, for example, which peaks are present only in the tumor but not in the normal epithelium.

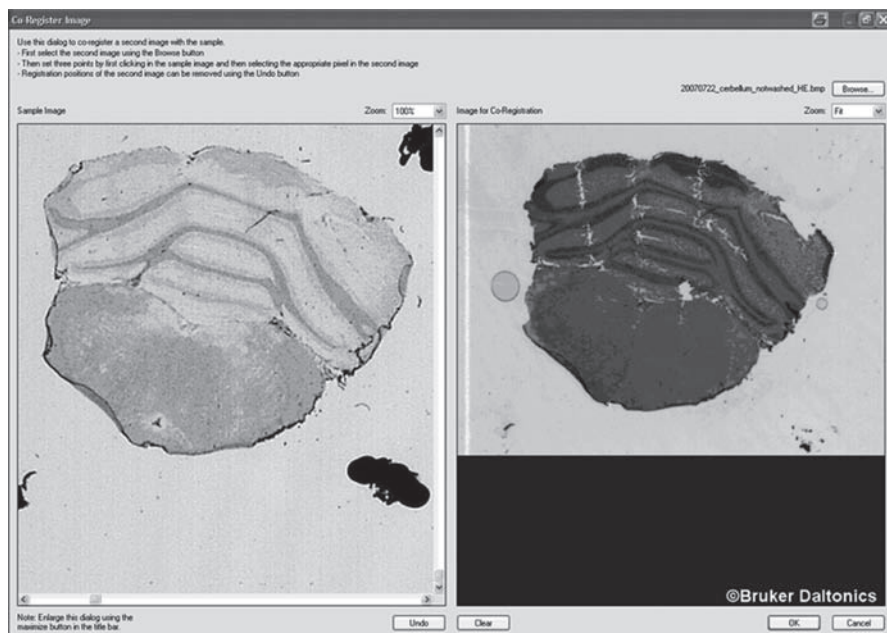


Fig. 15.3 The “co-register image dialogue.” The primary image is on the *left side*; on the *right side* is the histologically stained image that will be co-registered. The co-registration is done by selecting three corresponding spots on each image

In this case, one defines regions of interest. These regions can also be named, accordingly, “tumor” and “epithelium.” In the region pane in flexImaging, the average spectra for each region can be switched on via a checkbox. The average spectra of the selected regions can now be compared in the spectrum display, either as a pseudo-gel-like display or as an overlay of mass spectra traces. In this view, the differences between the regions become apparent. The peaks that are different can now be selected for display in the imaging display to confirm the localization.

Interesting signals can be added to a result filter list, for later reference. Each mass filter can be selected or deselected, so different masses can be overlaid. If two mass filters overlie, then a mixed color is created. For each mass filter, adjustments, such as the upper and lower intensity limits, can be selected.

15.4 Statistical Analysis

flexImaging interacts with ClinProTools for the statistical evaluation of data. The statistical analysis has two aspects: the analysis of one dataset and the comparison of different datasets.

15.4.1 *Statistical Analysis of Single Imaging Datasets*

15.4.1.1 **Principal Component Analysis**

Principal component analysis (PCA) is one way to reduce multidimensional data to lower dimensions while retaining as much information on the dataset as a whole as possible. In the context of MALDI imaging, this leads to a situation where similar spectra are displayed with similar color intensity. This is the preferable way of performing an unsupervised feature extraction, so a user of MS can generate images that reflect the overall variance of the dataset without having a complete understanding of the underlying histology. It can usually be assumed that the first principal components (PCs) will reflect the histology of the section, if a homogeneous sample preparation has been done.

Each mass spectrum with n -peaks can be considered a data point in an n -dimensional coordinate system. The PCA now transforms this coordinate system in such a way that the axis of the new coordinate system points in the direction of the largest variance in the dataset. The axes of the new coordinate system are called PCs. The first PC points in the direction of the largest variance in the dataset, the second PC points in the direction of the second-largest variance, and so on. The numerical value of a data point on a PC is called a score, and the score for each mass spectrometric data point can therefore be translated back to a color intensity on the MALDI image.

To perform PCA practically with ClinProTools, the raw spectra of the entire dataset must be loaded into ClinProTools. Once they are displayed, the calculation can be immediately initiated by clicking the “PCA” button. At this point, ClinProTools optionally aligns the spectra and performs the calculation. Once the calculation is finished, the result is displayed as a window with scores and loadings plots. More importantly, at the same time, a file is written that contains the results of the analysis; these results can now be imported into flexImaging by using the “add” button in the result filter list. A wizard then guides the user through the process. If individual PCs are selected, the rainbow mode is recommended for the display, because this will exaggerate the contrast on the image.

As a second option, it is also possible to import not the PC scores but rather the masses that contribute most to the variance in a dataset.

15.4.1.2 Hierarchical Clustering

Other unsupervised classification, such as hierarchical clustering, enhances the IMS data interpretation. Here, the underlying principle is that the more similar two spectra are, the closer they are in the multidimensional data space; classification is based on this distance. A hierarchical tree always contains all data points, and each branch in such a tree can be considered a class. The important question here is which classes are relevant; with flexImaging, this can be estimated in an interactive way. After the data were loaded into ClinProTools, a clustering calculation can be started by clicking on the “hierarchical cluster” button. Here, the option “create full tree” should be used. In the advanced options, it is possible to select different distance metrics. After the calculation is finished, the cluster is displayed; at the same time, the result of the clustering is written into a file. This file can now be imported into flexImaging using the “add” button in the result pane. A wizard guides the user through the process. In flexImaging, the top nodes of the tree are displayed in a tree browser. By selecting different branches on this tree, the respective data points on the image are highlighted. The tree is browsable, so it is possible to browse to the branch that contains the spectra specific for histological structures (Fig. 15.4).

15.4.1.3 Statistical Analysis on Many Datasets

Although the statistical analysis of single datasets is very useful in discovering the various features of a dataset, it will usually be necessary to compare different datasets. A common task would be to find markers that separate them – for example, one type of tumor from another type of tumor. In this case, it is necessary to generate MALDI images from a larger set of different patients first. Next, the regions that contain the tumor spectra need to be defined as regions of interest. The spectra for each region of interest can now be exported from a spectra list. It is necessary, for example, to group the spectra of tumor type one into one class and tumor type two into the second class; please refer to the ClinProTools manual on how to perform this.



Fig. 15.4 Display of hierarchical clustering result in flexImaging. Branches of the hierarchical tree can be selected in different colors, and the respective pixels in the image will be shown. In this example, the clustering is done on a rat kidney dataset

In ClinProTools, the two classes are then loaded. It is necessary to use the “support spectra grouping” switch in the data preparation settings of ClinProTools; each patient will then be considered one sample while each spectrum will then be considered a replicate of that sample.

ClinProTools does offer considerable information on the two classes. The average spectra for each class will be calculated and, for each peak, it is possible to display the average and standard deviation for each class. It is also possible to display a box-and-whiskers plot or a receiver operator characteristic for each peak. The spectra can also be visually compared in the pseudo-gel view. At this point, it is easy to look for those peaks that separate the two classes. Note that this “supervised” approach is very different from a PCA; with the supervised approach, the focus is set entirely on those peaks that separate the classes, whereas in a PCA, the focus is on variance in the dataset. Especially when samples from different patients are being compared, it can be expected that the overall variance in the dataset is a patient-to-patient variance; a PCA, therefore, would not separate the two classes but rather each patient from one another.

After the masses of interest are located as candidates, it becomes easy to validate findings on the imaging datasets by comparing that they are indeed specific, for example, to a cancer type in question.

15.4.2 Supervised Classification

As described in the previous section, a dataset can be used to generate a software model that can later be used to classify unknown spectra from different imaging runs.

Different algorithms can be used here: a support vector machine (SVM), a supervised neuronal network (SNN), a genetic algorithm (GA), or a quick classifier (QC).

The generated models can be used to classify unknown spectra. Here, it is useful to use the “classify in batch mode” option in ClinProTools, because an unlimited number of spectra can then be classified. The SVM, SNN, and GA algorithms will generate an output that will simply advise of the class to which a spectrum has been classified. With the QC algorithm, each spectrum can belong to all classes, each with a different weight.

After classification, ClinProTools asks the user where the classification results should be saved. These results can now be imported into flexImaging using the “add” button in the region pane; a wizard guides the user through the process. In the case of a SVM, SNN, or GA classification, the resultant image will display the classification, and each pixel in the imaging dataset will show the same color intensity. In the case of a QC classification, the weight value will be translated into a color intensity.