UPPSALA
UNIVERSITET

UPTEC X 15 038
Examensarbete 30 hp
Januari 2016

Automatic detection of protein degradation markers in mass spectrometry imaging

Stephanie Herman

UPPSALA
UNIVERSITET

Degree Project in Bioinformatics

Masters Programme in Molecular Biotechnology Engineering,
Uppsala University School of Engineering

UPTEC X 15 038 | **Date of issue 2016-01**

Author

Stephanie Herman

Title

Automatic detection of protein degradation markers in mass spectrometry imaging

Abstract

Today we are collecting a large amount of tissue samples to store for future studies of different health conditions, in hopes that the focus in health care will shift from treatments to early detection and prevention, by the use of biomarkers. To make sure that the storing of tissue is done in a reliable way, where the molecular profile of the samples are preserved, we first need to characterise how these changes occur. In this thesis, data from mice brains were collected using MALDI imaging mass spectrometry (IMS) and an analysis pipeline for robust MALDI IMS data handling and evaluation was implemented. The finished pipeline contains two reduction algorithms, catching images with interesting intensity features, while taking the spatial information into account, along with a robust similarity measurement, for measuring the degree of co-localisation. It also includes a clustering algorithm built upon the similarity measurement and an amino acid mass comparer, iteratively generating combinations of amino acids for further mass comparisons with mass differences between cluster members.
Availability: The source code is available at https://github.com/stephanieherman/thesis

Keywords

Proteolytic degradation, MALDI IMS, data reduction, spatial similarity, co-localisation

Supervisors

Mats Borén
Denator AB

Scientific reviewer

Andrew Palmer
EMBL

Project name -	Sponsors -
Language **English**	Security -
ISSN 1401-2138	Classification -
Supplementary bibliographical information -	Pages **56**

Biology Education Centre Biomedical Center Husargatan 3, Uppsala
Box 592, S-751 24 Uppsala Tel +46 (0)18 4710000 Fax +46 (0)18 471 4687

Automatic detection of protein degradation markers in mass spectrometry imaging

Stephanie Herman

Populärvetenskaplig sammanfattning

Idag lagras en mängd vävnadsprover från diverse diagnoserade patienter för att i framtiden kunna finna gemensamma nämnare för olika diagnoser. Förhoppningen är att kunna hitta unika molekyler som kan urskilja friska från sjuka. För att detta ska vara genomförbart krävs en robust lagringsmetod, där vävnadsprovernas molekylära profiler förblir oförändrade.

Även under kortare lagringsperioder som till exempel tidsintervall mellan provtagning och vävnadsanalys, krävs tillförlitliga metoder för att stabilisera den molekylära profilen, så att trovärdiga result kan produceras.

Syftet med detta examensarbete har varit att visualisera och konkretisera förändringar som sker efter provtagning, genom att hitta korrelerade peptider (kedjor av aminosyror) som genomgår degradering. När en peptid degraderas skapas peptidfragment som inte är närvarande i provets naturliga molekylära profil. Dessa peptidfragment antas skapas på samma position som dess föräldrapeptid och antas ha en massa som skiljer sig med ett antal aminosyror.

MALDI IMS (som står för matrix-assisted laser desorption/ionization imaging mass spectrometry) är en metod för att spatialt kartlägga ett vävnadsnitts molekylära profil, genom att generera masspektrum punktvis över hela vävnadssnittet. Mer specifikt genererar metoden en massbild för varje massvärde. Dessa bilder beskriver var molekyler är lokaliserade i vävnadsnittet.

I detta arbete har MALDI IMS använts för att generera massbilder från mushjärnor, varav hälften har värmestabiliserats för att förhindra degradering. Dessa bilder har sedan analyserats med en specialutvecklad analyspipeline, för att hitta massvärden som återfinns i samma områden av hjärnan och som skiljer sig med massan av en aminosyra. Dessa massvärden har sedan utvärderats för potentiell användning inom degraderingsindikation. Vilket är oerhört viktigt om vi ska etablera kliniska biobanker med tillförlitliga prover som framöver kan hjälpa oss att hitta biomarkörer för idag svåra diagnoser.

Acknowledgements

I would like to express special thanks to:

- My supervisor **Mats Borén**, for his support and guidance throughout the course of the project.

- My subject reader **Andrew Palmer**, for valuable discussions and inputs about MS data handling and algorithm design.

- **Malin Andersson**, for providing me with expert input and advise and for supervising me throughout the experimental data gathering.

- **Denator**, for the opportunity to do this project.

- **SCiLS**, for the use and support of SCiLS Lab 2014b.

Abbreviations

AA	Amino acid
CMC	Carboxymethyl cellulose
DR	Data reduction
ESI	Electrospray ionization
FTB	Fixed threshold based
HS	Heat stabilized
IMS	Imaging mass spectrometry
LC	Linear correlation
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
m/z	Mass-to-charge
NaN	Not-a-number
NCCN	National comprehensive cancer network
PCA	Principal component analysis
PC	Principal component
PM	Post mortem
RT	Room temperature
RTB	Ranged threshold based
SF	Snap frozen
TIC	Total ion count

Contents

Chapter 1

Introduction

Matrix-assisted laser desorption/ionization (MALDI) imaging mass spectrometry (IMS) is an emerging technology where one acquires mass spectra across tissue surfaces. The label-free and non prior knowledge setup makes it highly suitable for both explorative as well as comparative research of various tissue samples [1].

This powerful technology is highly used for proteome and peptidome imaging. By collecting mass spectra at discrete spatial points, one can generate ion images describing the localisation of biomolecules. With this information one can further map and study the 2D molecular profile of a tissue section.

Today we are collecting a large amount of tissue samples to store in clinical biobanks for future studies of different health conditions, in hopes that the focus in health care will shift from treatments to early detection and prevention, by the use of biomarkers.

Biomarkers is currently a hot topic in the clinical research community and remarkable progress has already been made which has initiated the start of a new era of health care. Personalized medicine and drug dosage optimization will soon be in our reach. Advances has for example been made in the cancer diagnostic field, where the national comprehensive cancer network (NCCN) reported tumor markers for six major malignancies [2, 3, 4].

However, a challenge this area is still facing is the struggle of preserving the true molecular profiles of the collected tissue samples. During tissue sampling, the fine tuned control of cellular processes and states, the homoeostasis, is severely disrupted. When a sample is removed from its natural environment a dramatic signaling cascade is initiated, which causes major shifts in the molecular profile. Highly active enzymes, triggered by the abnormal conditions, degrade proteins and peptides into smaller fragments, introducing new content in the molecular profile. When analysed, these tissues will thereby have a false profile, producing incorrect and deceiving results [5].

A recently developed method [the heat-based stabilization system from Denator] for preventing enzyme activity *post sampling* uses rapid and controlled heat denaturation of proteins and enzymes. The samples proteome will partly denature preventing any enzymatic activity. Therefore the true molecular profile of the sample will be preserved [6, 7].

All tissues are affected by *post sampling* changes, although enzyme rich tissues like pancreas, and tissues with low oxygen and energy storages, e.g. brain, are more affected and will benefit comparatively more from rapid heat denaturation, which was mainly the reason why brain tissue in particular was chosen as subject in this study [8, 9].

Assessment of tissue quality is currently performed by mass spectrometry experts using their previous experience and intuition to manually evaluate the tissues. This raises a large demand for experienced MS experts and since the MALDI IMS experiments are getting more automated and routinely used, this is not going to be a sustainable evaluation method in the future.

In this thesis, investigation of *post sampling* changes originating from degradation was assessed through automatic detection of protein degradation markers. The markers were assumed to be co-localised and differ in mass with the mass of a fixed set of amino acids. An automatic analysis pipeline was developed composed of a spatial similarity measurement, a customised density based clustering method and a cluster internal mass comparison algorithm. The output of this analysis pipeline fulfilled the criteria assumed of protein degradation markers. These markers can further be used to estimate the amount of proteolytic degradation that has occurred in a tissue. A standardized estimation that is not possible today.

Chapter 2

Background Theory

2.1 Mass Spectrometry

Mass spectrometry (MS) is an analytical chemistry technique that enables identification of the amount and type of molecules present in a samples, by measuring the mass-to-charge ratio (m/z) and abundance of gas phase ions. The mass spectrometer generates a mass spectrum, a plot of ion signals as a functions of mass-to-charge ratios. These plots will tell the masses and the abundances of the molecules present in a sample, which may be regarded as the elemental or isotopic signature of the sample [10].

2.2 Matrix-Assisted Laser Desorption/Ionization

Altough mass spectrometry has been a well used technique for many years, it can not be applied on macromolecules such as proteins and nucleic acids. In mass spectrometry the recording of m/z values are performed on molecules in gas phase and the heating or pre-treatment needed to transfer molecules to the gas phase is not compatible with macromolecules, since this will cause them to decompose [11]. In 1988 a technique called matrix-Assisted Laser Desorption/Ionization mass spectrometry (MALDI MS) was developed to overcome this problem. In MALDI MS the sample is mixed with a light-absorbing matrix and with a short pulse of laser light the macromolecules are ionized and desorbed from the matrix into gas phase [10, 11].

An alternative to MALDI MS is electrospray ionization mass spectrometry (ESI MS) where macromolecules in solution are forced directry from liquid to gas phase by using high electrical potential. A solution of analytes is passed through a charged needle that is kept at a high electrical potential, spraying a fine mist of charged microdroplets. The macromolecules will then be surrounded by these fine solvents which will rapidly evaporate, taking some of the macromolecules with them [10].

2.3 MALDI IMS

MALDI imaging mass spectrometry (IMS) is the use of MALDI as a imaging mass spectrometry technique, in which the sample is moved in two dimensions while the mass spectrum is recorded. The sample, often a thin tissue section coated with matrix, is analysed using a predefined pattern of coordinates where the laser hits. The laser ionizes some of the molecules present at the point of hit, which makes it possible to register their mass-to-charge values. The data is collected as lists of mass spectras, each of which has an associated coordinate. The final output is a three dimensional datacube with the coordinate information x and y on two axis and the m/z values on the third, see figure 2.1. The mass spectrum in a pixel represents the relative abundance of ionizable molecules with various m/z ratios. The m/z channels in the datacube represent maps of relative spatial abundance of molecular ions of a certain m/z value. These can be visualised as pseudocolored images also called m/z , ion or molecular images [11, 12, 16, 22].

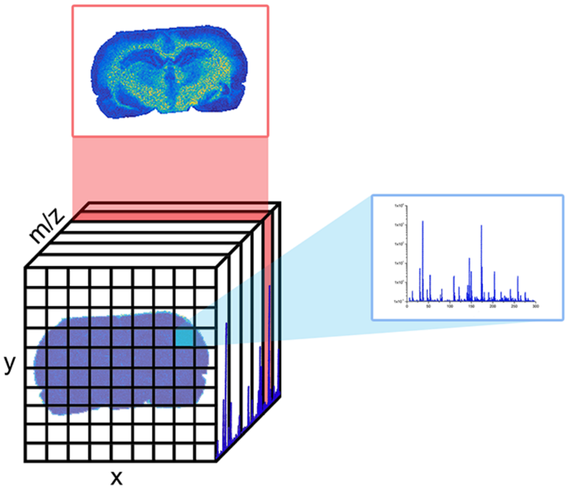


Figure 2.1: MALDI IMS generates a three dimensional datacube of position correlated mass spectras. In the first two dimensions, position coordinates x and y can be extracted. In the third dimension lie the recorded m/z values. The datacube can either be analysed in a pointwise manner where the corresponding mass spectrum is extracted from a specific pixel, as seen in light blue or a molecular image composed of intensity values for a specific m/z value can be obtained, seen in red.

The MALDI IMS data is typically very large, comprising 5000-50 000 pixels with 1000-100 000 m/z channels and can be as large as several gigabytes. This makes the data computationally heavy to analyse, even with sophisticated computational methods. Today the IMS data are analysed either manually, by looking at the images or mass spectras, or in a data mining manner, where the data of interest is extracted from the datacube. However, as IMS evolves into a routinely practice, manual analysis will become infeasible and as the data collections grow larger the need for effective data evaluation tools will increase.

2.4 Matrix Application

In MALDI IMS the molecular content in a thin tissue section is ionized by applying a matrix. This application can be done and optimised according to your needs. If one value image resolution, application and ionization using ESI might be the best option, since the matrix then will be applied as a uniform layer. While if one values ionization depth (in terms of how deep the matrix will penetrate and ionize the molecular content within the tissue) a better approach might be application by printing, since more matrix per area is added. That is, the smaller the point to point distances, the more visually correct images you achieve. But the higher ionization efficiency in each point, the larger proportion of the molecular content will be ionized. It is also argued that an additional disadvantage with application and ionization with ESI is that diffusion in the molecular content might occur throughout the tissue [11, 13, 22].

When performing IMS analysis this is a trade-off that has to be considered. Unfortunately, with todays technologies one can not reach perfection in all aspects.

2.5 Data Reduction Strategies

As previously stated, MALDI IMS generates a large amount of data which is extremely heavy to analyse and the need for a data reduction step is obvious, although the way of reduction might not be as obvious.

One strategy, used by J.M. Fonville *et al.* is to import the data with a bin size, merging m/z images inside an interval of choice [26] or representing the interval by the average m/z value and the maximum intensity image, as C.D. Wijetunge *et al.* [24]. The danger with unifying images without studying the images spatial information, is that images that originate from two different biomolecules might get merged. Even though it is rare that biomolecules m/z peaks overlap, it does occur. During the pre-analysis of the datacubes in this thesis, several cases of overlapping were seen. Another approach, similar to the bin size strategy is to define a list of peaks of interest to reduce the size of datacubes. This approach was taken by L.A. McDonnell *et al.* where they reduced approximately 72 0000 m/z images to around 100 peak images [27]. A mix of these strategies was also used by A. Palmer *et al.* where they generated images for every peak in a list containing peaks of interest, with a summation window of 10 Da (focusing on a m/z region between 2500 and 10 000) [23]. In both of these methods the risk of merging images of two different biomolecules remains, since the images spatial information is not considered.

In this thesis a new strategy for data reduction was developed, in an aim to overcome incorrect image merging and optimisation of kept images. Using the new strategy the data was reduced based on two properties of interest; images that corresponds to the local maximums within the m/z summation spectrum and images that feature at least one pixel that exceed ten times the intensity average within the image, while taking the spatial information into account.

2.6 Proteolytic Degradation

When tissue samples are removed from their natural *in vivo* environment, the homeostasis is brutally disrupted. A dramatic signaling cascade is initiated where the proteolytic activity is increased, which inflicts changes in the molecular profile of the sample [5]. These changes will be devastating for biomarker discorvery, explorative analysis and other profile studies, since an incorrect profile will be recorded.

Proteins and peptides, constructing the proteome and peptidome of tissue samples, consists of sequences of amino acids folded into an active shape. When the enzymatic activity increases within a sample, during and after sampling, the enzymes present cleaves these chains of amino acids into smaller fragments and thereby reducing the true biomolecules within the sample while introducing new ”false” molecules.

To detect degradation one can assume that proteins and peptides correlated through degradation 1) have the same localisation within the tissue (reasoning that the peptide fragments should be created at the same location as its parent peptide) and 2) differ with a mass corresponding to a fixed set of amino acids.

To prevent proteolytic degradation *post sampling*, Dentor has developed a heat-based stabilization system, the Stabilizor system, that partly denatures all proteins within the sample using heat [6]. The proteins unfold into an inactive and random, but stable configuration. By deactivating all proteins, the enzymes causing the degradation will simultaneously be deactivated. If the enzymes are deactivated, the degradation will be terminated and the molecular profile will be saved from further changes. An additional advantage with this strategy is that the proteins will still be intact in terms of mass, since the change is only affecting the shape of the proteins. [5, 7].

2.7 Biomarkers

A biomarker or a molecular marker is a distinct molecule or substance that is an indicator of a particular biological condition or process. In for example health care, a biomarker could be an indicator of a disease in early stage. The biomarker could enable early detection and thereby early treatment with better prognosis. In proteomics, proteins can serve as biomarkers in which their presence and abundance reveal the physiological basis of health and disease [18]. In our case the markers of degradation are assumed to be pairs of peptides that differ in mass with a fixed set of amino acids and have the same localisation wihtin the tissue.

Chapter 3

Aim

The aim of this thesis was to characterise the *post sampling* changes that occur when tissue is removed from the living and kept at room temperature and to validate the benefits of heat stabilization (HS). One or several potential markers for *post mortem* degradation were hoped to be found, which could be further used to evaluate the effectiveness of heat stabilization.

For an easier overview, the aim could be broken down into a set of substeps (illustrated in figure 3.1);

- Gather data from mouse brain using MALDI IMS
- Implement and evaluate subparts of analysis pipeline:
  - **Spatial similarity measurement**
    A similarity measurement which not only compares images based on spatial localisation but also solves the problematics of spatial signal detection.
  - **Customised clustering algorithm**
    A clustering algorithm which group images into co-localised and potentially correlated markers based on the similarity measurement.
  - **Amino acid mass matcher**
    A comparison algorithm that generates unique combinations of amino acids and compare these with the mass differences of all unique mass pairs within a cluster.

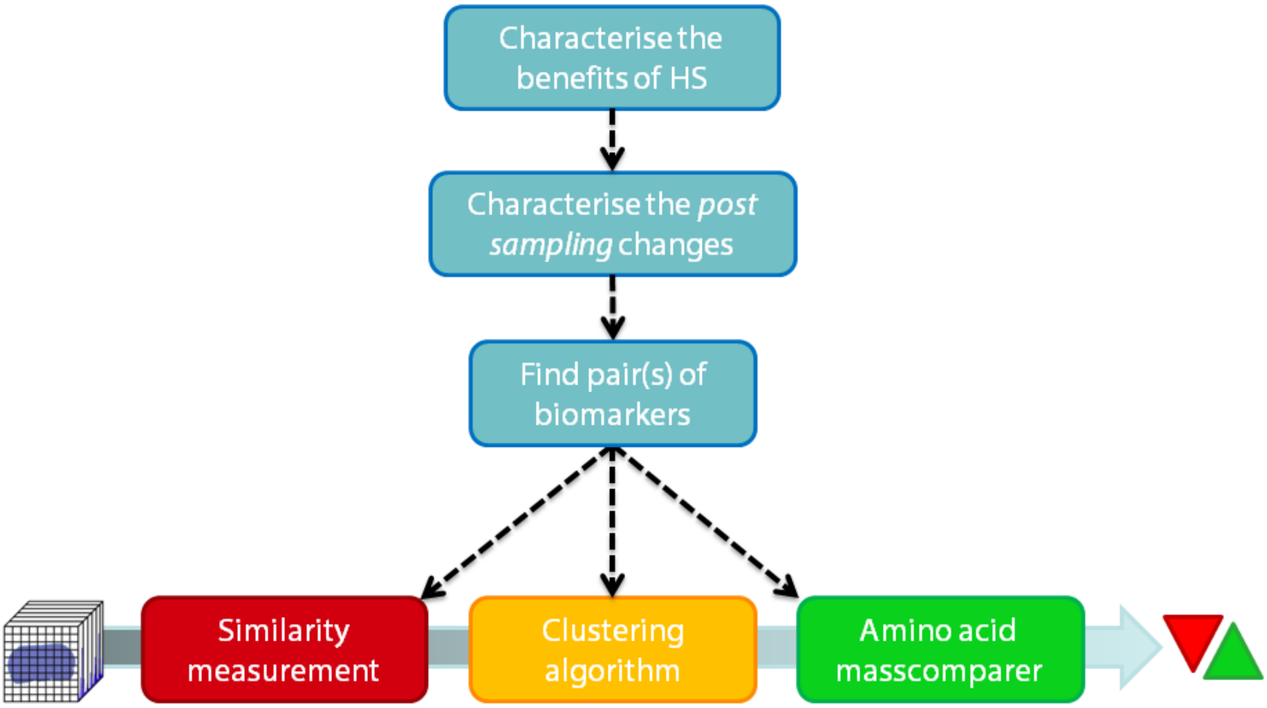

Characterise the benefits of HS
Characterise the *post sampling* changes
Find pair(s) of biomarkers
Similarity measurement
Clustering algorithm
Amino acid masscomparer

Figure 3.1: The aim of the project broken up into subgoals.

Chapter 4

Implementation and Algorithm Design

4.1 Pre-processing

Before applying the algorithms on the data, several steps of pre-processing have to be done. Baseline correction, to remove background noise, is done directly when recording the spectra. Spectral normalization is done in SCiLS Lab 2014b using the Root Mean Square method [19]. Automatic hot spot removal using quantile thresholding with a quantile value of 0.99 is done in Matlab using the predefined function *quantile* and scaling by the images maximum value is done. The effect of the hotspot removal and scaling can be seen in figure 4.1. As a final step the Not-a-number (NaN) values outside the tissues were replaced with zeros.

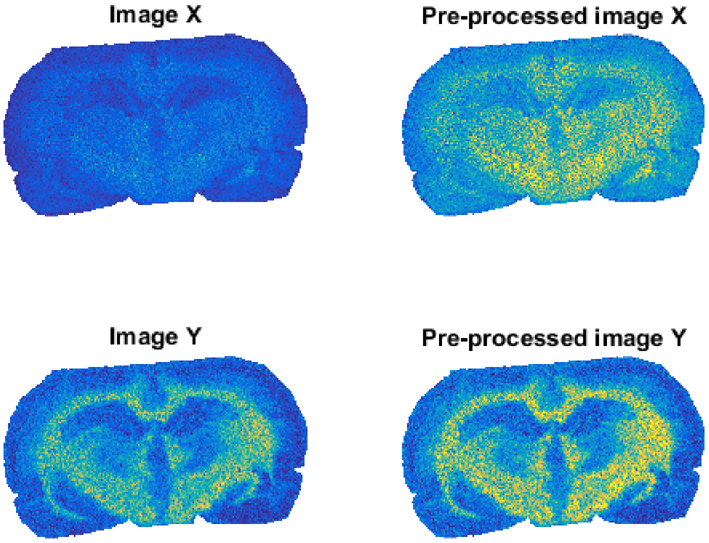


Figure 4.1: Two images with varying intensities before and after hot spot removal and normalization. After pre-processing the intensities within the images lie between zero and one.

4.2 Data Reduction

4.2.1 Justification

As previously stated, MALDI IMS data is extremely large. With a dataset that can hold up to 100 000 images a combinatorial comparison, comparing all possible image pairs, would require 1.25 billion comparisons to be made, which would scales exponentially. Using the suggested similarity measurement to score all pair of images in a dataset containing only 150 images (which requires 1100 unique comparisons) with 1100 pixels was timed to 2 min and 22 secs (Intel(R) Core(TM) i5-2500K CPU @ 3.30 GHz 3.60 GHz, 16,0 GB RAM). This made it clear that it would be completely infeasible to score all images in a raw datacube. A pre-filtering step, where non-interesting images (containing only background noise) and redundant images (images from the same m/z peak, with the same spatial distribution) are removed, is vital to reduce the datacube into a computational manageable size.

4.2.2 Filtering Conditions

To be able to remove non-interesting images, a rough definition of an interesting image had to be made. After exploring the neuropeptide Dynorphin B, a peptide present only in a small part of the mouse brain (its image seen in figure 4.2) which is of interest, a requirement of having at least one pixel that has an amplitude that exceed ten times the mean intensity in the image was determined (this requirement was specifically determined for the specific MALDI IMS settings/image resolution used in this study). If this is not fulfilled, the image is seen as non-interesting and should thereby be removed from the dataset.

A second group of interesting images was defined by examining the full datasets mean spectrum. The mean spectrum is the mean of all spectra from all pixels in the dataset. As can be seen in figure 4.3 there are several peaks throughout the m/z axis. Images of special interest are those that constitute the exact top of a peak. These can be found by calculating the spatial total ion count (TIC) for all images. The images representing the peak tops will be the local maximums of the total ion intensities.

As previously mentioned, there can exist images in a peak that have different localizations. The algorithm must therefore take the images spatial information into account when assigning which images the peak top image should represent. For this, the chosen similarity measurement score, explained in section 4.2.3 was used.

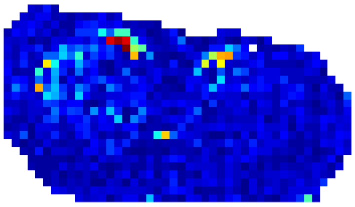


Figure 4.2: The image was generated in SCiLS Lab 2014 and shows the summed intensities for the m/z value 1571 ± 0.125%, which is the m/z value for the neuropeptide Dynorphin B.

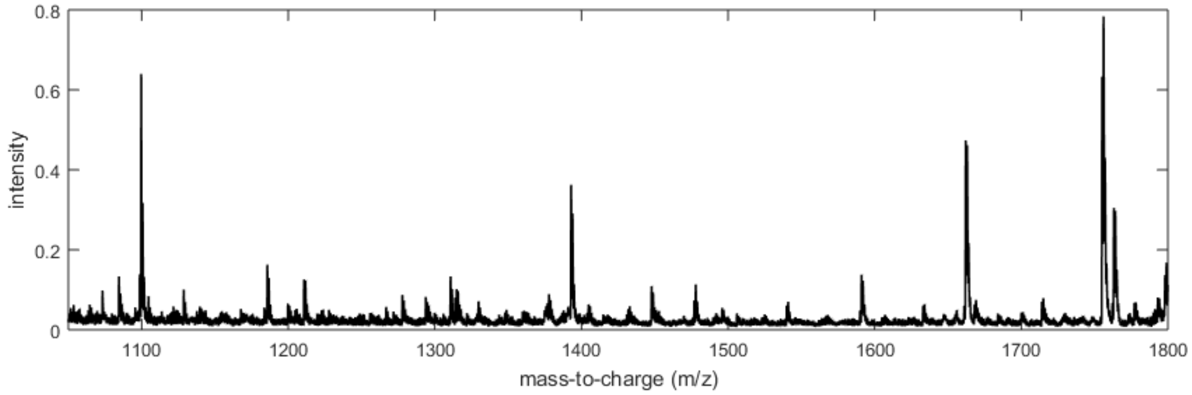


Figure 4.3: A part of the mean spectrum of mouse brain.

4.2.3 Local Maximum Detection by Spatial TIC

The data reduction is done in two steps, in which the initial aim is to save images that have a high spatial total ion count. Since the amplitude in the mean spectra of the dataset vary a lot throughout the m/z axis, the mean spectrum is handled in intervals of a sliding window of 20 Da (approximately half the size of the smallest amino acid). For each interval the following is done.

```
input  : Datacube $D$, TIC cutoff a, similarity cutoff b
output: Reduced datacube $R$
begin
    for each window do
        for each image in window do
            the spatial $TIC$ is calculated and stored in vector $\mathbf{t}$
        end
        if $TIC_{max} > a \times median( \mathbf{t} )$ then
            collect images where $TIC > a \times median( \mathbf{t} )$ in $R$
            create RTB similarity matrix $s$ for $R$
            if any value in $s \geq b$ then
                image groups that acquire a score ≥ b are removed from $R$ and
                represented by the image with the highest $TIC$
            end
        else
            images are regarded as background noise
        end
    end
end
```

4.2.4 Spatial Peak Detection

In the second data reduction step the aim was to find the images that have at least one peak that exceeds the value of ten times the peak intensity average in that image, thought of as a rough spatial peak detection. This is done as follows.

```
input  : Datacube $D$, intensity cutoff a, similarity cutoff b
output: Reduced datacube $R$
begin
    for each image in datacube $D$ $j \leftarrow 1$ to $J$ do
        mean intensity m is computed
        $peaks = \sum pixels > a \times m$
        if $peaks > 1$ then
            image is stored in $R$
        end
    end
    create RTB similarity matrix $s$ for $R$
    if any value in $s \geq b$ and their $\Delta m/z <$ the smallest mass of an aa then
        image groups that acquire a score $\geq$ b are removed from $R$ and represented by
        the image with the highest spatial $TIC$
    end
end
```

The two datacubes, regarded as R in the algorithms, are then combined and sorted in ascending m/z order. Any images present in both of the datacubes are detected and one of them removed for duplicate prevention.

4.3 Spatial Similarity Measurement

A challenging requirement for a co-localisation measurement, is the ability to discern between signals that are actual peaks and signals that arise from background noise. To deal with this, a fixed- and a ranged based threshold was evaluated for spatial signal detection. These were further compared with linear correlation, as a candidate measurement previously used by L.A. McDonnell *et al.* [27].

4.3.1 Fixed Threshold Based (FTB)

The peaks in two images were discerned from the background noise using a fixed threshold, the median of all intensities in the image. The shared peaks between the images were found by computing the intersection between the images peaks. To minimize the peak area size dependency, the score was normalized by division with the mean of the sum of peak areas in both images, see formula 4.1, where P is an image transformed into a binary vector; one for peak and zero for no peak.

$$Score_{(P_i,P_j)} = \frac{2 \times \sum_{i=1}^{n} (P_i \cap P_j)}{\sum_{i=1}^{n} P_i + \sum_{i=1}^{n} P_j} \qquad P = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ . \\ . \end{bmatrix} \qquad where\ P_k = \begin{cases} 1 & \text{if } P_k > threshold \\ 0 & \text{if } P_k \leq threshold \end{cases} \tag{4.1}$$

4.3.2 Ranged Threshold Based (RTB)

Instead of using a fixed threshold, a level-based threshold ranging from 0 to 1 with a step size of 0.02 was implemented, figure 4.4. The peaks in two images were detected for all levels and the intersection between the peaks were computed level-wise and normalized for peak area. The same algorithm as in FTB was used, formula 4.1, iteratively to generate one score per level.To compute a final score, the average of all level-wise scores was computed.

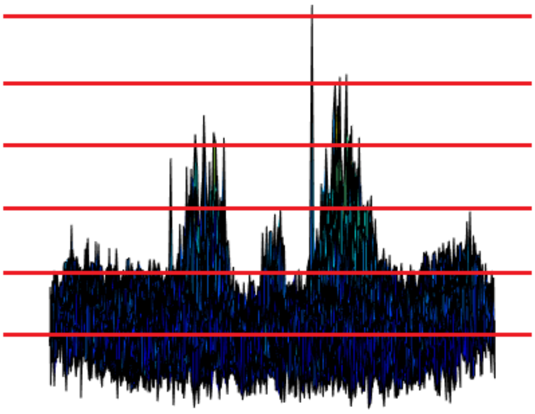


Figure 4.4: An image of a rat brain seen from the side. The signals over the tissue surface vary a lot in amplitude, which makes it hard to distinguish between actual peaks and background noise. To solve this issue, a level-based threshold was proposed. The image pairs are then scored several time, with an increasing threshold. The final co-localisation score is then computed by taking the average of all level-wise scores.

4.3.3 Linear Correlation (LC)

The correlation coefficient, R, was computed with the *coffcoef* function in Matlab and used as a co-localisation score between two images. The correlation coefficient is determined by formula 4.2, where C is the covariance of the two vectorised images A and B.

$$R(A,B) = \frac{C(A,B)}{\sqrt{C(A,A) \cdot C(B,B)}} \tag{4.2}$$

4.4 Clustering

After settling with the RTB similarity measurement, the next step was to implement a clustering algorithm that based its grouping on the RTB similarity matrix. The implemented clustering algorithm was inspired by the density based clustering algorithm introduced by Nanda *et al.* [20]. The clustering algorithm described by Nanda *et al.* comprised a similarity matrix generation step, which in this study was completely removed from the clustering session. Instead the similarity matrix was generated separately using the introduced RTB similarity measurement scoring algorithm, which was set as input for the clustering algorithm. A minor modification was also done in the threshold for merging clusters, in order to account for all unique pairs of cluster members.

4.4.1 Clustering Algorithm

Step 1 **Initial cluster creation**

```
input  : RTB similarity matrix $S$ and similarity threshold $t$
output : Vector $\mathbf{c}$ with initial cluster assignments
self-correlation values are replaced with zeros
the maximum value $S_{1max}$ in the first row of the similarity matrix $S$ is determined
if $S_{1max} > t$ then
    samples that achieve $S_{1max}$ is included in cluster $C_1$
end
for the rest of the samples $k \leftarrow 2$ to $K$ do
    if sample k already is included in a cluster then
        carry on to next sample
    else
        compute the maximum value $S_{kmax}$ in row $k$
        if $S_{kmax} > t$ then
            all samples which achieve $S_{kmax}$ are included in cluster $C_n$
        end
    end
end
```

Step 2 **Merging of initial clusters**

A similarity matrix of inter cluster RTB similarities is created by the similarity RTB scores between all possible combinations of members in $C_m = (c_1, c_2, \;, c_M)$ and $C_n = (d_1, d_2, \;, d_N)$. The inter cluster similarity score $\sigma_{m,n}$ is defined by the mean of all the RTB similarity scores $s_{c,d}$ (retrieved from the original RTB similarity matrix) between the members of the two different clusters, formula 4.3.

$$\sigma_{m,n}(C_m, C_n) = \frac{\sum_{i=1}^{M} \sum_{j=1}^{N} s_{c_i d_j}}{M \cdot N} \tag{4.3}$$

In the diagonal of the inter cluster similarity matrix, the self-correlation values are set to zero.

$$| \sigma_{m,n} |= 0, \forall m = n \tag{4.4}$$

The clusters are then merged using the following algorithm;

```
input  : Inter cluster similarity matrix $\sigma$, vector c and similarity threshold $t$
output: Vector c with updated cluster assignments
for each row in the inter cluster similarity matrix $\sigma$ $i \leftarrow 1$ to $I$ do
    the maximum value $\sigma_{imax}$ is computed
    if $\sigma_{imax} > t$ then
        clusters that achieve $\sigma_{imax}$ are merged into one cluster
    end
end
```

Step 3 **Termination condition**
Step 2 is repeated until no further merging can be done, i.e. when the difference between the number of clusters in the end of two computations of step 2 is zero. Any cluster containing only one member is removed from the vector **c**.

4.5 Amino Acid Mass Comparer

The aim of this thesis was to find correlated markers of protein degradation. Each cluster retrieved from the clustering algorithm holds images that have similar spatial localisation and thereby are potentially correlated. Since one of the markers in a pair should theoretically be a degradation product of the other, the difference in mass between the two markers should match a fixed combination of amino acids. Therefore, to find such pairs, an amino acid combinator was created, which generates all unique combinations of one or several amino acids. The masses of these combinations were then to be compared with the delta m/z values of all unique pairs of images for each computed cluster.

4.5.1 Algorithm

The m/z values corresponding to the members of the cluster of interest were extracted and the absolute value of the mass difference ($\Delta m/z$) of each unique m/z pair is calculated. A customised function called the *AAcombinator* generates all unique combinations of two input vectors holding single amino acids or/and combinations of amino acids, by using a combinatorial approach. By iteration, using the outputs of one computation as the input of the next, combinations of several amino acids can be generated. These combinations can further be compared with the $\Delta m/z$ values. The comparison is done with a function called the *AAmatcher*. It allows a m/z error of ±0.1 and can be manipulated to either output the hits in Matlabs command window or as a .csv-file.

4.5.2 Amino Acid Masses

The monoisotopic masses of the amino acids used in by the Amino acid mass comparer can be seen in below in table 4.1. To reduce redundant comparisons, amino acids with equal masses were treated as one in the comparison step.

Table 4.1: The monoisotopic masses of the amino acids used in this thesis.

Name	Short	Mass
Alanine	A	71.03711
Arginine	R	156.10111
Asparagine	N	114.04293
Aspartic Acid	D	115.02694
Cysteine	C	103.00919
Glutamic Acid	E	129.04259
Glutamine	Q	128.05858
Glycine	G	57.02146
Histidine	H	137.05891
Isoleucine	I	113.08406

Name	Short	Mass
Leucine	L	113.08406
Lysine	K	128.09496
Methionine	M	131.04049
Phenylalanine	F	147.06841
Proline	P	97.05276
Serine	S	87.03203
Threonine	T	101.04768
Tryptophan	W	186.07931
Tyrosine	Y	163.06333
Valine	V	99.06841

Chapter 5

Evaluation and Validation

5.1 Spatial Similarity Measurement

In order to decide which of the three spatial similarity measurements that was the most suitable candidate which gave the most desirable results, an evaluation procedure had to be done. Several testsets had to be constructed in order to test their robustness towards structurally different image and complete randomised noise.

5.1.1 Testsets for Measurement Evaluation

To evaluate the three candidate scoring algorithms several testsets were created. The first testset, figure 5.1, consists of 15 pairs of images that have co-localised expression. These pairs

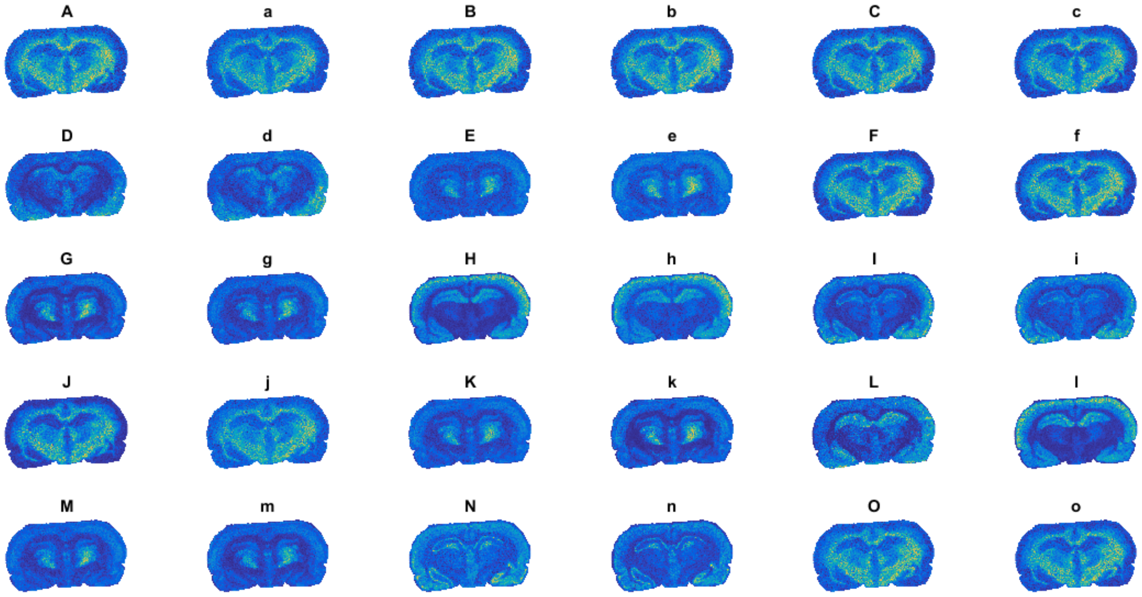


Figure 5.1: Testset 1 is composed of 15 co-localised pairs of images, hand-picked from a MALDI IMS dataset from rat brain tissue.

were hand-picked from the gold standard dataset created by A. Palmer *et al.*, which originated from rat brain [23] and were expected to get high co-localisation scores by the similarity measurements.

The second testset was composed of the first member of each pair in testset 1, paired with one of five noise images, figure 5.2, that were manually extracted from the same MALDI IMS dataset. The pairs in testset 2 were expected to get lower scores than those in testset 1.

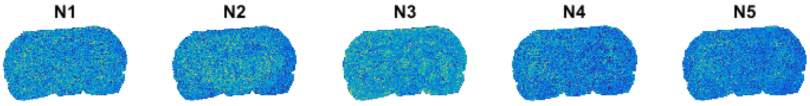


Figure 5.2: The five noise images used in testset 2.

A third testset was formed by pairing images from testset 1 based on delocalisation. This testset was created to evaluate how good the algorithms handle structurally different images, i.e. images that have structured patterns but are not co-localised.

5.1.2 Scoring of Testsets

The scorings for each testset were calculated, seen in table 5.1 and 5.2a. According to the results, all algorithms generated relatively stable and high scores (RTB and FTB scores range between 0 and 1 and LC scores between -1 and 1) for the pairs in testset 1, except from one outlier (pair Ll). It was also shown that the RTB algorithm generated lower scores for the pairs in testset 2, than the FTB algorithm. This indicates that using a ranging threshold is more efficient for distinguishing noise images than using a fixed one. The LC algorithm also showed lower scores for the pairs in testset 2. However, these corresponding scores (ranging from 0.0749 to 0.4866) were not as stable as FTB:s and RTB:s. Another difference worth mentioning with LC is that its scale is different (ranging from -0.2 to 1) from FTB:s and RTB:s (whose scales range from 0 to 1), which are percentage based and therefore more intuitive. It should also be noted that many of the pairs in testset 2 received higher scores than the delocalized pairs in testset 3, from all of the scoring algorithms. Indicating that the scoring algorithms score delocalised structured-to-structured image pairs (testset 3) higher than structured-to-noise image pairs (testset 2).

To quantify the ability to discern between pairs in testset 1 from pairs in testset 2, an average of the ratios were computed. As can be seen in table 5.1, the LC had the highest ratio, followed by RTB and FTB. A high average ratio indicates that the algorithm can distinguish between co-localised image pairs and structured-to-noise image pairs. In table 5.2b, the top 15 highest scored pairs (comparing the pairs in testset 1 and testset 2) are listed for each algorithm. As seen, all algorithms scored all the image pairs in testset 1 as the most co-localised pairs, which indicate that all these three candidate algorithms have the ability to find co-localised images.

Table 5.1: Computed FTB-, RTB- and LC-scores for all pairs in testset 1 (bold) and 2.

Pair	**FTB**	**RTB**	**LC**
Aa/AN1	**0.8466**/0.5235	**0.7443**/0.3994	**0.9073**/0.3677
Bb/BN2	**0.8882**/0.5994	**0.8428**/0.4366	**0.9558**/0.4866
Cc/CN3	**0.8941**/0.4791	**0.8401**/0.4060	**0.9566**/0.3612
Dd/DN4	**0.8721**/0.3906	**0.7155**/0.2826	**0.9131**/0.1260
Ee/EN5	**0.7447**/0.4520	**0.7365**/0.3033	**0.8620**/0.2941
Ff/FN1	**0.8978**/0.5177	**0.8336**/0.3812	**0.9548**/0.3356
Gg/GN2	**0.8664**/0.4256	**0.8325**/0.2817	**0.9415**/0.2368
Hh/HN3	**0.9371**/0.4897	**0.8433**/0.3335	**0.9712**/0.3136
Ii/IN4	**0.8982**/0.4164	**0.8014**/0.2803	**0.9452**/0.1601
Jj/JN5	**0.8782**/0.5372	**0.7959**/0.3487	**0.9369**/0.3202
Kk/KN1	**0.8724**/0.4644	**0.8625**/0.3105	**0.9546**/0.3311
Ll/LN2	**0.8029**/0.3676	**0.4683**/0.2829	**0.6873**/0.0749
Mm/MN3	**0.8572**/0.5101	**0.8605**/0.3350	**0.9475**/0.4334
Nn/NN4	**0.8390**/0.4272	**0.7971**/0.3166	**0.9260**/0.2215
Oo/ON5	**0.8875**/0.5492	**0.8467**/0.3806	**0.9567**/0.3831
Average ratio:	1.8430	2.3587	3.7953
Standard deviation:	0.2358	0.3756	2.0689

Table 5.2: a. Computed FTB-, RTB- and LC-scores for delocalized pairs (testset 3).

b. Top 10 highest scored pairs with the FTB, RTB and LC algorithm.

(a)

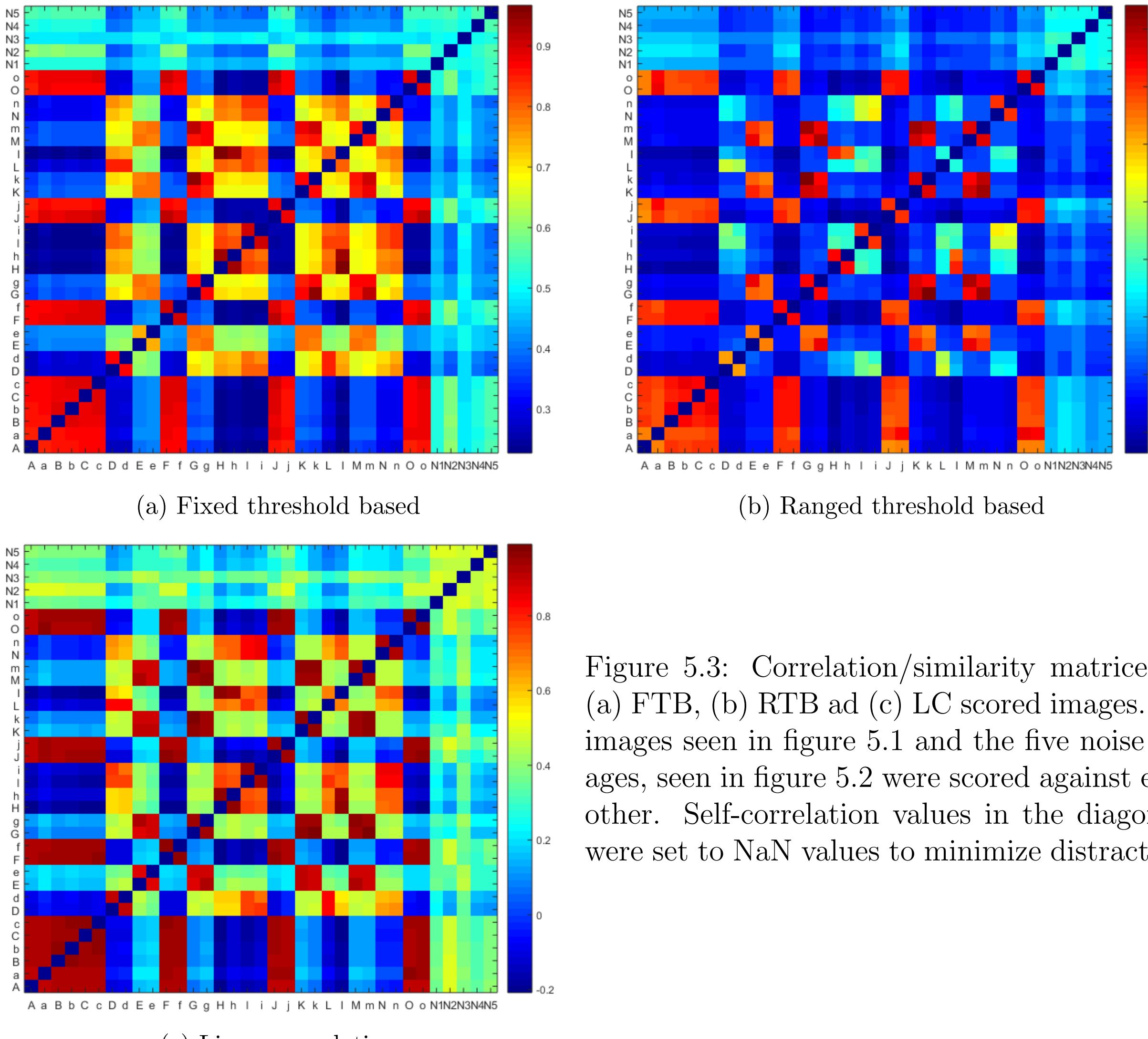
Pair	**FTB**	**RTB**	**LC**
AD	0.2704	0.2233	-0.1004
MN	0.6532	0.3130	0.4492
BE	0.4048	0.2649	0.1656
FL	0.2900	0.2050	-0.1216
KJ	0.1585	0.2444	0.1585
CH	0.2347	0.1795	-0.2007
GO	0.3806	0.2404	0.1227
Ia	0.2492	0.2105	-0.1128
kc	0.3732	0.2231	0.0801
lo	0.2382	0.1802	-0.1895

(b)

Top	**FTB**	**RTB**	**LC**
1	Hh	Kk	Hh
2	Ii	Mm	Oo
3	Ff	Oo	Cc
4	Cc	Hh	Bb
5	Bb	Bb	Ff
6	Oo	Cc	Kk
7	Jj	Ff	Mm
8	Kk	Gg	Ii
9	Dd	Ii	Gg
10	Gg	Nn	Jj

5.1.3 Similarity Matrices

To further explore the distinguishing power of each algorithm, a correlation/similarity matrix was computed and visualised for each algorithm, figure 5.3. These matrices show the scores (FTB, RTB and LC) for all unique pairs in testset 1 combined with the five noise images seen in figure 5.2. When examined, several interesting features can be seen. Both the similarity matrices of FTB and LC have an area in the middle of images that receive slightly higher scores than their surroundings, which RTB does not have. RTB seems to score the image pairs with less scoring value distribution. It generates either high or low scores, few in between values are distributed. While FTB and LC seem to use their full scoring scale.



(a) Fixed threshold based

(b) Ranged threshold based

(c) Linear correlation

Figure 5.3: Correlation/similarity matrices of (a) FTB, (b) RTB ad (c) LC scored images. All images seen in figure 5.1 and the five noise images, seen in figure 5.2 were scored against each other. Self-correlation values in the diagonals were set to NaN values to minimize distraction.

To study the algorithms scoring behaviors further, two images (image A and G) were chosen as fixed. Their rows in each similarity matrix were extracted and sorted in descending order, visualised in figure 5.4, 5.5 and 5.6 starting with the fixed image. The corresponding co-localisation score and pseudoname are shown above each image.

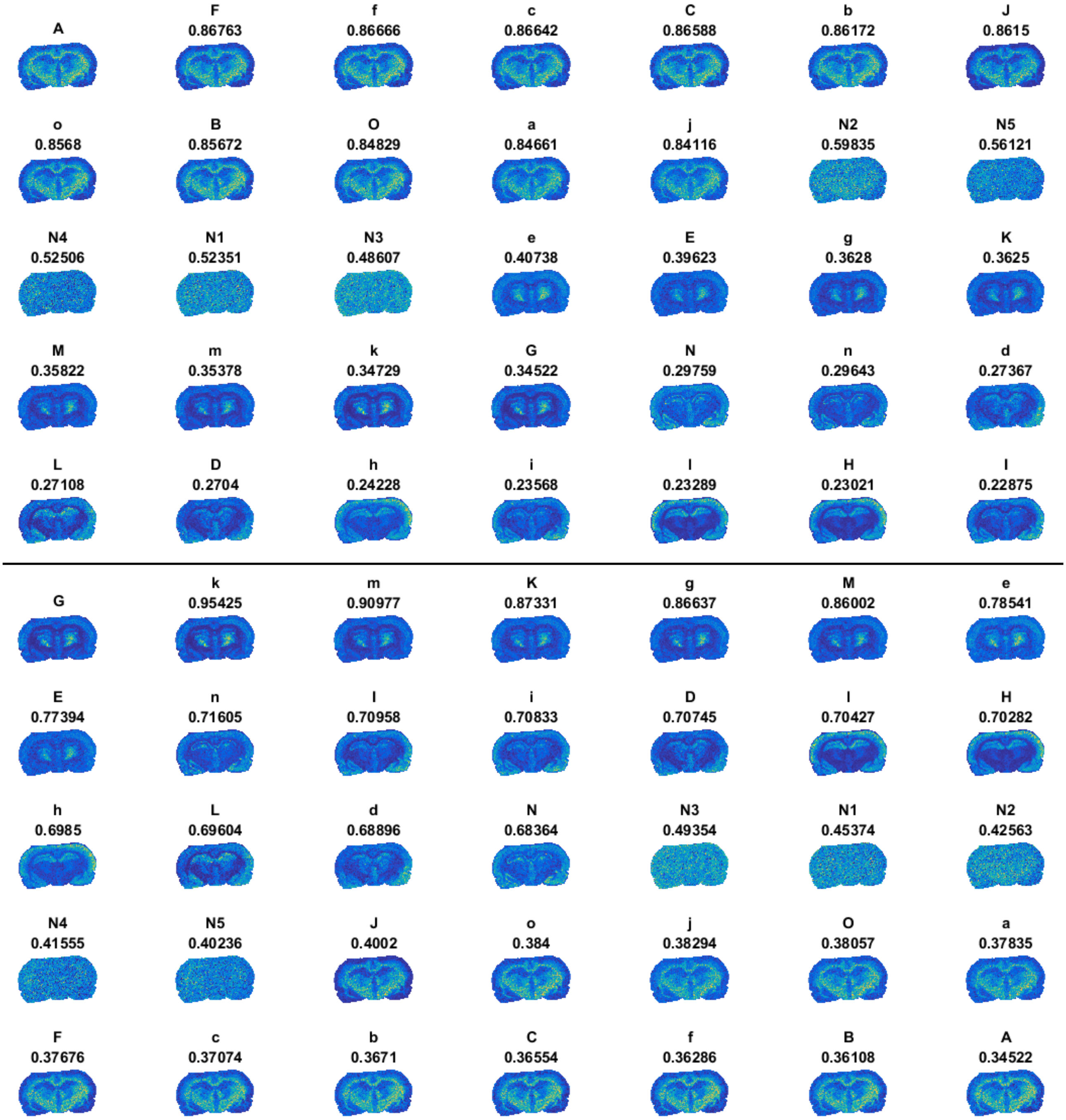


Figure 5.4: Image A and image G compared with all other images, showed in FTB scored descending order. The corresponding FTB scores are stated under the image name.

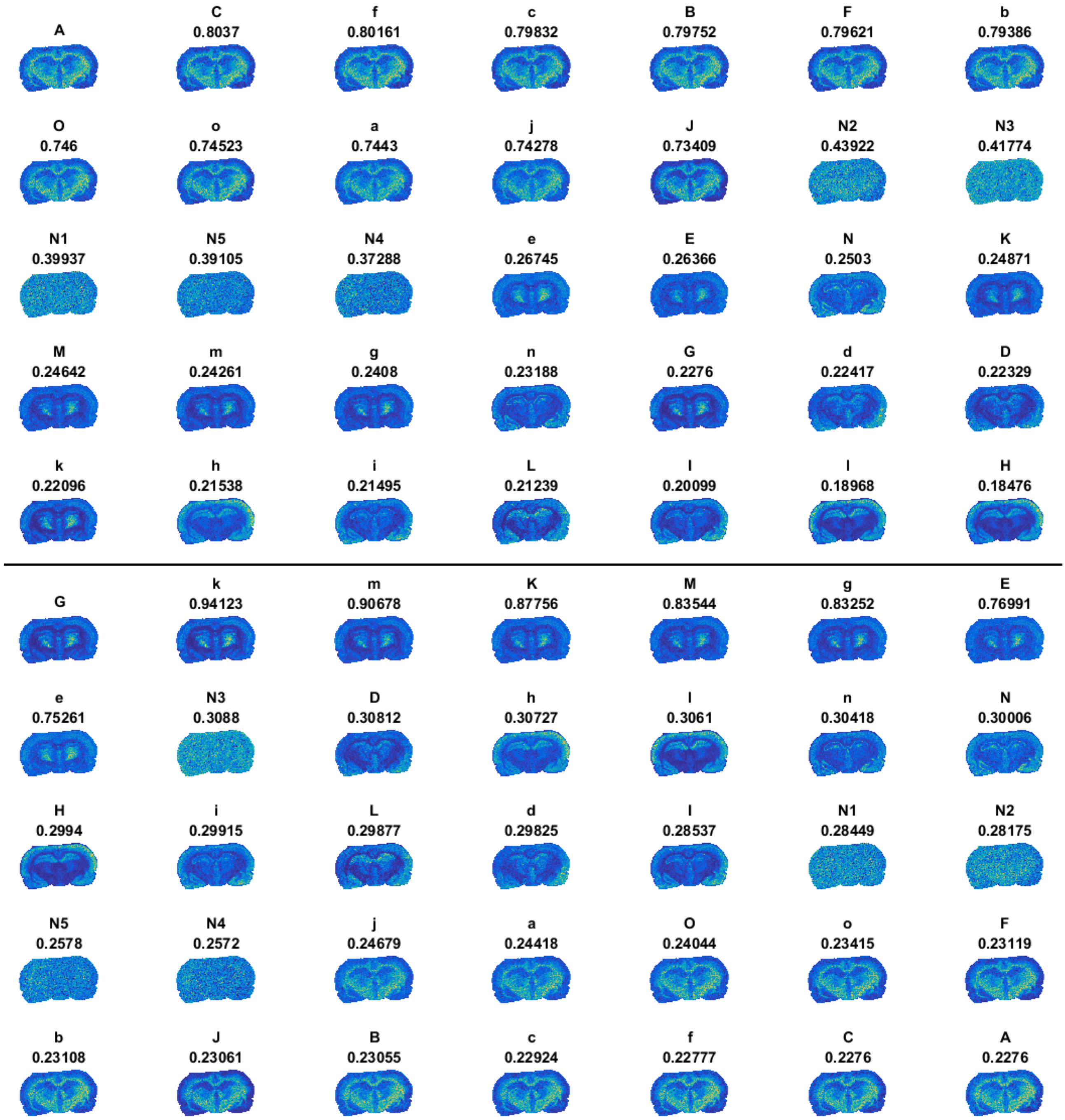


Figure 5.5: Image A and image G compared with all other images, showed in RTB scored descending order. The corresponding RTB scores are stated under the image name.

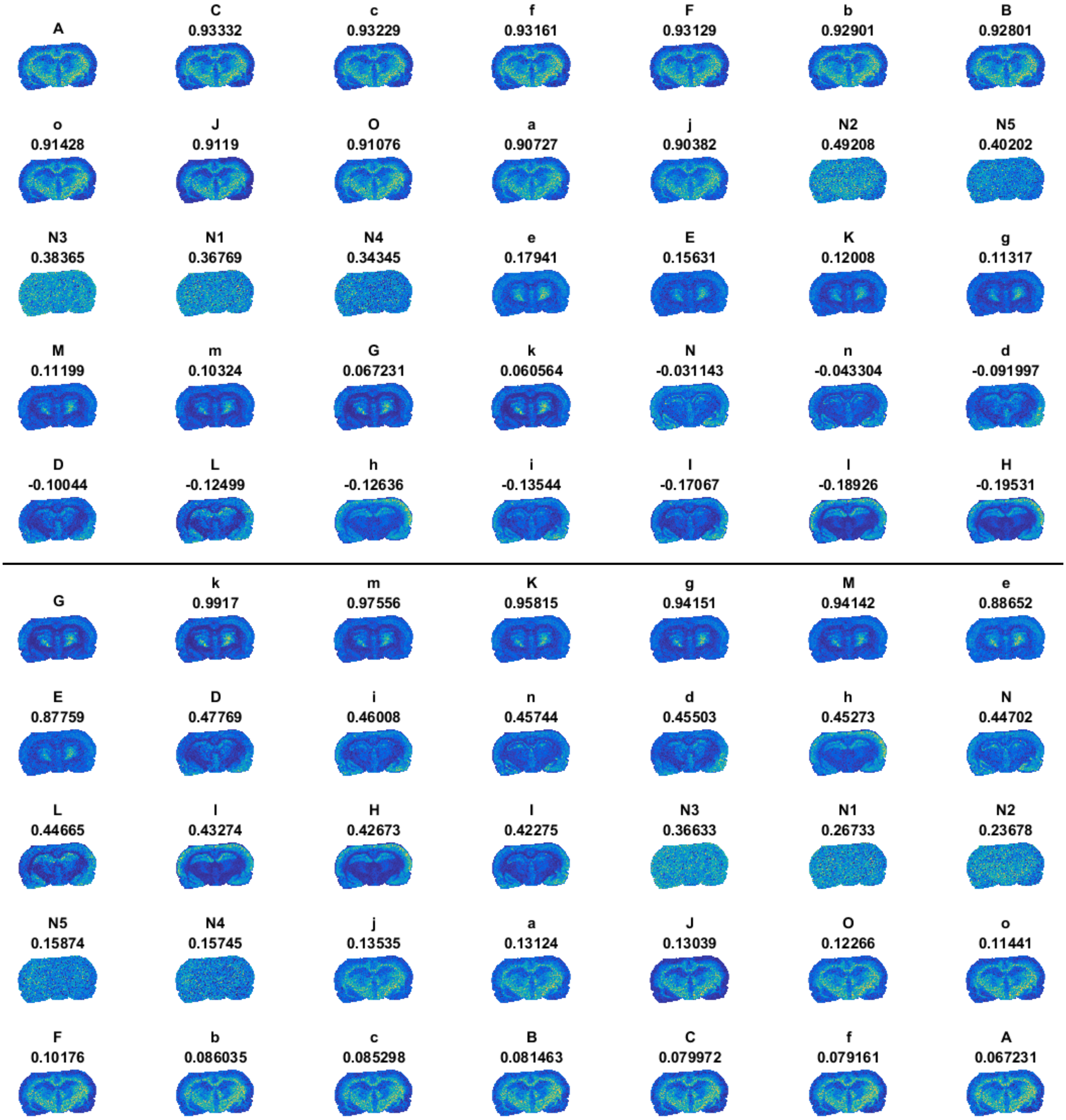


Figure 5.6: Image A and image G compared with all other images, showed in LC scored descending order. The corresponding LC scores are stated under the image name.

5.1.4 Determination of Measurement Algorithm

The result from the measurement evaluation indicates that all three candidate scoring algorithms can be used to discern co-localised images, with varying reliability. However, based on table 5.1, RTB is the most stable candidate in terms of the standard deviation, with a high average ratio between similar and non-similar pair scores. RTB scores noise images and delocalized images equally low. Therefore, using a cut off value of 0.6-0.7 would ensure co-localised pairs. Examining table 5.1, it can be seen that the RTB based algorithm has an almost binary-like behavior, scoring the images either high; they are co-localised or low; they are not co-localised.

A limitation of all of these algorithms and an assumption that we have to make is that the algorithms will only find peptides and potentially correlated degradation fragments that have the same spatial patterns, i.e. if the peptides are not degraded in the same degree everywhere in the tissue the given score will be lower. Biological systems are usually very complex and the outcome may not be as straight forward as we want it to be. This limitation would however be very difficult to avoid if the enzyme/degradation distribution is not known.

Another constraint to have in mind is that there exists a bias towards images with high intensity peaks. The images with very high intensity peaks compared to the background noise will get higher scores for more levels in the RTB scoring algorithm compared to those that have lower peaks versus the background noise.

5.2 Clustering

5.2.1 Threshold Optimisation

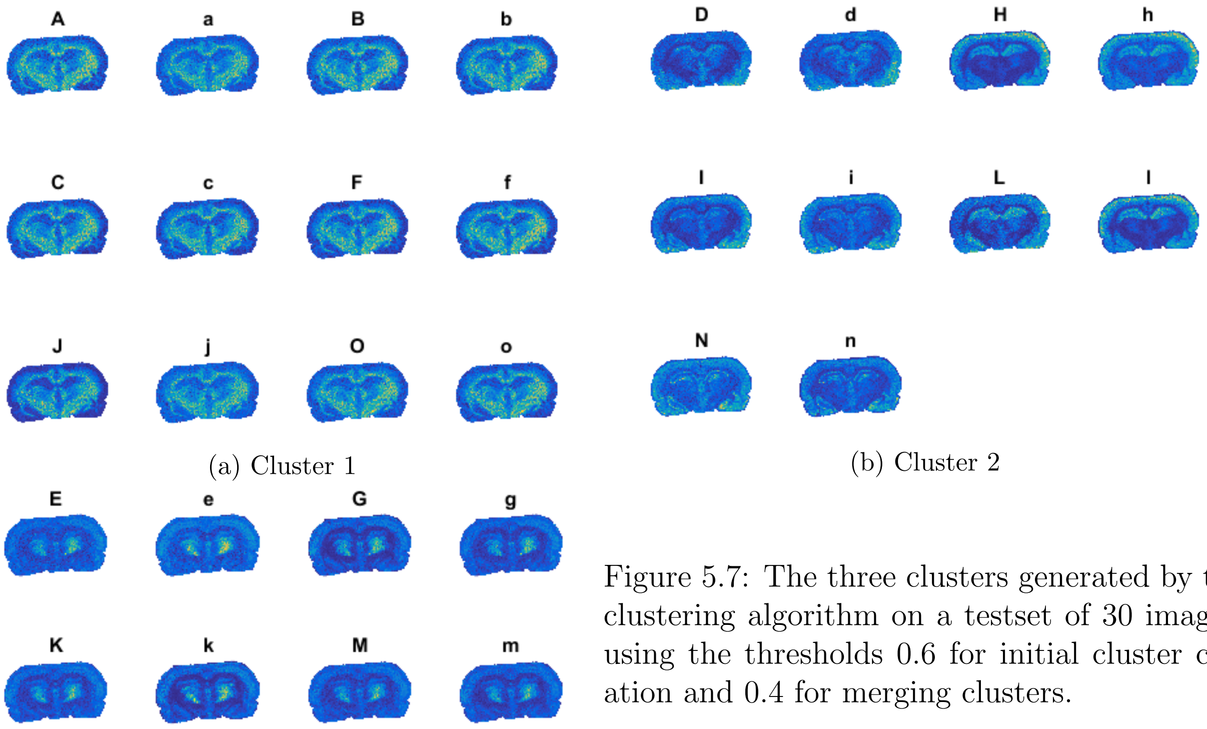
The fixed thresholds in the clustering algorithm were calibrated to optimize the clustering of a testset consisting of 30 images, figure 5.1. Using manual threshold optimization, the thresholds 0.6 for initial cluster creation and 0.4 for cluster merging was seen to generate the desired result. Namely the three clusters of images, seen in figure 5.7, showing three different spatial patterns.

5.2.2 Testset

The testset used to evaluate the customised clustering algorithm was composed of the top 150 most structured images from the same dataset as used in the similarity measurement evaluation (the MALDI IMS dataset from rat brain with 3045 images in total). These 150 images were found using the *spatial chaos score* described by T. Alexandrov and A. Bartels [21]. The testset images were thus extracted from the datacube using only the *spatial chaos score*, with no prior or post data reduction. Therefore images within the same peak exists in the testset.

5.2.3 Clustering

The images were pre-processed with automatic hot spot removal using quantile thresholding (quantile value 0.99) and normalized by their maximum values. Similarity scores of each pair



(a) Cluster 1

(b) Cluster 2

(c) Cluster 3

Figure 5.7: The three clusters generated by the clustering algorithm on a testset of 30 images, using the thresholds 0.6 for initial cluster creation and 0.4 for merging clusters.

of images in the testset were calculated using the RTB scoring algorithm described in section 4.3.2. and saved in a similarity matrix, seen in figure 5.9. The clustering algorithm previously described was then run with the computed similarity matrix as input and with a threshold of 0.6 for initial cluster creation and 0.4 for merging of clusters (optimised in section 5.2.1.). The result was four clusters with 25, 34, 68 and 23 members.

In order to visualise the clusters, a dimension reduction to three dimensions was done, by performing a principal component analysis (PCA) on the pre-processed testset. Each image was projected onto the linear space spanned by the top three principal components (PCs), accounting for the majority of the variation in the images (figure 5.10). The result was plotted in a 3D graph in Plotly, shown in figure 5.9, where the cluster memberships have been color-coded.

To identify the characteristics of each cluster, an average image for each cluster was computed by taking the mean of all images for each cluster, top left in figure 5.9. It can be seen that the average images have very diverse and distinct characteristics.

For further cluster characterisation the mean spectrum of the full dataset, containing 24 442 pixels and 3045 m/z images, was computed. This was retrieved by taking the average of all spectras in all pixels. The clustered images in the testset were then marked on the spectrum and color-coded for cluster membership, figure 5.9.

All images that have the same peak as origin, visualise the same peptide/protein and should thereby have the same spatial pattern (unless an overlap of two peaks has occurred). In figure 5.9 it can be seen that no peak has images that cluster in more than one cluster, which indicates that the cluster algorithm groups the images from the same origin, in the same cluster meanwhile finding other images with other peak origins that also has the same spatial pattern.

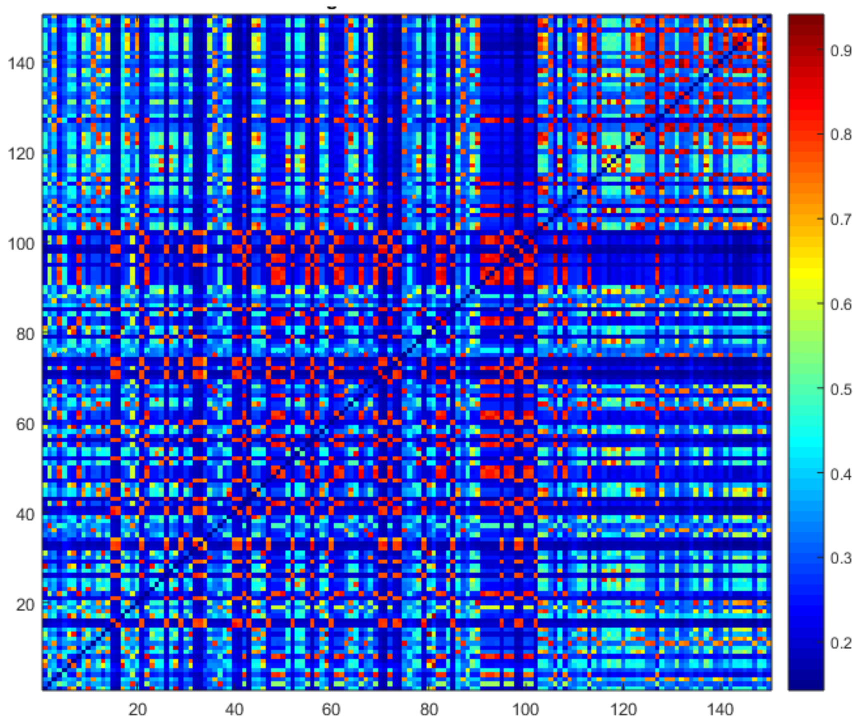


Figure 5.8: The similarity matrix containing the RTB similarity scores for each pair in the test-set. The self-correlation values in the diagonal have been set to NaN values for less distraction.

5.2.4 Clustering Advantages, Limitations and Performance

Since the PCA was performed independently, with no influence from the result of the clustering (except from the color-coding) and that the class members from the clustering result have high linear separability, we can conclude that the result of the PCA indicates that the clustering algorithm groups the images correctly.

Another conclusion that can be drawn from figure 5.9 is that the images in testset 1 should be discernible with their dimensions reduced to three dimensions (the top three PCs). Although, in figure 5.10 we can see that independent significant variance exists in the first 6-7 PCs, which indicates that potentially valuable information may be lost when reducing the images to three dimensions.

When running the clustering algorithm with a threshold of 0.5 instead of 0.4 for merging clusters, eight clusters were created. Cluster 3 from the previous run, figure 5.9, was divided into four distinct clusters, doing so the actual clustering was better (the clusters were tighter and held less variance). However, since the degree of similarity between a peptide and its degradation products localisation is unknown, the similarity requirements demanded by the clustering algorithm should not be too high.

The advantage of using an unsupervised analysis method is that no prior knowledge about the dataset is needed. The interpretation of the data is completely entrusted to the computer. The computer analyse the data iteratively, aiming to find hidden structures in the unlabeled data and thereby grouping the data based on the discovered structural patterns. With this design, a fixed number of clusters does not have to be known in advance, as is the case in for example

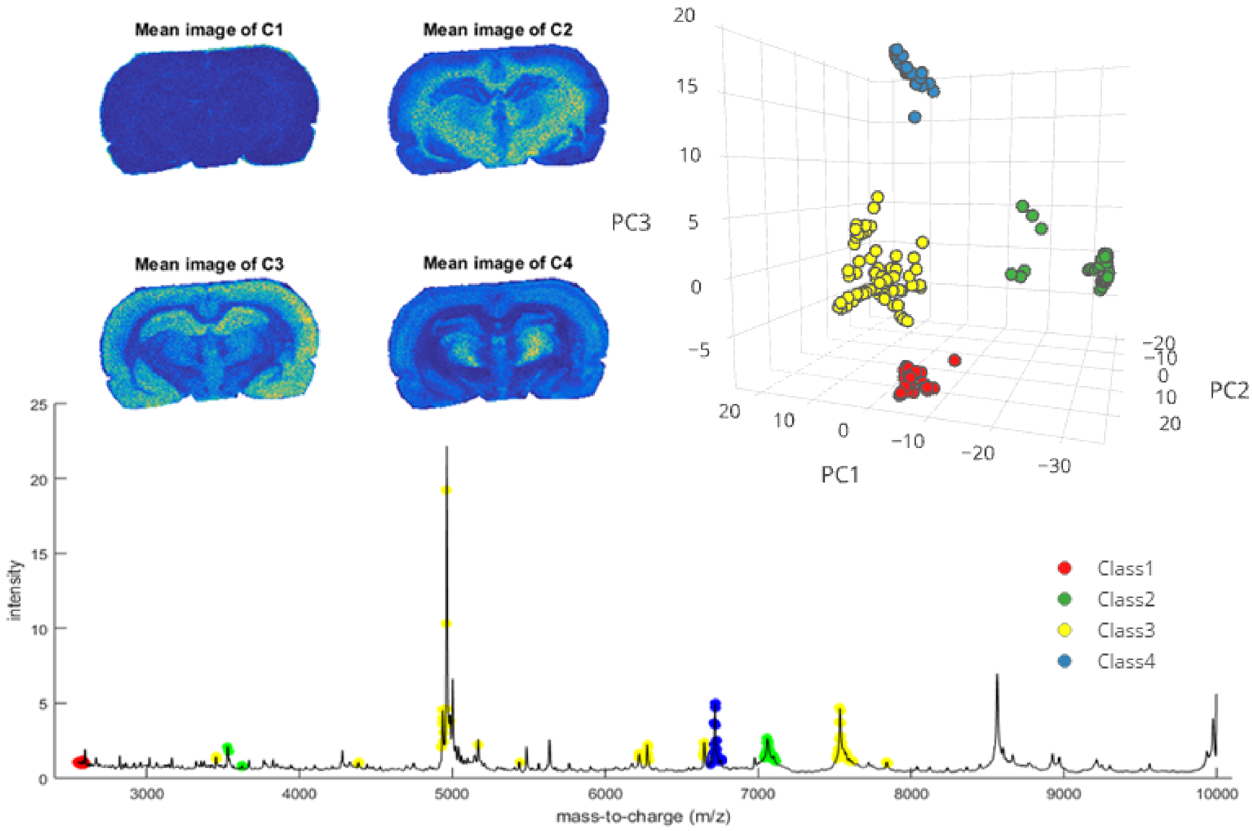


Figure 5.9: Cluster visualisation. *Top left,* the mean images of each cluster, showing the characteristics of the clusters. *Top right,* the 150 images in the testset projected onto the linear space spanned by the three first principal components. The clusters memberships (the classes) acquired from the clustering algorithm have been color-coded. The 3D graph can be viewed here; https://plot.ly/ stephanieherman/36. *Bottom,* mean spectrum of the full dataset, where the images in the extracted testset (the top 150 most structured images) have been marked with color-coded dots for cluster membership.

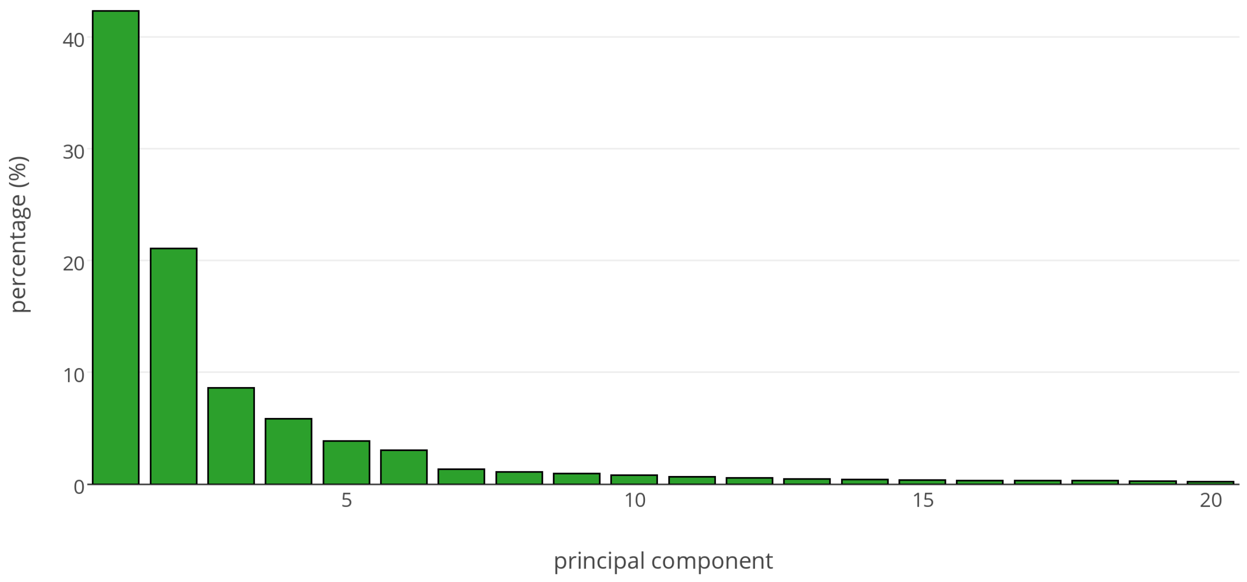


Figure 5.10: The percentage of variance explained by the top 20 principal components.

k-means clustering. With this clustering algorithm setup, only the threshold of the least allowed similarity has to be specified. The termination condition will terminate the iterations when no further scores above the threshold can be found.

5.3 Amino Acid Mass Comparer

5.3.1 Testset

To evaluate the inter cluster analysis pipeline, a testset was created based on an experimentally found degradation ladder of peptide fragments of the neuropeptide Copeptin, see table 5.3.

Table 5.3: The testset of peptide fragments from the neuropeptide Copeptin.

Peptide fragment of Copeptin	Mass-to-charge value
H-LRLVQLAGTQESVDSAKPRVY-OH	2311.24
H-RLVQLAGTQESVDSAKPRVY-OH	2198.16
H-LVQLAGTQESVDSAKPRVY-OH	2042.06
H-VQLAGTQESVDSAKPRVY-OH	1928.98
H-LAGTQESVDSAKPRVY-OH	1701.84
H-AGTQESVDSAKPRVY-OH	1588.76
H-GTQESVDSAKPRVY-OH	1517.72

The delta mass values of each unique pair of fragments were calculated and matched with the masses of single, double and triple combinations of amino acids, with an allowed margin of error of 0.2 Da.

5.3.2 Hits

Table 5.4: The matches of amino acid (AA) combinations found by the algorithm. The true combinations are highlighted in bold.

m/z-pair	Single AA matches	Double AA matches	Triple AA matches
2311/2198	**L**, I	-	-
2198/2041	**R**	G/V	-
2042/1929	**L**, I	-	-
1702/1589	**L**, I	-	-
1589/1518	**A**	-	-
2311/2042	-	**L/R**, I/R	A/P/T, A/V/V, G/L/V, G/I/V
2198/1929	-	**L/R**, I/R	A/P/T, A/V/V, G/L/V, G/I/V
1929/1702	-	**V/Q**, A/R, L/N, K/V, I/N	A/G/V, G/G/L, G/G/I
1702/1518	-	**A/L**, A/I, P/S	-
2311/1929	-	-	**L/L/R**, E/P/R, P/V/W, L/I/R, I/L/R, I/I/R
2042/1702	-	-	**L/Q/V**, A/L/R, D/P/Q, E/N/P, G/P/W, L/L/N, A/I/R, D/P/K, I/Q/V, L/K/V, I/L/N, L/I/N, I/I/N
1929/1589	-	-	**L/Q/V**, A/L/R, D/P/Q, E/N/P, G/P/W, L/L/N, A/I/R, D/P/K, I/Q/V, L/K/V, I/L/N, L/I/N, I/I/N
1929/1518	-	-	R/R/V, H/H/H, F/T/Y, P/Q/W, P/K/W

5.3.3 Limitations

Even though there only exist two pairs of amino acids that have similar masses (L has the same mass as I and K and Q have the same masses), there exist several two-combinations and three-combinations that have the same masses. This implies that if a pair of m/z values gets a match, it is highly likely that it will have several more matches. This phenomenon can be seen in table 5.4. When generating and comparing with larger combinations of amino acids, this phenomenon will be further increased. Comparing the testset used above with combinations of four amino acids, resulting in 5985 unique combinations, where only 3715 had unique mass values and the m/z-pair with the most matches had 16 matches in total.

This implies that when analysing large datasets with long amino acid combinations, the fact that there is a match between two m/z values will be of more importance than the actual matches, since there will be no computational way to tell which match is the real one. For this reason the comparison in this thesis was limited to one amino acid, producing m/z-pairs that consists of closely related degradation products differing with a single amino acid.

Chapter 6

Method

6.1 Experimental Method

The brains from nine different animals (wild-type mice) were extracted and saggitally cut in half, where one half was heat stabilized using the Stabilizor system and then rapidly frozen using dry ice and isopentane. The other half was directly snap frozen using the same freezing procedure as the heat stabilized half. This was done with three different *post mortem* times (0, 10 and 30 minutes) and the procedure was repeated three times to create three biological replicates. The experimental design can be seen in table 6.1. To prevent formation of large crystals, the tissue samples were repeatedly submerged into isopentane.

In order to prevent unnecessary death, pre-used mice were used which were obtained dead (Pre-usage was checked for treatments affecting the brain). Permission for usage of mice tissue was given by the Regional Ethical Review Board in Uppsala, Sweden (Dnr C 206/12). The mice brains were immediately extracted after death and further processed.

Table 6.1: The experimental setup.

Mouse	Left half	Right half	PM time	Replicate
1	HS 30:1	SF 30:1	30 min	1
2	HS 30:2	SF 30:2	30 min	2
3	HS 30:3	SF 30:3	30 min	3
4	HS 10:1	SF 10:1	10 min	1
5	HS 10:2	SF 10:2	10 min	2
6	HS 10:3	SF 10:3	10 min	3
7	HS 0:1	SF 0:1	0 min	1
8	HS 0:2	SF 0:2	0 min	2
9	HS 0:3	SF 0:3	0 min	3

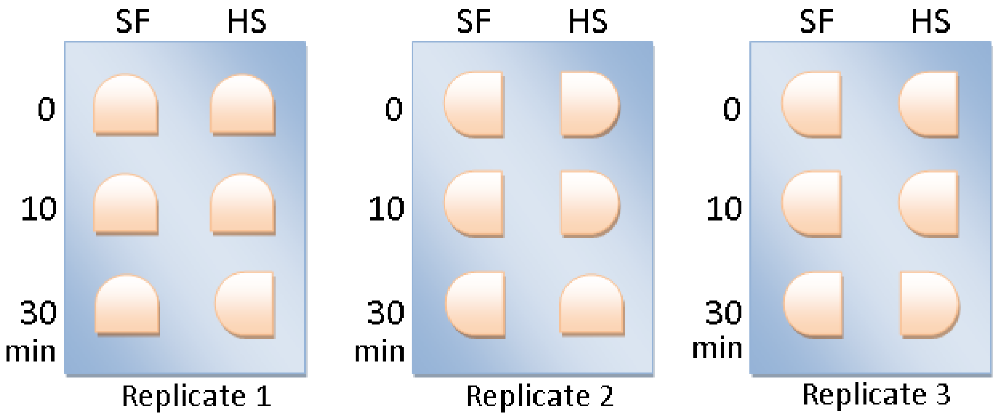


Figure 6.1: The experimental design (seen from above), where SF stands for snap frozen and HS for heat stabilized.

Six differently treated brain halves were then mounted onto an approximately 0.5 cm thick layer of carboxymethylcellulose (CMC) in a cube, seen in figure 6.1. The cube was then filled with CMC until the brain halves were totally covered, figure 6.2.

The sample cubes were then coronally cut using the Leica CM3050S instrument with a thickness of 12 μm which were prepared on conducted glass slides and artificially thaw mounted. The glass slides were then stored in a freezer until further work.

The protocol for matrix application, developed by J. Hanrieder *et al.* [22], was followed with modified matrix content proportions (25 mg/ml DHB, 50% methanol, 0.3% TFA and 10 % 0.15 AmAc) [22]. The matrix solution was printed onto the tissues using the CHIP-1000 Chemical Inkjet Printer with 15 drops per pass and 20 application passes, with a resolution of 400 μm (from spot to spot). This particular setup was chosen because the amount of ionized molecules in the samples were thought of as being of more importance than the resolution of the final images.

After matrix application the samples were run in an Ultraflex II TOF/TOF from Bruker with 800 shots with laser intensity of 50-65%.

Since the HS tissues behaved differently than the SF tissues, only one continous serie, HS replicate 3, was spotted using the CHIP-1000 Chemical Inkjet printer and imaged. The two other HS replicates were handspotted with 0,6 μl of matrix and two profile spectras were manually collected for each sample, with 800 shots of laser intensity of 65% with the Ultraflex II TOF/TOF.

The SF replicate 1 and 2 series and the HS replicate 2 serie were all spotted with matrix and run in the Ultraflex II TOF/TOF simultaneously, while the SF replicate 3 serie was spotted and run a week before them.

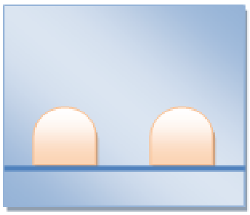


Figure 6.2: The brain halves (seen from the side) placed in a cube with a layer of approximately 0.5 cm CMC in the bottom. The surrounding areas were also, after placement, filled with CMC..

6.2 Data Processing

Baseline correction was performed in flexImaging and the raw data were normalized in SCiLS Lab 2014 using the Root Mean Square method and exported as .imzML files. The .imzML files were then imported and converted into three dimensional datacubes, as seen in figure 2.1, in Matlab using the imzMLConverter. The datacubes were then reduced by the the data reduction algorithms described in section 4.1.3 and 4.1.4.

6.3 Analysis Procedure

The reduced datacubes of each brain half were scored with the RTB similarity measurement described in section 4.3.2 and clusterings were performed on the computed similarity matrices, with a threshold of 0.6 for initial cluster creation and 0.4 for merging clusters. The clusters were then separately analysed by the amino acid mass comparer described in section 4.4. The full analysis pipeline can be seen in figure 6.3.

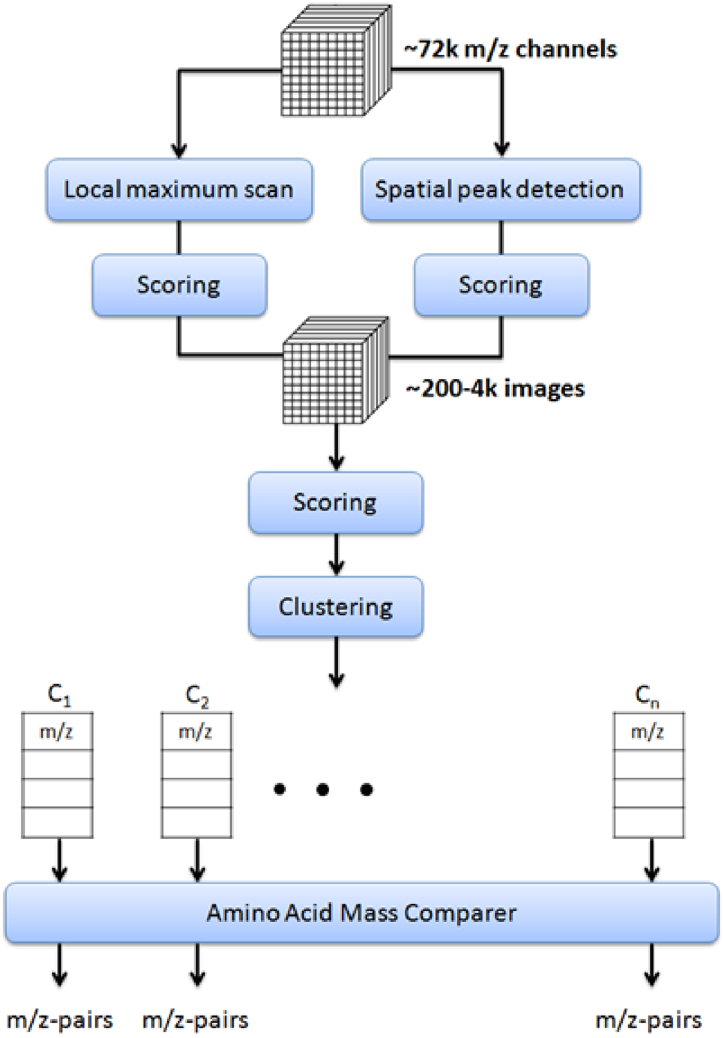


Figure 6.3: The full workflow in the analysis pipeline, where the scoring step is done using the RTB measurement.

Chapter 7

Result and Analysis

7.1 Analysis Output

After the data reduction, datacubes of 242, 476 and 701 images were generated from the SF tissues with a PM time of 30. These datacubes were scored using the RTB similarity measurement, generating similarity matrices. The subsequent clusterings upon the similarity matrices produced 26, 39 and 47 clusters. Information about the output of the data reduction, clusterings and amino acid mass comparisons from all PM times and replicates can be seen in table 7.1.

Table 7.1: Output from the three biological replicates with different PM times.

30 minutes PM	**Replicate 1**	**Replicate 2**	**Replicate 3**
Initial number of images	75062	75062	75062
Reduced number of images	242	476	701
Number of clusters	26	39	47
Number of pairs	30	70	70
10 minutes PM	**Replicate 1**	**Replicate 2**	**Replicate 3**
Initial number of images	75062	75062	75068
Reduced number of images	440	276	292
Number of clusters	40	29	17
Number of pairs	48	31	31
0 minutes PM	**Replicate 1**	**Replicate 2**	**Replicate 3**
Initial number of images	75062	75062	75062
Reduced number of images	225	773	304
Number of clusters	29	69	26
Number of pairs	34	105	48

In figure 7.1 all mass spectras from the SF tissues with a PM time of 30 minutes can be seen, where the images chosen by the spatial TIC based data reduction algorithm are marker in red

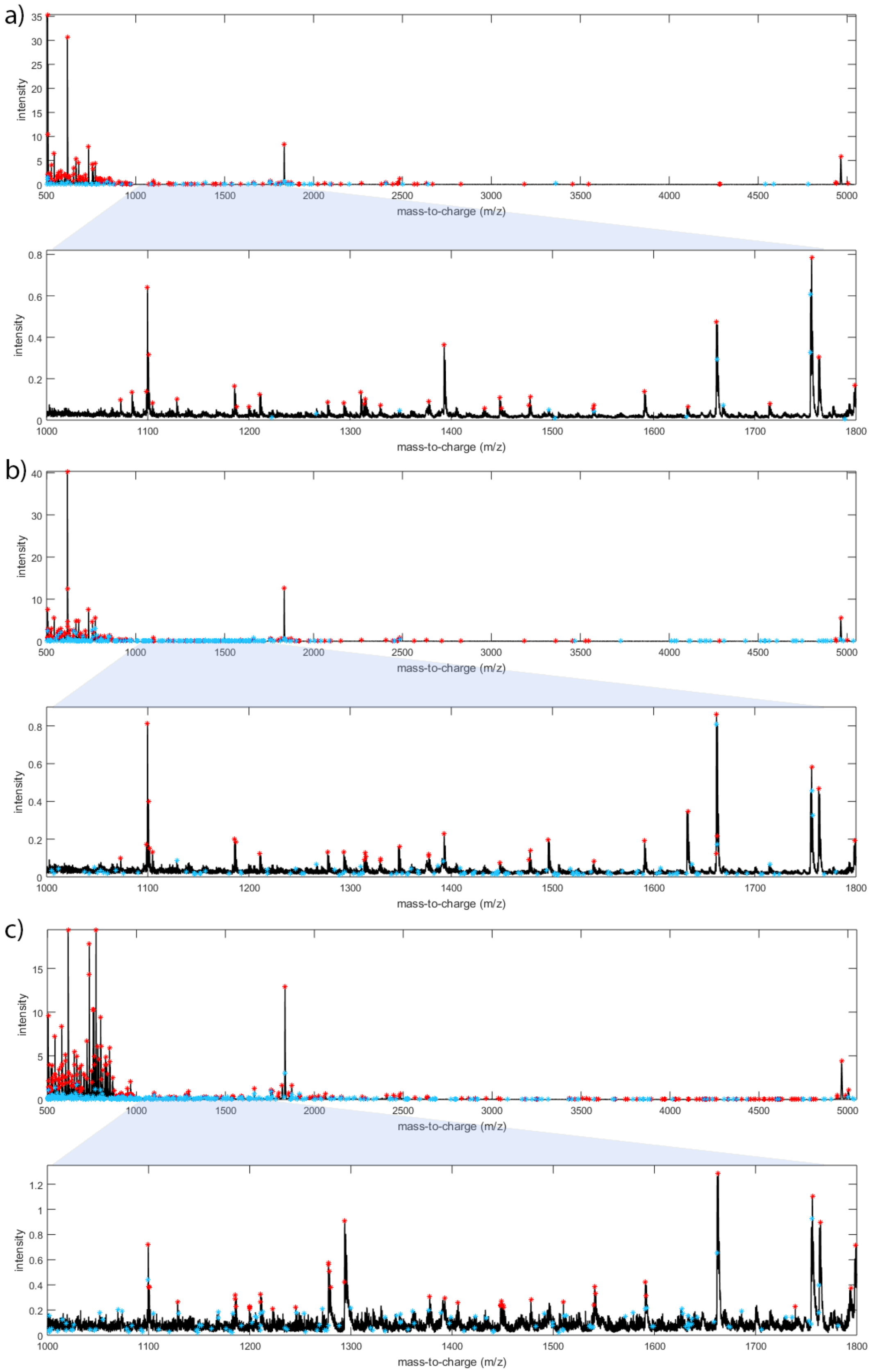


Figure 7.1: Mean spectras of each replicate with a PM time of 30 minutes in chronological order (*a:* R1SF30, *b:* R2SF30, *c:* R3SF30). Below each mean spectrum is a close up of the m/z region between 1000 and 1800. The red markers represent images chosen from the first DR step, the spatial TIC based DR and the light blue markers represent images from the second DR step, the spatial peak detection.

and the ones chosen from the spatial peak detection in blue. As depicted the spatial TIC based DR pick images representing local maximums in the mean spectrum while the spatial peak detection pick images that are not as obviously interesting, images that might hold very local spatial expression patterns.

Table 7.2: Output for replicate 2 of HS tissue.

	0 minutes PM	**10 minutes PM**	**30 minutes PM**
Initial number of images	75062	75062	75062
Reduced number of images	1558	1236	3748
Number of clusters	225	140	276
Number of pairs	103	88	382

The same analysis was further performed on the imaged HS tissues (replicate 2). Subresults can be seen in table 7.2. The HS tissues were reduced to a much larger number of images than the SF tissues. The large majority of these were extracted by the spatial peak detection DR, seen in the lower part of the heatmap shown in figure 7.2. This concludes that HS tissue contains more local peak areas than the SF tissue. Figure 7.2. also revealed that the spatial TIC based DR algorithm is much more reproducible than the spatial peak based DR. The spatial peak based DR might still be in need of further optimisation for higher reproducibility and reliability. Although an interesting observation is that the reproducibility of the spatial peak detection based DR is in general higher in the HS tissues. Boxplots were made of all reduced m/z values, seen in the top of figure 7.3, showing the distribution of reduced the m/z values in each sample. The boxplot showed that all the HS tissues have a higher m/z median than the SF tissues. While no striking trends can be seen for the SF tissues. Applying the same visualisation for only the reduced m/z values from the spatial TIC based DR, an interesting feature of the HS tissues can be seen, figure 7.3. While the SF tissues vary a lot in their m/z distribution, the HS tissues

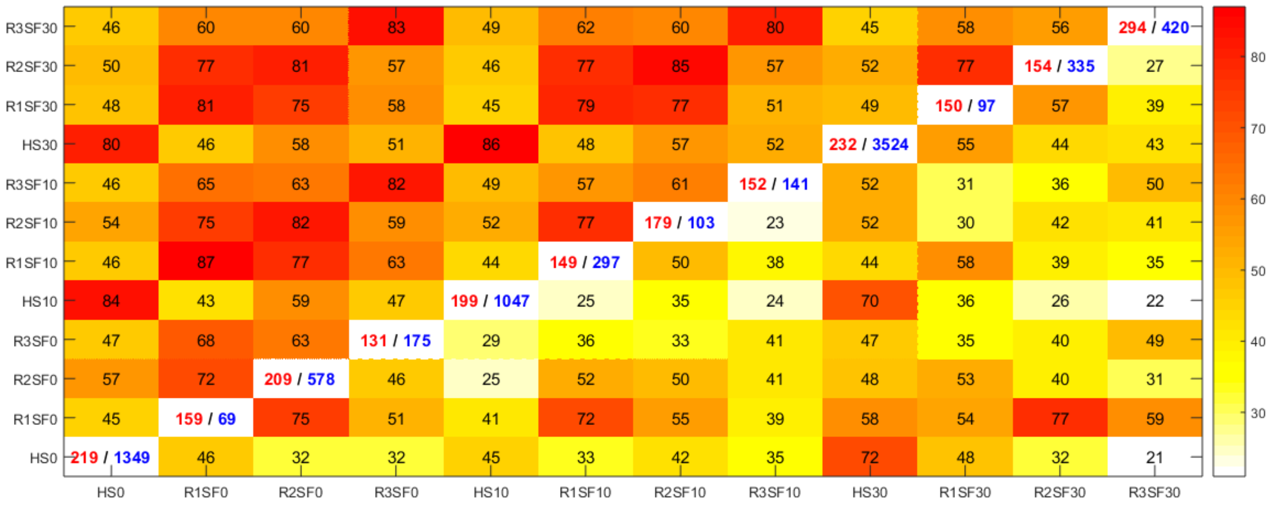


Figure 7.2: Heatmap showing percentage m/z values similar (allowing a shift of ±0.5 Da) between samples, where the upper half show the m/z values chosen by the spatial TIC based DR (1) and the lower half show the ones chosen by the spatial peak detection based DR step (2). The total amount of m/z values are showed in the diagonal, in red for spatial TIC and in blue for spatial peak detection.

m/z distributions are strikingly similar, indicating very similar peaks in the mean spectras for all PM time points. To confirm this observation the mean spectras of the HS tissues were plotted for two different m/z regions, figure 7.4.

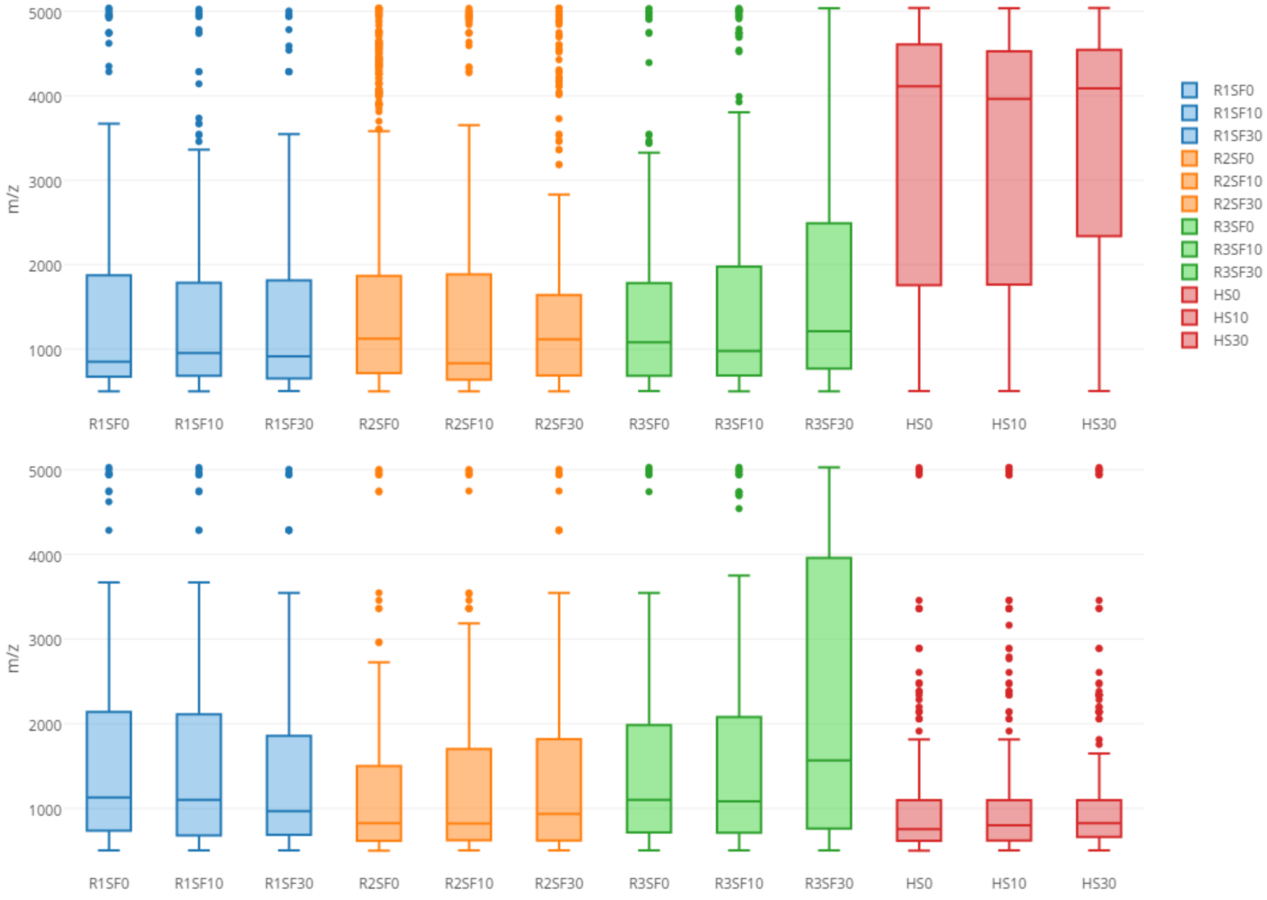


Figure 7.3: Boxplots showing the distribution of the reduced m/z values. *Top*, The distributions of the reduced m/z values. *Bottom*, the distribution of the m/z values chosen only by the more supported DR algorithm, the spatial TIC based.

Even though the HS tissues in the various PM time points have the same detected peaks, the intensity of these peaks should still vary over the time serie, due to the time of delay before heat stabilization. The close-ups in figure 7.4 do indeed confirm this belief. Degradation over time is occurring even in the heat stabilized tissues, due to the delay of time before heat stabilization.

A PCA was performed on the TIC normalized mean spectras of all samples (figure 7.5). The plot shows that replicate clusters are formed. Replicate 3, which was done seven days before the other SF replicates seem to be significantly different from the other SF samples, indicating that principal component two are related to experimental variation. While principal component one seems to be related to HS/SF differences.

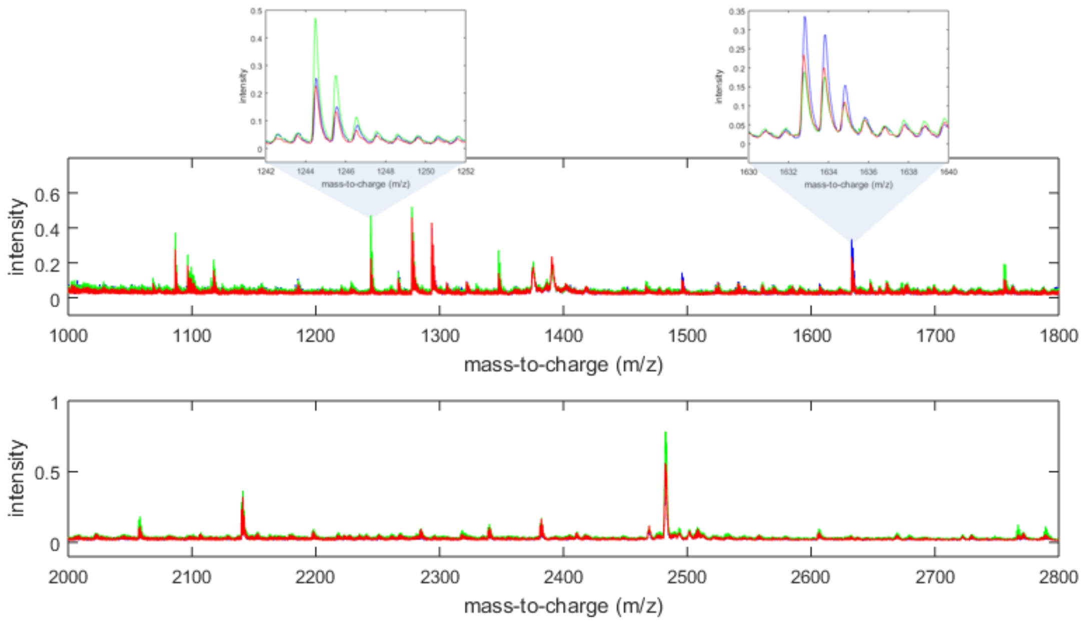


Figure 7.4: Mean spectras for the HS tissues where *blue*: PM time 0 minutes, *green*: PM time 10 minutes and *red*: PM time 30 minutes. The intensity varies over time indicating that degradation is happening even in this short time span. The two close-ups of the m/z region 1242-1252 and 1630-1640 give an even clearer image of what is happening. However, some peptides, like the peak at around m/z 1290, seem to be more stable.

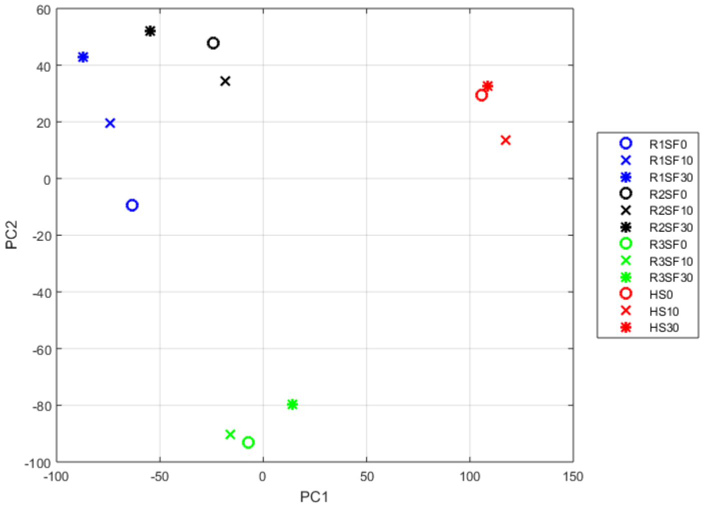


Figure 7.5: The normalized mean spectras projected onto the linear space spanned by their top two principal components. The three replicates have been color coded and the time points visualised by shapes.

7.2 Marker Candidates

Clearly there are too many potential marker candidates for manual evaluation. However, when a smaller cluster from R3SF30 was examined, the hits showed some very interesting features. The three m/z-pairs that differed with a single amino acid created a full degradation ladder. Too investigate if such features existed in more clusters, a small script was written which searched

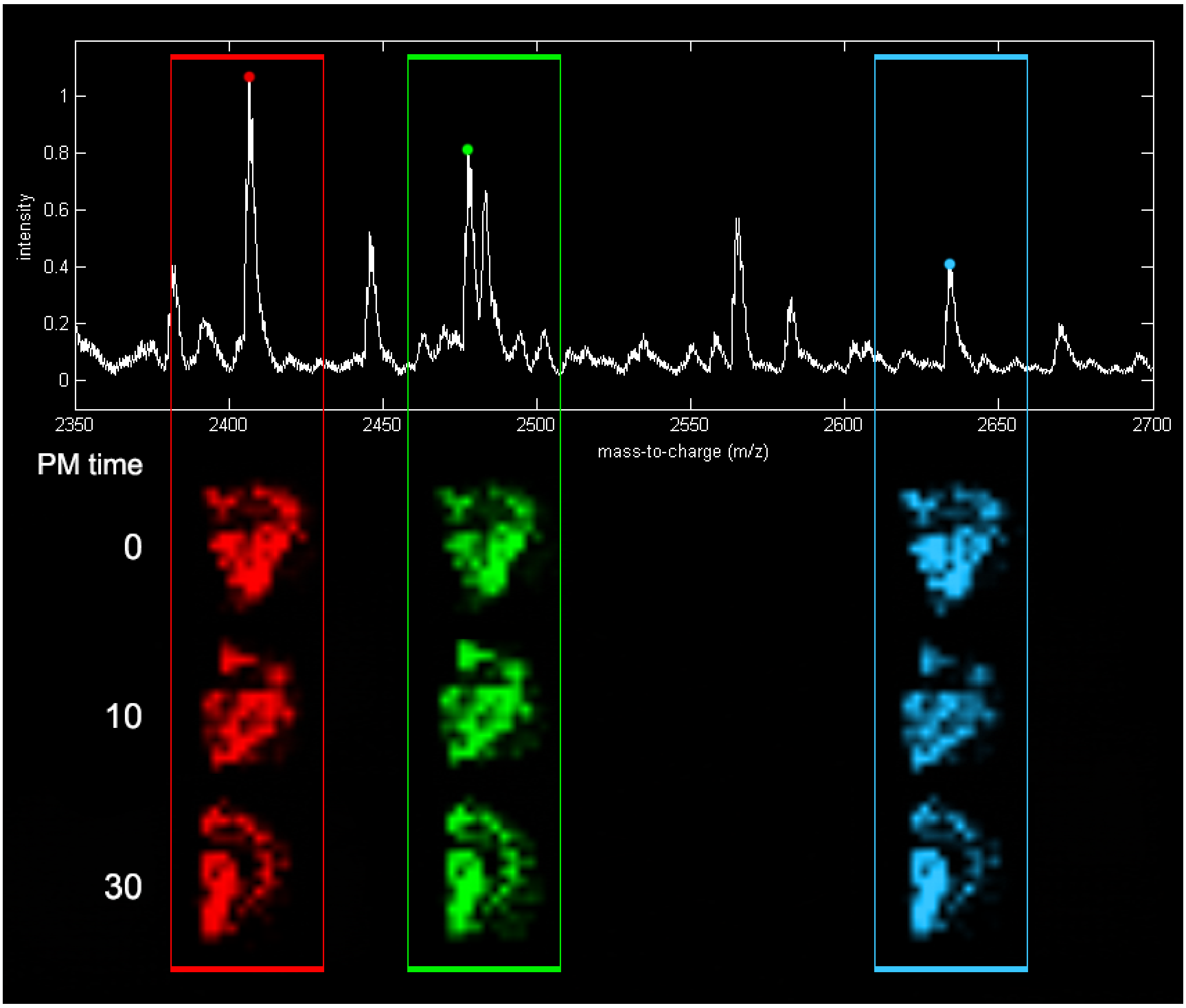


Figure 7.6: The mean spectrum of SF tissue where the members in the degradation ladder, A-AR, ion images are shown (m/z 2407: *red*, m/z 2478: *green*, m/z 2634: *blue*) for all PM times in replicate 3. The ion images were generated by the image with the highest intensity using flexImaging. The maximum color intensity is set to 40% of the intensity, in all ion images. Intensity and area decaying can be seen in m/z 2634, comparing with m/z 2478 and 2407 over the PM time serie.

through the m/z-pairs for sequences of pairs. The resulting degradation ladders can be seen in table 7.3, 7.4, 7.5 and 7.6. In these tables the m/z values have been rounded to nearest integer for easier display. As can be seen there are one degradation ladder, containing the amino acids Alanine and Arginine, present in all PM times and all replicates of SF tissue and several degradation ladders can be seen in more than one SF sample. In figure 7.6 ion images for all PM times in replicate 3 are shown with the corresponding mean spectrum. An intensity decaying can be seen over PM time in m/z 2634 shown in blue, when comparing with m/z 2407, showed in red and m/z 2478, showed in green. The mean spectras for all samples were plotted for this m/z region (m/z 2350-2700), figure 7.7, to visualise intensity differences between PM times within a replicate. Even in this m/z region the HS samples are stable between different PM time points. As seen in table 7.6 the A-AR degradation ladder is not present in the HS sample. Although the m/z value 2478 is seen decreasing from PM time 10 to 30 minutes, figure 7.7. These findings would suggest that two of the peptide fragments, m/z values 2407 and 2634, are created (in the SF tissues) after the HS tissues have been heat stabilized, in other words during the experimental workflow.

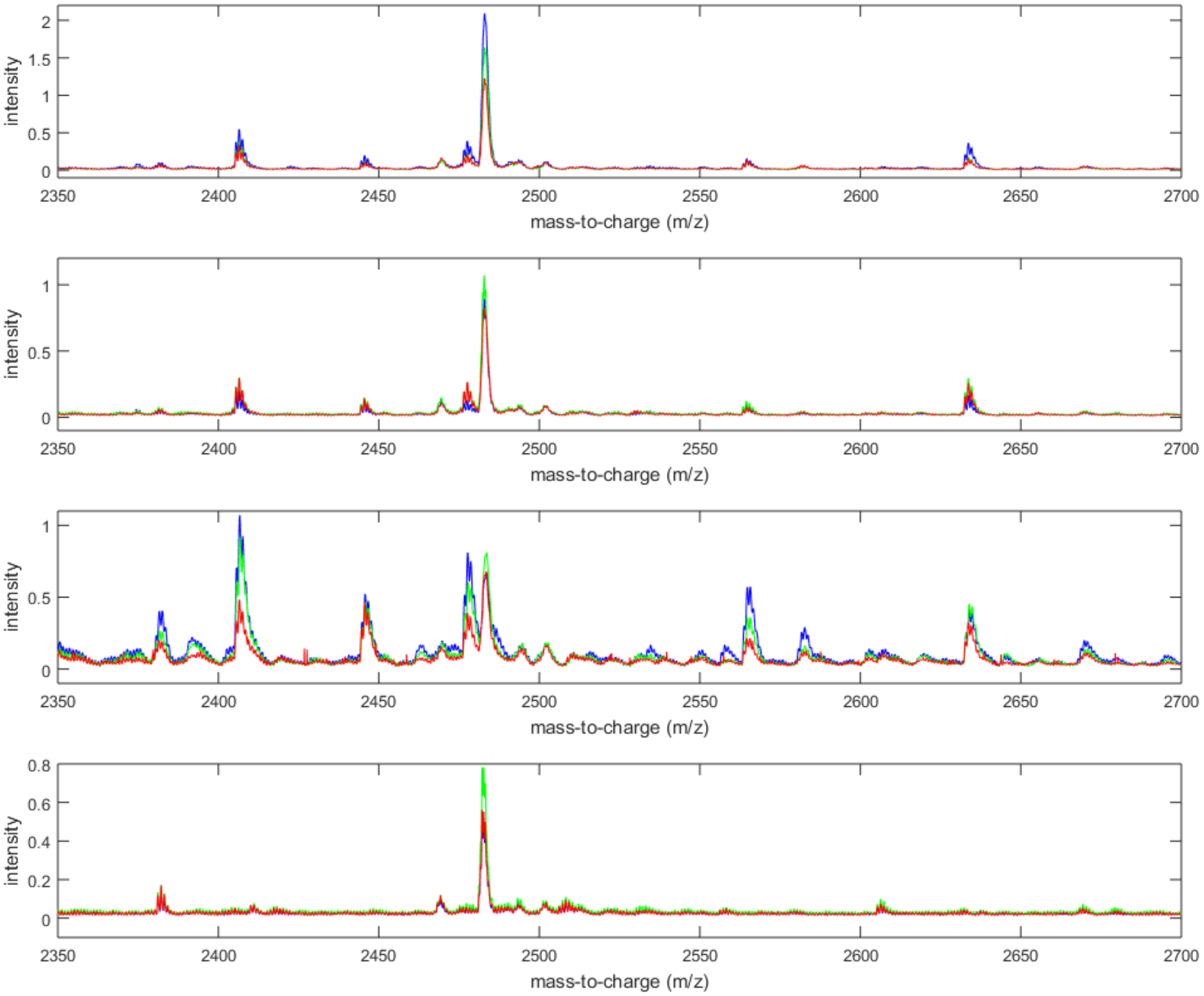


Figure 7.7: All mean spectras for the m/z region where the A-AR degradation ladder lies, where the SF samples are in replicate order followed by the HS samples (*From top:* SF Replicate 1, SF replicate 2, SF replicate 3, HS). Where the *blue* spectras are at PM time o minutes, *green* 10 minutes and *red* 30 minutes.

Table 7.3: Observed degradation ladders in the SF samples with a PM time of 30 minutes, where Q can be either Q or K and L could be L or I (because of similar masses). Degradation ladders present in any other SF sample are highlighted in blue.

(a) Replicate 1

m/z-pair	**AA**
Cluster 15	
1349/1478	E
1349/1591	EL
1349/1662	ELA
1349/1763	ELAT
alt.	
1349/1478	E
1349/1591	EL
1349/1662	ELA
1349/1799	ELAH
alt.	
1349/1478	E
1349/1634	ER
1349/1763	ERE
2407/2478	A
2407/2565	AS
alt.	
2407/2478	A
2407/2634	AR
Cluster 19	
943/1099	R
943/1186	RS
943/1314	RSQ
alt.	
943/1099	R
943/1186	RS
943/1315	RSE
Cluster 20	
1541/1669	Q
1541/1756	QS

(b) Replicate 2

m/z-pair	**AA**
Cluster 7	
568/715	F
568/818	FC
alt.	
584/715	M
584/818	MC
Cluster 10	
678/792	N
678/891	NV
Cluster 29	
1478/1591	L
1478/1662	LA
1478/1763	LAT
alt.	
1478/1591	L
1478/1662	LA
1478/1799	LAH
1478/1896	LAHP
2407/2478	A
2407/2634	AR
Cluster 38	
973/1129	R
973/1285	RR
973/1448	RRY
Cluster 39	
815/943	Q
815/1099	QR
815/1186	QRS
815/1314	QRSQ
815/1477	QRSQY
alt.	
815/943	Q
815/1099	QR
815/1186	QRS
815/1315	QRSE
2022/2153	M
2022/2268	MD

(c) Replicate 3

m/z-pair	**AA**
Cluster 17	
1540/1627	S
1540/1730	SC
Cluster 26	
1663/1764	T
1663/1835	TA
2407/2478	A
2407/2634	AR
1591/1662	A
1591/1763	AT
1591/1834	ATA
Cluster 44	
512/611	V
512/767	VR
781/894	L
781/997	LC
781/1183	LCW
1739/1895	R
1739/2032	RH
2352/2423	A
2352/2522	AV
Cluster 46	
945/1048	C
945/1234	CW
945/1333	CWV
1120/1234	N
1120/1333	NV
1120/1276	R
1120/1333	RG
1162/1276	N
1162/1333	NG

Table 7.4: Observed degradation ladders in the SF samples with a PM time of 10 minutes, where Q can be either Q or K and L could be L or I (because of similar masses). Degradation latters observed in any other SF sample are highlighted in blue. Note that the ladder in R3 cluster 14 has been shifted with 1 Da and thereby starts with a N instead of a L.

(a) Replicate 1

m/z-pair	AA
Cluster 19	
735/849	N
735/948	NV
Cluster 27	
1478/1591	L
1478/1662	LA
1478/1763	LAT
1478/1834	LATA
1478/1990	LATAR
2406/2478	A
2406/2634	AR

(b) Replicate 2

m/z-pair	AA
Cluster 6	
718/815	P
718/944	PE
alt.	
718/815	P
718/943	PQ
718/1099	PQR
718/1186	PQRS
718/1314	PQRSQ
718/1447	PQRSQY
Cluster 20	
1315/1478	Y
1315/1591	YL
Cluster 21	
1857/1971	N
1875/2099	NQ
1921/2022	T
1921/2153	TM
1921/2268	TMD
2406/2478	A
2406/2634	AR

(c) Replicate 3

m/z-pair	AA
Cluster 9	
556/658	T
556/757	TV
alt.	
556/658	T
556/773	TD
Cluster 14	
943/1099	R
943/1186	RS
1478/1592	N
1478/1663	NA
1478/1764	NAT
1478/1835	NATA
1756/1857	T
1756/1972	TN
1756/2100	TNQ
alt.	
1756/1857	T
1756/1972	TN
1756/2100	TNE
2270/2407	H
2270/2478	HA
2270/2634	HAR

Table 7.5: Observed degradation ladders in the SF samples with a PM time of 0 minutes, where Q can be either Q or K and L could be L or I (because of similar masses). Degradation ladders observed in any other SF sample are highlighted in blue. Note that a shift of 1 Da has occured in two degradation ladders in cluster 23.

(a) Replicate 1

m/z-pair	AA
Cluster17	
502/688	W
502/803	WN
502/906	WNC
Cluster 23	
1100/1187	S
1100/1316	SE
1405/1591	W
1405/1662	WA
alt.	
1478/1591	L
1478/1662	LA
1764/1835	A
1764/1922	AS
1764/2023	AST
1857/1971	N
1857/2099	NQ
2407/2478	A
2407/2565	AS
alt.	
2407/2478	A
2407/2634	AR
2407/2771	ARH
2407/2828	ARHG
alt.	
2407/2478	A
2407/2634	AR
2407/2771	ARH
2407/2908	ARHH
Cluster 24	
943/1099	R
943/1186	RS

(b) Replicate 2

m/z-pair	AA
Cluster 1	
712/769	G
712/856	GS
Cluster 10	
1024/1081	G
1024/1267	GW
1107/1164	G
1107/1267	GC
Cluster 17	
511/568	G
511/655	GS
Cluster 21	
718/815	P
718/944	PE
Cluster 36	
1478/1591	L
1478/1662	LA
1478/1763	LAT
2406/2477	A
2406/2634	AR
Cluster 40	
577/634	G
577/731	GP
577/844	GPL
Cluster 58	
1409/1496	S
1409/1633	SH

(c) Replicate 3

m/z-pair	AA
Cluster 3	
636/783	F
636/897	FN
Cluster 15	
827/926	V
827/983	VG
Cluster 20	
1663/1764	T
1663/1835	TA
1663/1922	TAS
1663/2023	TAST
1756/1857	T
1756/1971	TN
1756/2100	TNQ
2407/2478	A
2407/2634	AR
Cluster 21	
1591/1662	A
1591/1763	AT
1591/1834	ATA
alt.	
1591/1662	A
1591/1763	AT
1591/1850	ATS
Cluster 24	
2383/2484	T
2383/2583	TV
2383/2670	TVS

Table 7.6: Observed degradation ladders in the HS samples, with PM times of 0, 10 and 30 minutes. For the HS section with a PM time of 30 minutes, only a selection of the found ladders are shown. Degradation ladders found in any SF sample are highlighted in blue and those similar between HS samples in red.

(a) 0 minutes

m/z-pair	AA
Cluster 128	
3961/4118	R
3961/4246	RQ
Cluster 212	
1098/1245	F
1098/1376	FM
2141/2198	G
2141/2285	GS
2141/2382	GSP
2141/2469	GSPS
Cluster 214	
580/651	A
580/780	AE
548/651	C
548/780	CE
586/673	S
586/787	SN
526/689	Y
526/803	YN
526/874	YNA
526/972	YNAP
526/689	Y
526/805	YD
732/803	A
732/874	AA
732/972	AAP
912/1099	W
912/1186	WS
943/1099	R
943/1186	RS

(b) 10 minutes

m/z-pair	AA
Cluster 14	
799/912	N
799/1099	NW
803/874	A
803/972	AP
803/900	P
803/972	PA
803/940	H
803/1097	HR
840/897	G
840/955	GG
800/897	P
800/955	PG
Cluster 20	
2340/2469	E
2340/2606	EH
2341/2469	Q
2341/2606	QH
2382/2469	S
2382/2606	SQ

(c) 30 minutes

m/z-pair	AA
Cluster 3	
510/567	G
510/624	GG
alt.	
510/567	G
510/666	GV
Cluster 6	
688/802	N
688/931	NE
732/802	A
732/931	AE
678/806	Q
678/935	QQ
Cluster 7	
514/651	H
514/764	HL
580/651	A
580/764	AL
526/689	Y
526/803	YN
526/940	YNH
alt.	
526/689	Y
526/853	YY
Cluster 10	
2141/2198	G
2141/2285	GS
2141/2382	GSP
Cluster 50	
3802/3873	A
3802/4029	AR
3802/4160	ARM
3895/4024	E
3895/4081	EG
Cluster 195	
4267/4395	Q
4267/4558	QY

Chapter 8

Discussion and Conclusion

In this thesis an automated analysis pipeline for MALDI IMS data handling and evaluation was developed and implemented. The pipeline included a pre-phase of data reduction, a mid-phase of co-localisation comparison and clustering and a post-phase of within-cluster comparative analysis. A novel data reduction strategy was proposed in order to account for the spatial information to extract images with expression over very small areas and prevent incorrect merging of peaks. Lastly an attempt to identify differences between the degradomes of heat stabilized versus non-heat stabilized mouse brain tissue was done.

8.1 Experimental Optimisation

Spotting matrix upon HS tissue has never been done before, which is why a lot of methodology optimisation had to be done. The HS tissue is much more fragile and behaves differently than the ordinary SF tissue. The printed matrix droplets expanded to approximately an area of three times the size of the droplet area on SF tissue. Which is mainly why a low resolution of 400 μm was chosen in the experimental setup.

Attempts to increase the surface tension of the matrix droplets were performed by increasing the concentration of water, while decreasing the concentration of methanol (methanol concentrations of 10%, 25% and 50% were tested). However, the initial methanol concentration of 50% seemed to work best, resulting in three times the area size of droplets on SF tissue but with good crystal formation.

A consequence of having a larger droplet area is that the probability that the laser hits the crystals is smaller than with a tighter area, since the laser is collecting the 800 shots with a movement of random walk. To account for this, larger steps for the random walk in the Ultraflex II TOF/TOF were chosen for the HS tissues.

An alternative way to avoid this issue is to distribute the matrix with an electrospray, which sprays an even layer of matrix over the tissues. However, using an electrospray will enable less peptide extraction from the tissues since the matrix will not penetrate the tissue as much as when being printed upon (Since the matrix is allowed to be absorbed between the passes when printing, we can add more matrix per area unit and thereby get higher penetration). There might also be a risk of spatial molecule diffusion when applying with an electospray, which will imply less reliable and accurate m/z images.

Another fundamental conclusion drawn for the experimental setup, is that all experiments have to be performed simultaneously and preferably, if possible, placed on the same glass slide. As concluded by figure 7.5 the SF replicate 3 does differ a lot from replicate 1 and 2. Replicate 1 and 2 does contain a bit variability, probably arisen from the fact that they were placed on different glass slides and by that not washed simultaneously. But the variability between those are minor in comparison to the variability between those and replicate 3. The PCA, figure 7.5, also showed that replicate 3 was very separated from the other SF replicates, projected onto the linear space spanned by the two first principal components. The technique related variability, depicted in principal component 2, also seem to be larger than the time-to-time differences, which makes it very hard to draw any reliable conclusions for the time-to-time differences. This demonstrates the major importance of having a standardised experimental setup in MALDI IMS, where everything is performed simultaneously for all samples to be compared.

8.2 Future Work

The result of this thesis is just to be taken as an indicator of potential markers. It is a first step towards finding a robust and reliable way to measure proteolytic degradation. L. Anderson stated in his article about proteomics ”While observing 10,000 proteins in 10 samples may be considered good discovery technology, measuring 10 proteins in 10,000 samples is likely to be far more effective in finding a real biomarker (assuming of course that one chooses the 10 proteins wisely or luckily)” [25]. This thesis provides a reasonable way to choose those 10 proteins wisely, relying on the wisdom of computer algorithms. The next step for the candidate markers found in this thesis would be to be further validated, either with a reproducibility test using a greater number of biological replicates or/and by investigating the candidates in mouse brain tissue with ordinary mass spectrometry. A less demanding validation method could also be to BLAST a sequence of amino acids (or m/z values) provided by the degradation ladders seen in table 7.3, 7.4 and 7.5, to see if they have been observed in previous studies or are part of a larger known polypeptide. One might also perform a mass matching, comparing the masses of the interesting m/z values, m/z-pairs and m/z-pair chains with the masses of known neuropeptides. For further SF/HS difference evaluation one can compute the difference in the PM time point mean spectras within a replicate. This will show which and how many peaks that decrease/increase in intensity in the different replicates (SF and HS) and by that one can arbitrary quantify the changes in the HS versus the SF samples.

Because of time constraints, the profile spectras for HS replicate 1 and 3 were never interrogated. If adding these data to the HS versus SF difference evaluation one can draw stronger conclusions.

Another approach that could be taken, is to run this analysis on a computer cluster, leaving out the data reduction step. Although some local clusterings have to be performed to remove image redundancies that derive from the same MS peak. Taking this approach will however most probably lead to overwhelming results. The option of running the analysis in this study on a cluster was considered, but since the aim was to demonstrate proteolytic degradation not to find all products of degradation, it was dismissed. But yet, we were aware of the risk of losing true interesting images in the data reduction step.

Considering the data reduction, more work and validation needs to be done on the spatial peak detection algorithm. The images chosen by this algorithm have to be further interrogated to reveal if these are truly interesting or if this approach might need higher resolution images or more rigorous conditions. As depicted in figure 7.2, the data reduction algorithm based on spatial TIC is more reproducible than the data reduction based on the spatial peak detection with current settings. There are more similar m/z values between replicates with the spatial TIC based than for the spatial peak detection based, although these similarities should be taken with a grain of salt since the replicates are not biological replicates. Another notable phenomena seen in figure 7.2 is that the HS tissues seem to have more local spatial peaks (derived from the total number of chosen images for each sample). What the reason for this might be is still to be investigated.

To sum up, there might be more to the datasets then were generated in this thesis. But the fact that the ionization does not occur in a standardised and controllable way is still unavoidable. It is well known that the full molecular content in a tissue sample is not ionized, partly because of differential ionization efficiency and ion suppression, due to varying sample complexity and composition. Today there is no method to estimate the ionized proportion, which makes it hard to amend the datasets with normalization. It is also known that a side effect of the MALDI process is that the applied matrix is detected in the mass spectrum and interfere with the m/z spectrum, especially in the m/z regions below 1000 [26]. With todays technologies we will never be sure that three datasets in a time serie will be completely comparable. Hence there is no guarantee that more computer power, enabling analysis of more data, will lead to better and more accurate results.

All in all, this thesis provides a new and innovative strategy of extracting m/z values from MALDI IMS data related to degradation, using the full range of information provided by MALDI IMS. It also suggests a progressive approach to tackle the challenges of MALDI IMS data handling and data reduction by compressing the hyperspectral datacubes into non-redundant, interesting images either featuring a local maximum of the spatial TIC (a peak top in the mean spectra) or containing very local spatial peaks/peak areas. A data reduction scheme that take the images spatial information into account, to avoid any biomolecules with overlapping MS peaks to be merged or missed.

8.3 Future Applications

In the near future, ten years from now, L. Anderson foretell that medical effort will be redirected from disease treatment to disease prevention based on early detection. The research funding will shift from drug development to biomarker development, with the argument that prevention and early stage treatment is much less expensive than late stage and potentially lifelong treatment [25]. Major efforts will be placed on biomarker development for various areas of use. At this stage, in early development, the analysis procedure and the taken strategy in this thesis could be of use. The analysis procedure provides a rough and straight forward way to view and extract potentially interesting m/z values from a dataset. These can then be further investigated with various methods, step-wise decreasing the amount of biomarker candidates.

Bibliography

[1] Cazares, L. H. *et al.* (2011) MALDI Tissue Imaging: From Biomarker Discovery to Clinical Application. *Analytical and Bioanalytical Chemistry* , **401**, 17-27.

[2] Table of Pharmacogenomic Biomarkers in Drug Labels. (2015, May 20) *U.S. Food and Drug Administration.* Table available at http://www.fda.gov/Drugs/ScienceResearch/ResearchAreas/Pharmacogenetics/ucm083378.htm

[3] Febbo P.G. *et al.* (2011) NCCN Task Force Report: Evaluating the Clinical Utility of Tumor Markers in Oncology. *Journal of National Comprehensive Cancer Network*, **9**, S1-S32.

[4] Sethi S. *et al.* (2013) Clinical Advances in Molecular Biomarkers for Cancer Diagnosis and Therapy. *International Journal of Molecular Scince*, **14(7)**, 14771-14784.

[5] Goodwin, R. J. A. *et al.* (2010) Stopping the Clock on Proteomic Degradation by Heat-Treatment at the Point of Tissue Excision. *Proteomics*, **10**, 17511761.

[6] Svensson M. *et al.* (2009) Heat Stabilization of the Tissue Proteome: A New Technology for Improved Proteomics. *Journal of Proteomics*, **8**, 974981.

[7] Borén, M. (2013) Improving Molecular Distribution Analysis. *GEN*, **33**.

[8] Scholz B. *et al.* (2011) Impact of Temperature Dependent Sampling Procedures in Proteomics and Peptidomics–A Characterization of the Liver and Pancreas Post Mortem Degradome. *Molecular & Cellular Proteomics*, **10(3)**, M900229-MCP200.

[9] Zhang X. *et al.* (2012) High Identification Rates of Endogenous Neuropeptides from Mouse Brain. *Journal of Proteome Research*, **11**, 2819-2827.

[10] Bruice, P. Y. (2010) *Essential Organic Chemistry*, Pearson Education, 2nd edition, 370-372.

[11] Norris, J. L. and Caprioli, R. M. (2013) Analysis of Tissue Specimens by Matrix-Assisted Laser Desorption/Ionization Imaging Mass Spectrometry in Biological and Clinical Research. *Chemical reviews*, **113**, 2309-2342.

[12] Jones, E. A. *et al.* (2012) Imaging Mass Spectrometry Statistical Analysis. *Journal of Proteomics*, **75**, 4962-4989.

[13] Aerni, H. R *et al.* (2006) Automated Acoustic Matrix Deposition for MALDI Sample Preparation. *Analytical Chemistry*, **78**, 827-834.

[14] Cornett, D. S. *et al.* (2006) A Novel Histology-Directed Strategy for MALDI-MS Tissue Profiling That Improves Throughput and Cellular Specificity in Human Breast Cancer. *Molecular & Cellular Proteomics*, **5**, 1975-1983.

[15] Goodwin, R. J. A. (2012) Sample preparation for mass spectrometry imaging: Small mistakes can lead to big consequences. *Journal of Proteomics*, **75**, 4893-4911.

[16] Alexandrov, T. (2012) MALDI imaging mass spectrometry: statistical data analysis and current computational challenges. *Bioinformatics*, **13**, S11.

[17] Hanrieder, J. *et al.* (2012) MALDI Imaging Mass Spectrometry of Neuropeptides in Parkinson's Disease *Vis. Exp.*, **60**, e3445, doi:10.3791/3445

[18] Sun, C. S. *et al.* (2010) Recent Advances in Computational Analysis of Mass Spectrometry for Proteomic Profiling. *Journal of Mass Spectrometry*, **46**, 443-456.

[19] SCiLS (2016, January 18) *SCiLS Lab 2014b.* Software available at http://scils.de/software/download/

[20] Nanda, S. J. *et al.* (2014) A New Density Based Clustering Algorithm for Binary Data Sets. *High Performance Computing and Applications (ICHPCA), 2014 International Conference on*, 1-6.

[21] Alexandrov, T. and Bartels, A. (2013) Testing for Presence of Known and Unknown Molecules in Imaging Mass Spectrometry. *Bioinformatics*, **29**, 2335-2342.

[22] Hanrieder, J. *et al.* (2012) MALDI Imaging Mass Spectrometry of Neuropeptides in Parkisonś Disease. *Journal of Visualized Experiments*, **29**, e3445.

[23] Palmer, A. *et al.* (2014) Using Collective Expert Judgements to Evaluate Quality Measures of Mass Spectromtry Images. *Bioinformatics*, **31**, i375-i384.

[24] Wijetunge C. D. *et al.* (2015) EXIMS: An Improved Data Analysis Pipeline Based on a New Peak Picking Method for EXploring Imaging Mass Spectrometry Data. *Bioinformatics*, **31**, 1-9.

[25] Anderson, L. (2014) Six Decades Searing for Meaning in the Proteome. *Journal of Proteomics*, **107**, 24-30.

[26] Fonville, J. M. *et al.* (2012) Robust Data Processing and Normalization Strategy for MALDI Mass Spectrometric Imaging. *Analytical Chemistry*, **84**, 1310-1319.

[27] McDonnell, L. A. *et al.* (2008) Mass Spectrometry Image Correlation: Quantifying Colocalization. *Journal of Preoteome Research*, **7**, 3619-3627.