

# 07–MathDAMP–Reference

This notebook provides reference documentation for MathDAMP functions.

## DAMPAlign

DAMPAlign[msdata,shiftfunction,timepoints,options] aligns msdata according to timeshift function shiftfunction and selects (by interpolation) time points identical to the ones in the timepoints list.

Options:

SampleNameSuffix - string to be added to the SampleName from the msdata to keep track of modifications performed on the dataset (default: "a")

## DAMPAlignAnnotationTable

DAMPAlignAnnotationTable[peaklist,annotationtable,options] aligns annotationtable to peaklist and returns the new annotation table. Options for DAMPFitShiftFunction, which is used internally, may be passed directly via options. Additional option Resolution determines the rounding of m/z values in the annotation table. If the resulting annotation table is intended to be used on a density plot where the underlying data were binned to 0.1 m/z resolution, the resolution option should be set to 0.1 as well (default: 1)

## DAMPAlignPeakList

DAMPAlignPeakList[peaklist,timeshiftfunction] applies the retention/migration time shift function timeshiftfunction to all retention/migration times in the peaklist (expected to have format as returned by the function DAMPPickPeaks)

## DAMPAnnotationTableToPeakList

DAMPAnnotationTableToPeakList[annotationtable,options] converts an annotation table (execute ?DAMPLoadAnnotation for format description) to a peaklist with format as returned by the function DAMPPickPeaks.

Options:

Resolution - determines to which resolution the m/z values in the annotation table will be rounded (default: 1)

## DAMPAppendToSampleName

DAMPAppendToSampleName[msdata4,string] appends string to the SampleName in msdata4. msdata4 is expected to be a list of rules (format as the fourth element in the msdata format). The added string is enclosed inside two vertical bars. If previous modifiers are present, the newly added string will be separated from these by a single vertical bar.

Options:

ShareBars - if set to False, the newly added string is separated from previous modifiers by two vertical bars with a space in between (default: True)

## DAMPApplyFunctionToGroup

DAMPApplyFunctionToGroup[msdatas,function,options] applies function to either all corresponding signal intensities in the datasets or to the msdatas (decided by option).

Options:

ApplyToIntensitiesOnly - if set to true, the function is applied to the corresponding signal intensities in the msdatas (instead of to the whole msdatas list) (default: True)

ResultSampleName - string to be set as the SampleName of the resulting dataset. If set to Automatic, the SampleName of the first dataset from msdatas is used (default: Automatic)

SampleNameSuffix - string to be added to the SampleName to keep the track of the modifications performed on the dataset (default: "")

## DAMPApplyFunctionToSingle

DAMPApplyFunctionToSingle[msdata,function,options] applies a pure function function to msdata.

Options:

SampleNameSuffix - string to be added to the SampleName from the msdata to keep track of modifications performed on the dataset (default: "")

## DAMPBinChromatogram

DAMPBinChromatogram[chromatogram,binsize,options] bins the chromatogram by dissecting it into intervals of size binsize (minutes) and applying the binning function passed as an option to signal intensities and timepoints in every interval.

Options:

BinningFunction - function to be applied to signal intensities and timepoints in every interval (default: Mean)

## DAMPCheckDir

DAMPCheckDir[directory] checks if directory exists, creates it if it does not. The parent directory must exist, error message is shown otherwise.

**DAMPCrop**

DAMPCrop[msdata,options] reduces the msdata dataset to datapoints falling within the timerange and m/z value range specified by options:  
 mzRange - two element list specifying the cropping m/z value range (default: All)  
 TimeRange - two element list specifying the cropping time range (default: All).

**DAMPDensityPlot**

DAMPDensityPlot[msdata,options] plots the msdata using the ListDensityPlot function. A gradient palette is used for representing the signal intensities and the plot may be annotated to allow easier identification of peaks.  
 Options:  
 MaxScale - determines the extent of the signal intensity scale (default: Automatic)  
 LogScale - determines whether the signal intensities should be displayed using a logarithmic scale (default: False)  
 FrameTickFreqs - frequencies at which tickmarks are placed on the time axis (frequency in minutes) and on the m/z axis (in terms of the number of elements in the msdata's mz list) (default: {1,1})  
 FrameTickOffsets - determines the positions with respect to the origin of msdata's dimensions at which the tickmarks are started to be placed (default: {0,0})  
 mzTickShift - shift of tickmark positioning on the m/z axis (default: -0.5)  
 mzFrameTicks - a list of custom frame tickmarks to be shown on the m/z axis (default: Automatic)  
 Palette - list of color specifications to be used for representing the signal intensity values (default: DAMPGradientPalette[])  
 mzGridLineFreq - frequency of horizontal gridlines in terms of the number of elements in the msdata's m/z value list. Enter a list of values to place the gridlines at certain specified positions (default: 5)  
 mzGridLineStyle - style options for horizontal gridlines (default: {{AbsoluteThickness[0.25],GrayLevel[0.5]},{AbsoluteThickness[0.25],GrayLevel[0.5],Dashing[{0.01,0.01]}}})  
 AnnotationTables - list of annotation tables to be overlaid on the plot (default: None)  
 PlotOptions - options for the ListDensityPlot function (default: {Mesh->False,ImageSize->930,AspectRatio->0.35,FrameLabel->{"Time (min)", "m/z"},TextStyle->DAMPTextStyle})  
 AnnotationOptions - lists of options for the DAMPDrawAnnotation function for each annotation table. The number of lists of options must correspond to the number of passed annotation tables (default: Automatic)

**DAMPDPSScore**

DAMPDPSScore[peaklist1,peaklist2,options] calculates the dynamic programming (DP) score between two peaklists having format as returned by the DAMPPickPeaks function.  
 Options:  
 GapPenalty - gap penalty (in minutes) for DP (default: 0.5)

**DAMPDrawAnnotation**

DAMPDrawAnnotation[annotationtable,options] draws the annotation for density plots.  
 (For details about the format of the annotationtable, execute ?DAMPLoadAnnotationTable.)  
 Options:  
 TimeRange - time range for which the annotation should be drawn (default: {0,50})  
 MZs - list of m/z values for which the annotation should be drawn. The order of m/z values in the mzs list determines the positioning of annotation labels along the m/z scale (default: Range[50.,1000.,1.])  
 Resolution - resolution to which the underlying data were binned. The m/z values from the annotation table will be rounded accordingly to ensure the appropriate appearance of the annotation labels (default: 1)  
 ScaleCoefficients - coefficients used to convert the scales (currently used for timescale only, the second item is 1 by default) of the annotation table to the scales of the ListDensityPlot (default: {62.2488,1})  
 TextStyle - style options for the annotation label's text (default: Join[{FontSize->4},DAMPTextStyle])  
 LabelStyle - style options for the annotation label's mark (default: {AbsoluteThickness[.25]})  
 LabelShape - pure function to draw the annotation label's mark.  
 Two parameters are passed to the function, first determining the position {xpos, ypos}, and second determining the size {xsize,yysize}. (default: Circle[#1,#2]&)  
 LabelSize - size of the annotation label's mark (relative units). When relative size {1,1} is used by the Circle function as radius, an ellipse is drawn which appears to appropriately capture the dimensions of peaks on the ListDensityPlot (default: {1,1})

**DAMPDropMZs**

DAMPDropMZs[msdata,mzs] reduces the msdata dataset by eliminating datapoints corresponding to m/z values specified in the mzs list.

**DAMPExportBDT**

DAMPExportBDT[filename,msdata] exports data in a MathDAMP format {matrix of signal intensities, list of m/z values, list of timepoints, additional information (list of rules)} into a BDT format file specified by filename. The m/z values in the data are not saved one by one but as a range specified by min, max and step. Because of this, only data having a regular list of m/z values which can be converted to the range representation should be exported.

**DAMPFilter**

DAMPFilter[msdata,criteriamsdata,threshold,options] sets to 0 those signal intensities in msdata for which the absolute values of corresponding signal intensities in the criteriamsdata dataset are not equal to or greater than threshold.

Options:

SampleNameSuffix - string to be added to the SampleName from the msdata to keep track of modifications performed on the dataset (default: "f")

DAMPFilter[msdata,criteriamsdata,filterfunction,options] pure function filterfunction is applied to the signal intensity matrix of criteriamsdata. For zero signal intensities in the result, the corresponding signal intensities in the msdata dataset are set to 0 as well.

Options:

SampleNameSuffix - string to be added to the SampleName from the msdata to keep the track of modifications performed on the dataset (default: "f")

### DAMPFitShiftFunction

DAMPFitShiftFunction[peaklist1,peaklist2,options] optimizes the parameters of a retention/migration time shift function so that when applied to peaklist2 the optimum peak alignment (as measured by DAMPDPScore) is achieved.

Options:

ShiftFunction - pure function to be used as a retention/migration time shift function (default:  $(1/(1/(\alpha \#)+\gamma/2)\&)$ )  
 ShiftFunctionParameters - list of parameters for optimization. If Automatic is specified, these are extracted automatically from ShiftFunction. The parameters may be also specified explicitly with seek ranges (default:  $\{\{\alpha,0.8,1.2\},\{\gamma,-0.04,0.04\}\}$ )  
 GapPenalty - gap penalty value to be passed to the DAMPDPScore function for the scoring of the alignment. A list of gap penalty values may be passed to perform the fitting of the retention/migration time shift function iteratively (default: {3,0.5})  
 NMinimizeOptions - list of options to be passed to the NMinimize function used for optimization (default: {MaxIterations->1000})  
 TimeRange - specifies the selection time range of peaks from peaklist1 to be used for scoring (default: {0, $\infty$ })

### DAMPGenColors

DAMPGenColors[number] generates a list of color specifications with a number of elements determined by the parameter number. If number $\leq$ 6, the colors are assigned from the following sequence: blue, red, green, yellow, cyan, magenta. If number $>$ 6, the hue range is split proportionally and the colors are assigned from this sequence.

### DAMPGetChromatogram

DAMPGetChromatogram[msdata,mz] returns a chromatogram/electropherogram ( a list of {time,signal intensity} elements) from msdata corresponding to m/z mz.

### DAMPGetIntensities

DAMPGetIntensities[msdata,mz] returns a list of signal intensities from msdata corresponding to m/z mz.

### DAMPGradient

DAMPGradient[pos1,pos2,color1,color2,backgroundgraylevel] returns a function which calculates the parameters for the Hue function of a specific point within a gradient. The relative location  $\langle 0;1 \rangle$  of the point of interest within the gradient is the only parameter of the returned function. The color gradient consists of three subgradients defined by the parameters of the DAMPGradient function. The gradient starts with backgroundgraylevel  $\langle 0;1 \rangle$  at relative position 0, then proceeds to color1 (numeric value corresponding to hue specification  $\langle 0;1 \rangle$ ) at pos1, and color2 at pos2, and finally to darkened color2 at the relative position 1.

### DAMPGradientPalette

DAMPGradientPalette[options] - generates a color palette for assigning colors to intensity values on density plots. It uses the DAMPGradient function to create the gradients for both positive and negative ranges. Both gradients consist of three subgradients (execute ?DAMPGradient for details).

Options:

ColorPositions - list with two elements specifying the relative location within the gradient of endpoints of the middle subgradient (default: {.075,.4})  
 PositiveColors - hue color specification of the endpoints of the middle subgradient in the positive range color gradient (default: {1/6,0})  
 NegativeColors - hue color specification of the endpoints of the middle subgradient in the negative range color gradient (default: {1/2,2/3})  
 BackgroundGrayLevel - gray level of the density plot's background area (default: .7)  
 PaletteSize - number of palette entries (subdivisions) of both positive and negative range color gradients, palette entry for value 0 is shared between the two (default: 100)

### DAMPImportBDT

DAMPImportBDT[filename,samplename] reads and processes a binary BDT file specified by filename into a Math DAMP format {matrix of signal intensities, list of m/z values, list of timepoints, additional information (list of rules)}. The BDT is a binary datafile used internally by The Institute for Advanced Biosciences (Keio University). Separate in house software for preprocessing of the Analyst QS generated csv files stores the data in the BDT format. Since the original csv file does not contain information about the sample name, this has to be specified as a second parameter (samplename).

### DAMPImportCDF

**DAMPImportCDF**[filename,options] reads and processes a CDF file specified by filename into a *MathDAMP* format {matrix of signal intensities, list of m/z values, list of timepoints, additional information (list of rules)}. This function was not thoroughly tested with CDF files generated by different software and may not be universal.

Options:

Resolution - specifies the resolution to which the data will be binned along the m/z axis (default: 1)

#### **DAMPImportCSV**

**DAMPImportCSV**[filename,samplename,options] reads and processes a csv file generated by Analyst QS software for Agilent TOFMS and specified by filename into a *MathDAMP* format {matrix of signal intensities, list of m/z values, list of timepoints, additional information (list of rules)}. Since the csv file does not contain information about the sample name, this has to be specified as the second parameter (samplename).

Options:

Resolution - specifies the resolution to which the data will be binned along the m/z axis (default: 1)

Verbose - if set to True, prints a notification along with current time every 100 processed chromatograms/electropherograms (default: True)

#### **DAMPImportMS**

**DAMPImportMS**[filename] reads and processes an Agilent ChemStation MS file specified by filename into a *MathDAMP* format {matrix of signal intensities, list of m/z values, list of timepoints, additional information (list of rules)}. This function is universal for SIM and scan mode datafiles. The data are binned to 1 m/z unit resolution by averaging.

#### **DAMPImportMZXML**

**DAMPImportMZXML**[filename,samplename,options] reads and processes the first level MS scans from an mzXML data file specified by filename into a *MathDAMP* format {matrix of signal intensities, list of m/z values, list of timepoints, additional information (list of rules)}. The sample name is specified via the second parameter (samplename).

Options:

Resolution - specifies the resolution to which the data will be binned along the m/z axis (default: 1)

ImportMode - determines the m/z dimension elements. If set to Sequential, the m/z dimension is represented by a discrete range of values determined by the smallest and the largest m/z value in the imported dataset (rounded according to the resolution). Resolution determines the stepsize of the discrete range. If set to Selective, the m/z dimension is represented by only those m/z values which are present in the imported dataset (rounded according to the resolution). Sequential mode is recommended for scan data, Selective mode is recommended for SIM data. (default: Sequential)

#### **DAMPIntegrate**

**DAMPIntegrate**[chromatogram,options] calculates the area below the signal intensities of a chromatogram/electropherogram within the retention/migration time range specified by the option TimeRange (default: {0,∞}). Baseline may be calculated as an average of signal intensities within the timerange specified by the options BaselineFromTimeRange (default: None). If set to None, baseline is set to signal intensity value 0.

#### **DAMPLoadAnnotationTable**

**DAMPLoadAnnotationTable**[filename] loads a list of annotation labels for plots from a csv file. The organization of columns in the csv file is assumed to be as follows: (1) m/z value, (2) retention/migration time, (3) short compound name/id (density plot label text), (4) full compound name (chromatogram/electropherogram label text), and (5) label text position on the density plots relative to the label (1 - right, 2 - top, 3 - left, 4 - bottom, 1.5 - top right, etc.).

#### **DAMPMultiGroups**

**DAMPMultiGroups**[msdatas,replicates,options] performs the comparison of multiple groups of datasets containing identical number of replicates determined by the replicates parameter. msdatas is expected to contain an ungrouped list which will be later grouped according to the number of replicates. The results are returned as a list of rules {NormalizedDatasets->...,AlignedAnnotationTables->...,FRatios->...,GroupNames->...}

Options:

NormalizeGroupOptions - a list of options for the **DAMPNormalizeGroup** function which is used internally to normalize the datasets (default: {})

GroupNames - names to assign to groups (will be combined into the SampleName of the resulting FRatios dataset). If set to Automatic, the SampleName of the first dataset from every group is used (default: Automatic)

#### **DAMPNormalize**

**DAMPNormalize**[msdata,coefficient,options] multiplies the signal intensities in msdata (msdata[[1]]) with coefficient.

Options:

SampleNameSuffix - string to be added to the SampleName from the msdata to keep the track of modifications performed on the dataset (default: "n")

#### **DAMPNormalizeGroup**

**DAMPNormalizeGroup**[msdatas,options] aligns msdatas (a list of datasets) and normalizes them according to the areas of the peaks of the internal standard and external normalization coefficients (optional). The results are returned as a list of rules: {NormalizedDatasets->...,AlignedAnnotationTables->...}

Options:

- Reference - position of the reference dataset within msdatas to which the remaining msdatas will be normalized (default: 1)
- AlignmentTimeRange - peak picks from the reference dataset falling within this timerange only (specified as {starttime,endtime} in minutes) will be used for alignment (default: All)
- RepresentativePeakOptions - options to be passed to the DAMPSelectRepresentativePeaks function to filter the initial peak picks (default: {PeaksPerChromatogram->5,PeaksPerInterval->8,IntervalSize->.5})
- PeakPickingOptions - options to be passed to the DAMPPickPeaks function (default: {Threshold->5000})
- PeakLayoutPlotOptions - options to be passed to the DAMPPlotPeakLayout function (default: {})
- FitShiftFunctionOptions - options to be passed to the DAMPFitShiftFunction function (default: {})
- AnnotationTables - a list of annotation tables to be aligned to the reference msdata (default: None)
- OutputTimeRange - time range to which the resulting normalized datasets should be cropped (default: All)
- ExternalNormalizationCoefficients - list of coefficients by which the signal intensities in msdatas will be multiplied. The number of coefficients in the list must equal the number of datasets in msdatas (default: None)
- Resolution - resolution to which the datasets were binned along the m/z dimension. The annotation tables passed through the AnnotationTables option will be binned the same way (default: 1)
- InternalStandard - the internal standard for signal intensity normalization may be specified in one of two ways: 1) short name (3rd column) from the first annotation table in the list passed via the AnnotationTables option. In this case the position of the internal standard will be extrapolated from the aligned annotation table and the vicinity blindly integrated 2) specifying the m/z and integration time range (as in the reference dataset) explicitly: {mz,{starttime,endtime}}. (default: None)
- AutoISIntegrationVicinity - if the location of the internal standard is extrapolated from the aligned annotation table, this option determines the vicinity (in minutes) of the predicted retention/migration time to be blindly integrated (default: {-.25,.25})
- SaveMemory - if set to true, signal intensities are rounded to integers in internal calculations and results (default: True)

### DAMPOutliers

**DAMPOutliers**[msdatas,options] highlights the presence of outlying signal intensities in msdatas (a list of datasets). Two types of result datasets are generated: a z-score map and a quartile-based result. These datasets show the result for the most outlying signal intensity from within the set of corresponding signal intensities. Optionally, the results may be generated individually for every dataset in msdatas. The results are returned as a list of rules: {NormalizedDatasets->...,AveragedDatasets->...,ZScores->...,ZScoresIndividual->...,QuartileResult->...,QuartileResultIndividual->...,AlignedAnnotationTables->...}

Options:

- NormalizeGroupOptions - a list of options for the DAMPNormalizeGroup function which is used internally to normalize the datasets (default: {})
- OutliersToDrop - specifies the number of outliers to be dropped from the set of corresponding signal intensities prior to calculating the mean and standard deviation for the z-score calculation (default: 1)
- IndividualZs - if set to true, z-score map will be calculated for every dataset in msdatas in addition to the overall result (default: False)
- IndividualQs - if set to true, results based on the interquartile range will be generated for every dataset in msdatas in addition to the overall result (default: False)

### DAMPParallelPlot

**DAMPParallelPlot**[msdatas,options] plots msdatas (a list of datasets) on a density plot in a parallel format, so that chromatograms/electropherograms from the datasets corresponding to the same m/z value appear next to each other. The datasets do not have to have chromatograms corresponding to an identical set of m/z values, neither do the datasets have to be aligned. In the latter case, the time axis is labeled according to the first dataset in msdatas (please note that the time axis may be misleading for the remaining datasets in this case). The options are passed directly to the DAMPDensityPlot function which is used internally for plotting the data.

### DAMPPeakListToAnnotationTable

**DAMPPeakListToAnnotationTable**[peaklist,options] converts the peaklist returned by DAMPPickPeaks to an annotation table format for plots. The peak names are assigned from sequential numbering.

Options:

- TextLabelPosition - specifies the position of the text label relative to the annotation label mark (default: 1.5)

### DAMPPickChromatogramPeaks

**DAMPPickChromatogramPeaks**[chromatogram,options] picks peaks from chromatogram and returns them as list {{Retention Time,Signal Intensity},{Retention Time,Signal Intensity},...}.

Options:

- Threshold - minimum vertical distance of a new strategic point from the line defined by the two neighbouring strategic points. Please refer to the MathDAMP.nb notebook for a more detailed description of the peak picking algorithm (default: 1000)
- NodeRestrictedVicinity - neighbourhood of a strategic point (in minutes) from which new strategic points can not be picked (default: 0.05)
- CentroidVicinity - determines the time interval (in minutes) around the peak-top strategic point for the calculation of centroided retention/migration time (default: 0.05)

### DAMPPickPeaks

DAMPPickPeaks[msdata,options] picks peaks from all chromatograms/electropherograms in msdata and returns them in a list {{m/z value,peaklist},{m/z value,peaklist},...}. The function does not have any default options, the options passed to the DAMPPickPeaks function are passed further to the DAMPPickChromatogramPeaks function which is used internally to pick peaks from individual chromatograms/electropherograms.

#### DAMPPlotCandidates

DAMPPlotCandidates[msdatas,criteriamsdata,options] plots a ranked list of overlaid chromatograms/electropherograms from msdatas datasets in the vicinities of the most significant signals in the criteriamsdata dataset.

Options:

PlotCount - number of chromatograms/electropherograms to plot (default: 12)  
 TimeRange - the most significant signals will be searched in the criteriamsdata dataset in this retention/migration time range only (default: All)  
 PlotVicinity - retention/migration time range to show on the chromatogram/electropherogram around the candidate difference in minutes (default: {-0.6,.6})  
 DropVicinity - after a particular datapoint in the criteriamsdata is selected as the most significant signal, the vicinity of this datapoint determined by this option (in minutes) will no longer be searched in subsequent iterations (default: .2)  
 PlotChromatogramOptions - options to be passed to the DAMPPlotChromatogram function which is used internally to plot the chromatograms/electropherograms (default: {})

#### DAMPPlotChromatogram

DAMPPlotChromatogram[{msdata,msdata,...},mz,options] plots multiple overlaid chromatograms/electropherograms from the list of msdatas corresponding to m/z mz (mz may be a list of m/z values as well). The DAMPGenColors function is used by default to assign colors to individual chromatograms/electropherograms. Custom colors may be specified via the PlotOptions option (and the enclosed PlotStyle option).

Options:

PlotOptions - list of options for the MultipleListPlot function which is used internally for plotting (default: see the output of Options[DAMPPlotChromatogram]),  
 AnnotationTable - annotation table for the labeling of the chromatogram/electropherogram (default: {})  
 Resolution - resolution to which the msdata were binned along the m/z dimension. The annotation table passed through the AnnotationTable option will be rounded accordingly to ensure the appropriate appearance of annotation labels on the chromatogram/electropherogram (default: 1)  
 LegendData - a list of legend elements {{color,label},...}. If Automatic is specified, the SampleName from each msdata is used as a label (default: Automatic)

#### DAMPPlotPeakLayout

DAMPPlotPeakLayout[peaklists,options] plots the positions of peaks from a list of peaklists which are expected to have a format as returned by DAMPPickPeaks. Symbols' colors are by default assigned using the DAMPGenColors function. Custom colors may be passed via the PlotOptions option (and the enclosed SymbolStyle option).

Options:

mzGridLineFreq - frequency of horizontal gridlines on plot (default: 1)  
 mzTickFreq - frequency of labeled tickmarks on the m/z axis (default: 1)  
 LegendItems - list of strings to be shown as legend labels (default: Automatic)  
 PlotOptions - list of options to be passed to the MultipleListPlot function used internally to generate the plot (default: {TextStyle→DAMPTextStyle,Frame→True,ImageSize→600,SymbolShape→{PlotSymbol[Diamond,5,Filled→False],PlotSymbol[Box,5,Filled→False],PlotSymbol[Diamond,3,Filled→False],PlotSymbol[Box,3,Filled→False]})

#### DAMPRemoveNoise

DAMPRemoveNoise[msdata,options] removes noise from every chromatogram/electropherogram in msdata by leveling to 0 all signal intensities, absolute values of which are smaller than a threshold. The threshold is calculated as a specific multiple of the standard deviation of signal intensities from a specified time range of every chromatogram/electropherogram.

Options:

TimeRange - time range (in minutes) from which to calculate the standard deviation of signal intensities (default: {1,3})  
 SDThreshold - specifies the multiple of the standard deviation of signal intensities from the selected time range to be used as the noise discrimination threshold (default: 5)  
 LevelNegativeSignals - if set to True, all negative signal intensities in every chromatogram/electropherogram will be leveled to 0 (default: True)  
 SampleNameSuffix - string to be added to the SampleName from the msdata to keep the track of modifications performed on the dataset (default: "nr")

#### DAMPRemoveSpikes

DAMPRemoveSpikes[msdata,options] levels to 0 all signal intensity values both neighbours of which have 0 signal intensity value in every chromatogram/electropherogram in msdata.

Options:

SampleNameSuffix - string to be added to the SampleName from the msdata to keep the track of modifications performed on the dataset (default: "sr")

#### DAMPRobustPolynomialFit

DAMPRobustPolynomialFit[chromatogram,options] fits a polynomial to chromatogram data by robust nonlinear regression and returns a dataset in which the signal intensity is calculated from the polynomial for each timepoint. The function is intended for baseline fitting. Please refer to the MathDAMP.nb notebook for details about implementation.

Options:

PolynomialDegree - degree of polynomial which is going to be fitted to the data (default: 1)  
 ConvergenceCriteria - convergence criteria for nonlinear regression (default: 0.001)  
 MaxIterations - maximum number of outer (recalculating the weights) iterations (default: 10)  
 b - parameter for calculating weights (default: 4.05)  
 BinningFunction - function to bin the chromatogram data before polynomial fitting. This may increase the speed considerably without considerable loss of precision. Use BinningFunction->None not to bin the data for polynomial fitting (default: (DAMPBinChromatogramFast[#,0.25]&))

#### DAMPSelectMZs

DAMPSelectMZs[msdata,mzs] reduces the msdata dataset to datapoints corresponding to m/z values specified in the mzs list.

#### DAMPSelectRepresentativePeaks

DAMPSelectRepresentativePeaks[peaklist,options] reduces the peaklist returned by the DAMPPickPeaks function to a selected number of highest peaks in every chromatogram/electropherogram and a selected number of highest peaks in every time interval of selected size.

Options:

PeaksPerChromatogram - number of highest peaks to select from every chromatogram/electropherogram (default: All)  
 PeaksPerInterval - number of highest peaks to select from every time interval (default: All)  
 IntervalSize - size of the time interval (in minutes) for selection of highest peaks when PeaksPerInterval is not set to All (default: 1)  
 TimeRange - select peaks from this retention/migration time range only (default: All)

#### DAMPSmooth

DAMPSmooth[msdata,options] applies a smoothing function to all chromatograms/electropherograms in msdata.

Options:

SmoothingFunction - pure function to be applied to the list of signal intensities from a chromatogram/electropherogram (default: DAMPMovingAverageFast[#,9]&)  
 SampleNameSuffix - string to be added to the SampleName from the msdata to keep track of modifications performed on the dataset (default: "s")

#### DAMPSubtractBaselines

DAMPSubtractBaselines[msdata,options] subtracts baselines from all chromatograms/electropherograms in msdata dataset and returns the new processed dataset.

Options:

BaselineFittingFunction - pure function which is expected to return the baseline points for every timepoint in a chromatogram/electropherogram passed as an argument (default: DAMPRobustPolynomialFit)  
 SampleNameSuffix - string to be added to the SampleName from the msdata to keep the track of modifications performed on the dataset (default: "bs")

#### DAMPThreshold

DAMPThreshold[msdata,threshold,options] levels to 0 all signal intensity values in msdata which are within  $\pm$ threshold.

Options:

SampleNameSuffix - string to be added to the SampleName from the msdata to keep track of modifications performed on the dataset (default: "t")

#### DAMPTrendFilter

DAMPTrendFilter[group1msdatas,group2msdatas,tofiltermsdata,filtnum,options] filters out those signal intensities from a result (tofiltermsdata) of two averaged groups of msdatas where at least a certain number of corresponding signal intensities from the individual datasets specified by filtnum follows the same trend as their averages.

Options:

SampleNameSuffix - string to be added to the SampleName from the tofiltermsdata to keep the track of modifications performed on the dataset (default: "tf")

#### DAMPTwoDatasets

DAMPTwoDatasets[msdata1,msdata2,options] generates datasets representing the absolute, relative, and absolute $\times$ relative differences between msdata2 and msdata1 and returns them along with the normalized datasets and aligned annotation tables as a list of rules: {NormalizedDatasets->..., Absolute->...,Relative->...,AbsoluteRelative->...,AlignedAnnotationTables->...}

Options:

NormalizeGroupOptions - a list of options for the DAMPNormalizeGroup function which is used internally to normalize the datasets (default: {})  
 ThresholdForRelative - relative difference in the relative result will be set to 0 if neither of the two corresponding signal intensities from msdata1 and msdata2 are equal to or greater than this threshold (default: 0)

#### DAMPTwoGroups

DAMPTwoGroups[msdatas1,msdatas2,options] generates datasets representing the absolute, relative, and absolute×relative differences between the averaged datasets of msdatas2 datasets and msdatas1 datasets. Additionally, a dataset representing t-scores between the groups of corresponding signal intensities from msdatas1 and msdatas2 datasets is generated. The results are returned as a list of rules: {NormalizedDatasets->..., AveragedGroup1->...,AveragedGroup2->...,Absolute->...,Relative->...,AbsRel->..., FilteredAbsRel->...,TScores->...,AlignedAnnotationTables->...,GroupCounts->...,GroupNames->...}

Options:

NormalizeGroupOptions - a list of options for the DAMPNormalizeGroup function which is used internally to normalize the datasets (default: {})

ThresholdForRelative - relative difference in the relative result will be set to 0 if neither of the two corresponding signal intensities (from the averaged datasets) is equal to or greater than this threshold (default: 0)

GroupNames - names to assign to groups (will be combined into the SampleName of the results). If set to Automatic, the SampleName of the first dataset from every group is used (default: Automatic)

AbsRelTrendFilter - determines the minimum number of individual corresponding signal intensities from every group which must follow the same trend as their averages to remain in the absolutexrelative result. This is intended to filter out results originating from individual spikes or outliers (default: 2)