

Contents

Chapter 1. Introduction	3
Chapter 2. Defining the Standard	5
Chapter 2.1. Defining the ATC parameters	5
Chapter 2.2. Troubleshooting ATC	7
Chapter 3. Analysis of the Unknown	9
Chapter 3.1. Converting to a Base Pair axis	11
Chapter 3.2. Normalising the data	11
Chapter 4. Trace Types in DAX	13
Chapter 4.1. Assigning trace types	13
Chapter 4.1.1. How DAX recognises trace types	15
Chapter 5. Conclusion	17

Chapter 1. Introduction

DAX can be used to read and analyse trace files such as the files that are created by the Amersham MegaBACE™ and ABI Genescan® systems.

The MegaBACE™ and ABI systems, like other DNA analysis systems, can analyse four or five traces in a single run, using four (five) different (coloured) labels. One way to use this feature is to include a standard sample with known DNA fragment sizes as one of the traces, and apply this standard to the other traces. The standard is used to ascertain the fragment sizes of the components in the unknown samples.

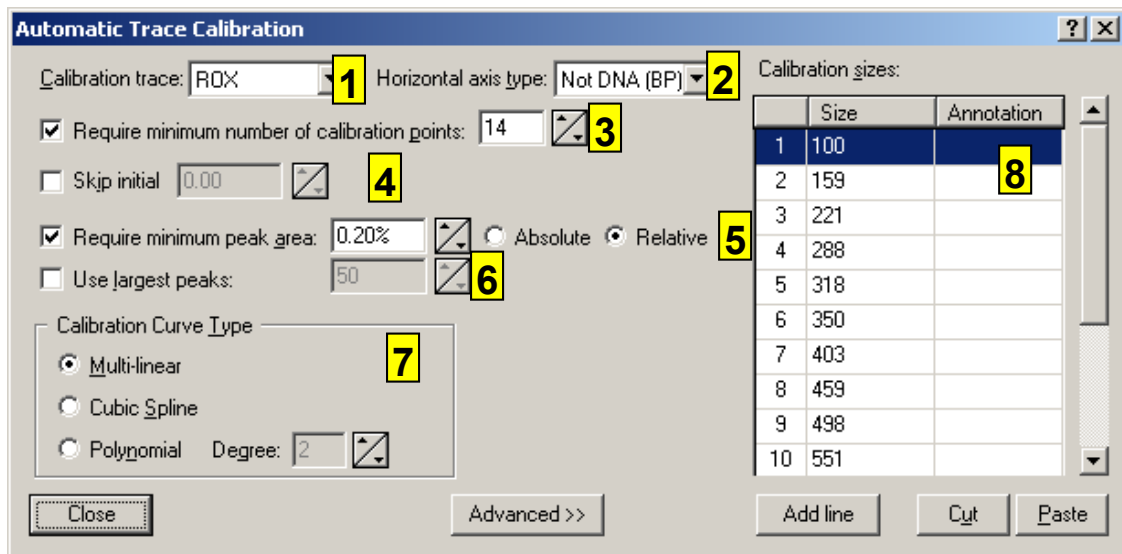
Chapter 2. Defining the Standard

This chapter describes the steps needed to set up the analysis of the sizing standard contained in trace files. Typically, you will go through this process once for a particular sizing standard.

DAX uses a heuristic method called Automatic Trace Calibration to recognise the sizing standard's fragments¹.

Chapter 2.1. Defining the ATC parameters

Use the **Analysis | Edit ATC ..** menu option to invoke the Automatic Trace Calibration, or ATC, dialog box. The dialog box looks like this:



The items in this dialog box have the following meaning:

1. The **Calibration trace** defines the trace that contains the sizing standard, in this instance ROX. For more on trace types, see chapter 4.
2. The **Horizontal axis type** can be used to exclude some measurements from being analysed using ATC. The most common setting is **Not DNA**.

¹ Earlier versions of DAX used a process that recognised the sizing standard's peaks, essentially, by their positions; refer to daxtrace.pdf for an explanation of this procedure.

3. **Require minimum number of calibration points** will often be set to the number of standard sizes, but can be set to a lower value if the sample may not have been analysed completely (i.e. if the analysis was cut short).
4. **Skip initial ...** can be used to disregard the early part of a measurement, which speeds up the analysis, but is not required.
5. **Require minimum peak area ...** can be used to disregard small peaks, which again speeds up the analysis, but is not required. Values of 0.2% or 0.4% generally work well.
6. The **Use largest peaks** option is an alternative way of disregarding small peaks. If n standard peaks are present, a value of $2n$ generally works well.
7. The **Calibration curve type** defines what sort of curve will be drawn through the sizing calibration points.
8. The **Calibration sizes** are a list of the sizes of the fragments in the sizing standard. For the example file L0083.rsd, these sizes are: 100, 159, 221, 288, 318, 350, 403, 459, 498, 551, 600, 633, 700, 726

The **Advanced** items have the following meaning.

- **Start looking for highest size** runs the ATC algorithm “in reverse”, which can sometimes be useful.
- **Allow missing points at both start and end** allows for low end as well as high end calibration points to be missing.
- **Reward higher numbers of calibration points** tells the ATC algorithm that searches that turn up more calibration points are to be considered better.
- **Reward higher total calibration peak area percentage** tells the ATC algorithm that calibration points are likely to have peaks with significant sizes.
- **Reward calibration peak areas being similar** instructs the ATC algorithm to make use of the fact that calibration peaks probably have similar sizes.
- **Maximum allowed discontinuities** can be used if the measurement is known to contain a discontinuity in the horizontal axis. It is rarely necessary to allow discontinuities.
- The **low and high allowed direction ratios** determine how forgiving the ATC algorithm will be. More forgiving settings (further away from 1.0) will result in more inclusive searches, which may take longer to run. Generally, values of 0.7 and 1.5 give excellent results.

Once all fields have been filled, click the **Close** button to close the ATC dialog. You should now use **Analysis | Save ATC parameters** to store this setup on disk.

The file ATC.atc also contains these settings.

Chapter 2.2.Troubleshooting ATC

In the next chapter, Automatic Trace Calibrations will be used to find the standard peaks in a ROX trace. This is normally a very straightforward process. However, should the procedure run into a snag, there are some steps that should be followed.

The first step is to make sure that the calibration trace being looked at does indeed contain all (or some) expected calibration peaks. Here are some reasons why this might not be the case:

- the most common reason is that too large an interval at the start of the trace is being excluded from peak finding, causing one or more initial calibration peaks to be skipped. Make sure not to skip too large an interval at the start of the trace when finding peaks.
- peak find parameters might have too high a detection limit.
- the trace may actually not contain the calibration peaks, or not all of them. If 10 calibration peaks are expected, but only 8 peaks are present, ATC will fail if the **require minimum number of calibration points** parameter was set higher than 8.

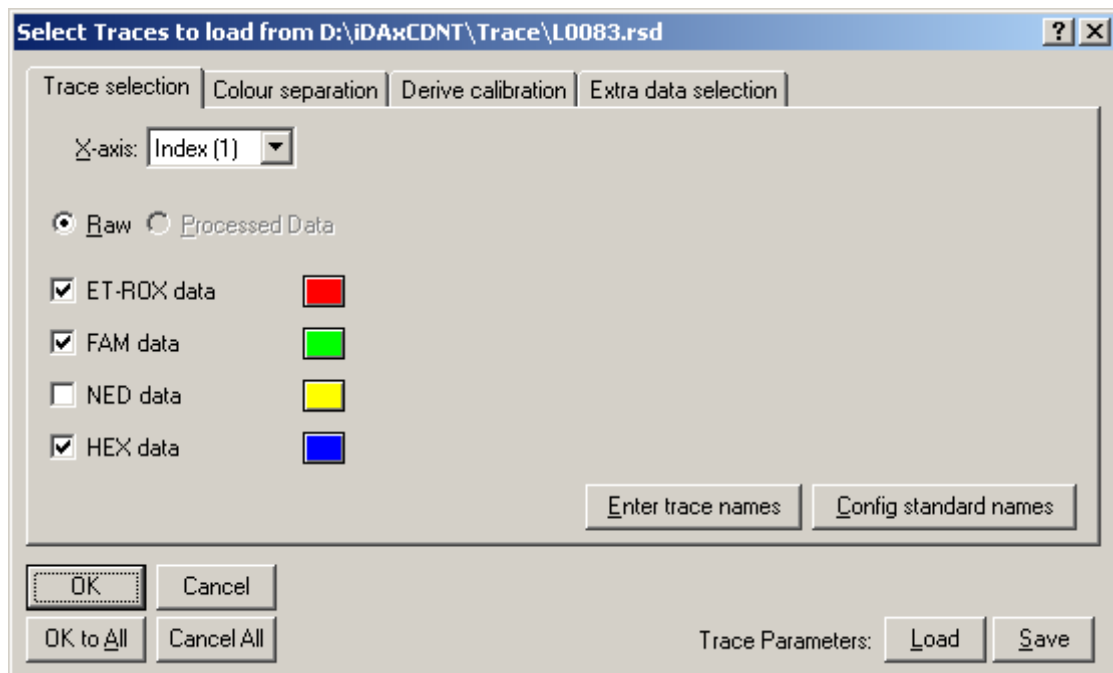
If the calibration peaks are present and found correctly, but ATC still fails, some parameters may have been given values that are too limiting. It is important to understand that most ATC parameters (3 .. 6 in the dialog box above) are used to speed up the ATC algorithm, without being strictly necessary for the algorithm to work. Therefore, if the ATC fails to find a calibration, these parameters can be relaxed, advisably in this order:

- skip a smaller interval (4)
- require a smaller peak area (5)
- use more peaks (6)
- require fewer points to be present (3)

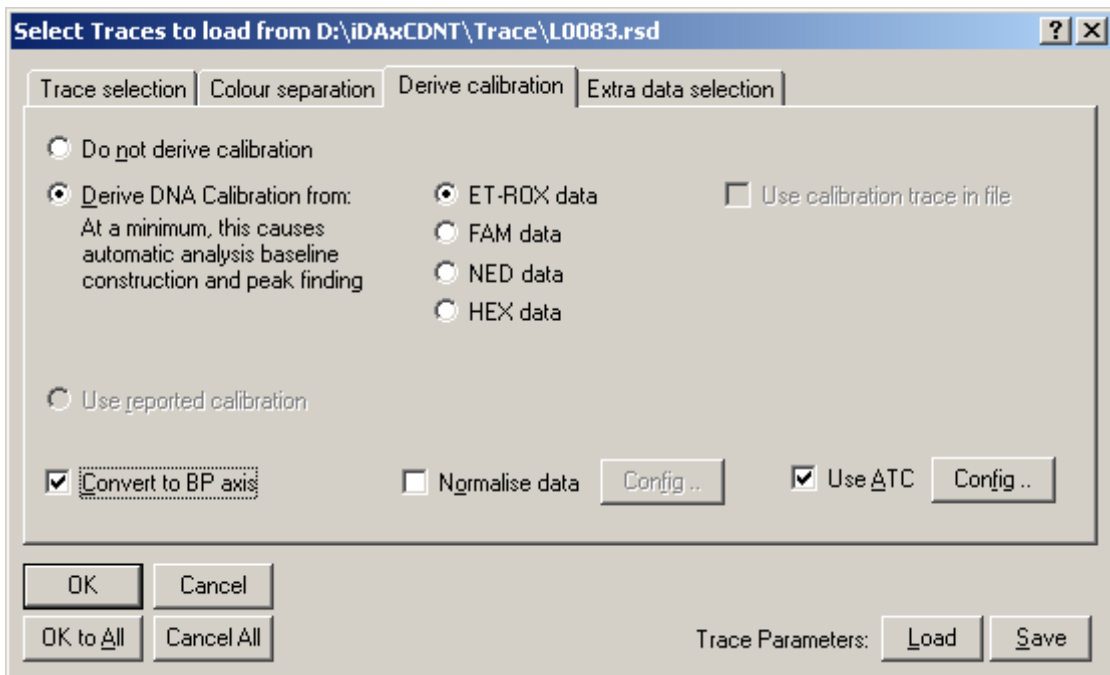
Chapter 3. Analysis of the Unknown

To analyse a file, follow the following procedure.

1. Invoke the **File | Open** command. Select **Files of Type MegaBACE™**. Select the file or files you want to analyse.
2. Check the **AutoAnalyse** button.
3. Click the **Config** button next to the AutoAnalyse button. The AutoAnalysis setup dialog box appears.
4. Make sure that at least **Construct Baseline** and **Detect Peaks** are checked. Additional items can be checked also. However, **Derive DNA Calibration** should not be checked. For now, do not check **Convert Axis**.
5. Click the **OK** button.
6. Click the **Open** button. This will start loading the trace file.
7. The Trace Analysis Parameters dialog box is displayed. Set it up as shown below. Alternatively, the settings can be loaded from the file UNKNOWN.TAP.



On the **Trace selection** tab, the ROX, FAM and HEX traces have been checked to be loaded. For more on trace types, see chapter 4.



On the **Derive calibration** tab, **Derive DNA Calibration** has been checked, and **ROX** has been selected as the trace to derive the size standard from².

8. Click the **OK** button.

Now, the following will happen in succession.

- the ET-ROX trace is loaded. It is analysed using the current settings. Since the Automatic Trace Calibration parameters define ROX as the calibration trace, the sizing standard will be recognised, and a calibration relating time coordinates to fragment sizes is derived from the ROX trace.
- the FAM and HEX traces are loaded and analysed using the current settings.

You can now click any of the peak list buttons at the left side of the data tags for the ET-ROX, FAM, or HEX traces. You will note that the correct fragment sizes are listed in the Base Pairs column.

It is also possible to display the fragment sizes of the peaks that were found in the graphics window. To do this, invoke the **File | Customise** menu option. Select the **Plotting Peaks** tab. Under

² The trace you select here must match the trace selected in the ATC dialog box (cf. Chapter 2). If it does not, a warning will be given. You can use the **Config** button next to **Use ATC** to re-configure ATC.

Peak Labeling, select **Component Name** as the first peak label. Select **Base Pairs** as the second peak label (it is near the bottom of the list of possible peak labels). Click the **OK** button.

The graphic window will now display base pair counts for all the peaks. There will be very many peaks displayed. You can click and drag the mouse on the graph to zoom in on the peaks, and see more detail.

Chapter 3.1. Converting to a Base Pair axis

In the example above, fragment sizes (base pair counts) are labeled for the peaks, but the horizontal axis displays time coordinates. For gene analysis, this often does not make much sense. Often, the horizontal axis does not correspond to specific time coordinates.

For this reason, DAX can convert the horizontal axis to fragment size units. Use the **Calibration | Axis Conversion** menu option to do this. Make sure to check the **New Window** check box, and then make sure that all data lines are selected.

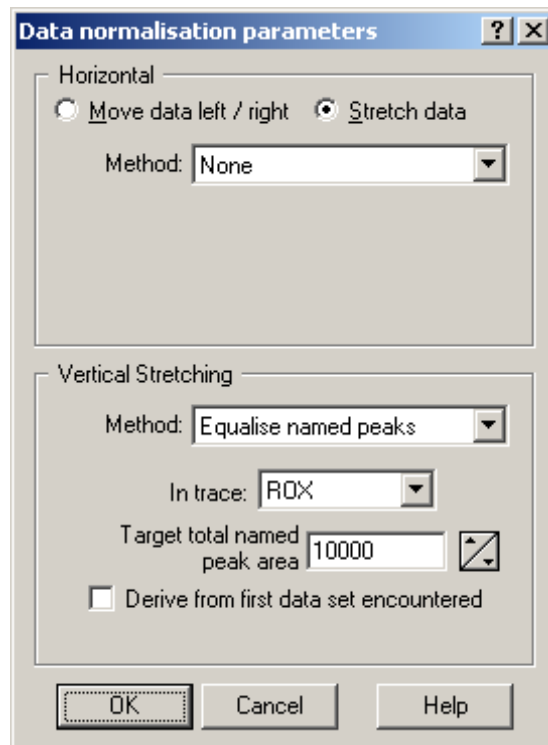
A more convenient way to convert the axis is to check the **Convert to BP Axis** box in the Trace Analysis Parameters dialog box (see above).

The analysis will be performed as before, but the horizontal axis will be converted to fragment units at the end of the analysis.

Chapter 3.2. Normalising the data

When loading trace data, it can be desirable to normalise the data after they have been loaded. One common form of normalisation stretches the data vertically in such a way that the total area of the size calibration peaks is equal across several measurements.

Normalisation can be achieved after data have been loaded using the **Data | Overlay** menu option. A more convenient way to normalise is to check the **Normalise data** check box on the **Derive calibration** tab in the Trace Analysis Parameters dialog box (see above). Click the **Config** button to set up normalisation.



In this example, the horizontal axis is not affected.

The vertical axis will be stretched in such a way that the total area of all named peaks in the ROX trace will become 10000. *Named* peaks are used because the named peaks in the ROX trace will be the calibration standard (as called by ATC).

The other traces (FAM, HEX) will be stretched an equal amount as the ROX data.

Chapter 4. Trace Types in DAX

Trace types are used in DAX for a number of purposes.

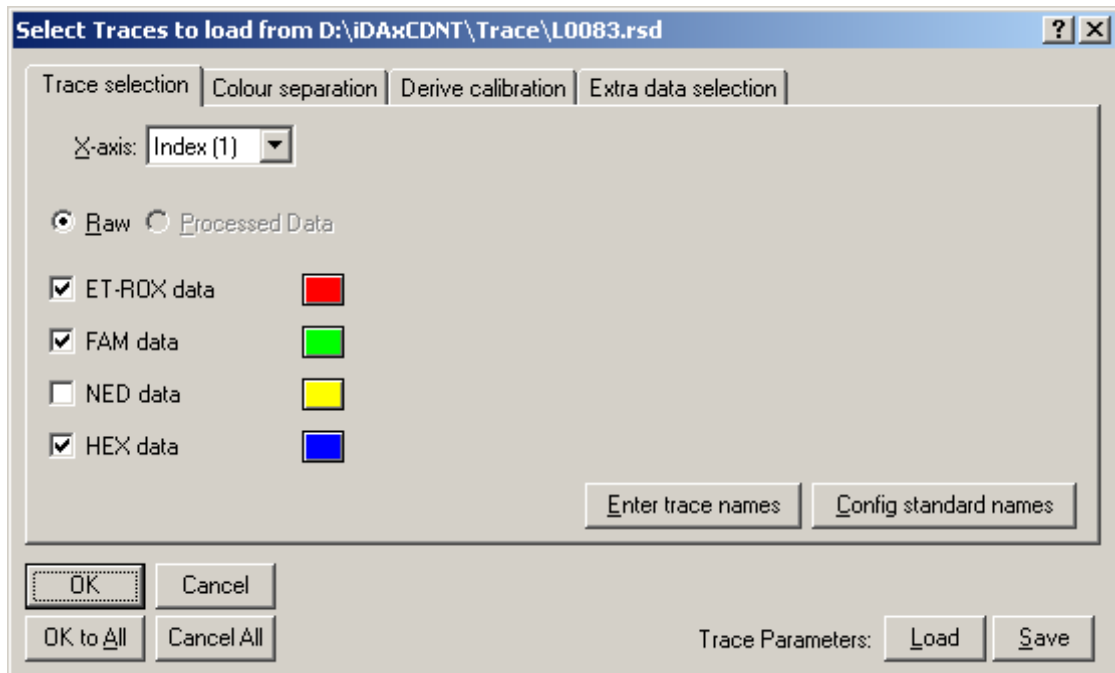
Their principal use is to distinguish between the up to five traces that can be contained in a single measurement file, where obviously the file name alone would be insufficient.

A secondary use is to limit certain operations to specific trace types. These operations include:

- Automatic Trace Calibrations. The trace type is used to indicate which trace contains the sizing calibration.
- Identification Database and Marker Peak items. The trace type is used to limit the detection of certain components to a specified trace type.

Chapter 4.1. Assigning trace types

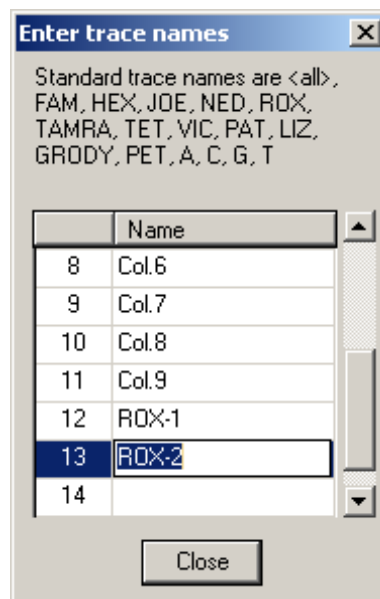
DAX attempts to retrieve the trace type of each trace from the measurement file, which often (but not always) contains trace names. The trace names are displayed on the first tab of the TAP dialog (see also chapter 3).



In this case, the measurement file L0083.rsd is recognised to contain ET-ROX, FAM, NED and HEX trace names³. If no trace types are manually specified (see below), DAX will recognise these traces as ROX, FAM, NED and HEX⁴.

The standard trace types that DAX recognises are FAM, HEX, JOE, NED, ROX, TAMRA, TET, VIC, PAT, LIZ, GRODY, and PET. Mainly for use with SCF files, there are also A, C, G and T trace types.

The **Config standard names** button allows additional trace types to be added. A dialog box like this one is displayed:

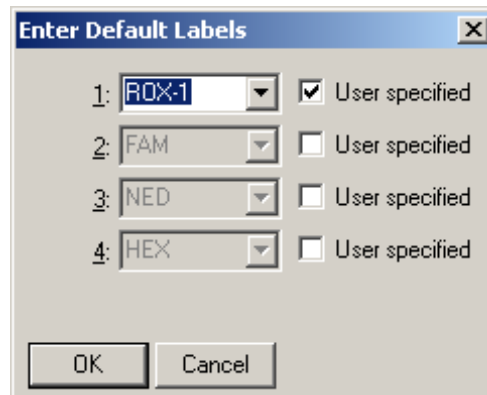


One way to use this feature is to enter types such as ROX-1 and ROX-2, so that two distinct ROX trace types are added.

³ The trace name is “what the trace is called in the measurement file”.

⁴ ROX is recognised because trace type ROX is part of trace name ET-ROX.

The **Enter trace names** button allows the trace names specified in the measurement file to be overridden with a user selected name or type. Clicking the button displays this dialog box:



To override a trace type, check the **User specified** box, then select a trace type or type a trace name.

Chapter 4.1.1. How DAX recognises trace types

As was mentioned above, an ET-ROX trace name will be recognised as a ROX trace type. This is true even if ROX-1 and ROX-2 trace types are manually added, because “ET-ROX” does not contain “ROX-1” or “ROX-2”, but does contain “ROX”. In order for a trace name to be recognised as a certain type, the name must contain the entire trace type.

The exception are the A, C, G and T trace types, which are only recognised if the entire trace name equals A, C, G or T.

Chapter 5. Conclusion

DAX can be used to perform highly complex analyses of trace files with relatively little effort.

There are a large number of additional features that have not been discussed in this manual. Examples are:

- peak width standardisation. If a sample is known to contain mostly peaks that have a width of 1 fragment, DAX can be told to preferentially assign peak widths of 1 fragment. If a peak is *clearly* wider, its width will not be affected.
- command line analyses. Typically, trace analysers will be used to perform dozens to hundreds of analyses per day. To make it possible to analyse such large quantities of data, DAX can be run from batch files. All the data on an entire hard disk can be analysed with a single command, while highly flexible report files are created.

These features exceed the scope of this manual. Refer to the DAX User's Manual for details.