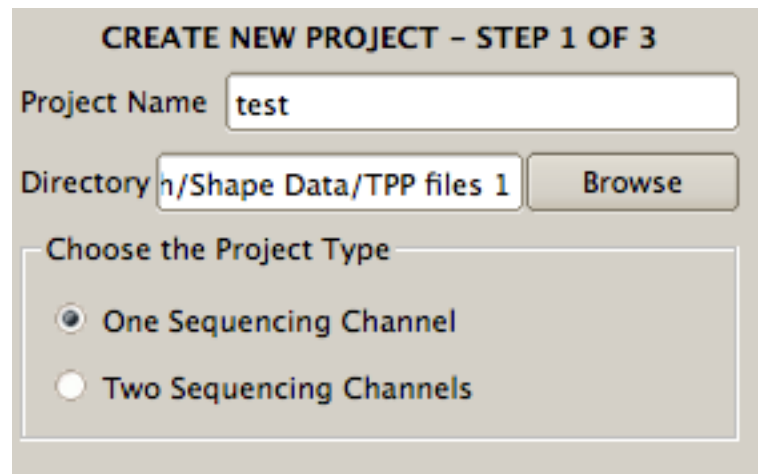# TPP QuShape Analysis

Note: We include this particular TPP experiment specifically because its automated QuShape analysis generated several errors in its sequence alignment. The presence of these errors allows us to illustrate the efficient manual sequence correction features of QuShape. A more typical TPP experiment has fewer errors.

Creating the project

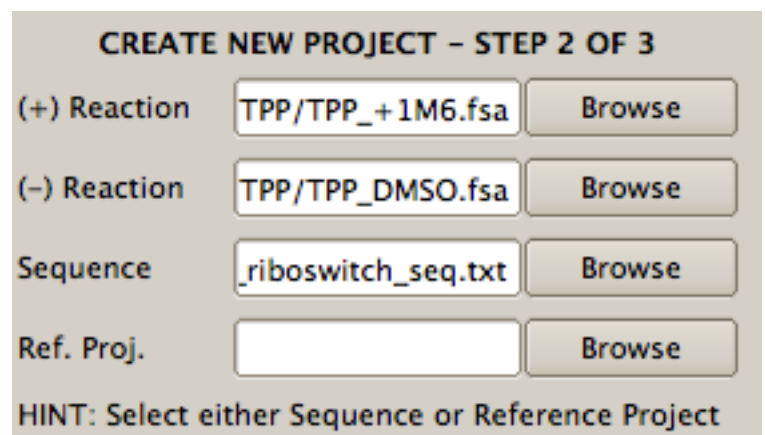A detailed description can be found in Section 3 of the Tutorial.

Step 1. Click **New Project** from the **File** menu. Enter the name of the project and select a working directory ("Step 1 of Creating New Project" figure). We recommended that you select the directory that contains raw data files.

> **CREATE NEW PROJECT – STEP 1 OF 3**
>
> Project Name [test]
>
> Directory [h/Shape Data/TPP files 1] [Browse]
>
> Choose the Project Type
>
> ⦿ One Sequencing Channel
>
> ◯ Two Sequencing Channels

Step 2. Select (+) reagent, (–) reagent, and sequence files ("Step 2 of Creating New Project" figure). For this dataset:

    i.   TPP_riboswitch_seq.txt: Official sequence file for TPP riboswitch RNA.

    ii.   TPP_+1M7.fsa: This file contains the raw (+) reagent data that have not been processed in any way. The second column is (+) reagent, the third column is ddC sequencing.

    iii.   TPP_DMSO.fsa: This file contains the raw (–) reagent data that have not been processed in any way The second column is (–) reagent, the third column is ddC sequencing.

> **CREATE NEW PROJECT – STEP 2 OF 3**
>
> (+) Reaction [TPP/TPP_+1M6.fsa] [Browse]
>
> (–) Reaction [TPP/TPP_DMSO.fsa] [Browse]
>
> Sequence [_riboswitch_seq.txt] [Browse]
>
> Ref. Proj. [ ] [Browse]
>
> HINT: Select either Sequence or Reference Project

Step 3. Select the column number ("Step 3 of Creating New Project" figure). For this dataset, Channel 2 is used for reagent and Chanel 3 is used for sequencing.
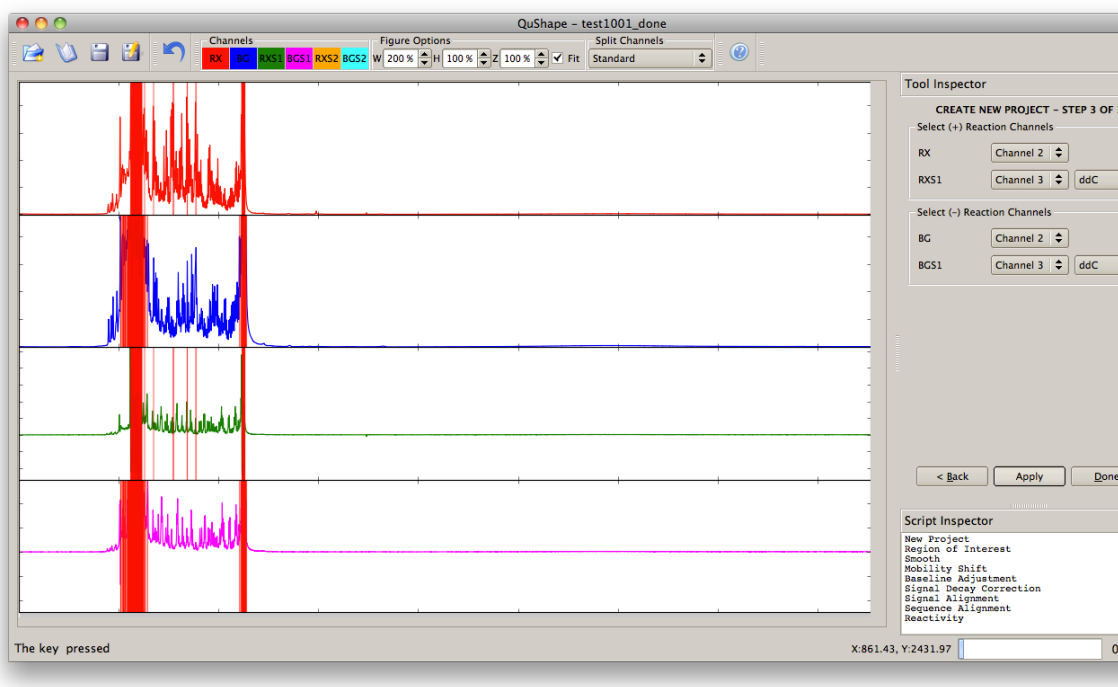

The raw data after Create New Project steps are completed.

Region of Interest

Region of interest (ROI) is selected by clicking on the plots (it might help to zoom in on the plots by unchecking the Fix box in the top bar) or by typing the numbers in the dialog boxes. In the "Result of ROI" figure, the region of interest for (+) reagent data is between 1329 and 2195. The region of interest for (–) reagent data is between 1334 and 2201. A detailed description can be found in Section 4.1 of the Tutorial.
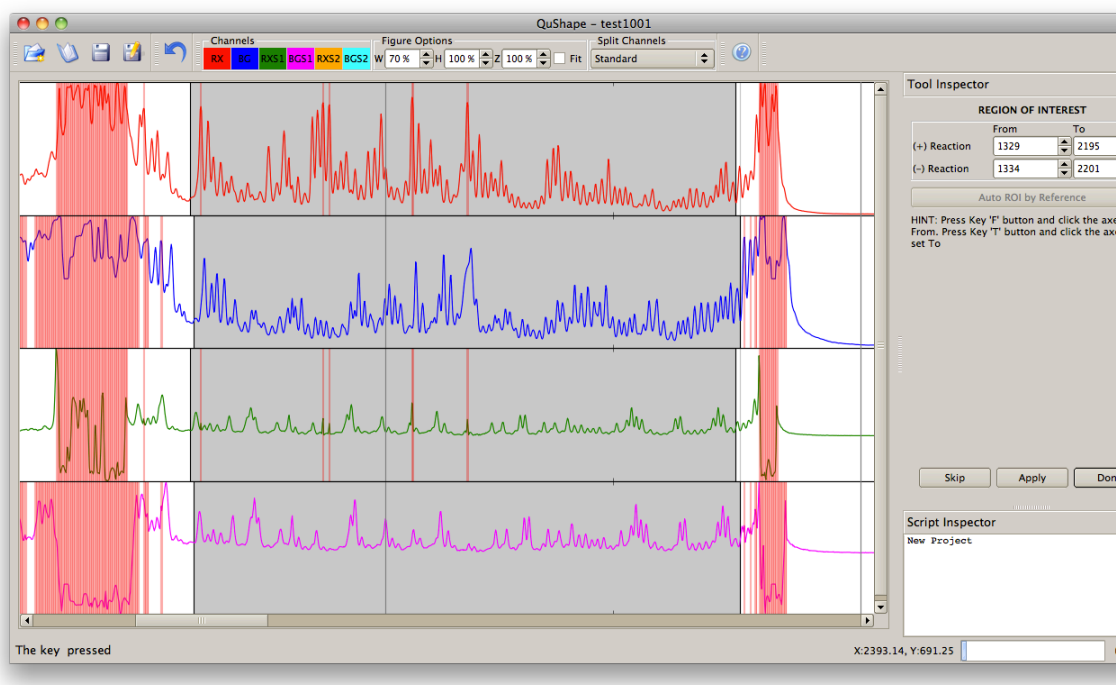

Result of ROI selection.

Smoothing

The **Smoothing** tool filters out high-frequency noise in the data and finds and corrects saturated data points. A detailed description can be found in Section 4.2 of the Tutorial.
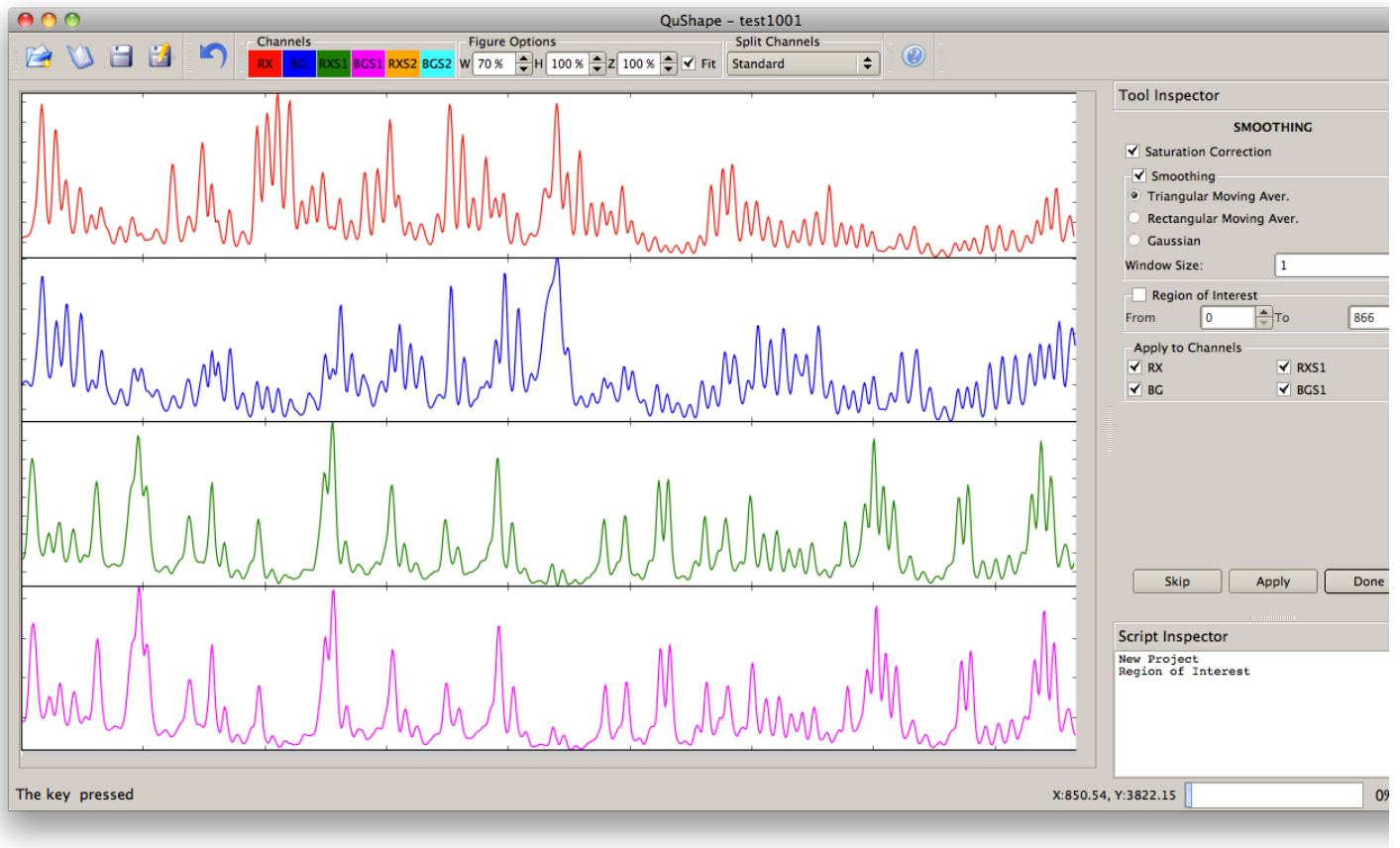
Result of Smoothing.

Mobility Shift

All data traces within each capillary are aligned relative to each other by time-shifting and time-scaling them along the elution time axis. A detailed description can be found in Section 4.3 of the Tutorial.
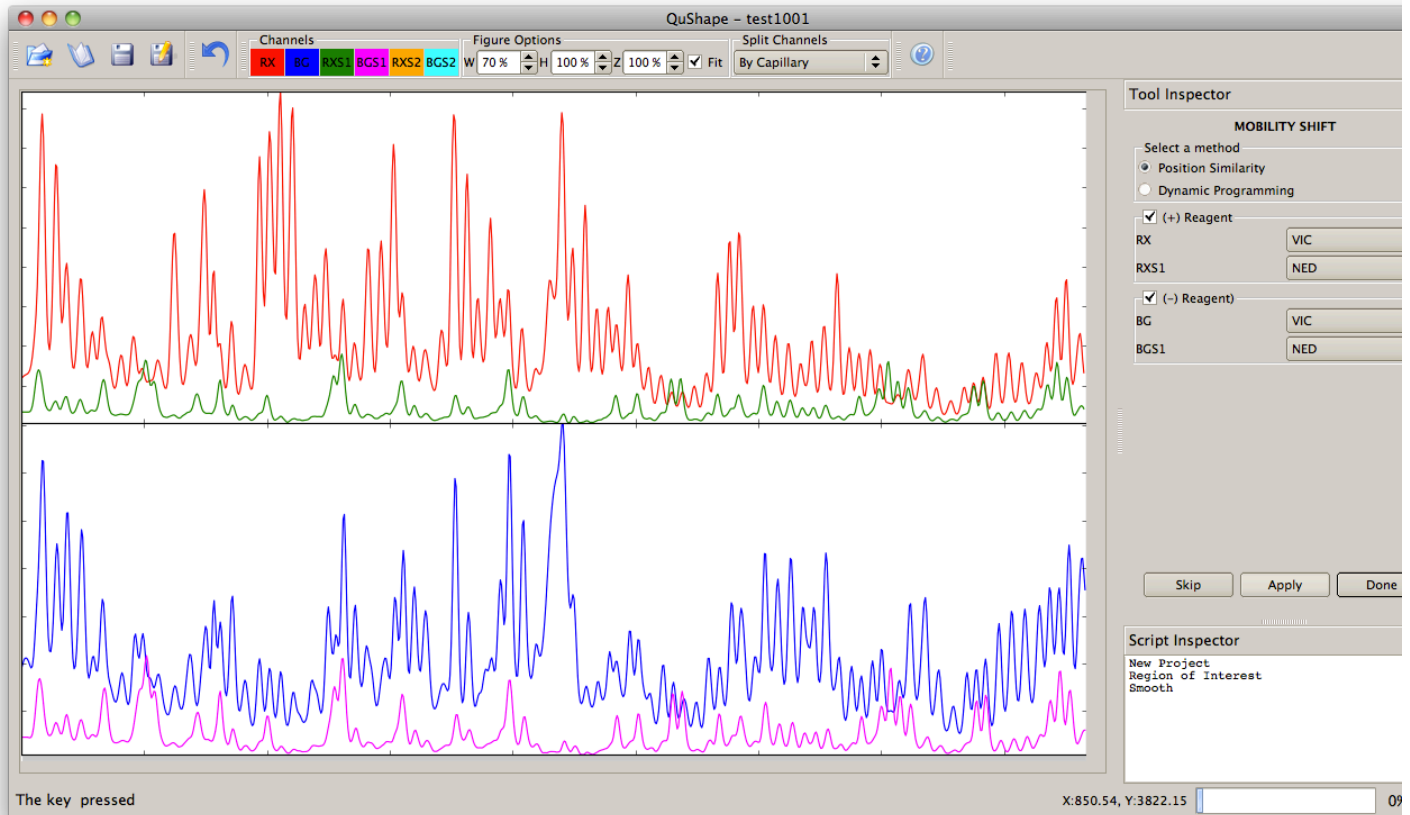
Result of Mobility Shift.

Baseline Adjustment

The **Baseline Adjustment** tool removes baseline offset. A detailed description can be found in Section 4.4 of the Tutorial.
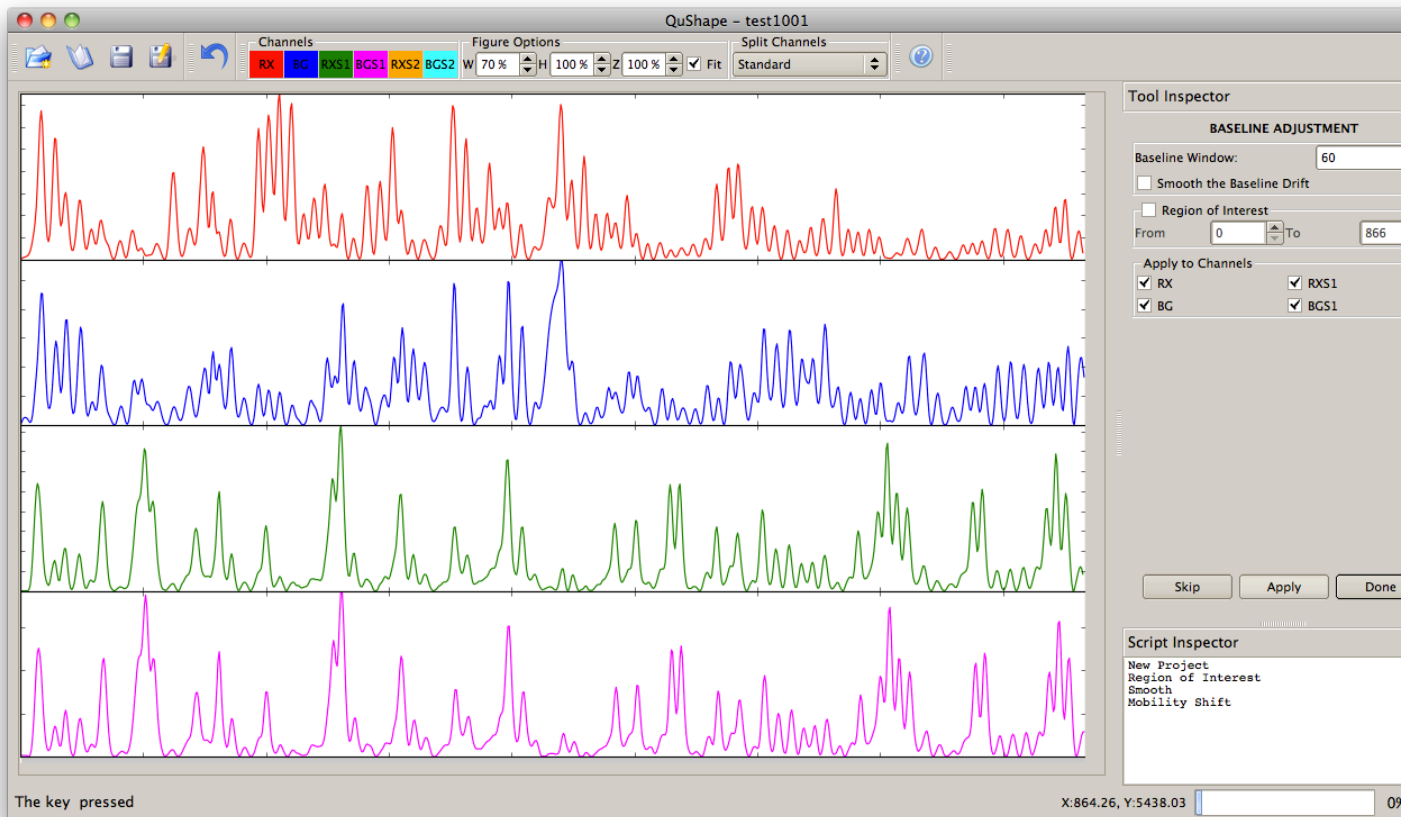
Result of Baseline Adjustment.

Signal Decay Correction

A characteristic feature of fluorescent signals in SHAPE electropherograms is that intensity gradually declines as a function of the elution time. This gradual signal decay is corrected by the **Signal Decay Correction** tool. A detailed description can be found in Section 4.5 of the Tutorial.
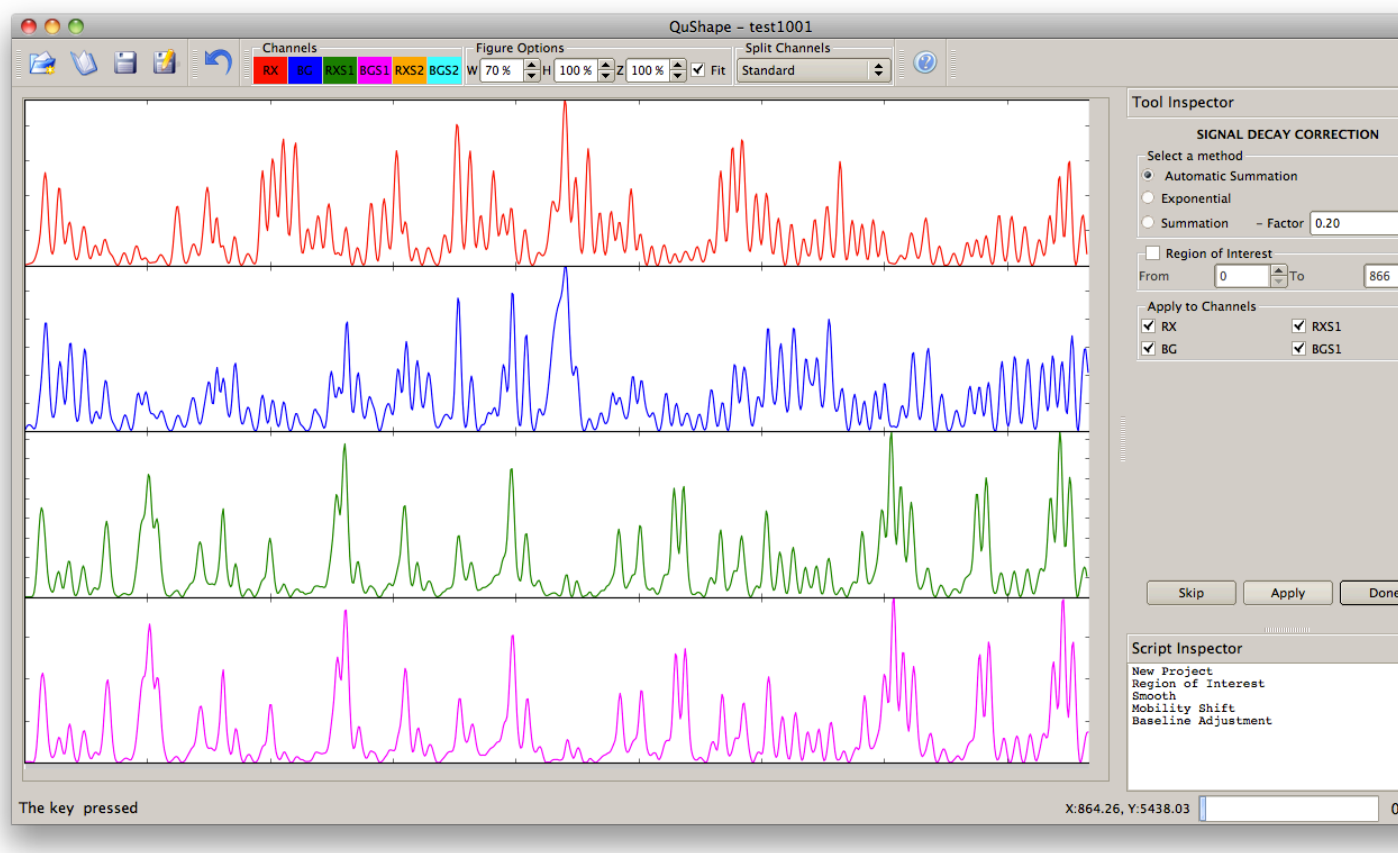

Result of Signal Decay Correction.

Signal Alignment

The **Signal Alignment** tool aligns pairs of signals across two capillaries. A detailed description can be found in Section 4.6 of the Tutorial.
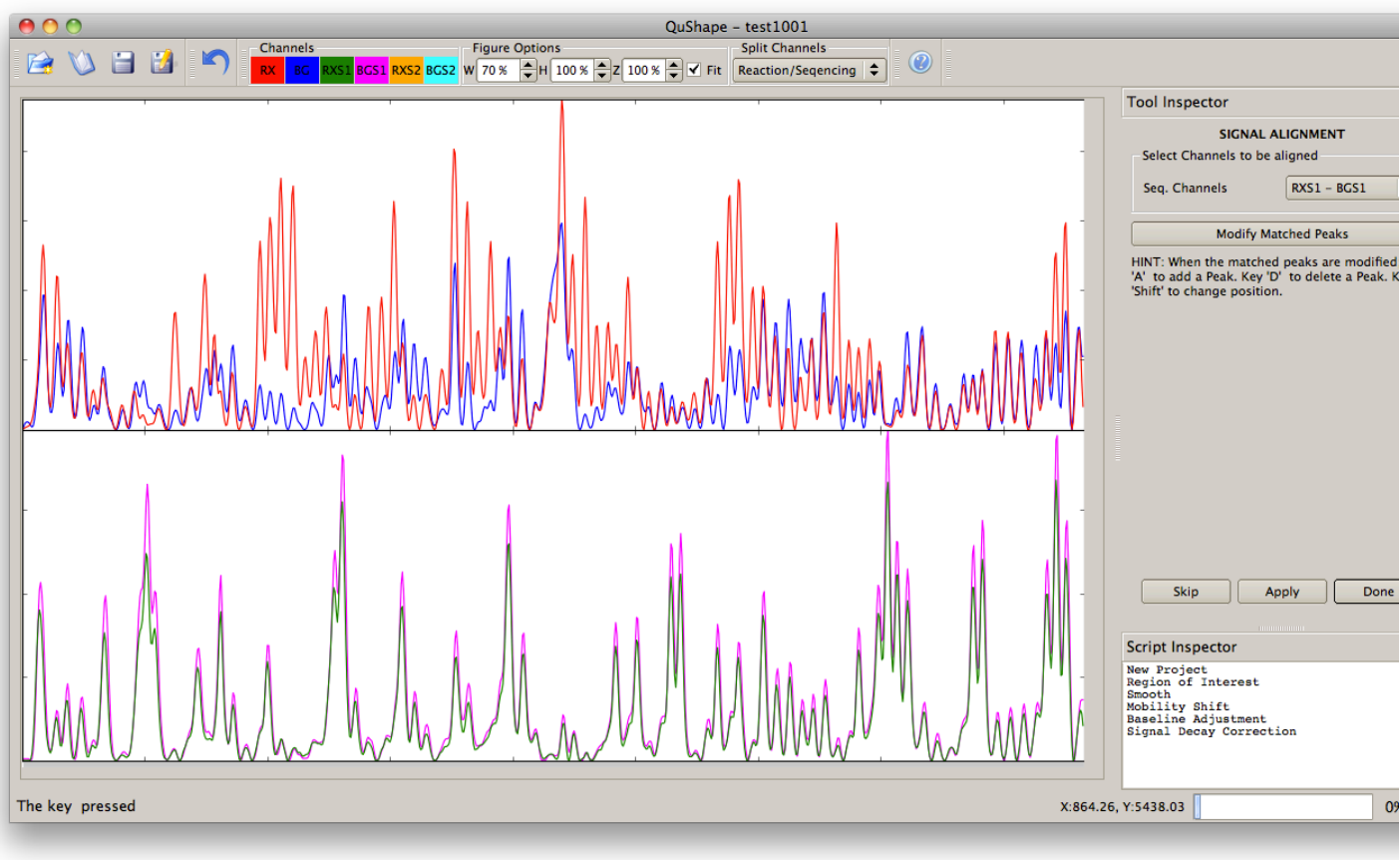

Result of Signal Alignment.

After performing signal alignment, the accuracy of this alignment can be evaluated by clicking the "Modify Matched Peaks" button, thus switching to a new display mode, in which some of the matched peaks are shown linked by vertical arrows. If some of the links are wrong, they can be manually corrected (see Section 4.6 in the Tutorial). Clicking the "Apply" button will realign all the peaks accordingly. In this example, no manual correction is required.
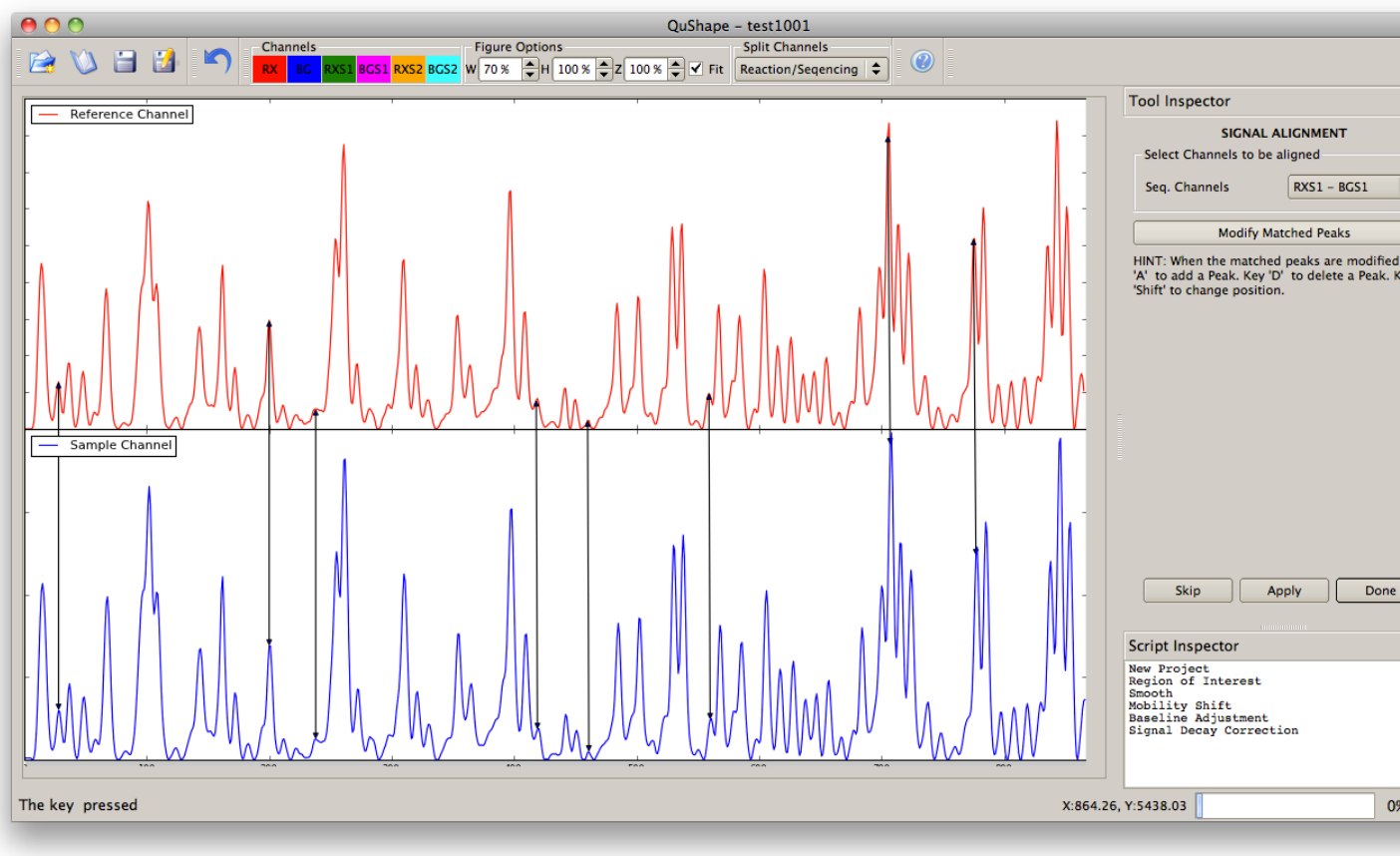

"Modify Matched Peaks" display.

Sequence Alignment

This tool performs three operations. First, it performs base calling, an operation that classifies all the peaks in the sequencing signal as either *specific* peaks produced by ddNTP-paired nucleotides or *non-specific* or background peaks corresponding to nucleotides of the other three bases. Next, the tool aligns peaks in the sequencing signal with the RNA nucleotide sequence. Finally, this tool assigns nucleotide-matched peaks in the sequencing signal to the corresponding peaks in the (+) reagent and (−) reagent signals, thus assigning each peak to its corresponding RNA position. All three operations are performed automatically once the **Apply** button is clicked. A detailed description can be found in Section 5.1 of the Tutorial.
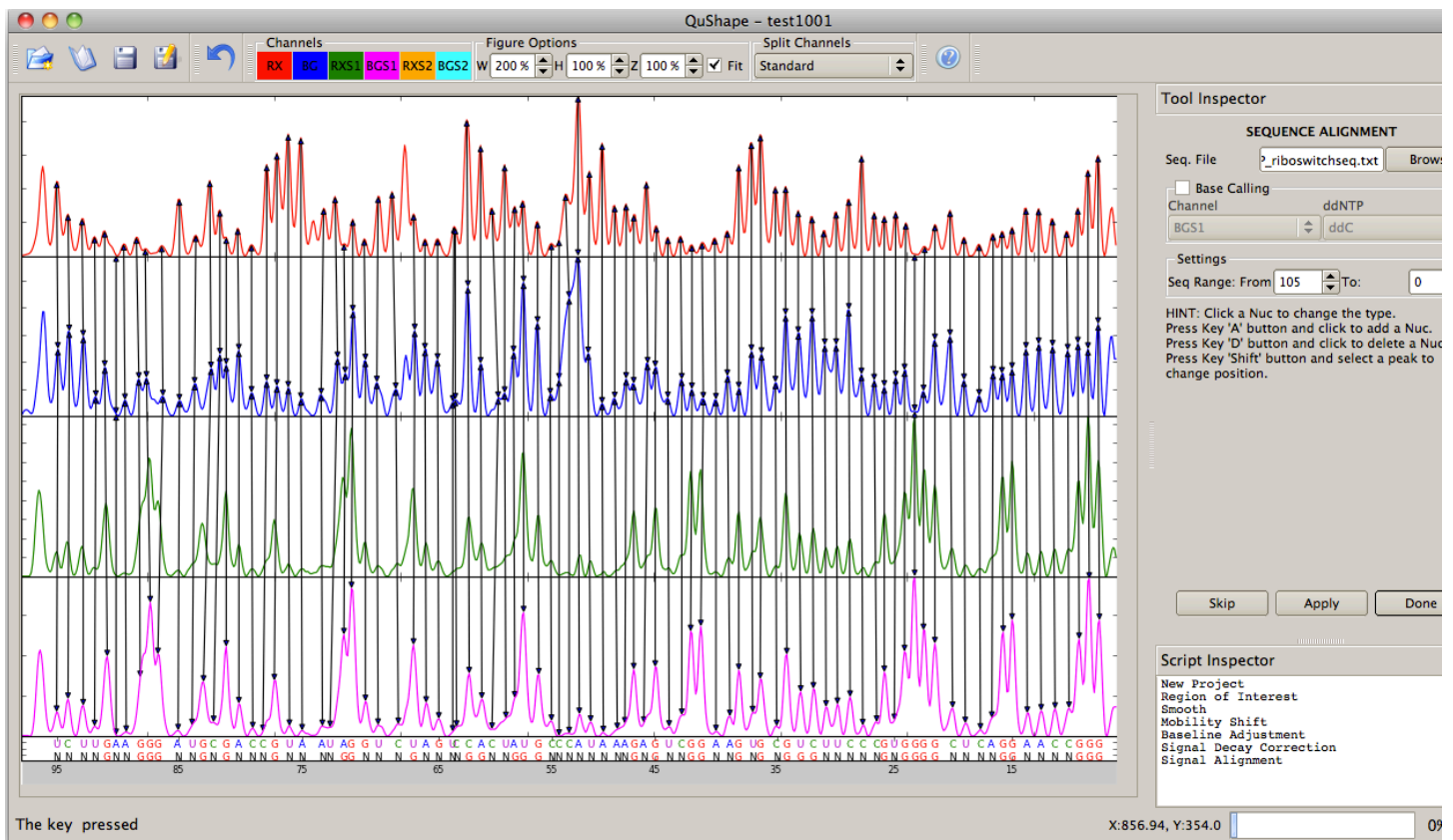

Initial sequence alignment

After pressing the **Apply** button, the initial alignment will be shown. You should check the alignment results for accuracy (it might help to zoom in on the plots by unchecking the Fix box in the top bar). If you find a problem, it can be fixed by adding or deleting nucleotides or changing the type (see Section 5.1 in the Tutorial). In this example, four errors must be corrected, as explained below.

Errors in sequence alignment are highlighted with arrows. These must be fixed manually.
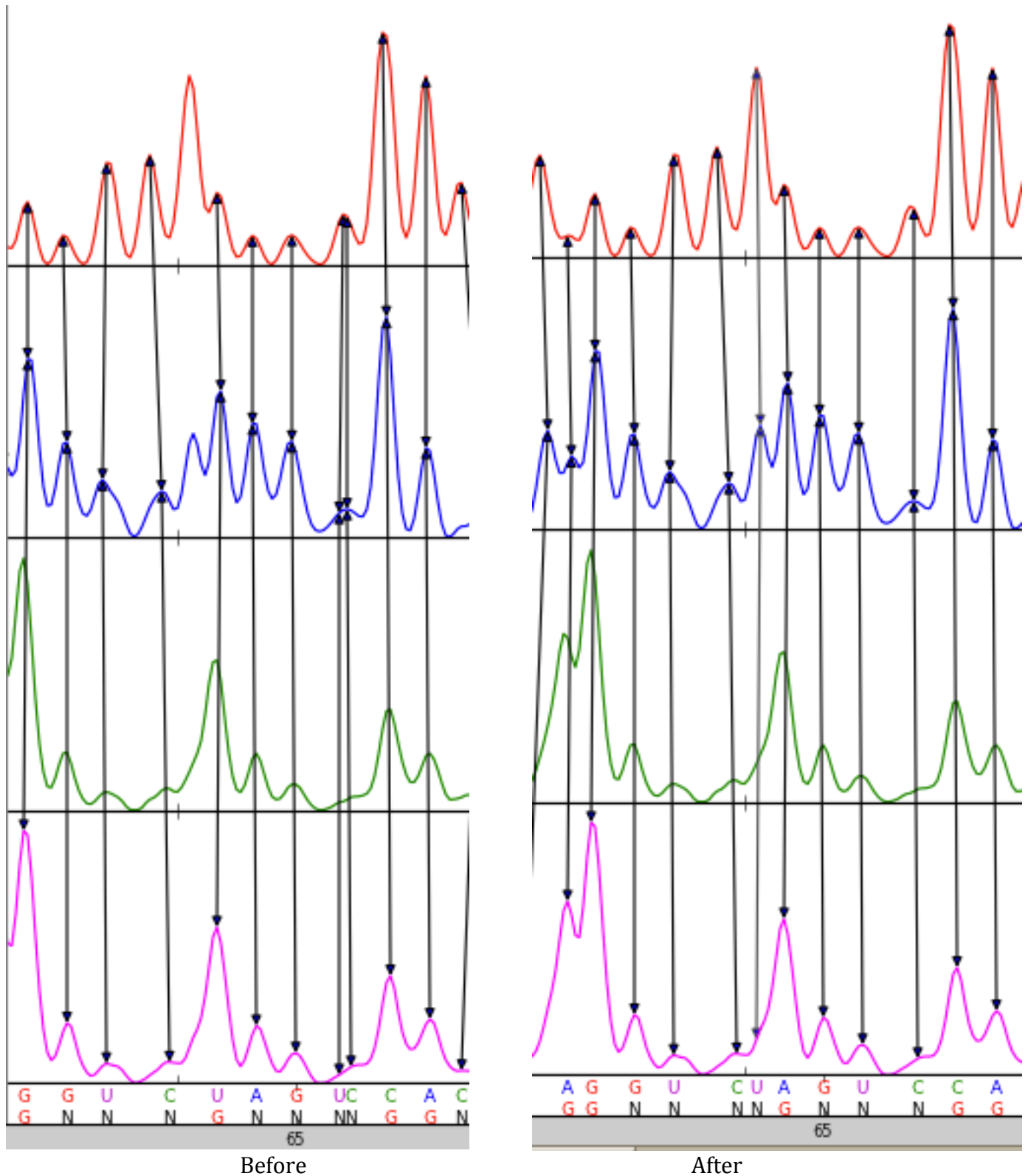
Before

After

There is an extra peak at position 66. This peak should be deleted by pressing and holding the **'D'** key while clicking at the base.
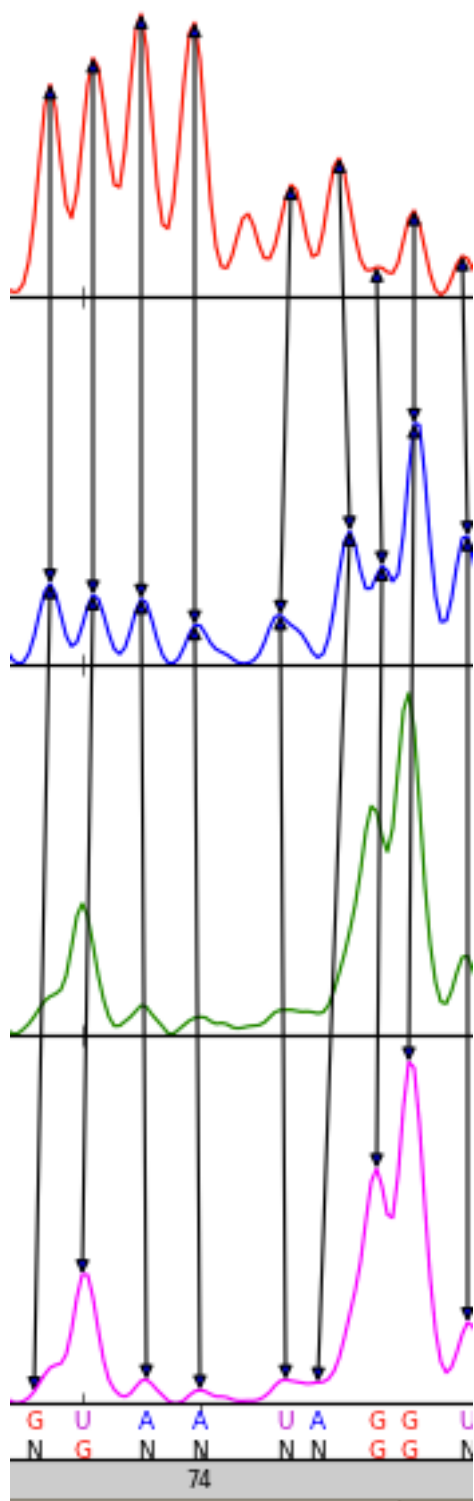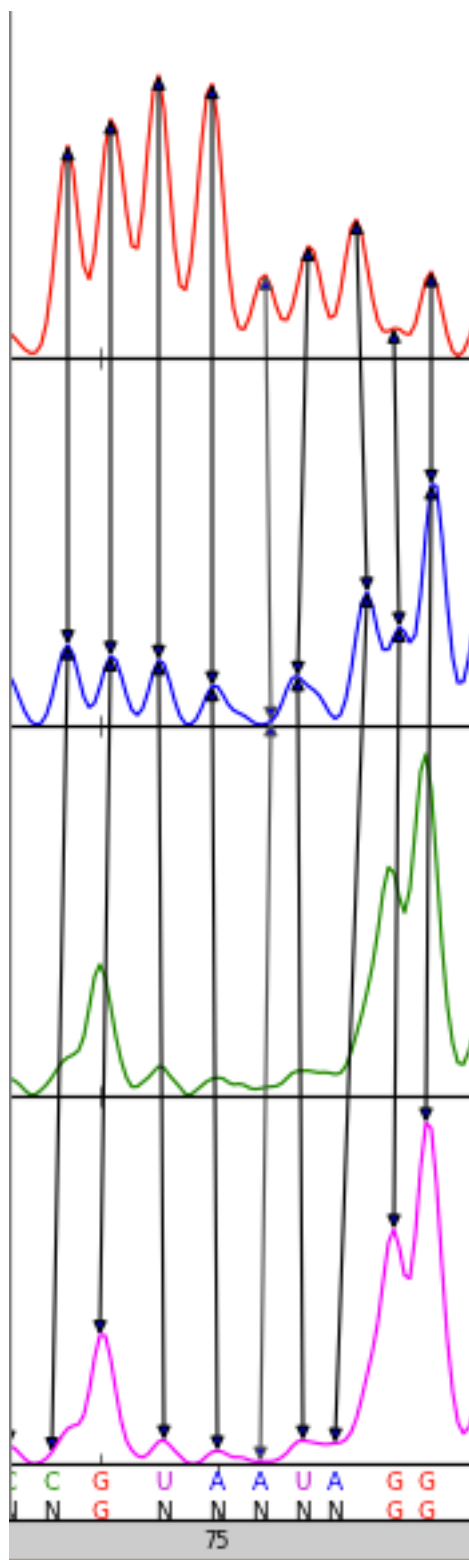
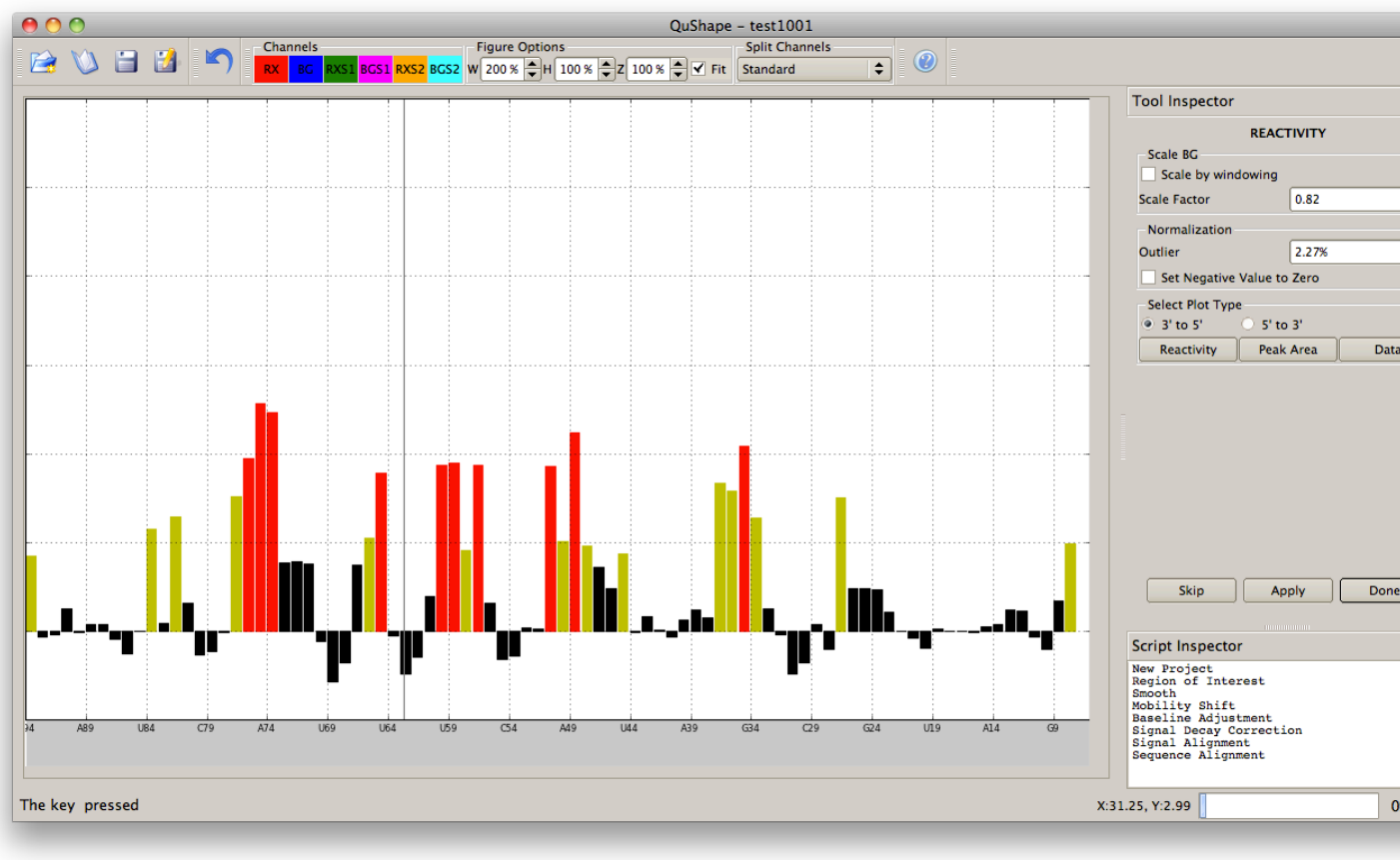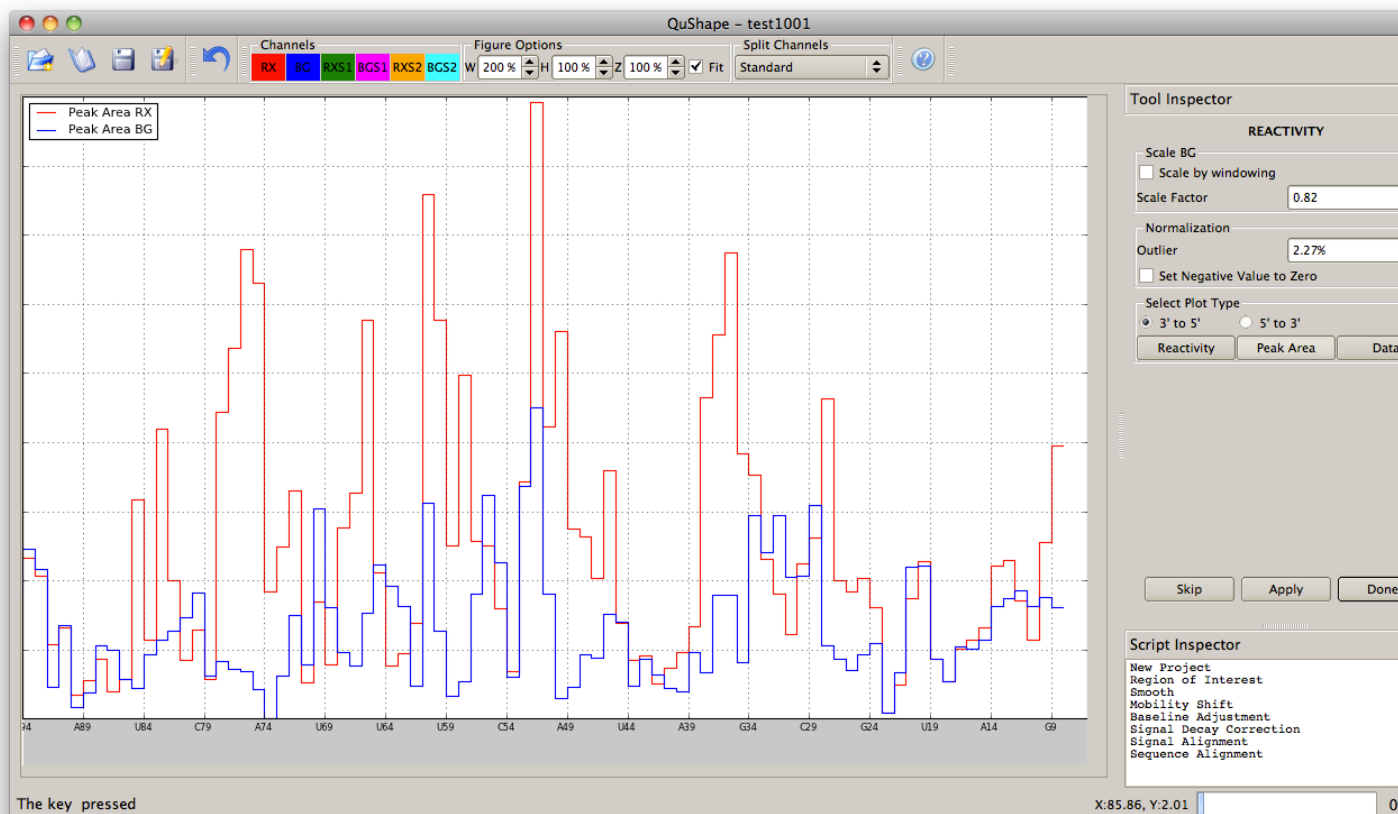Before   After

There is a missing peak at position 75. A peak should be added by pressing and holding the '**A**' key while clicking at a particular location in the bottom row with the mouse.
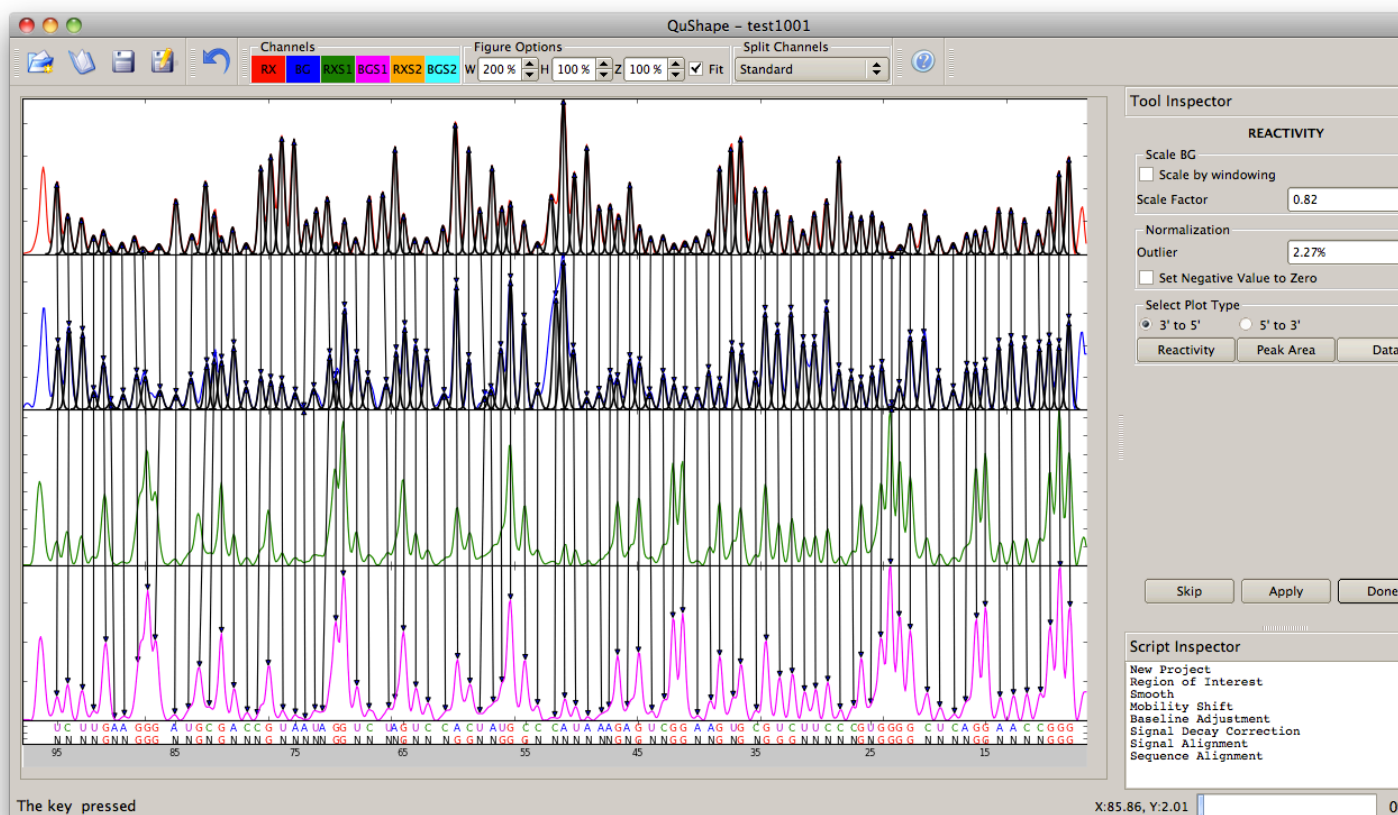
Reactivity

A detailed description can be found in Section 5.2 of the Tutorial.
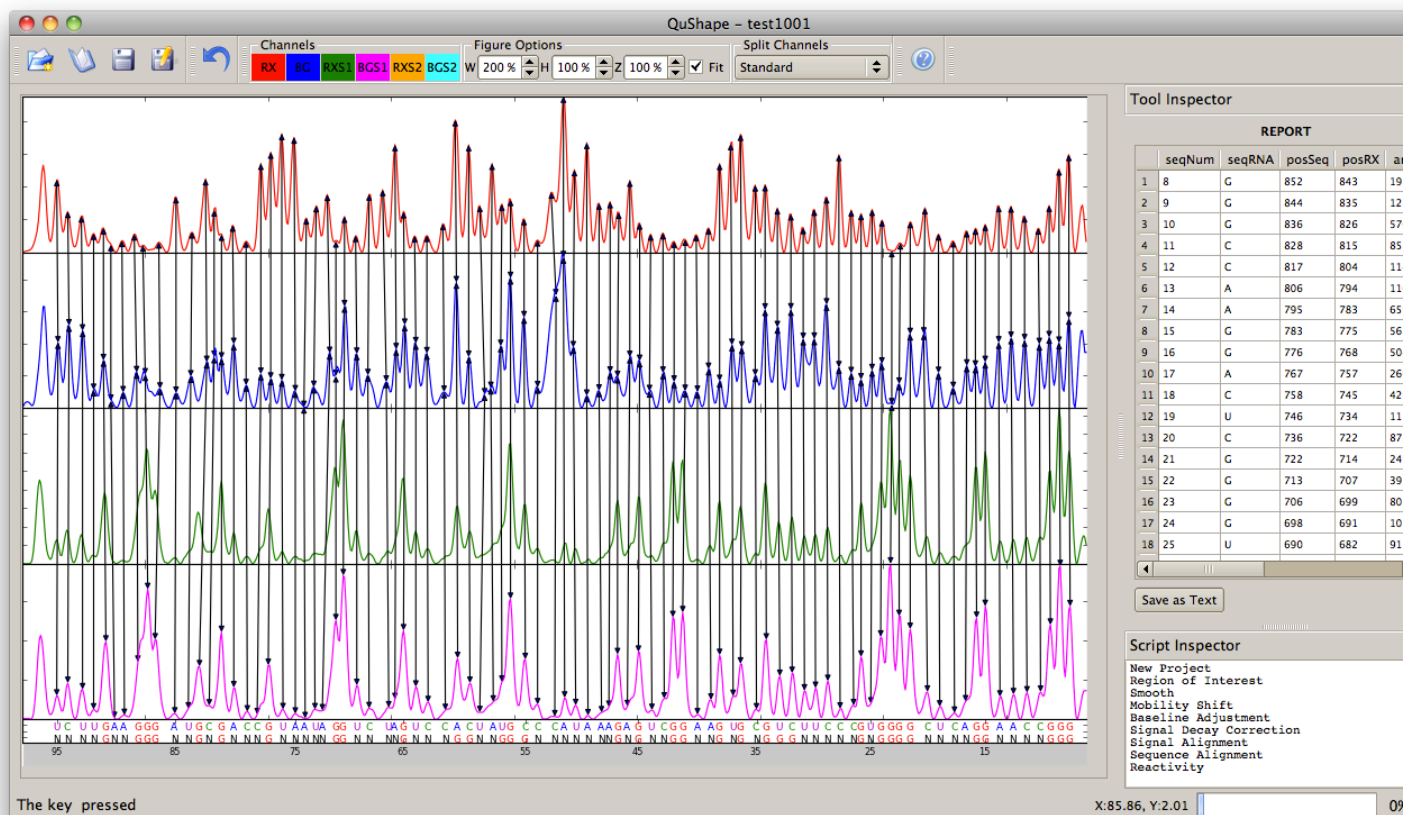

Reactivity of TPP.

Peak area of (+) and (−) reagent peaks.

Gaussian fit is shown in black.

15

Report

See Section 5.3 of the Tutorial.

The final report is shown as a table. It can be saved as a text file.