

Chemometrics

Product Note



Aligning Chromatograms - Part I

Summary

Chromatographic pattern matching using Infometrix® products has traditionally involved extraction of tabulated results (area, height) from processed chromatograms. This approach works well on closed systems with a small number of well-characterized peaks (<100). However, some issue may arise with this approach:

- Chromatograms to be matched may have many different peaks or few peaks in common
- Peaks may be misidentified because they shift outside expected retention time windows
- Closely eluting peaks may not show sufficient resolution and the smaller peak is not integrated
- In new samples, compounds not already in the defined peak table may appear but would not be reported

To minimize the impact of these issues, it may be better to process the whole chromatographic profile instead. Peaks with incomplete resolution or peaks from unexpected components which would be overlooked with a peak table approach, are routinely characterized with whole profiles. Such issues become features within the chromatographic fingerprint and are processed along with any other features.

The primary problem in working with whole chromatograms is the lack of retention time stability, which affects multivariate data processing. Infometrix offers two solutions for dealing with RT fluctuations: the Align transform within Pirouette® and the standalone software package LineUp™. This note will discuss the former tool.

Data Preparation

To exploit the alignment feature requires import of whole profiles into Pirouette. This is accomplished most easily by export from the chromatography system of an AIA format file. Note that in the case of the Agilent ChemStation, which does not store marker information in the AIA file, a special macro will save a Pirouette formatted ASCII file.

Suitable alignment markers in each file are entered into the Pirouette spreadsheet before processing. Marker identification and time entry can be automated by including marker peaks in the peak ID table of the chromatographic method; an AIA file exported after reprocessing the data should contain information about these marker peaks.

Different scenarios for processing chromatograms are:

Using native peaks as alignment markers - *If compounds that appear in all of your chromatograms, you can use them as alignment markers. An example is the n-paraffins usually present in petroleum-related data.*

Using internal standards as alignment markers - *If suitable markers cannot be designated among native peaks, particularly if no common peaks appear in all chromatograms, you can incorporate internal standards. Choose such internal standards so that their retention times bracket other diagnostic peaks.*

Using external standards as surrogate markers - *When native markers are not found and peak density precludes the use of internal standards, an external standard approach may be required. In this case, you usually analyze a QC sample that contains appropriate alignment markers, followed by analysis of one or more analytical samples. This presumes that RT drift after running the QC sample will be negligible in the short term, a presumption that needs to be tested and verified.*

Alignment Processing

Pirouette uses the marker retention times (in scan units) to perform interpolative alignment, modifying the retention time of intervening peaks. In each case, Pirouette aligns sample chromatograms to a designated target chromatogram. The actual peak positions used by Pirouette in the alignment step are derived in one of three ways.

1. Estimated retention times (these are peak tops, in scan units) are added to the data table. Marker positions for all peaks in all samples are entered. During processing, Pirouette estimates true peak tops for markers in all samples. True peak tops are found via a polynomial fit on the tallest peak around each estimated marker position.
2. Estimated retention times are added to the data table only for the target; for samples, retention times may be entered as zero. Pirouette estimates peak tops for markers from the target. In this case, Pirouette will use the target's marker times in each sample that has these zero values. It will then proceed as above to find true peak tops.
3. Marker peaks may be absent in samples but estimated marker positions are available; these positions are added to the sample marker cells in the data table. This approach of using external standards assumes there is negligible retention time drift between samples. The Pirouette option for alignment window size is set to zero and no attempt is made to find true peak tops. Rather, the explicit time entered for a marker peak is used whenever the sample has non-zero values. If a marker value is zero, the target's marker position is used.