Supplementary Information

Chromophoreasy, an Excel-Based Program for Detection and Integration of Peaks from Chromatographic and Electromigration Techniques

Fernando A. S. Vaz,* Leandra N. O. Neves, Rafael Marques, Renata T. Sato and Marcone A. L. Oliveira

Departamento de Química, Universidade Federal de Juiz de Fora, 36036-330 Juiz de Fora-MG, Brazil

Experimental

Chemicals and reagents

All reagents used in this work were of analytical grade. Tris(hydroxymethyl)aminomethane (Tris), organic acids (oxalic, pyruvic, tartaric, citric, formic, malic, lactic, succinic, aspartic, acetic and phthalic), thiourea, ammonium acetate, zinc acetate, potassium hexacyanoferrate and ethanol were obtained from Vetec (Rio de Janeiro, Brazil). Cetyltrimethylammonium bromide (CTAB), lactulose, lactose, 2,6-pyridine dicarboxylic acid (PDC), hexane, naphthalene and acenaphthene were obtained from Sigma (St. Louis, USA). Fluorene was obtained from Aldrich (Steinheim, Germany). Phenanthrene and anthracene were obtained from Fluka (Buchs, Switzerland). Acetonitrile was acquired from JT Baker (Center Valley, USA). The aqueous solutions were prepared with purified water from Milli-Q system (Millipore, Bedford, USA).

Sample treatment

For liquid chromatography (LC) analysis, a milk sample from a local store was deproteinized, clarified and ultrafiltrated in a 0.45 μm membrane, before injection. For gas chromatography (GC) analysis, biodiesel samples were obtained from transesterification reactions, such as basic catalysis of soybean, sunflower and food frying commercial oils, acid catalysis of soybean oil, as well as acid pre-treatment followed by basic catalysis of food frying oil. These samples were diluted into hexane (1% v/v) and subjected to vortexing for 2 min, prior to injection into the GC equipment. The major fatty acids methyl esters were then separated.

Instrumental conditions: liquid chromatography

The LC system used was a Breeze Modular high performance liquid chromatograph (HPLC) (Waters, Milford, USA) equipped with refractive index detector for modular systems, binary pump system and amide modified column

*e-mail: nando.simas.vaz@gmail.com

(3.5 µm; 4.6 mm × 150 mm) with temperature controlled at 35 ^{o}C. An isocratic elution mobile phase of acetonitrile-water (75:25, v/v), with flow-rate of 1.8 mL min^{-1} and sample injection of 20 µL were used.

Instrumental conditions: capillary electrophoresis

For capillary zone electrophoresis (CZE) analysis of organic acids standards, an Agilent 1600 capillary electrophoresis (CE) system (HP^{3d} CE, Palo Alto, USA) equipped with a diode array detector (DAD) was used. The conditions were: running electrolyte composed of phthalic acid (20 mmol L^{-1}), Tris (16 mmol L^{-1}) and CTAB (0.8 mmol L^{-1}), pH 3.4; hydrodynamic injection: 25 mbar for 2 s; reverse flow voltage: –15 kV; cartridge temperature: 25 ºC; and indirect UV detection: 220 nm. A polyimide-coating fused-silica capillary (Polymicro Technologies, Phoenix, USA) with 48.5 cm (40 cm of effective length) and 75 µm inner diameter was used.

The experiments involving CZE analysis of lactose and lactulose standards and capillary electrochromatography (CEC) analysis of polycyclic aromatic hydrocarbons (PAH) standards were performed in an Agilent 7100 CE system equipped with a DAD and a capacitively coupled contactless conductivity detector (C^{4}D). For CZE analysis of lactose and lactulose the conditions were: electrolyte composed of PDC (20 mmol L^{-1}), CTAB (0.5 mmol L^{-1}), pH 12.5; hydrodynamic injection: 50 mbar for 6 s; reverse flow voltage: –11 kV; cartridge temperature: 20 ºC. The C^{4}D was employed. A polyimide-coating fused-silica capillary (TSP, Polymicro Technologies) with 48.5 cm (40.0 cm of effective length) and 50 µm of internal diameter was used.

For analysis of PAH standards the conditions were: mobile phase composed of ammonium acetate aqueous solution (16.7 mmol L^{-1}) and acetonitrile (40% v/v), pH 7.3; hydrodynamic injection: 25 mbar for 5 s; voltage: 20 kV; cartridge temperature: 20 ^{o}C; DAD set at 220 (for thiourea, naphthalene, acenaphthene and fluorine detection) and 250 nm (for phenanthrene and anthracene detection); and simultaneous pressurization to both inlet and outlet vials: 5 bar. A TSU capillary 36.0 cm long (8.5 cm of effective length), 100 µm of internal diameter and 8.0 cm of monolithic stationary phase was used.

Instrumental conditions: gas chromatography

The experiments involving GC analysis were performed in a GC 2010-Plus equipment (Shimadzu, Kyoto, Japan), equipped with split/splitless autoinjector AOC 20-I and flame ionization detector (FID). An SLB-IL 111 column with liquid ionic stationary phase (1,5-di (2,3-dimethylimidazolium) pentane bis(trifluoromethylsulfonyl)imide), 14 m long, 0.10 mm of inner diameter and 0.08 µm of film thickness (Supelco, Bellefonte, USA) was used. Biodiesel samples were analyzed according to Delmonte *et al.*[1] parameters, which consisted of injection and FID temperature of 250 ^{o}C; hydrogen carrier gas set at 26 cm s^{-1} and 1 µL of injection volume, in split mode (1:100). Other FID conditions were: nitrogen (make up gas, 30 mL min^{-1}); hydrogen (40 mL min^{-1}); and synthetic air (400 mL min^{-1}).

Additional tools

Initial considerations

The macros, developed in Visual Basic for Applications (VBA), required for operation of a command box, peaks extraction to distinct spreadsheets, fine adjustments of the integration limits, split coeluted peaks and grouping the integration results are described in the following sections. Figure S1a shows a data set obtained from organic acids standards analysis through CZE, before any treatment. Note that each detected peak with its integration results in Figure S1b is placed in a row and different parameters (area, height, etc.) of a peak are placed in columns. The red baselines are a single plot of the first and last point of each peak, separated into distinct segments. The chromatogram and baseline plots are also performed via macros.

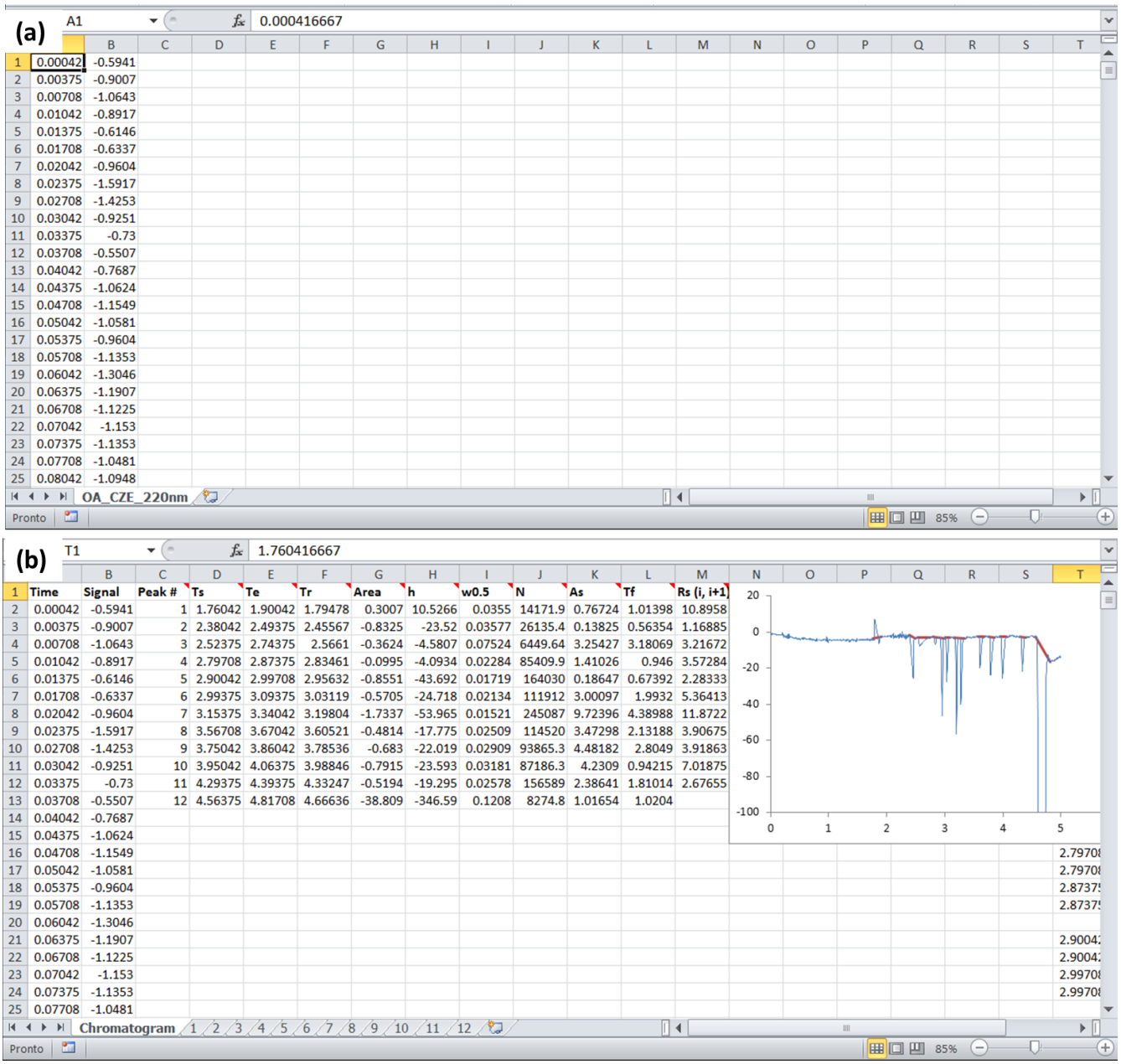

Figure S1. Main worksheet of an electropherogram file (a) before and (b) after recognition and integration. Each tab (1-12) contains an integrated peak.

Command box

All functions are placed in the command box, which should be opened in the chromatogram file via keyboard shortcut (Ctrl + a). The first page (Figure S2a) gathers functions related to the recognition and extraction of peaks. Options regarding time range definition, peak type (positive/negative/both), chromatogram smoothing modes and smoothing window, sensitivity for determining the slope and curvature thresholds, among others, are available. Command buttons for peak recognition, chromatogram and baseline plots, and peak extraction are present. At the bottom of the command box, a text box

for information about the program status, “Do everything” and “Exit” buttons are available. The “Do everything” button is responsible for performing a sequence of basic operations: (*i*) recognition of peaks; (*ii*) extraction of the detected peaks to distinct worksheets; (*iii*) integration of each extracted peak; (*iv*) grouping the results in the main worksheet; and (*v*) plotting of the chromatogram and baselines.

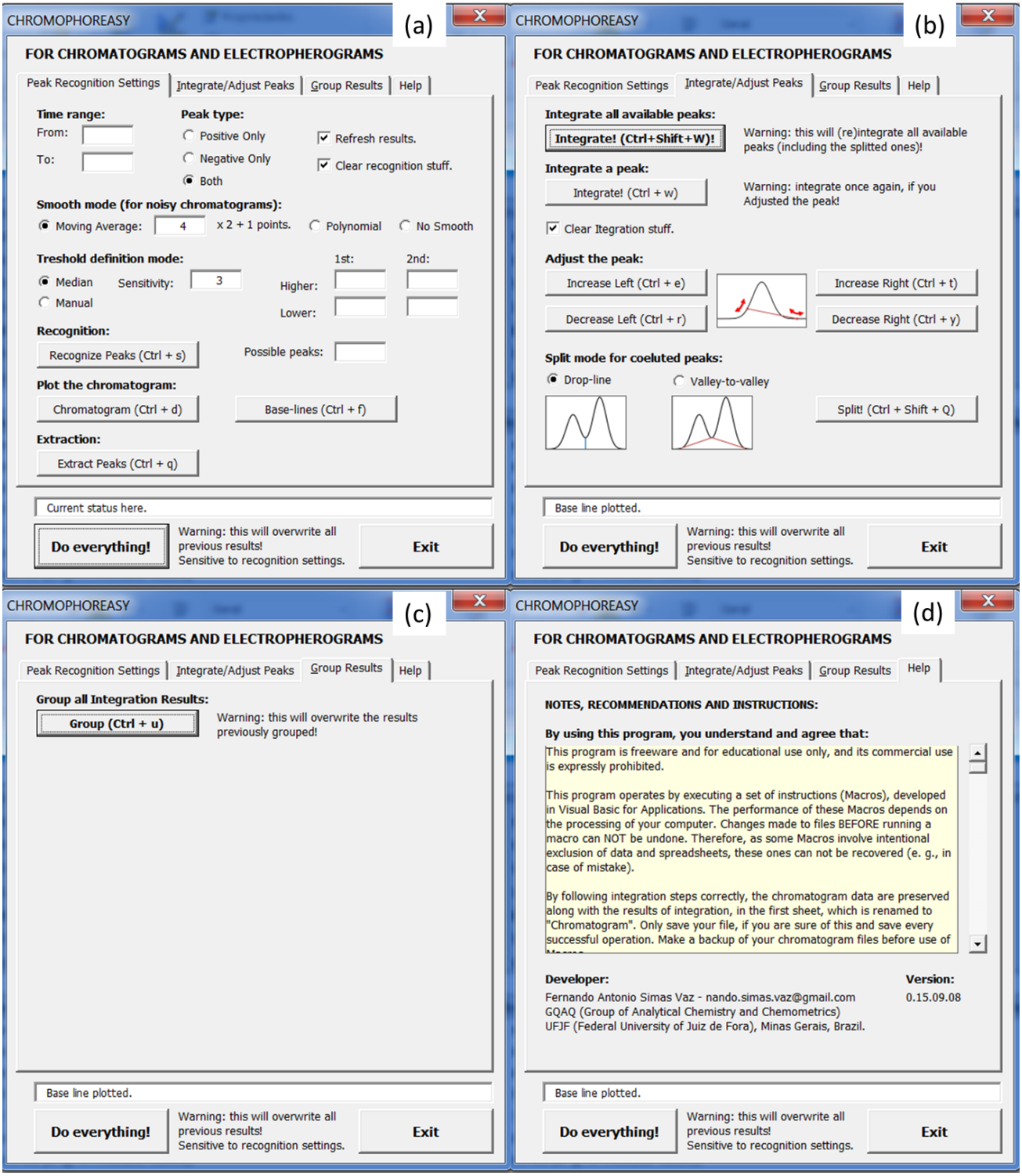

Figure S2. Command box showing (a) peak recognition settings; (b) integrate/adjust peaks; (c) group results; and (d) help pages of options.

The second page (Figure S2b) shows the commands related to the integration, side adjustments and splitting of peaks. The “Integrate all” button (re)integrates every peak of a workbook. The “Integrate” button integrates only the actual

peak. The buttons at the "Adjust the peak" area correct manually the limits of a peak. At the "Split mode for coeluted peaks" area drop-line and valley-to-valley options can be used to separate partially coeluted peaks. The "Group Results" page (Figure S2c) contains a command to collect the integration results into the first worksheet and the "Help" tab (Figure S2d) contains some important notes, recommendations and information to the user.

Extraction

If any peak was recognized or manually inserted, the peak extraction can be performed. In this procedure, the chromatogram data between the limits of each peak are copied and pasted into separate worksheets, which are more convenient for peak treatment. The original data are kept intact in the main worksheet. If this function was performed manually, the extracted data should be integrated (the "Integrate all" button is a good option when several peaks are present).

Manual adjustments of peak limits

If the baseline was not adequately defined, manual adjustments of the peak limits can be made, through adding or removing a chromatogram point to the left (before) or to the right (after) of the peak. Four macros were developed for these operations. If any point has to be added, the macro searches that point in the main chromatogram data. These functions are easily done/repeated/undone through keyboard shortcuts. For example, the peak in Figure S3a (peak 3, from the electropherogram of Figure S1) was not properly recognized by the program and needs some manual adjustments. After some adjustment (Figure S3b), a new "integration" and "grouping" (further section) should be done, in order to update the integration results.

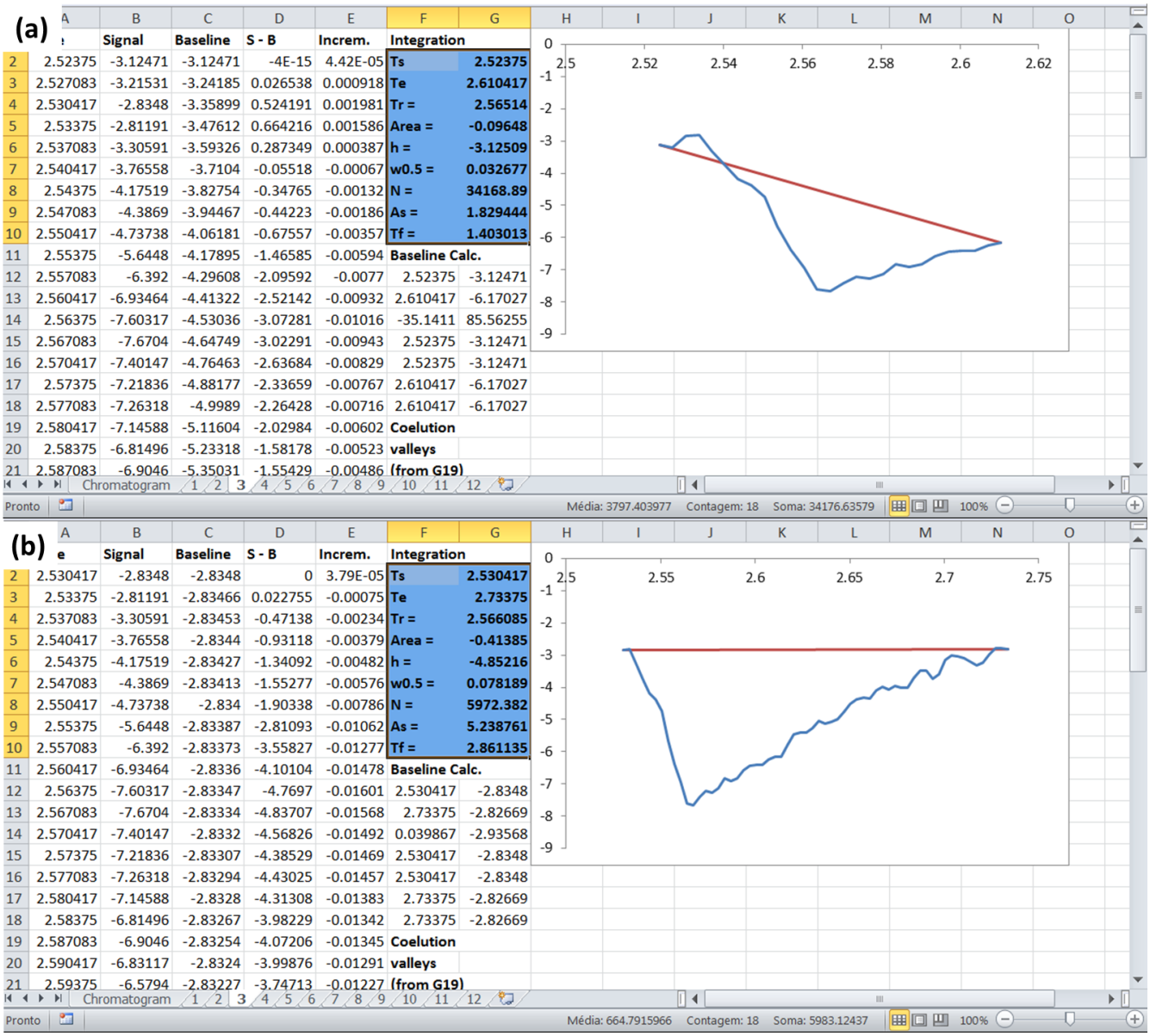

Figure S3. (a) Integration of a peak with limits not properly defined; (b) the same peak manually adjusted at both sides ("Decrease at left" and "Increase at right" functions) and reintegrated.

Splitting coeluted peaks

Peaks in partial coelution/comigration can be split into separate sheets for individual treatments. Once the group of coeluted peaks is integrated a macro searches for valleys (local minima) in the data set of the adjusted peak. If valleys are found, each segment delimited by valleys is extracted to new worksheets (along with integration results) and the original worksheet is deleted. Thus, a group of two peaks (Figure S4a), for instance, will be split into two parts. Two splitting modes are available: the "drop-line", where the baseline is kept intact (Figure S4b); and the "valley-to-valley", where a new baseline, touching the limits of each peak, is defined (Figure S4c). If the macro did not recognize any valley or if the peaks

are highly overlapped, the time coordinates of valleys or shoulders can be manually entered in specific cells. The splitting procedure is not listed in the "Do everything" commands package. Thus, if this process was applied, the results in the main worksheet should be updated with the grouping function.

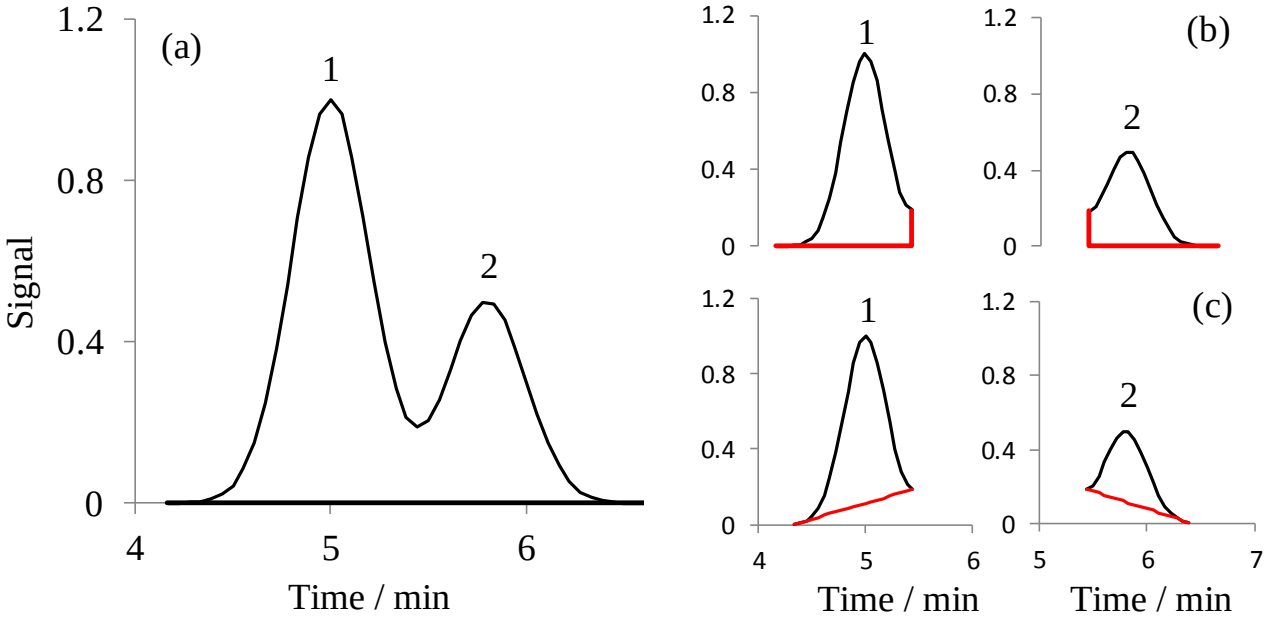

Figure S4. (a) Group of coeluted peaks (1 and 2). Peaks were split in (b) drop-line and (c) valley-to-valley modes.

Grouping the results and resolution calculation

Because integrations are made in specific worksheets for each peak the grouping of all results into a main worksheet is convenient (Figure S1b). In fact, this tool is useful if more than one peak is present in the chromatogram. Basically, a macro copies the integration results from every worksheet and pastes them in the main chromatogram worksheet. The data copied are: starting, ending and retention times, A, h, $w_{0.5}$ and N. If there are two or more peaks, the resolutions between adjacent peaks are calculated by $R_{i,i+1} = 1.175 \times (t_{Ri+1} - t_{Ri}) / (w_{0.5,i} + w_{0.5,i+1})$.

Recognition efficiency

The chromatogram in Figure S5 was used to test the recognition efficiency of the program and the results are shown in Table 1 of the main manuscript.

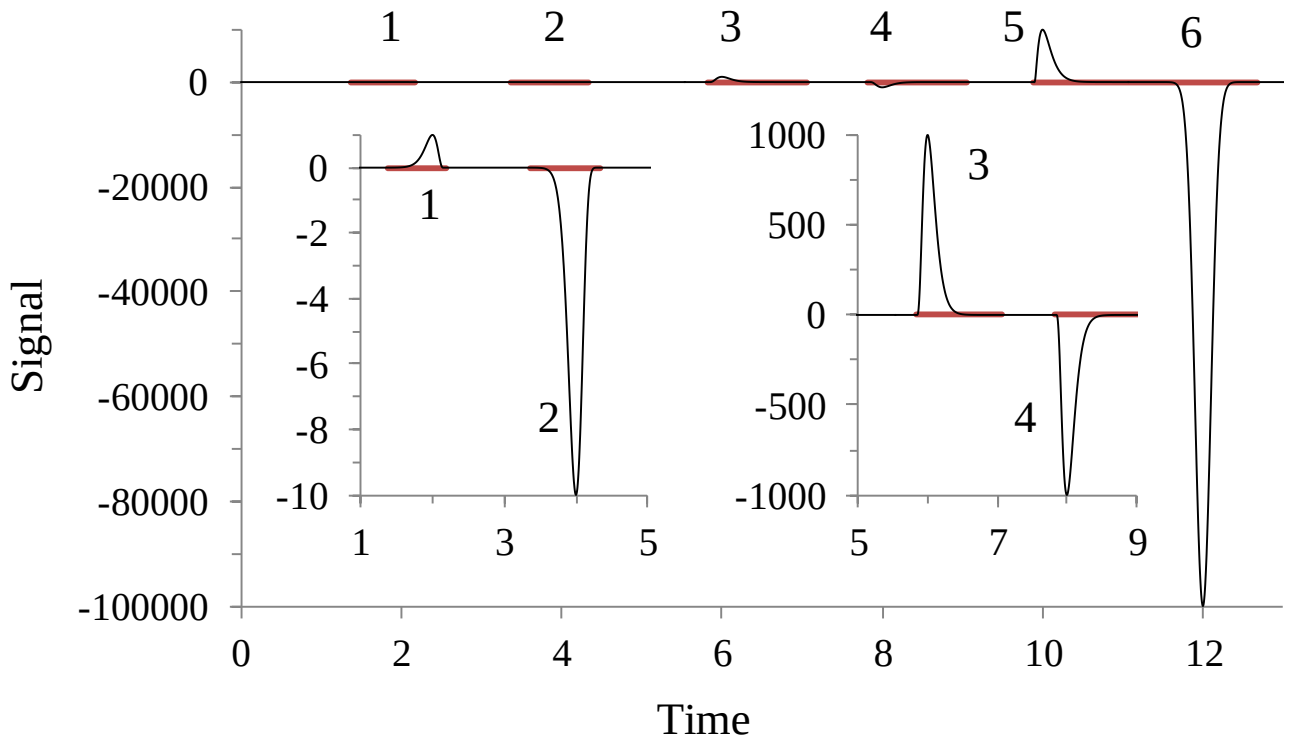

Figure S5. Simulated chromatogram with 14 min, sampling rate of 2 Hz (1681 points) containing six peaks with varied heights (h) and asymmetry terms (a, from equation 1): (1) h = 1, a = –1; (2) h = –10, a = –0.5; (3) h = 1000, a = 1; (4) h = –1000, a = 1; (5) h = 10000, a = 1.25; (6) h = 100000, a = 10^{-4}.

Reference

1. Delmonte, P.; Kia, A.-R. F.; Kramer, J. K. G.; Mossoba, M. M.; Sidisky, L.; Rader, J. I.; *J. Chromatogr. A* **2011**, *1218*, 545.