1. Evaluation of chromatographic resolution

1.1. Peak prominence

The peak prominence is an elementary resolution criterion recently developed [T. Alvarez-Segura, A. Gómez-Díaz, C. Ortiz-Bolsico, J. R. Torres-Lapasió, M.C. García-Alvarez-Coque, A chromatographic objective function to characterise chromatograms with unknown compounds or without standards available, *J. Chromatogr. A* 2015, **1409**, 79–88], which can be classified in the group of valley-to-peak functions, but with a significant difference: in conventional valley-to-peak functions, the maxima of two adjacent peaks are compared to a property of the valley that lies between them [R. D. Caballero, S. J. López-Grío, J. R. Torres-Lapasió, M. C. García-Alvarez-Coque, Single-peak resolution criteria for optimization of mobile phase composition in liquid chromatography, *J. Liq. Chromatogr. Rel. Technol.* 2001, **24**, 1895–1919], whereas in the new function, the area of a peak is delimited between the two valleys that separate it from other peaks (Fig. S1).

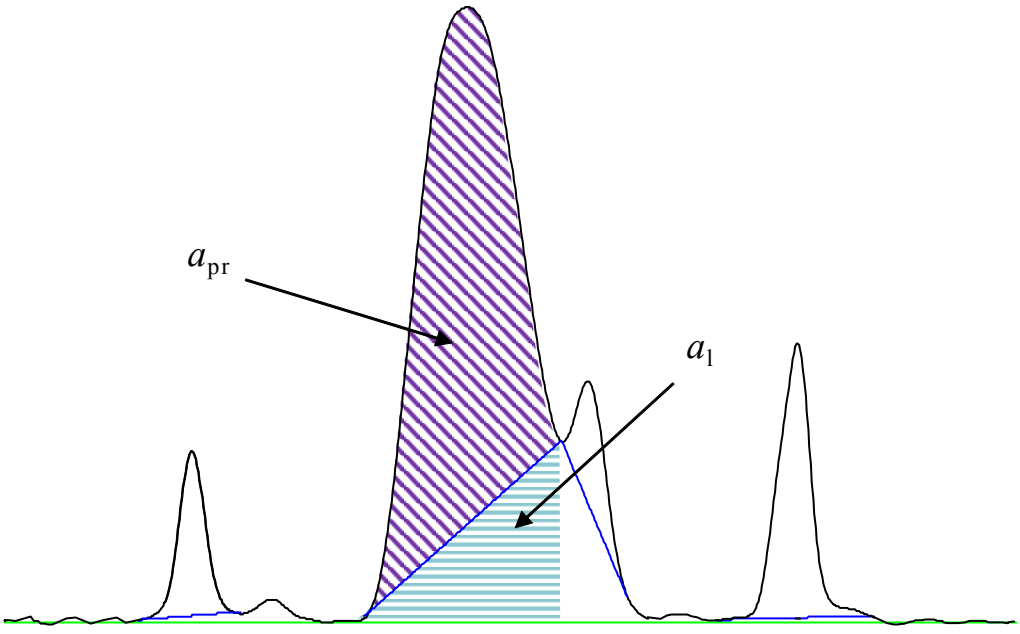


Fig. S1. Peak prominence criterion (see Eq. (S1)). The line that joins the valleys delimiting each peak divides it in two regions of area a_{pr} and a_l. The total peak area (a_T) is approximated to the sum of a_{pr} and a_l.

The aim of this new resolution criterion is to quantify the relationship between the size of the peak area that is above the valleys that define it (or above the baseline) and its total area (Fig. S1):

$$pr_i = \frac{a_{\mathrm{pr},i}}{a_{\mathrm{T}}} = \frac{a_{\mathrm{pr},i}}{a_{\mathrm{pr},i} + a_{\mathrm{l},i}} \qquad \text{(S1)}$$

The peak prominence has several advantageous features for measuring the chromatographic resolution:

(i) It is a normalized function, which facilitates its interpretation.

(ii) It allows the inclusion of the size ratio between neighbouring peaks.

(iii) It qualifies individual peaks, instead of peak pairs.

(iv) It does not require the measurement of the properties of the peaks obtained from standards.

The latter feature differentiates the peak prominence from the peak purity, which requires the information of individual signals obtained through retention and peak profile models, established through design of experiments. Hence, the peak prominence is an ideal criterion for the measurement of the resolution of chromatographic fingerprints and, in general, of the experimental chromatograms of any sample.

1.2. Peak purity

A reliable measurement of the resolution requires information not only about the position of the chromatographic peaks, but also about their full profile. In 1986, Schoenmakers wrote a pioneering work in the field of chromatographic optimization, where the use of the "overlapped peak fraction" to measure accurately the resolution was described [P. J. Schoenmakers, *Optimisation of Chromatographic Selectivity: A Guide to Method Development*, Elsevier, Amsterdam, 1986]. However, for many years it was no more than a proposal, as it required knowledge of the position and profile for each peak, and complex and laborious numerical calculations with the assistance of software. Fortunately, the proposal of new more practical peak models, together with the development of computers in the last decades and its widespread use

in laboratories, have returned the interest in this criterion. Thus, a function that measures the peak purity (the complement of the overlapped fraction) was proposed [M. C. García-Alvarez-Coque, J. R. Torres- Lapasió, J. J. Baeza-Baeza, Models and objective functions for the optimisation of selectivity in reversed-phase liquid chromatography, *Anal. Chim. Acta* 2006, **579**, 125–145].

The peak purity quantifies the percentage of peak area for a given analyte free of interference, considering as such all other peaks in the chromatogram (see Fig. S2). The peak purity may be expressed as follows:

$$p_i = \frac{a_\mathrm{f}}{a_\mathrm{T}} = \frac{a_\mathrm{f}}{a_\mathrm{f} + a_\mathrm{o}} = 1 - \frac{a_\mathrm{o}}{a_\mathrm{f} + a_\mathrm{o}} \qquad \text{(S2)}$$

where a_o is the area under the analyte peak overlapped by a hypothetical chromatogram built with the peaks of the accompanying compounds in the sample (the overlapped area), and a_f is the peak area free of interference (the free area). The resolution value obtained in this way tends to zero when the overlap of the analyte peak with the peaks of the interferences is total, and reaches $p_i = 1$ when the peak is fully resolved. However, it should be noted that the peak purity depends on the relative peak areas.

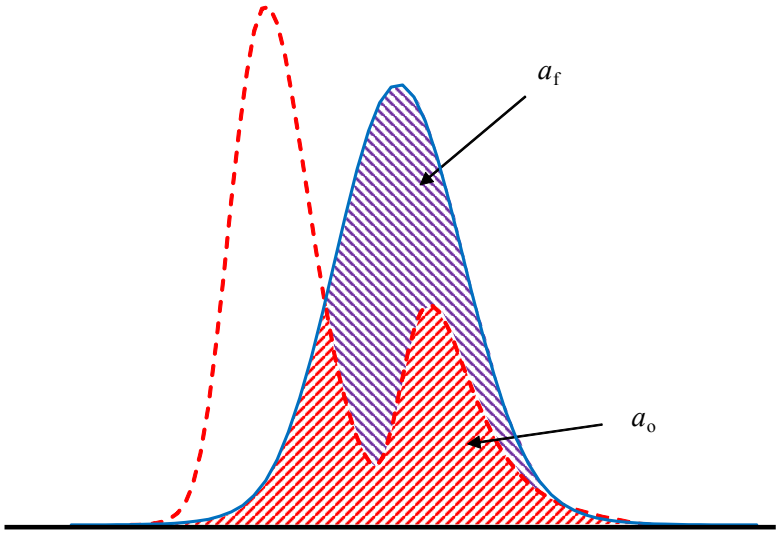


Fig. S2. Peak purity criterion (see Eq. (S2)). The free fraction area (a_f) and overlapped area with other peaks in the chromatogram (a_o) is shown. The total peak area (a_T) is the sum of a_f and a_o.

Even in situations where the chromatograms contain peaks remarkably deformed and largely overlapped, the peak purity shows an excellent correlation with the assessment of the resolution of expert analysts. For this reason, it has been considered as the best measurement of the resolution [E. Tyteca, G. A. Desmet, Universal comparison study of chromatographic response functions, *J. Chromatogr. A* 2014, **1361**, 178–190]. It also has a number of features, which generally make it the most appropriate criterion.

(i) Its meaning is very intuitive: it correlates with the information the analyst is interested on, that is, the interference level. For example, a value of 0.98 peak purity simply means that 98% of the peak of interest is free of interference (in other words, it shows 2% of interference or overlap).

(ii) It provides a realistic evaluation of the separation capability of the system, and can be easily applied to situations of diverse complexity, taking into account the full signal (peak profile, size and noise).

(iii) It is an inherently normalized measurement, which facilitates the combination of elementary resolution values into a single global measurement and the combination with other quality criteria.

(iv) One of the most important features, due to the consequences that it entails, is the qualification of individual peaks rather than peak pairs, so there is no possibility of unambiguous relationships between the identities of the peaks and the numerical resolution values. In addition, knowledge of the identity of the neighbouring peaks is not as important as it is for the criteria related to peak pairs, like the classical R_S criterion. All this allows operations such as peak weighting or exclusion easier, which avoids problems associated to peak identities in situations of peak reversals.

The concept of peak purity has allowed the development of new optimization strategies. On the one hand, the fact that it is able to anticipate the maximal resolution capability of the separation system is particularly useful for dealing with situations of low resolution, where conventional resolution criteria fail [A. Ortín, J. R. Torres-Lapasió, M. C. García-Alvarez-Coque, Finding the best separation in situations of

extremely low chromatographic resolution, *J. Chromatogr. A* 2011, **1218**, 2240–2251]. On the other hand, it allows the simultaneous optimization of two or more mobile phases, eluents and/or columns, or even separation techniques (complementary situations) [C. Ortiz- Bolsico, J. R. Torres-Lapasió, M. C. García-Alvarez-Coque, Approaches to find complementary separation conditions for resolving complex mixtures by high performance liquid chromatography, *J. Chromatogr. A* 2012, **1229**, 180–189]. The only drawback is that it is designed for the evaluation of the resolution through simulations and it is hard to apply directly to experimental chromatograms of mixtures, since individual contributions are not available.

2. Additional information about the comparison of peak prominences versus peak purity

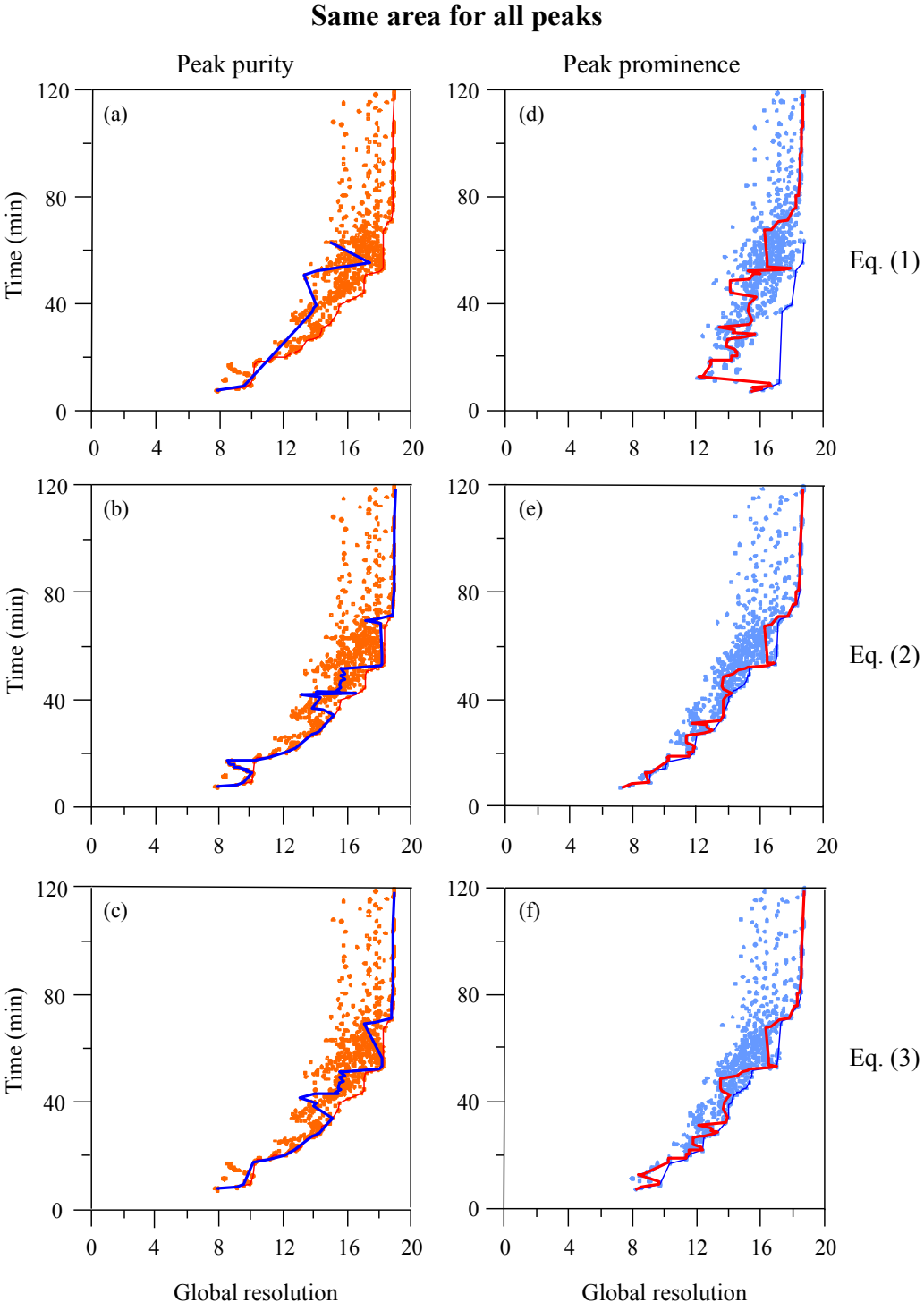


Fig. S3. Pareto plots according to different global resolution criteria, corresponding to the separation of the OPA-NAC derivatives of the 19 amino acids assuming normalized peak areas. See Fig. 2 and text for other details.

3. Automatically processed fingerprints of medicinal herbs

Figs. S4 and S5 show MATLAB screenshots of chromatograms for extracts of medicinal herbs after being processed by the developed application. The peaks are numbered according to their elution order, and only those that exceed a relative peak area of (0.05%) are shown. Blue tangents define the optimal protruding region for each peak, which is marked in red.

The inserted tables in each chromatogram show some of the parameters that the developed application provides: peak index, start, maximum and end peak times, area of the protruding part of the peaks ($a_{\mathrm{pr},i}$), total area ($a_{\mathrm{pr},i}$ + $a_{\mathrm{l},i}$) (see Fig. S1), and prominence (Eq. (S1)) expressed as percentage.

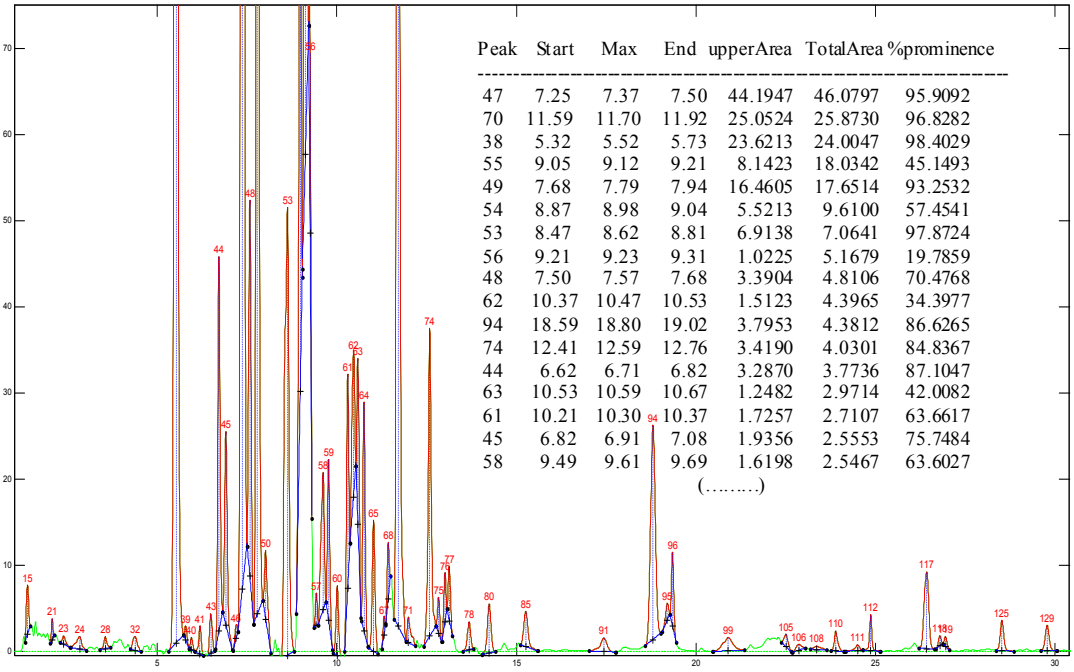
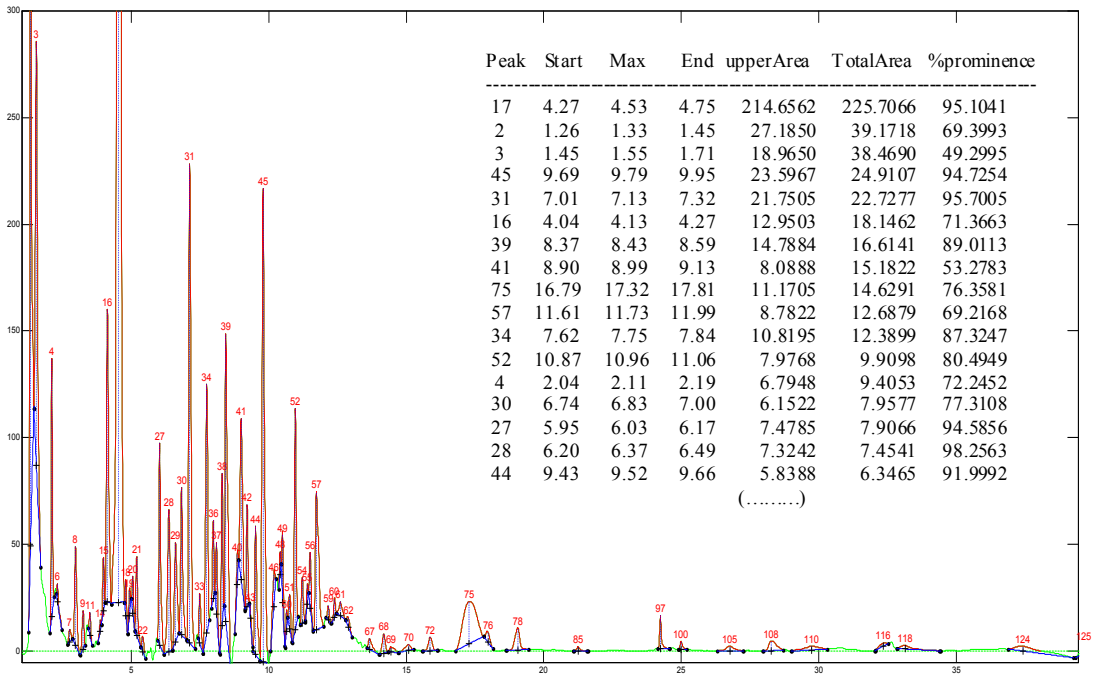


Fig. S4. Screenshot for horsetail tea.



Peak	Start	Max	End	upperArea	TotalArea	%prominence
17	4.27	4.53	4.75	214.6562	225.7066	95.1041
2	1.26	1.33	1.45	27.1850	39.1718	69.3993
3	1.45	1.55	1.71	18.9650	38.4690	49.2995
45	9.69	9.79	9.95	23.5967	24.9107	94.7254
31	7.01	7.13	7.32	21.7505	22.7277	95.7005
16	4.04	4.13	4.27	12.9503	18.1462	71.3663
39	8.37	8.43	8.59	14.7884	16.6141	89.0113
41	8.90	8.99	9.13	8.0888	15.1822	53.2783
75	16.79	17.32	17.81	11.1705	14.6291	76.3581
57	11.61	11.73	11.99	8.7822	12.6879	69.2168
34	7.62	7.75	7.84	10.8195	12.3899	87.3247
52	10.87	10.96	11.06	7.9768	9.9098	80.4949
4	2.04	2.11	2.19	6.7948	9.4053	72.2452
30	6.74	6.83	7.00	6.1522	7.9577	77.3108
27	5.95	6.03	6.17	7.4785	7.9066	94.5856
28	6.20	6.37	6.49	7.3242	7.4541	98.2563
44	9.43	9.52	9.66	5.8388	6.3465	91.9992
				(……..)		

Fig. S5. Screenshot for decaffeinated tea.