Journal of Chromatography A, 1218 (2011) 5819–5828

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Relationship between selectivity and average resolution in comprehensive two-dimensional separations with spectroscopic detection

Joe M. Davis[a,*], Sarah C. Rutan[b], Peter W. Carr[c]

[a] Southern Illinois University at Carbondale, Department of Chemistry and Biochemistry, 1245 Lincoln Drive, Carbondale, IL 62901-4409, USA
[b] Department of Chemistry, Virginia Commonwealth University, 1001 W. Main Street, Box 842006, Richmond, VA 23284-2006, USA
[c] Department of Chemistry, University of Minnesota, Smith and Kolthoff Halls, 207 Pleasant Street SE, Minneapolis, MN 55455-0431, USA

ARTICLE INFO

Article history:
Received 31 March 2011
Received in revised form 5 June 2011
Accepted 22 June 2011
Available online 30 June 2011

Keywords:
Multivariate selectivity
Resolution
Peak-broadening factor
Comprehensive two-dimensional separation
Diode array detection
Statistical-overlap theory

ABSTRACT

The average value of the multivariate selectivity (*SEL*) of randomly positioned peaks in a multi-component separation is shown to equal the average fraction of peaks that are singlets, as predicted by statistical-overlap theory (SOT). This equality is the basis for proposing a useful metric, specifically the average minimum resolution of nearest-neighbor peaks, for the performance of comprehensive two-dimensional (2D) separations. Furthermore this metric was computed both without ancillary spectroscopic information and with the assistance of such help, specifically multi-wavelength UV–vis spectra, acquired during the separation. Separations are simulated with randomly positioned peaks over wide ranges of total number of peaks, first- and second-dimension peak capacity, dimensionless first-dimension sampling time, and spectral diversity. The specific version of the general multivariate selectivity concept that is used here – identified as *SEL* – gives the relative precision of quantification when using the PARAFAC (parallel factor analysis) method, a popular curve resolution algorithm. The *SEL* values of all peaks were calculated, averaged, and compared to the predictions of SOT. In the absence of auxiliary spectral data, the *SEL*-based average minimum resolution required to separate two peaks in a 2D separation is 0.256, compared to resolution of 0.5 if no chemometric assistance is available. This was found to be valid over a wide range of conditions and is essentially independent of peak crowding. With the assistance of the spectral data, the requisite minimum resolution substantially improves, that is, it decreases, especially when peak crowding is severe. The requisite minimum resolution decreases even further, up to a limit, as the spectral diversity is increased. In contrast, the *SEL*-based average under-sampling correction factor is virtually independent of the presence of the additional spectral data, and additionally is about the same as calculated with SOT from the average number of maxima in closely analogous simulations. The use of selectivity greatly increases the fraction of peaks that are singlets, relative to the number of singlet maxima, especially when spectral assistance is added. The insensitivity of the under-sampling correction factor to either the use of selectivity or added spectral data simplifies optimization of the corrected peak capacity in on-line comprehensive 2D separations.

1. Introduction

The popularity of multivariate analysis, e.g., the use of absorbance measurements at many wavelengths, has been accompanied by the development of various formalized metrics. One of these is selectivity. For traditional analytical methods, selectivity is defined as the ratio of the slope of the calibration curve for the target analyte to the slope of the calibration curve for a specified interferent [1]. It is desirable to generalize this metric for use with multivariate methods, where multiple interferents occur and multiple sensors (e.g., wavelengths) are used. Lorber proposed a metric for multivariate selectivity based on the net analyte signal [2,3]. It is computed as that part of the target analyte signal that is *orthogonal* to the total signal. The multivariate selectivity (or simply selectivity) is the normalized net analyte signal, equaling that fraction of the total signal that is *unique* to the analyte [2,4,5]. Thus, the selectivity of a perfectly selective method is unity, whereas that of a totally unselective method is zero. Selectivity measures the expected accuracy and precision of an analysis, relative to one where no interferents are present. For example, a target analyte with a selectivity of 0.5 can be quantified with an accuracy and precision only one-half as good as that in the absence of any interferents. The calculation assumes a fixed noise level throughout.

Recent studies have shown the suitability of such metrics for selectivity for the characterization of the performance of multi-

* Corresponding author. Current address: 733 Schloss Street, Wrightsville Beach, NC 28480, USA. Tel.: +1 910 256 4235.
E-mail address: chimicajmd@ec.rr.com (J.M. Davis).

0021-9673/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.chroma.2011.06.086

dimensional chromatographic separations [6–8]. Our interest lies in applying these metrics to assess various attributes of comprehensive two-dimensional liquid chromatography (LC × LC), with and without the benefits of auxiliary spectroscopic data, more specifically by use of a diode array detector (DAD). An attribute of particular importance in characterizing the performance of LC × LC methods is the average resolution of nearest-neighbor peaks. Although relationships between selectivity and resolution have been reported [6,9–11], they were based only on small numbers of overlapping peaks. In contrast, a relationship based on the partial overlap of hundreds of peaks could serve as a more representative measure of the effect of selectivity on the entire separation. It would also be useful to know how the average resolution varies with the diversity of the auxiliary spectra.

Another characteristic of interest in 2D separations is the expected increase in first-dimension peak width due to the almost unavoidable problem of under-sampling the first-dimension effluent, prior to its injection into the second column. This effect is quantified by a peak-broadening factor, which can be quite large (e.g., more than an eight-fold broadening [12]), with a concomitant loss of peak capacity by the same factor [13]. Although in principle the broadening factor is unique, its value varies with the model used to regenerate the inherently Gauss-like first-dimension peaks from the data obtained at the end of the second-dimension separation, e.g., as histograms [14], digital pulses [15], or linear interpolations [16–18]. This is so because a unique regeneration does not exist when under-sampled peaks do not satisfy the Nyquist theorem [16]. Since the multivariate selectivity is based on only the orthogonal part of the total analyte signal, no regeneration formalism is required and the broadening factor that results from under-sampling could well depend on the amount of help that comes from the spectral data, as previously hypothesized by Potts et al. [19].

In the absence of additional spectral data, these attributes – the average minimum resolution and the peak-broadening factor – have been characterized in multi-component separations [16,20,21] by the use of statistical-overlap theory (SOT), which describes the expected amount of peak overlap in a large ensemble of separations. In this work we will show the important unifying result that the *SEL* can be interpreted as the fractional singlet character of a peak. The average *SEL* of randomly positioned peaks is equal to the fraction of singlet peaks (singlets) predicted by SOT. In light of this result, the average resolution and peak-broadening factor were predicted from *SEL* values calculated from simulations of comprehensive two-dimensional (2D) separations containing randomly positioned peaks.

The effect of additional information provided by a DAD on these attributes was evaluated by randomly assigning UV–vis spectra to chromatographic peaks from a pool of different UV–vis spectra. Although our interest is in LC × LC, with and without help from DAD, our methods are general; therefore, our results are relevant to other comprehensive 2D separations and multi-sensor detection methods, e.g., GC × GC with mass-spectral detection.

2. Theory

2.1. Review of multivariate selectivity

Under the assumption that all analytes present in the unknown samples are also present at known concentrations in calibration samples, the value of the multivariate selectivity *SEL* of the *i*th peak in a 2D chromatogram is [22]

$$SEL = [(\mathbf{A}_1{}^T\mathbf{A}_1 * \mathbf{A}_2{}^T\mathbf{A}_2)^{-1}]_{i,i}^{-1/2} \quad (1)$$

where $\mathbf{A}_1$ and $\mathbf{A}_2$ are matrices containing the first- and second-dimension chromatographic peak profiles for each peak, respectively. In Eq. (1), '*' indicates the Hadamard or element-wise matrix product, '*T*' indicates the matrix transpose, '−1' indicates the matrix inverse, and '*i, i*' indicates the *i*th diagonal element of the resultant matrix. This calculation is based on the formalism of Messick et al. [9,22], which determines the selectivity of the PARAFAC algorithm. Eq. (1) can be extended to include additional spectroscopic information by [22]

$$SEL = [(\mathbf{A}_1{}^T\mathbf{A}_1 * \mathbf{A}_2{}^T\mathbf{A}_2 * \mathbf{A}_3{}^T\mathbf{A}_3)^{-1}]_{i,i}^{-1/2} \quad (2)$$

where $\mathbf{A}_3$ is a matrix containing the spectra, e.g., UV–vis or mass spectra, of all chromatographic profiles in $\mathbf{A}_1$ and $\mathbf{A}_2$. In this paper, *SEL* is calculated only as a metric: no curve resolution is actually made.

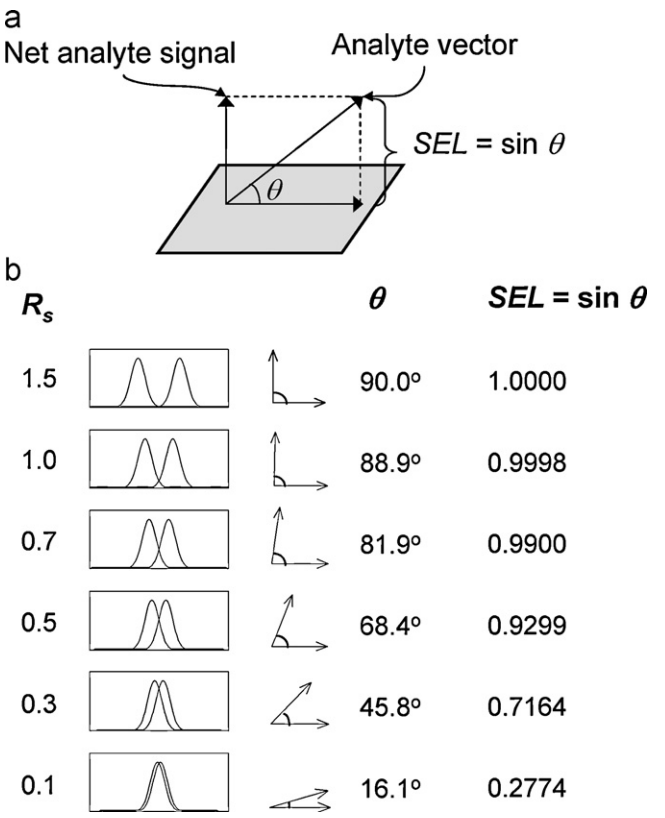

Fig. 1. Schematic illustration of net analyte signal and multivariate selectivity value, *SEL*, for one-dimensional chromatography. (a) The gray plane indicates the (hyper)plane of all signals except for the analyte. The analyte vector is at an angle θ to that plane, such that the *SEL* value is given as $\sin\theta$. (b) Relationship among resolution (R_s), angle (θ), and selectivity (*SEL*) for two peaks.

As pointed out previously, the *SEL* value corresponds to that part of the signal of interest that is *orthogonal* to the signals for all other peaks that contribute to the overall signal. This is simply illustrated using the vector interpretation of *SEL* based on the formalism of Messick et al. [9]. As shown in Fig. 1a, the vector representing an analyte signal can be decomposed into two vectors, one that lies in a (hyper)plane spanned by the vectors representing all other analyte signals, and a second vector – the net analyte signal – that is orthogonal to this (hyper)plane. Only the net analyte signal carries unique analyte information, and the magnitude of its associated metric of multivariate selectivity, *SEL*, equals the sine of the angle between the analyte vector and the (hyper)plane. To illustrate the meaning and significance of *SEL*, we use a simulation of two overlapped peaks in one-dimensional chromatography, as shown in Fig. 1b. Here we show two chromatographic peaks with varying degrees of overlap, whose resolutions range from 0.1 to 1.5. If each peak was described by 100 points, then we could also describe

each peak by a vector in a 100-dimensional space. Reducing this space to two dimensions for ease of visualization, we can show the corresponding angle between the two different chromatographic peaks as the angle between the two vectors. The vector component of the analyte signal that is not selective lies in the (hyper)plane; the selective vector component for the analyte of interest lies perpendicular to it (see Fig. 1a). As the resolution of the two peaks decreases in Fig. 1b, the vectors become more oriented in the same direction (i.e., each carries less unique information relative to the other), the angle between them decreases, and *SEL* decreases. Thus large *SEL* values are associated with well resolved peaks, and small *SEL* values are associated with poorly resolved peaks. In Fig. 1b, both peaks have the same *SEL* value, but in multi-component separations different peaks will have different *SEL* values depending on the amount of peak overlap. It should be pointed out that the *SEL* value is independent of the noise level and the number of data points in each peak (except in the limit of under-sampling).

The geometric interpretation of *SEL* shown in Fig. 1 is not changed on addition of subsequent dimensions of data; only the numerical value of *SEL* is changed. The effect of using spectral information to aid in the resolution of the signals (e.g., LC-DAD) is shown in Fig. 2. Here a pair of severely overlapped chromatographic peaks (R_s = 0.1, *SEL* = 0.277, top of figure) is used to illustrate the improvement in SEL upon using DAD detection. Four different spectral pairs (shown on the right in the figure) are shown with *SEL* values of the spectral information alone ranging from 0.227 to 0.541. The *SEL* values for the combined information are reported as well, where the *SEL* values range from 0.384 to 0.589. It can be seen that the largest improvement in *SEL* is seen when the *SEL* values of the individual techniques are similar, at least for the two-component case.

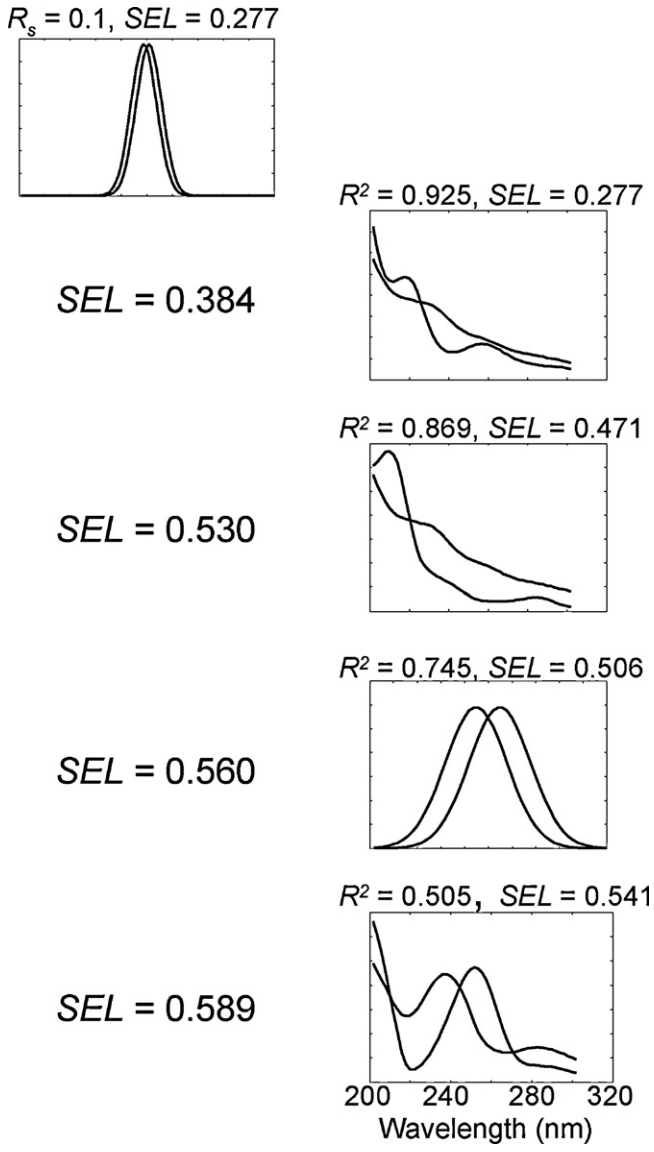

Fig. 2. Illustration of the effect of added spectral information on the value of *SEL*. This figure shows the increase in selectivity provided for a chromatogram with two peaks having a resolution of 0.1 and *SEL* of 0.277 (top panel), upon the addition of spectral information. The squared correlation coefficient R^2 and *SEL* values are shown for each pair of spectra on the right, and the *SEL* of the combined information is shown to the left of the spectra.

2.2. Review of SOT

Consider a series of comprehensive 2D separations with first and second dimension durations ${}^1D$ and ${}^2D$. On average they contain $\bar{m}$ randomly positioned peaks having (prior to sampling) first- and second-dimension standard deviations ${}^1\sigma$ and ${}^2\sigma$. Upon sampling, the average first-dimension peak width is effectively increased by a factor known as $\langle\beta\rangle$ [14]. The SOT prediction of the expected number of singlets s in these separations is [23]

$$s = \bar{m}p_1 \tag{3}$$

where p_1 is the singlet probability [23]

$$p_1 = w_i p_{1,i} + w_e p_{1,e} + w_c p_{1,c} \tag{4}$$

In Eq. (4), $p_{1,i}$, $p_{1,e}$, and $p_{1,c}$ are the independent probabilities of forming singlet peaks in the interior (i), edge (e), and corner (c) regions of a rectangular 2D space. They are given elsewhere [23] as functions of the saturation α. The saturation is a metric of peak crowding, which in comprehensive 2D separations is [16]

$$\alpha = \frac{\pi\bar{m}(R_s^*)^2}{4n'_{c,2D}} \tag{5}$$

where R_s^* is the average minimum resolution sufficient to separate a singlet from its nearest-neighbor peak and $n'_{c,2D}$ is the corrected 2D peak capacity, i.e., the traditional 2D peak capacity [24–26] corrected for the sampling-induced peak broadening [13]

$$n'_{c,2D} = \frac{{}^1n_c\,{}^2n_c}{\langle\beta\rangle} \tag{6a}$$

In Eq. (6a), ${}^1n_c$ and ${}^2n_c$ are the conventional (uncorrected) first- and second-dimension peak capacities at unit resolution

$${}^1n_c = \frac{{}^1D}{4\,{}^1\sigma} \tag{6b}$$

$${}^2n_c = \frac{{}^2D}{4\,{}^2\sigma} \tag{6c}$$

The factors w_i, w_e, and w_c in Eq. (4) are the fractions of the rectangular 2D space associated with the interior, edge, and corner regions. They have been expressed elsewhere in geometric terms [23] but also can be written as

$$w_i = 1 - 2R_s^*\left(\frac{\langle\beta\rangle}{{}^1n_c} + \frac{1}{{}^2n_c}\right) + \frac{4(R_s^*)^2}{n'_{c,2D}} \tag{7a}$$

$$w_e = 2R_s^*\left(\frac{\langle\beta\rangle}{{}^1n_c} + \frac{1}{{}^2n_c}\right) - \frac{8(R_s^*)^2}{n'_{c,2D}} \tag{7b}$$

$$w_c = \frac{4(R_s^*)^2}{n'_{c,2D}} \tag{7c}$$

Eqs. (3)–(7a–c) are exact as long as $2R_s^* \langle \beta \rangle / {}^1n_c$ and $2R_s^*/{}^2n_c$ are less than one [23]. They correct for the "edge effect", that is, the reduced probability of peak overlap that exists near the edges and corners (i.e., the boundaries) of a rectangular 2D separation [23]. The edge effect exists because there are no peaks outside the boundaries available to overlap with peaks just inside and near the boundaries [27]. In the present study, the edge effect can be appreciable, especially when ${}^1n_c$ is small and the sampling rate is slow, causing peaks to become very elongated in the first dimension. The edge effect becomes negligible as ${}^1n_c$ and ${}^2n_c$ become increasingly large, with weight w_i approaching one, weights w_e and w_c approaching zero, and Eq. (4) approaching the asymptotic result [28]

$$p_{1,i} = \exp(-4\alpha) \tag{8}$$

Eq. (8) is the classic equation for the singlet probability in a hypothetical 2D separation in a plane of infinite extent, i.e., an infinite traditional 2D peak capacity.

2.3. Connection between SEL and SOT

In accordance with our earlier discussion, a vector representing the signal of an analyte peak can be resolved into two components. One component lies in the (hyper)plane spanned by vectors representing signals from all other peaks, and the other is orthogonal to the (hyper)plane. With vector mathematics, the component of the analyte signal in the hyper(plane) can be expressed by combining vectors representing signals from all other peaks. In other words, this part of the analyte signal overlaps with other signals. Chromatographically, it behaves as a peak in a multiplet. In contrast, the orthogonal part of the analyte signal does not overlap with other peak signals; it behaves like a singlet peak of a single chemical species.

By this interpretation, all peaks in chromatographic separations have both fractional singlet and multiplet characters. Furthermore, in our view, *the fractional singlet character of any peak equals SEL*. Clearly a pure singlet peak would then have *SEL* equal to unity whereas a badly overlapped peak would have a much lower *SEL*. This is consistent with the variation of *SEL* with chromatographic resolution, as shown in Fig. 1b. It follows that the expected fractional singlet character of the entire separation is the average selectivity $\langle SEL \rangle$, taken as an average of the *SEL* values of all peaks. Thus, for randomly positioned peaks in a multi-component separation, this closely corresponds to the singlet probability p_1, Eq. (4)

$$\langle SEL \rangle \approx p_1 = \frac{s}{\tilde{m}} \tag{9}$$

where the final identity comes from Eq. (3). Eq. (9) is the central operating premise of this work.

2.4. Protocol

Our computational protocol determines R_s^* and $\langle \beta \rangle$ independently. First, the relationship between the average minimum resolution R_s^* and saturation α is established by using SOT to interpret the average selectivities $\langle SEL \rangle$ from simulated comprehensive 2D separations so rapidly sampled that $\langle \beta \rangle \approx 1$; thus, all parameters but R_s^* are known. Second, this $\alpha - R_s^*$ relationship is used to establish $\langle \beta \rangle$ by using SOT to interpret $\langle SEL \rangle$ values from simulations of under-sampled separations. In these simulations, all parameters but $\langle \beta \rangle$ are known. The influence of the DAD data is evaluated by randomly assigning UV–vis spectra to different peaks. The amount of additional spectroscopic data is varied by changing the number of distinct spectra available to be assigned to the different peaks.

Table 1
Combinations of $\tilde{m}$, ${}^1n_c$, and ${}^2n_c$ used to determine average minimum resolution R_s^* as function of saturation α.

$\tilde{m}$	${}^1n_c$	${}^2n_c$	$\tilde{m}/({}^1n_c{}^2n_c)$[a]	${}^1\sigma/{}^2\sigma$[b]
100	40	80	0.03125	2
100	40	40	0.0625	1
100	40	20	0.125	0.5
100	20	20	0.25	1
100	20	6.67	0.75	0.33
100	20	3.33	1.5	0.17
100	20	2	2.5	0.1
100	10	2.86	3.5	0.29
200	20	20	0.5	1
200	10	20	1	2
200	10	10	2	1
200	5	20	2	4
200	3.33	20	3	6
200	2.5	20	4	8
200	6.67	6.67	4.5	1
200	2	20	5	10
200	1.67	20	6	12
200	5	5	8	1
300	6.55	6.55	7	1
300	2.51	12.57	9.5	5
300	15.49	1.55	12.5	0.1
400	6.03	6.03	11	1
400	11.95	2.39	14	0.2
400	1.61	16.06	15.5	10

[a] Equal to effective saturation at $\langle \beta \rangle = 1$ (see Eq. (10)).
[b] Equal to ${}^2n_c/{}^1n_c$.

Thus no spectral help is provided when all chromatographic peaks share a single spectrum.

3. Procedures

3.1. Simulations

Comprehensive 2D separations of $\tilde{m}$ peaks ($100 \le \tilde{m} \le 400$) were simulated based on equal-height Gaussian profiles for both the first- and second-dimension peaks. Uniformly random retention times were distributed over the durations, ${}^1D = {}^2D = 0.8$, and standard deviations ${}^1\sigma$ and ${}^2\sigma$ were determined by ${}^1n_c$ and ${}^2n_c$ values in Tables 1 and 2, and by Eqs. (6b) and (6c). Equal-height chromatographic peaks were used in this study, even though peak heights in multi-component separations often have an exponential-like [29–33] or log–normal [34] distribution, because the multivariate-selectivity calculations are based on peaks normalized to a vector length of one. The first-dimension profiles (contained in matrix $\mathbf{A}_1$) were generated by integrating sections of the first-dimension Gaussian over sampling times of duration t_s and dividing by t_s, with the times at the interval centers. The second-dimension profiles (contained in matrix $\mathbf{A}_2$) were simply the simulated Gaussian peaks, with points spaced at intervals of 0.001. For each combination of $\tilde{m}$, ${}^1n_c$, ${}^2n_c$, and $t_s/{}^1\sigma$, the simulations were repeated 50 times with different random peak coordinates. The *SEL* values of peaks in each 2D chromatogram were computed

Table 2
Combinations of ${}^1n_c$ and ${}^2n_c$ used to determine average peak-broadening factor $\langle \beta \rangle$.

${}^1n_c$	${}^2n_c$	${}^1\sigma/{}^2\sigma$[a]
20	160	8
80	40	0.5
80	20	0.25
40	40	1
20	40	2
80	10	0.125
20	20	1

[a] Equal to ${}^2n_c/{}^1n_c$.

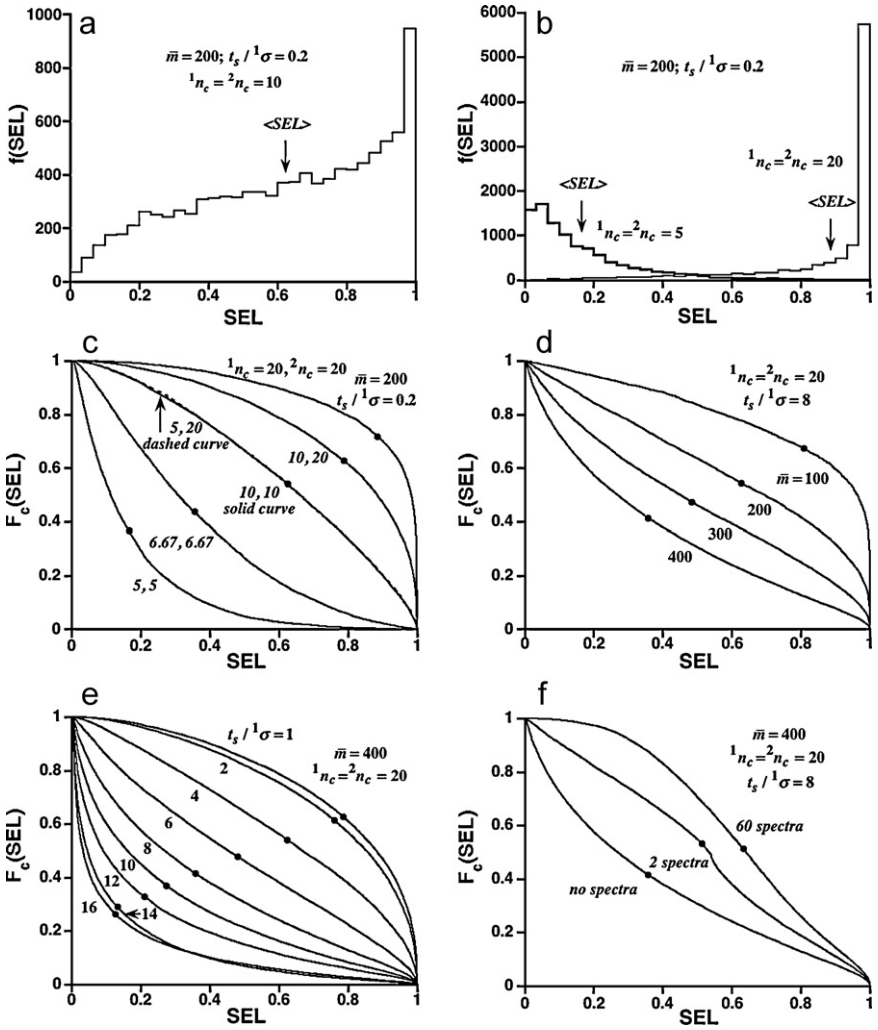

Fig. 3. (a) and (b) Graphs of observed frequency of occurrence $f(SEL)$ vs. selectivity SEL for $\bar{m}=200$, $t_s/{}^1\sigma=0.2$, and different ${}^1n_c$, ${}^2n_c$ values. Average selectivities $\langle SEL\rangle$ are shown. Distributions contain 10,000 SEL values (200 values per simulation times 50 simulations). (c) Graphs of survival function $F_c(SEL)$ vs. SEL for $\bar{m}=200$, $t_s/{}^1\sigma=0.2$, and different ${}^1n_c$, ${}^2n_c$ values identified as coordinate pairs. $\langle SEL\rangle$ values are represented by solid circles. (d) As in panel c, but for ${}^1n_c={}^2n_c=20$, $t_s/{}^1\sigma=8$, and different $\bar{m}$. (e) As in panel c, but for $\bar{m}=400$, ${}^1n_c={}^2n_c=20$, and different $t_s/{}^1\sigma$. (f) As in panel c, but for $\bar{m}=400$, ${}^1n_c={}^2n_c=20$, $t_s/{}^1\sigma=8$, and different spectral diversity. The square of the correlation coefficient for the two-spectra case is 0.505.

from Eq. (1). The computation of the average *SEL* value is described in detail below.

To evaluate the influence of the DAD data, chromatographic peaks were randomly assigned to 60 UV–vis spectra generated from drugs of abuse and toxicological relevance that were obtained elsewhere [35,36]. The correlation coefficients of the spectra ranged from −0.30388 to 0.99998. Because the number of peaks greatly exceeded the number of spectra, a number of peaks were assigned the same spectrum. Spectra were randomly selected from either the main database of 60 spectra, from a subset of 30 spectra in the database, or from three pairs of spectra having a squared correlation coefficient of 0.505, 0.869 and 0.925, respectively. The spectral profiles (contained in matrix $\mathbf{A}_3$) consisted of absorbances taken at intervals of 2 nm from 201 to 301 nm. The *SEL* values of peaks in each 2D chromatogram were computed from Eq. (2).

3.2. Calculation of multivariate selectivity SEL

For each simulation, the *SEL* values of all peaks were calculated from Eqs. (1) and (2) using the chromatographic peak profiles and (when desired) the spectral profiles described above. Calculations were implemented using Matlab 7.5.0 (Mathworks, Natick, MA) using code described previously [7].

3.3. Calculation of average value of SEL

For each combination of $\bar{m}$, ${}^1n_c$, ${}^2n_c$, and $t_s/{}^1\sigma$, the 50 simulations produced $50\bar{m}$ values of *SEL*, (i.e., $\bar{m}$ values per simulation, multiplied by 50 simulations) from which the average value $\langle SEL\rangle$ was computed from survival functions [37]. We point out that although $\langle SEL\rangle$ could have been calculated simply by averaging the $50\bar{m}$ *SEL* values, the survival function contains information on the distribution of *SEL*, which is also of interest. These consisted of the fraction of *SEL* values that exceeded selectivity thresholds between 0 and 1, with thresholds spaced by 0.001. As shown in the Supplementary Material to this paper, the area under this survival function is $\langle SEL\rangle$.

3.4. SOT calculations

3.4.1. General considerations

The average value $\langle SEL\rangle$ was equated to p_1 in Eq. (4). For a given saturation α, the probability $p_{1,i}$ in Eq. (4) was calculated from Eq. (8), whereas the probabilities $p_{1,e}$ and $p_{1,c}$ were linearly interpo-

lated from previously calculated results for $\alpha = 0$, 0.01, 0.02, ..., 1.99, and 2.00 (results available on request).

3.4.2. Determination of R_s^*

The average minimum resolution R_s^* was determined as a function of saturation α with $\langle SEL \rangle$ values calculated from simulations based on the 24 combinations of $\bar{m}$, ${}^1n_c$, and ${}^2n_c$ in Table 1 at the very fast dimensionless sampling time, $t_s/{}^1\sigma = 0.2$. For this $t_s/{}^1\sigma$, the average peak-broadening factor $\langle \beta \rangle$ is essentially unity [14–18]. Consequently, the factor $\langle \beta \rangle$ in Eqs. (6a) and (7a–c) was set to one, leaving R_s^* as the only unknown. Thus R_s^* was determined by an iterative bisection procedure satisfying Eqs. (3)–(8). On completion, each of the 24 values of R_s^* was associated with one of the 24 values of α, thus establishing the relationship between these variables.

3.4.3. Determination of $\langle \beta \rangle$

For $t_s/{}^1\sigma$ values between 1 and 16, 28 estimates of $\langle \beta \rangle$ were made using $\langle SEL \rangle$ values calculated from simulations based on the four $\bar{m}$ values (100, 200, 300, and 400) and the seven ${}^1n_c$, ${}^2n_c$ combinations in Table 2. The appropriate value of R_s^*, consistent with the relationship between α and R_s^* discussed in the preceding paragraph, was used (details are given below). For each estimate, a small trial value of $\langle \beta \rangle$ was assigned to Eqs. (6a) and (7a–c), such that the singlet probability p_1, Eq. (4), exceeded $\langle SEL \rangle$. The trial value sequentially was incremented by 0.001 (with R_s^* corrected if needed) until p_1 was less than $\langle SEL \rangle$. This final trial value was averaged with the preceding trial value and assigned to $\langle \beta \rangle$. The average and standard deviation of the 28 $\langle \beta \rangle$ values so determined were calculated.

4. Results and discussion

4.1. SEL distributions

Fig. 3a–e shows several distributions of *SEL* calculated from Eq. (1). Panels a and b are plots of the observed frequency of occurrence $f(SEL)$ as a function of the observed *SEL* value for $\bar{m} = 200$, $t_s/{}^1\sigma = 0.2$, and various combinations of ${}^1n_c$ and ${}^2n_c$. These asymmetrical distributions and their average values $\langle SEL \rangle$ shift to greater values as ${}^1n_c$ and ${}^2n_c$ increase, and are quite different depending on the peak capacities. In panel b, with ${}^1n_c$ and ${}^2n_c$ both equal to 5, there is virtually no chance of observing a high *SEL* as the probability of a singlet peak is extremely small. Panel c shows these same results as in panels a and b, and others as well, but they are now plotted as the survival functions $F_c(SEL)$ vs. *SEL*. As discussed in the Supplementary Material, the survival function $F_c(SEL)$ has three important properties: it decreases rapidly where the frequency $f(SEL)$ is large, it decreases slowly where the frequency $f(SEL)$ is small, and the area under the graph of $F_c(SEL)$ vs. *SEL* is the average value of *SEL*, denoted as $\langle SEL \rangle$. The coordinate pairs and filled circles in panel c represent ${}^1n_c$, ${}^2n_c$ combinations and $\langle SEL \rangle$, respectively. Inspection of panel c shows that the survival function and $\langle SEL \rangle$ value for ${}^1n_c = 10$ and ${}^2n_c = 10$ (solid curve) are almost the same as for ${}^1n_c = 5$ and ${}^2n_c = 20$ (dashed curve), showing that the *product* of the two peak capacities (${}^1n_c\,{}^2n_c$) is controlling, not ${}^1n_c$ or ${}^2n_c$ individually. The small difference is due to the edge effect. Panel c also shows for fixed $\bar{m}$, the distribution and $\langle SEL \rangle$ shift to higher values as the product ${}^1n_c\,{}^2n_c$ increases, that is, as the saturation decreases.

Other important trends are shown by the survival functions in panels d–f. Panel d shows that for fixed ${}^1n_c\,{}^2n_c$ and $t_s/{}^1\sigma$, the selectivity distribution and $\langle SEL \rangle$ shift to smaller values as $\bar{m}$ increases, that is, as the saturation increases. Panel e clearly shows that for fixed ${}^1n_c\,{}^2n_c$ and $\bar{m}$, the distribution and $\langle SEL \rangle$ shift to smaller values, i.e., they deteriorate, as $t_s/{}^1\sigma$ increases, that is, as under-sampling of the first-dimension separation becomes more severe. All of the trends observed for *SEL* are precisely what we anticipate for the average resolution and for the average number of singlet peaks according to SOT.

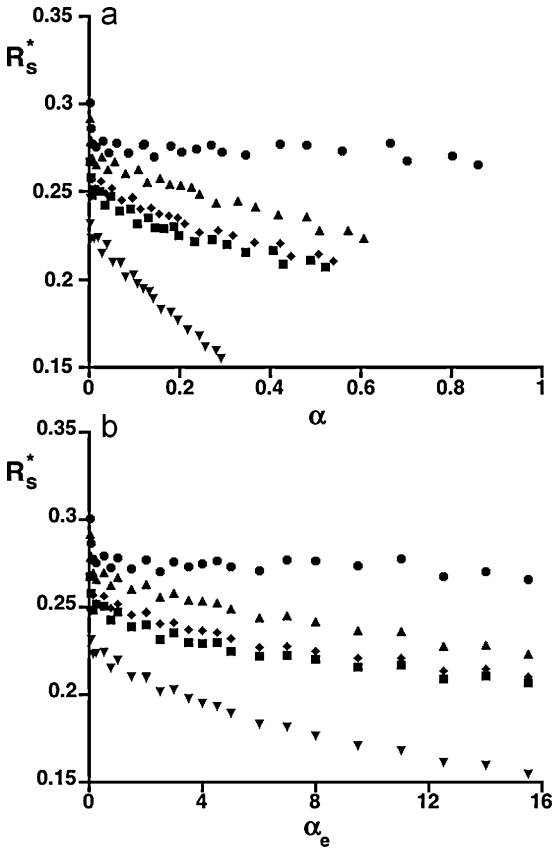

Fig. 4. (a) Graph of average minimum resolution R_s^* vs. saturation α, as calculated from $\langle SEL \rangle$ values determined with simulations having the $\bar{m}$, ${}^1n_c$, and ${}^2n_c$ values in Table 1. $t_s/{}^1\sigma = 0.2$, with $\langle \beta \rangle$ assigned the value, 1.000. Number of UV–vis spectra: none (●), two (■, $R^2 = 0.505$; ♦, $R^2 = 0.869$; ▲, $R^2 = 0.925$), and sixty (▼). (b) As in panel a, but with α replaced by the effective saturation α_e.

Note all of these results were obtained without help from spectral data. Panel f (computed using Eq. (2)) shows that for fixed ${}^1n_c$ ${}^2n_c$, $\bar{m}$, and $t_s/{}^1\sigma$, the distributions and average $\langle SEL \rangle$ shift to larger values as the amount of spectroscopic information is increased. On first glance, $\langle SEL \rangle$ has a complicated dependence on $\bar{m}$, ${}^1n_c\,{}^2n_c$, $t_s/{}^1\sigma$, and the number of spectra. We shall show, however, that $\bar{m}$, ${}^1n_c\,{}^2n_c$, and $t_s/{}^1\sigma$ are linked simply to the saturation, as defined by Eq. (5).

4.2. Dependence of R_s^* on α

The average minimum resolution R_s^* (determined as per Section 3.4.2) of singlet peaks is a metric of the average interval of closest approach between a singlet and its nearest-neighbor peak [38,39]. It differs from the traditional resolution R_s reported in Fig. 1b, which is a freely chosen parameter. Fig. 4a is a graph of R_s^* vs. the saturation α, as calculated from $\langle SEL \rangle$ values determined by simulations having the $\bar{m}$, ${}^1n_c$, and ${}^2n_c$ values in Table 1, for which $t_s/{}^1\sigma = 0.2$ (very fast sampling) and thus $\langle \beta \rangle$ is assigned the value 1.000. The assignment is reasonable, since one estimate of $\langle \beta \rangle$ at $t_s/{}^1\sigma = 0.2$ is 1.004 [16]. The simulation parameters are robust, with $\bar{m}$ varying over a 4-fold range, the aspect ratio ${}^1\sigma/{}^2\sigma$ varying from 0.1 to 12, and the ratio $\bar{m}/({}^1n_c\,{}^2n_c)$ varying from 0.03125 to 15.5 (see Table 1). Some ${}^1n_c$ values are unusually small (e.g., 1.61) in order to mimic extreme cases of under-sampling induced broadening. Despite these variations, the results are described by simple curves (see Fig. 4a) that differ only in the amount of spectral assistance. *The simplicity of these results strongly supports our concept that the aver-*

age multivariate selectivity equals the probability of observing singlets as predicted by statistical-overlap theory.

Without spectral help (see solid circles in Fig. 4a), R_s^* is almost independent of α and has a value of 0.276 ± 0.007. This is almost one-half the value, 0.5, needed to separate two bi-Gaussians of equal height and standard deviation as two distinct maxima, and the low value shows the average improvement of separation achievable when using selectivity. As spectral assistance is increased by increasing the number of distinct spectra that can be assigned to different chromatographic peaks, the limiting R_s^* at zero saturation decreases because the spectra provide more information, leading to better resolution. Upon increasing the saturation, R_s^* decreases further because the spectral information becomes a more powerful way to increase the selectivity, when peak overlap is high than when it is low. As the variety of included spectra is increased R_s^* improves even further, because the randomly assigned spectra of nearest-neighbor peaks are more likely to differ. However, the improvement has a limit; the α–R_s^* relationship for 60 spectra (filled triangles, with downward apices) is statistically no different than that for 30 spectra (results not shown for graphical simplicity). This happens because even though the number of differing spectral responses in the whole data set is small, the probability of adjacent, overlapping peaks having the same spectral response is even smaller (e.g., if n spectra are randomly assigned, the probability of two adjacent peaks having the same spectrum is $1/n^2$, which is very small even for modest n). Thus, the addition of more spectra produces only minor improvements. Furthermore, the limit on R_s^* occurs because the UV spectra provide rather limited information, e.g., the average width of UV absorption bands is relatively large (~20 nm) and a relatively narrow spectral range (201–301 nm) is used here. An improvement of R_s^* is found even with only two spectra, but the numerical value of R_s^* at any α varies with the square of the correlation coefficient R^2 of the spectra (and therefore the *SEL* of the spectra). Clearly, the less correlated are the spectra the greater is their information content, thus improving the *SEL* and also R_s^*. Three cases are shown where only two spectra are assigned randomly to each of the peaks, with R^2 equaling 0.505, 0.869, and 0.925. The further improvement of R_s^* (up to a limit) is expected as R^2 approaches zero. As R^2 approaches unity (i.e., the less is the information), the α–R_s^* relationship begins to approach that without spectral help.

The saturation ranges (i.e., the magnitude of the horizontal coordinate in Fig. 4a) of these relationships vary, because the saturation (α) itself depends on R_s^* (see Eq. (5)). A recently introduced metric, the effective saturation α_e, absorbs R_s^* and allows one to express SOT predictions by the traditional metrics of resolution and peak capacity [27]. For a comprehensive 2D separation with first- and second-dimension peak capacities having traditional resolutions of unity (e.g., as in Eqs. (6b) and (6c))

$$\alpha_e = \frac{4\alpha}{\pi(R_s^*)^2} = \frac{\bar{m}}{n'_{c,2D}} = \frac{\bar{m}\langle\beta\rangle}{{}^1n_c{}^2n_c} \tag{10}$$

Fig. 4b shows the same R_s^* values in Fig. 4a but expressed relative to α_e. Here, $n'_{c,2D}$ equals ${}^1n_c{}^2n_c$, since $\langle\beta\rangle = 1$. The α_e range is the same for all α_e–R_s^* relations, because the same ${}^1n_c$ and ${}^2n_c$ values are used. Fig. 4b shows that R_s^* is useful for interpreting the chemometrically enhanced separations having peak numbers $\bar{m}$ exceeding the corrected 2D peak capacity $n'_{c,2D}$ (defined at unit resolution) by up to 15.5-fold. We consider these to be rather highly effective saturation factors and very crowded chromatograms.

4.3. Determination of ⟨β⟩

The expectation that the α–R_s^* relationships in Fig. 4a apply to under-sampled separations was justified in another study, where a similar relationship was used to predict average numbers of peak maxima [16]. The α–R_s^* relationships for help from zero, two ($R^2 = 0.505$), and sixty spectra were used to calculate the average peak-broadening factor ⟨β⟩ from ⟨*SEL*⟩ values determined by simulations having the $\bar{m}$, and ${}^1n_c$ and ${}^2n_c$, values in Table 2. The value of $t_s/{}^1\sigma$ ranged from 1 and 16. As before, the simulation parameters are robust, with $\bar{m}$ and ${}^1n_c$ varying over 4-fold ranges, ${}^2n_c$ varying over a 16-fold range, and the aspect ratio ${}^1\sigma/{}^2\sigma$ varying (prior to sampling) from 0.125 to 8. The values of ${}^1n_c$ and ${}^2n_c$ are representative of those in LC × LC. For the case of no spectra, R_s^* was assigned the constant value, 0.276; for two and sixty spectra, R_s^* was described by quadratic fits to the graphs in Fig. 4a, with the appropriate (α, R_s^*) coordinate determined by a bisection algorithm. Fig. 5 is the graph of ⟨β⟩ vs. $t_s/{}^1\sigma$ so determined. Each ⟨β⟩ is the average of 28 determinations. Standard deviations are not shown to avoid clutter but the data are precise; the largest RSD is 10.1%. Also shown are ⟨β⟩ values calculated using SOT from the number of peak maxima in simulations of comprehensive 2D separations with randomly positioned peaks of equal height. These ⟨β⟩ values were determined in a similar (but not identical) manner to that reported elsewhere [16]; the details of the determination are lengthy and will be described in a subsequent publication [40]. The results are described well by the empirical equation [16]

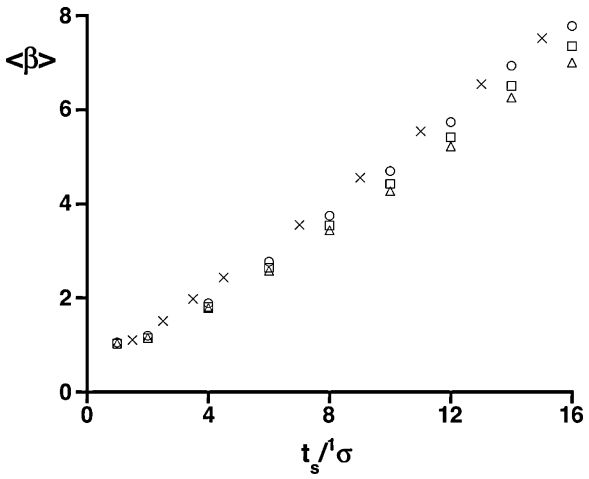

Fig. 5. Graph of average peak-broadening factor ⟨β⟩ vs. dimensionless sampling time $t_s/{}^1\sigma$, as calculated from ⟨*SEL*⟩ values determined from simulations containing 100–400 peaks and having the ${}^1n_c$, ${}^2n_c$ values in Table 2. Number of UV–vis spectra: none (○), two (□, $R^2 = 0.505$), and sixty (△). Over the range, $4 \le t_s/{}^1\sigma \le 16$, the slopes of lines determined by weighted fits are 0.48 ± 0.03, 0.45 ± 0.02, and 0.43 ± 0.02 for 0, 2, and 60 spectra, respectively. Also shown are results calculated with SOT from the number of maxima in simulations of randomly positioned peaks with equal heights (×).

$$\langle\beta\rangle = \sqrt{1 + \kappa\left(\frac{t_s}{{}^1\sigma}\right)^2} \tag{11}$$

with $\kappa = 0.23 \pm 0.01$. For simplicity, the results are called maxima-based ⟨β⟩ values to distinguish them from those determined from ⟨*SEL*⟩, i.e., ⟨*SEL*⟩-based ⟨β⟩ values.

At any $t_s/{}^1\sigma$, the ⟨*SEL*⟩-based ⟨β⟩ values are strikingly similar to the maxima-based ⟨β⟩ values in our original study [16]; furthermore, they vary only slightly with the amount of spectroscopic assistance (i.e., number of spectra). However, subtle differences do exist. Eq. (11) gives a poorer fit to the ⟨*SEL*⟩-based ⟨β⟩ values, which are slightly smaller than the maxima-based ⟨β⟩ values at any $t_s/{}^1\sigma$, as shown in Fig. 5. The poorer fit is not surprising. As shown by Blumberg and co-workers, Eq. (11) has a theoretical basis if the sampled first-dimension peak profile is represented by line segments connecting the average sample concentrations and if κ

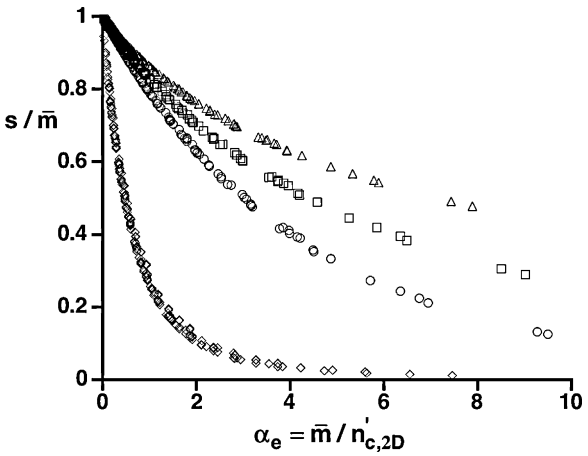

Fig. 6. Graph of fraction of peaks that are singlets, $s/\bar{m}$, vs. effective saturation α_e, as calculated from 252 ⟨*SEL*⟩ values determined by simulations containing $\bar{m} = 100$ to $\bar{m} = 400$ peaks and having the ${}^1n_c$, ${}^2n_c$ values in Table 2. Number of UV–vis spectra: none (○), two (□, R^2 = 0.505), and sixty (△). Also shown are theoretical results for expected fraction of peaks appearing as maxima of single compounds (◊).

has the value, 1/4 [17,18]. However, the multivariate-selectivity calculation is not based solely on the first-dimension profile.

Another difference is that at large ${}^1t_s/\sigma$ the limiting slopes of graphs of ⟨β⟩ vs. ${}^1t_s/\sigma$ decrease slightly but significantly with increasing numbers of spectra. These slopes are reported in the caption of Fig. 5. However, the differences among the slopes are small, such that at high values of ${}^1t_s/\sigma$ the number of spectra has only a small effect on ⟨β⟩. The principal conclusion obtained from Fig. 5 is that the κ value does not vary substantially with the amount of spectroscopic help and thus the hypothesis of Potts et al. [19] that ⟨β⟩ might be less dependent on ${}^1t_s/\sigma$ is essentially false.

4.4. Improvement of separation by selectivity

Fig. 6 gives the fraction of peaks that are singlets, $s/\bar{m}$, vs. the effective saturation α_e. The three data sets represented by triangles, squares, and circles are equal to the 252 ⟨*SEL*⟩ values used to determine the ⟨*SEL*⟩-based ⟨β⟩ values, with and without spectral data assistance. As shown by Eq. (9), ⟨*SEL*⟩ is approximately equal to $s/\bar{m}$. The data sets are grouped according to the number of spectra; there is a slight scatter about otherwise smooth curves because of the edge effect. First and foremost we see that as the amount of spectral help is increased the fraction of the peaks that can be interpreted as singlets increases. The increase is greater at higher effective saturation. However, as the spectral diversity increases, $s/\bar{m}$ decreases less rapidly with α_e, because R_s^* decreases with spectral assistance (see Fig. 4) and causes α_e (which varies inversely with $(R_s^*)^2$ per Eq. (10)) to increase rapidly.

Also shown are SOT predictions of $s/\bar{m}$ for these 252 combinations of ${}^1n_c$, ${}^2n_c$, and ${}^1t_s/\sigma$, when $s/\bar{m}$ is interpreted as the fraction of peaks appearing in the separation as maxima produced by single compounds (diamond symbols in Fig. 6). A singlet maximum is the classic way of identifying singlets in chromatography. This calculation was made by approximating ⟨β⟩ with Eq. (11) and using the recently predicted α–R_s^* relation for singlets in a 2D separation of randomly positioned peaks having an exponential distribution of heights [39] (for this height distribution, the coefficient κ in Eq. (11) is 0.21 [16]). These predictions would be more relevant if all peak heights were equal, but SOT currently can make such predictions only at very low effective saturation.

Relative to the fraction of singlet maxima, it is apparent that ⟨*SEL*⟩ quantifies the vastly improved number of signals that can be interpreted as singlets. As an example, for an effective saturation of 2 (i.e., $n'_{c,2D} = 100$ and $\bar{m} = 200$), selectivity with no added spectra allows detection of about 120 singlets out of 200 peaks; in contradistinction based on counting the peak maxima only 20 or so singlets are detected. Even more improvement is achieved when spectral information is added.

4.5. Interpretation and importance of results

Fig. 4 shows that the application of multivariate selectivity to comprehensive 2D separations greatly reduces the average minimum resolution required for a separation; it reduces this resolution even further upon addition of spectral data. The result of this reduction of resolution is an improvement of separation beyond that possible when only peak maxima are detected, as shown in Fig. 6 (see diamonds). In the present work, we describe the improvement in the separation upon addition of spectral assistance as taking place through the reduction of the average minimum resolution (R_s^*) needed to detect singlets. Previously, the addition of mass-spectral data to 2D separations was interpreted as equivalent to adding a third dimension of peak capacity [41–43]. However, this interpretation is not possible with our two-dimensional model of peak overlap. We believe that ours is the more physically realistic interpretation, since the collection of spectra does not actually transport peaks into different regions of space (i.e., does not cause separation). In our view the spectral data simply allow individual peaks to be identified as such at lower values of the average minimum resolution. The improvement of separation by the spectrally induced reduction in resolution, i.e., the consequences of the decrease of R_s^*, is evident in Fig. 6.

In contrast, Fig. 5 shows that selectivity has only a small – not quite negligible – effect on the average first-dimension peak-broadening factor resulting from under-sampling, either with or without spectral assistance. Consequently, it also has only a small effect on the corrected 2D peak capacity, Eq. (6a). This is useful to know, since the corrected peak capacity is a critical issue in optimizing on-line comprehensive 2D separations. For example, it causes the existence of the optimal second-dimension cycle time [19,44]. Several studies have shown that increasing the sampling time (i.e., the collection time of first-dimension effluent, which in on-line 2D separations is equal to the second-dimension cycle time) initially increases and ultimately limits the second-dimension peak capacity. However, it always decreases the corrected first-dimension peak capacity ${}^1n_c/\langle\beta\rangle$, such that an optimal sampling time and maximum in the plot of corrected 2D peak capacity vs. sampling time must exist in on-line comprehensive 2D separations [19,44,45].

Fig. 6 shows that the fraction of singlets always increases as the effective saturation decreases, although the fractions differ for *SEL* and peak-maxima singlets, and as the amount of spectral assistance varies. Since the corrected peak capacity varies inversely with the effective saturation (see Eq. (10)), the fraction of singlets always increases with increasing values of the corrected peak capacity. Therefore, the optimization of the corrected peak capacity, with a peak-broadening factor approximated by Eq. (11), should give the greatest fraction of singlets in a comprehensive 2D separation, at least when peak positions are random and the fractional coverage [46] (the fraction of the corrected peak capacity that is used) is constant. Neither the means of singlet evaluation (*SEL* or maxima) nor the amount of spectral assistance (some or none) significantly changes the value of the optimal corrected peak capacity; they only change the fraction of singlets found at that optimum.

While the selectivity values calculated in the present work reflect the specific results expected with an analysis of the data with the PARAFAC algorithm under ideal conditions, we believe that the trends observed here would still be observed with different selectivity metrics. The actual form of the selectivity calculation

used here is quite robust, unlike some of the other selectivity metrics reported by Olivieri [22], so that as a relative measure, if not an absolute one, we feel that the *SEL* offers a simple measure of the performance of a chemometrically assisted separation. This is not to say that *SEL* is the only metric by which the interpretation of multi-component separations can be improved. Other metrics, e.g., ones based on information theory [47] and separation quality [48], are also useful. However, *SEL* is unique in its interpretation as a singlet probability, which fosters a close connection with SOT and the parameters (e.g., R_s^* and $\langle\beta\rangle$) on which it depends.

Our results should not be overinterpreted; they are based only on randomly positioned peaks of equal height. Although Eq. (11) appears to work well with weakly correlated peaks having exponentially distributed heights [49], it is unlikely that the average minimum resolutions in Fig. 4 work well with non-randomly positioned peaks. Rather, they are specific to randomly positioned peaks. It is probable that with non-randomly positioned peaks the value of (but not the trends in) R_s^* will change, with its exact value determined by the amount of retention-time and spectral correlation. Also, it is clear from the three different degrees of correlation for the two-spectra cases shown in Fig. 4 that R_s^* varies with the amount of spectral correlation. From a chemical perspective it is possible that the spectra of nearest-neighbor peaks in experimental 2D separations will be more similar than considered here, because the neighboring peaks are probably chemically more related than are more remote peaks. However, it is not the exact value of these metrics that is relevant here but their trends: the average minimum resolution measured by using *SEL* values is much more dependent on the spectral assistance than is the broadening caused by undersampling.

An alternative approach for assessing the reliability of detection of a target compound in a mixture has been described by Stoev et al. [50–52]. In that work, the peak capacity and the singlet probability were used to calculate a probability of successful identification of a target compound in a specified mixture. The power of DAD and mass spectrometric (MS) detection was compared using a grid approach for finding characteristic points in the spectra. It was found that for the specific example of detection of antifungual agents, the probabilities of identification of LC–DAD and LC–MS/MS (low resolution) were very similar [52]. Based on the data provided, since the chromatographic peaks were very well resolved for the most part, it is likely that *SEL* values would be close to one in both cases, leading to a similar conclusion. The real challenge comes in evaluating more complex mixtures, such as described here. It should be noted that this calculation does not take into account other factors, such as ionization interferences and irreproducibility of signals that can affect electrospray MS detection.

Because MS is so commonly coupled to chromatography, we think it is important to extend the multivariate selectivity concept and SOT-based interpretation to simulated chromatograms assisted by MS, with an emphasis on understanding the differences among R_s^* and $\langle\beta\rangle$ values for MS and DAD. Unlike DAD, the information content of MS varies greatly with the instrumentation, whose resolution can be low (e.g., mass selective detection) to extraordinarily high (e.g., Fourier transform ion cyclotron resonance). Of particular interest is evaluating the dependence of the asymptotic lower limit of R_s^*, which is found at high spectral diversity, on the resolution of the MS instrumentation.

5. Conclusions

Unless otherwise stated, these conclusions apply to chromatograms with randomly positioned peaks of equal height, with and without spectral assistance from DAD.

1. We have shown that our results are consistent with the concept that all chromatographic peaks have both singlet and multiplet character, and the fractional singlet character of any peak is equal to the multivariate selectivity *SEL*.
2. The average multivariate selectivity $\langle SEL\rangle$ of all peaks closely corresponds to the fraction of singlet peaks predicted by SOT, i.e., $\langle SEL\rangle \approx s/\bar{m}$ (Eq. (9)).
3. Without UV–visible spectral assistance, the multivariate selectivity of 2D separations provides an average minimum resolution R_s^* to be used in SOT that is about twofold smaller (0.276) than that needed to see two maxima (0.5).
4. The addition of spectral assistance when using multivariate selectivity as measured by the *SEL* metric improves R_s^*, especially when the chromatogram is crowded and peak overlap is severe (i.e., when the saturation α is high).
5. Increasing the spectral diversity improves R_s^* even further but there is a limit that cannot be exceeded dependent on the characteristics of the spectra.
6. The average peak-broadening factor $\langle\beta\rangle$ is about the same, whether it is determined by counting peak maxima or by using multivariate selectivity. Spectral assistance has only a small effect on $\langle\beta\rangle$.
7. The near independence of $\langle\beta\rangle$ from its means of evaluation simplifies the optimization of the corrected 2D peak capacity in on-line comprehensive 2D separations.
8. The general trends, if not the exact values, found here for R_s^* and $\langle\beta\rangle$ apply to chromatograms with non-randomly positioned peaks of unequal height.
9. The use of multivariate selectivity greatly increases the likelihood of evaluating peaks as singlets. For example, with a corrected 2D peak capacity of 100, 120 or more of 200 peaks can be identified as singlets, whereas only 20 maxima can be so identified.

In future work we will compare the improvement brought about by DAD spectra and by mass spectroscopy of various resolving power.

Acknowledgments

The authors thank John H. Kalivas (Department of Chemistry, Idaho State University) for helpful discussions. S.C.R. and P.W.C. acknowledge financial support for this research from the National Science Foundation, grant number CHE-0911330.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2011.06.086.

References

[1] D.A. Skoog, F.J. Holler, S.R. Crouch, Principles of Instrumental Analysis, 6th ed., Thomson-Brooks/Cole, Florence, KY, 2007.
[2] A. Lorber, Anal. Chem. 58 (1986) 1167.
[3] A. Lorber, B. Kowalski, J. Chemom. 2 (1988) 67.
[4] K.S. Booksh, B.R. Kowalski, Anal. Chem. 66 (1994) 782A.
[5] A. Lorber, K. Faber, B.R. Kowalski, Anal. Chem. 69 (1997) 1620.
[6] A.E. Sinha, J.L. Hope, B.J. Prazen, C.G. Fraga, E.J. Nilsson, R.E. Synovec, J. Chromatogr. A 1056 (2004) 145.
[7] M.T. Cantwell, S.E.G. Porter, S.C. Rutan, J. Chemom. 21 (2007) 335.
[8] M. Dumarey, Y. vander Heyden, S.C. Rutan, Anal. Chem. 82 (2010) 6056.
[9] N.J. Messick, J.H. Kalivas, P.M. Lang, Anal. Chem. 68 (1996) 1572.
[10] K. Faber, A. Lorber, B.R. Kowalski, J. Chemom. 11 (1997) 419.
[11] G. Vivo-Truyols, J.R. Torres-Lapasio, M.C. Garcia-Alvarez-Coque, J. Chromatogr. A 991 (2003) 47.
[12] K. Wagner, T. Miliotis, G. Marko-Varga, R. Bischoff, K.K. Unger, Anal. Chem. 74 (2002) 809.
[13] K. Horie, H. Kimura, T. Ikegami, A. Iwatsuka, N. Saad, O. Fiehn, N. Tanaka, Anal. Chem. 79 (2007) 3764.

[14] R.E. Murphy, M.R. Schure, J.P. Foley, Anal. Chem. 70 (1998) 1585.
[15] J.V. Seeley, J. Chromatogr. A 962 (2002) 21.
[16] J.M. Davis, D.R. Stoll, P.W. Carr, Anal. Chem. 80 (2008) 461.
[17] L.M. Blumberg, F. David, M.S. Klee, P. Sandra, J. Chromatogr. A 1188 (2008) 2.
[18] L.M. Blumberg, J. Sep. Sci. 31 (2008) 3358.
[19] L.W. Potts, D.R. Stoll, X. Li, P.W. Carr, J. Chromatogr. A 1217 (2010) 5700.
[20] J.M. Davis, Anal. Chem. 69 (1997) 3796.
[21] S. Liu, J.M. Davis, J. Chromatogr. A 1126 (2006) 244.
[22] A.C. Olivieri, Anal. Chem. 77 (2005) 4936.
[23] J.M. Davis, J. Sep. Sci. 28 (2005) 347.
[24] B.L. Karger, L.R. Snyder, C. Horvath, An Introduction to Separation Science, Wiley and Sons, New York, 1973.
[25] G. Guiochon, L.A. Beaver, M.F. Gonnord, A.M. Siouffi, M. Zakaria, J. Chromatogr. 255 (1983) 415.
[26] J.C. Giddings, Anal. Chem. 56 (1984) 1258A.
[27] J.M. Davis, P.W. Carr, Anal. Chem. 81 (2009) 1198.
[28] J.M. Davis, Anal. Chem. 63 (1991) 2142.
[29] L.J. Nagels, W.L. Creten, P.M. Vanpeperstraete, Anal. Chem. 55 (1983) 216.
[30] D.P. Herman, M.-F. Gonnord, G. Guiochon, Anal. Chem. 56 (1984) 995.
[31] L.J. Nagels, W.L. Creten, Anal. Chem. 57 (1985) 2706.
[32] F. Dondi, Y.D. Kahie, G. Lodi, M. Remelli, P. Reschiglian, C. Bighi, Anal. Chim. Acta 191 (1986) 261.
[33] M.C. Pietrogrande, A. Cavazzini, F. Dondi, Rev. Anal. Chem. 19 (2000) 123.
[34] C.G. Enke, L.J. Nagels, Anal. Chem. 83 (2011) 2539.
[35] D.R. Stoll, C. Paek, P.W. Carr, J. Chromatogr. A 1137 (2006) 153.
[36] S.E.G. Porter, D.R. Stoll, C. Paek, S.C. Rutan, P.W. Carr, J. Chromatogr. A 1137 (2006) 163.
[37] M. Evans, N. Hastings, B. Peacock, Statistical Distributions, 3rd ed., Wiley, New York, 2000.
[38] J.M. Davis, J. Chromatogr. A 831 (1999) 37.
[39] J.M. Davis, Talanta 83 (2011) 1068.
[40] P.W. Carr, J.M. Davis, S.C. Rutan, D.R. Stoll, Adv. Chromatogr., in press.
[41] K.C. Lewis, G.J. Opiteck, J.W. Jorgenson, D.M. Sheeley, J. Am. Soc. Mass Spectrom. 8 (1997) 495.
[42] Y. Shen, N. Tolic, R. Zhao, L. Pasa-Tolic, L. Li, S.J. Berger, R. Harkewicz, G.A. Anderson, M.E. Belov, R.D. Smith, Anal. Chem. 73 (2001) 3011.
[43] S.J. Valentine, M. Kulchania, C.A. Srebalus Barnes, D.E. Clemmer, Int. J. Mass Spectrom. 212 (2001) 97.
[44] X. Li, D.R. Stoll, P.W. Carr, Anal. Chem. 81 (2009) 845.
[45] K. Horvath, J.N. Fairchild, G. Guiochon, Anal. Chem. 81 (2009) 3879.
[46] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Anal. Chem. 77 (2005) 6426.
[47] F. Gong, Y.-Z. Liang, P.-S. Xie, F.-T. Chau, J. Chromatogr. A 1002 (2003) 25.
[48] F. Gong, Y.-Z. Liang, Q.-S. Xu, F.-T. Chau, K.-M. Ng, Anal. Chim. Acta 450 (2001) 99.
[49] J.M. Davis, D.R. Stoll, P.W. Carr, Anal. Chem. 80 (2008) 8122.
[50] G. Stoev, A. Mihailova, J. Chromatogr. A 869 (2000) 275.
[51] G. Stoev, A. Michailova, J. Chromatogr. A 1031 (2004) 11.
[52] G. Stoev, Al. Stoyanov, J. Chromatogr. A 1150 (2007) 302.