Principle of i-PDeA II and Demonstration Procedures

An i-PDeA II (Intelligent Peak Deconvolution Analysis II) data analysis method is for extracting target peaks from unseparated peaks by analyzing photodiode array (PDA) detector data using the chemometrics multivariate curve resolution alternating least squares (MCR-ALS) technique.

1. Features

1. Fast and accurate quantitative analysis is possible even if components are not fully separated in the column.
2. i-PDeA II can even be used to analyze isomers with identical molecular weights.
3. Spectral data can be analyzed even after peak separation.

2. Fundamental Theoretical Basis

2.1 Modeling PDA Detector Data

Given peak profiles and spectra for each component in a three-component mixture, $c_1(t)$, $s_1(\lambda)$, $c_2(t)$, $s_2(\lambda)$, $c_3(t)$ and $s_3(\lambda)$, then measurement data in an ideal system $d(t, \lambda)$ can be described by the following expression.

$$d(t,\lambda) = c_1(t)s_1(\lambda) + c_2(t)s_2(\lambda) + c_3(t)s_3(\lambda)$$

Then spectra $d(t_i,\lambda)$ measured as a function of time t_i can be expressed as follows:

$$d(t_i,\lambda) = c_1(t_i)s_1(\lambda) + c_2(t_i)s_2(\lambda) + c_3(t_i)s_3(\lambda)$$

Assuming spectral components are vectors with discrete values λ_j(where j = 1 to m), then spectra can be described as follows:

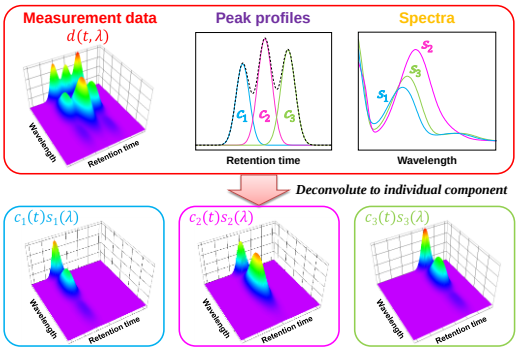


Fig 1. Three-Component Mixture Sample

$$\boldsymbol{d}_i^T = \alpha_i\boldsymbol{s}_1^T + \beta_i\boldsymbol{s}_2^T + \gamma_i\boldsymbol{s}_3^T = \begin{pmatrix}\alpha_i & \beta_i & \gamma_i\end{pmatrix}\begin{pmatrix}\boldsymbol{s}_1^T \\ \boldsymbol{s}_2^T \\ \boldsymbol{s}_3^T\end{pmatrix}$$

Where,

$$\boldsymbol{d}_i^T = \begin{pmatrix}d(t_i,\lambda_1) & \dots & d(t_i,\lambda_m)\end{pmatrix}$$

$$\alpha_i = c_1(t_i),\ \beta_i = c_2(t_i),\ \gamma_i = c_3(t_i)$$

$$\boldsymbol{s}_1^T = \begin{pmatrix}s_1(\lambda_1) & \dots & s_1(\lambda_m)\end{pmatrix},\ \boldsymbol{s}_2^T = \begin{pmatrix}s_2(\lambda_1) & \dots & s_2(\lambda_m)\end{pmatrix},\ \boldsymbol{s}_3^T = \begin{pmatrix}s_3(\lambda_1) & \dots & s_3(\lambda_m)\end{pmatrix}$$

By summarizing each spectrum measurement at time t_i (where i = 1 to n), measurements can be expressed in matrix form, as follows:

$$\begin{pmatrix}\boldsymbol{d}_1^T \\ \vdots \\ \boldsymbol{d}_n^T\end{pmatrix} = \begin{pmatrix}\alpha_1 & \beta_1 & \gamma_1 \\ \vdots & \vdots & \vdots \\ \alpha_n & \beta_n & \gamma_n\end{pmatrix}\begin{pmatrix}\boldsymbol{s}_1^T \\ \boldsymbol{s}_2^T \\ \boldsymbol{s}_3^T\end{pmatrix}$$

or by direct product (outer product), as follows:

$$D = \boldsymbol{c_1 s_1^T} + \boldsymbol{c_2 s_2^T} + \boldsymbol{c_3 s_3^T} \quad \text{Eq. (1)}$$

or alternatively

$$D = CS^T \quad \text{Eq. (2)}$$

where,

$$\boldsymbol{c_1} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix},\ \boldsymbol{c_2} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix},\ \boldsymbol{c_3} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}$$

$$D = \begin{pmatrix} d(t_1,\lambda_1) & \cdots & d(t_1,\lambda_m) \\ \vdots & \cdots & \vdots \\ d(t_n,\lambda_1) & \cdots & d(t_n,\lambda_m) \end{pmatrix},\ C = \begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \vdots & \vdots & \vdots \\ \alpha_n & \beta_n & \gamma_n \end{pmatrix},\ S^T = \begin{pmatrix} s_1(\lambda_1) & \cdots & s_1(\lambda_m) \\ s_2(\lambda_1) & \cdots & s_2(\lambda_m) \\ s_3(\lambda_1) & \cdots & s_3(\lambda_m) \end{pmatrix}$$

The data can be expressed schematically as follows:

$$\begin{pmatrix} Meaurement\ spectrum\ 1 \\ \vdots \\ Meaurement\ spectrum\ n \end{pmatrix} = \begin{pmatrix} \text{Component 1 profile} & \text{Component 2 profile} & \text{Component 3 profile} \end{pmatrix} \begin{pmatrix} Component\ 1\ pure\ spectrum \\ Component\ 2\ pure\ spectrum \\ Component\ 3\ pure\ spectrum \end{pmatrix}$$

Considering measurement error, noise, and unpredictable factors, and given a remainder R, the measurement data can be modeled as follows:

$$D = CS^T + R$$

This relational expression is valid for any number of components.

2.2 Solutions Using MCR-ALS Technique

The MCR-ALS technique estimates the peak profile or the spectrum with the closest fit to measurement data by repeatedly approximating C (peak profiles) or S (spectra) in equation (2) using least squares approximation.

The following is the typical method for determining solutions by the MCR-ALS technique.

Step 1 Specify the number of components in measurement data D.
Step 2 Calculate initial estimate
(for example, by specifying the initial value for C).
Step 3 Using the estimate of C, calculate the S^T matrix under appropriately chosen constraints.
Step 4 Using the estimate of S^T ,calculate the C matrix under appropriately chosen constraints.
Step 5 From the product of C and S^T found in the above steps of an iterative cycle, calculate an estimate of the original data matrix, D.
Step 6 Repeat steps 3, 4, and 5 until convergence is achieved.

Equation (2) generally does not give a unique solution. Therefore, to determine the optimal solution, constraints must be specified based on problem characteristics. Consequently, by specifying appropriate constraints, MCR-ALS can provide valid solutions even without prior information.

2.3 i-PDeA II Peak Separation Algorithm

If equation (1) is expanded for N components, the measurement signal D can be expressed by the following equation.

$$D = \boldsymbol{c_1 s_1^T} + \boldsymbol{c_2 s_2^T} + \cdots + \boldsymbol{c_N s_N^T}$$

This algorithm determines a solution by minimizing the following squared errors, with the chromatogram vector $\boldsymbol{c_k}$ substituted by the chromatogram model function $\boldsymbol{f_k}$.

$$E = \left|D - \Sigma \boldsymbol{f_k s_k^T}\right|^2$$

In this case, a bidirectional exponentially modified Gaussian (BEMG) function is used as the chromatogram model function. BEMG is the reciprocal of the delay time component of the exponentially modified Gaussian (EMG) function, as defined by the following equations.

$$bemg(t,a,b) = \int_{-\infty}^{0} e^{ax} * emg(t-x,b)dx$$

$$emg(t,b) = \int_{0}^{\infty} e^{-bx} * exp(-(t-x)^2)\,dx$$

This algorithm applies the MCR-ALS technique by using an estimated value as the initial value and the BEMG model function as the chromatogram constraint. Since the number of components after separation is unknown, the initial condition starts with a single component and then successively adds components as the presence of unseparated peaks are determined in the residual signal to determine the optimal solution.

3. Data Analysis Using LabSolutions

The i-PDeA II peak separation algorithm is included in LabSolutions data analysis functionality. Data for separated peaks can be displayed as chromatograms for individual peaks, and also as separated spectra, in the LabSolutions PDA data analysis window.

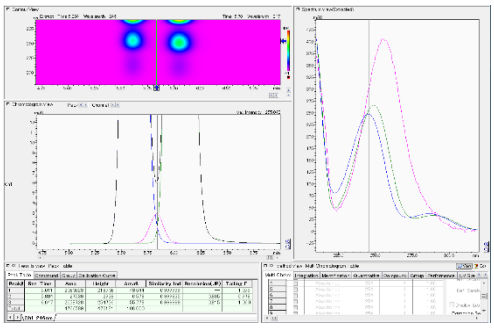


Fig 2. PDA Data Analysis Window in LabSolutions

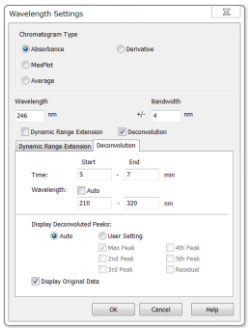


Fig 3. i-PDeA II Setting Window

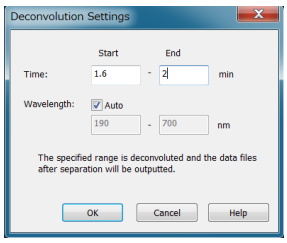


Fig 4. Export separated 3D data

Peaks can be separated using the i-PDeA II function by simply specifying the wavelength and time ranges. By using the data analysis functionality in LabSolutions, the entire process of separating peaks, integrating the areas under separated peaks, and calculating quantitative values can be performed seamlessly without any data conversion and spectra can be identified and libraries searched based on peak-top spectra.

The separated 3D data files can be exported from the [File]-[Export Data]-[Export Deconvolution Results] menu.

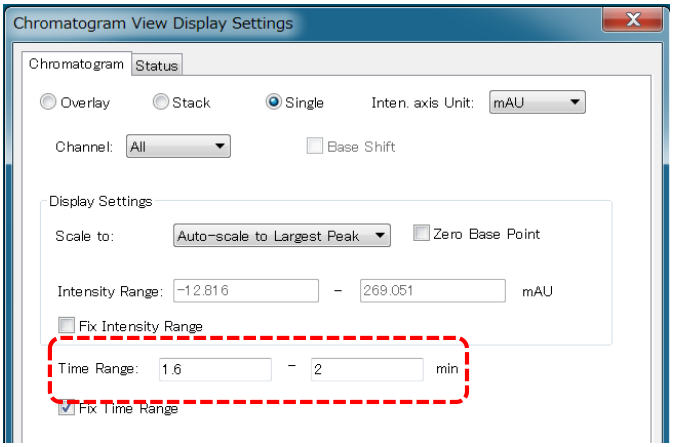
4. Demonstration procedures

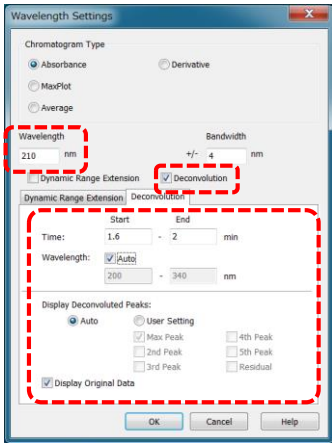
Use the following files under the LabSolutions¥Sample¥LC¥i-PDeA folder to perform demonstration.

PDA_DerivQuan_DFBP100_VP1.lcd
PDA_DerivQuan_DFBP100_VP10.lcd
PDA_DerivQuan_DFBP100_VP50.lcd
PDA_DerivQuan_DFBP100_VP100.lcd
PDA_DerivQuan_DFBP100_VP200.lcd
Spectrum of DFBP(100).jcm
Spectrum of VP(100).jcm
PDA_Deconvolution_SummaryReport.lsr

Step1. Deconvolution settings

1) Open PDA_DerivQuan_DFBP100_VP100.lcd in the PDA Data Analysis window of Postrun program.
2) Set Time Range to 1.6 - 2 min in the Ch1 Chromatogram View Display Settings to expand the chromatogram plot.
3) Click the Type column of Ch1 on the Multi Chromatogram table of Method View and set Wavelength and Deconvolution conditions. Disable other channels if they are set.





Step 2. Integration parameter settings

1) Set the Ch1 integration parameters on the Integration tab of Method View as follows.:

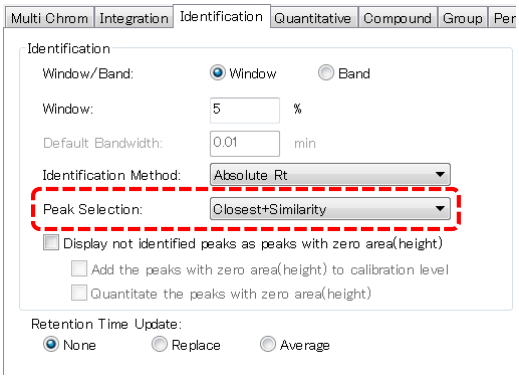
Width	5
Slope	1000
Drift	0
T.DBL	1000
Min. Area/Height	1000 (Calculated by Area)

2) Chang Method View from Edit to View to perform deconvolution and integration.

Step 3. Spectrum identification and quantitation settings

1) On the Identification tab of Method View, select Peak selection as Closest and Similarity.

2) On the Quantitative tab of Method View, set parameters as follows:

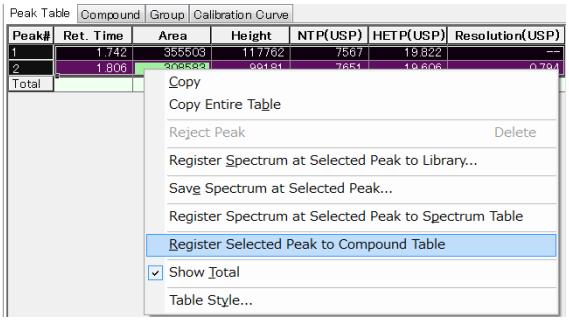


Quantitative Method	External Srandard
Calculated by	Area
# of Calib. Levels	5
Curve Fit Type	Linear
Zero	Not Forced
Weighting Method	None
X Axis of Calib. Curve	Conc.

Step 4. Compound table settings

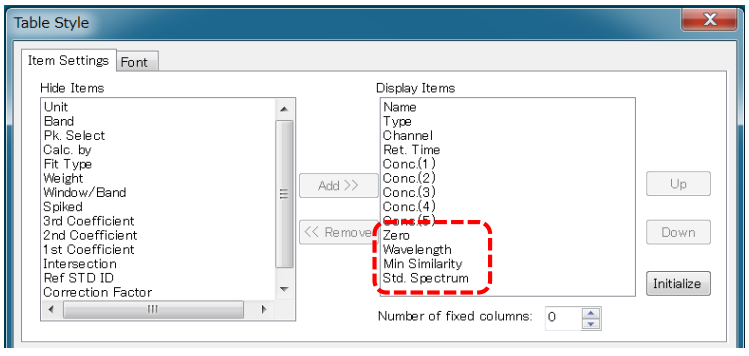
1) Select the Compound table of Method View.
If the compound table are already set, delete all settings.

2) Select two peaks on the Peak table of Results View and perform [Register selected peak to Compound Table] in the right click menu.



3) Change Name RT:1.742 to DFBP, and RT:1.806 to VP.

4) Perform [Table style…] in the right click menu and move Zero, Wavelength, Min. Similarity and Std. Spectrum to Display Items.



5) On the Std. Spectrum cell, perform [Register Standard Spectrum]-[UV Spectrum File] in the right click menu and set “Spectrum of DFBP(100).jcm” as DFBP and “Spectrum of VP(100).jcm” as VP. Set Wavelength to 200- 340nm and Min. Similarity to 0.90000.

6) Set Conc. and Zero of DFBP and VP as follows:

Name	Conc.(1)	Conc.(2)	Conc.(3)	Conc.(4)	Conc.(5)	Zero
DFBP	100	100	100	100	100	Force Through
VP	1	10	50	100	200	Not Forced

Step 5. Save method file

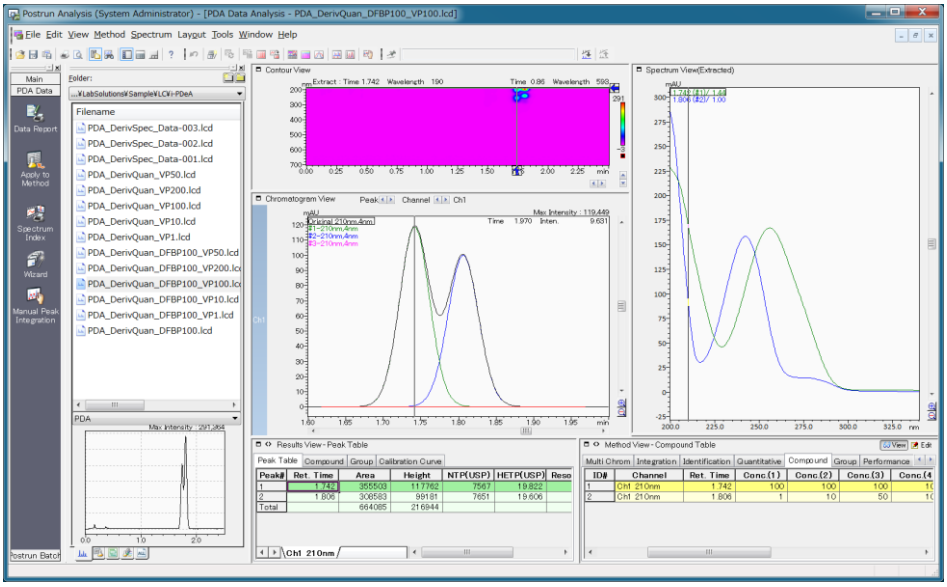
1) Set Method View to the View mode and perform [File]-[Save Method As…] to save the method as “PDA_Deconvolution.lcm”.

Step 6. Batch table settings

1) Open Postrun Batch window from the assistant bar.
2) Drag and drop the data files on the batch table from Data Explorer in the following order.
   PDA_DerivQuan_DFBP100_VP1.lcd
   PDA_DerivQuan_DFBP100_VP10.lcd
   PDA_DerivQuan_DFBP100_VP50.lcd
   PDA_DerivQuan_DFBP100_VP100.lcd
   PDA_DerivQuan_DFBP100_VP200.lcd
3) Set Sample Type to “1:Standard(I)” (Standard, Initialize Calibration Curve) on the first row and “1:Stardard” on the rest of the following rows.
4) Set Method File to “PDA_Deconvolution.lcm”.
5) Set Level# to 1, 2, 3, 4 and 5 in ascending order from the first row.
6) Save batch file as “PDA_Deconvolution_Batch.lcb”.

Step 7. Review data

1) Run the post run batch.
2) Review the reprocessed data in the PDA Data Analysis window.



Step 8. Report Output

1) Open Report window from the assistant bar.
2) Drag and drop the report format file “PDA_Deconvolution_SummaryReport.lsr” on the report editor from Data Explorer.
3) Drag and drop the data files on the report editor from Data Explorer in the following order.

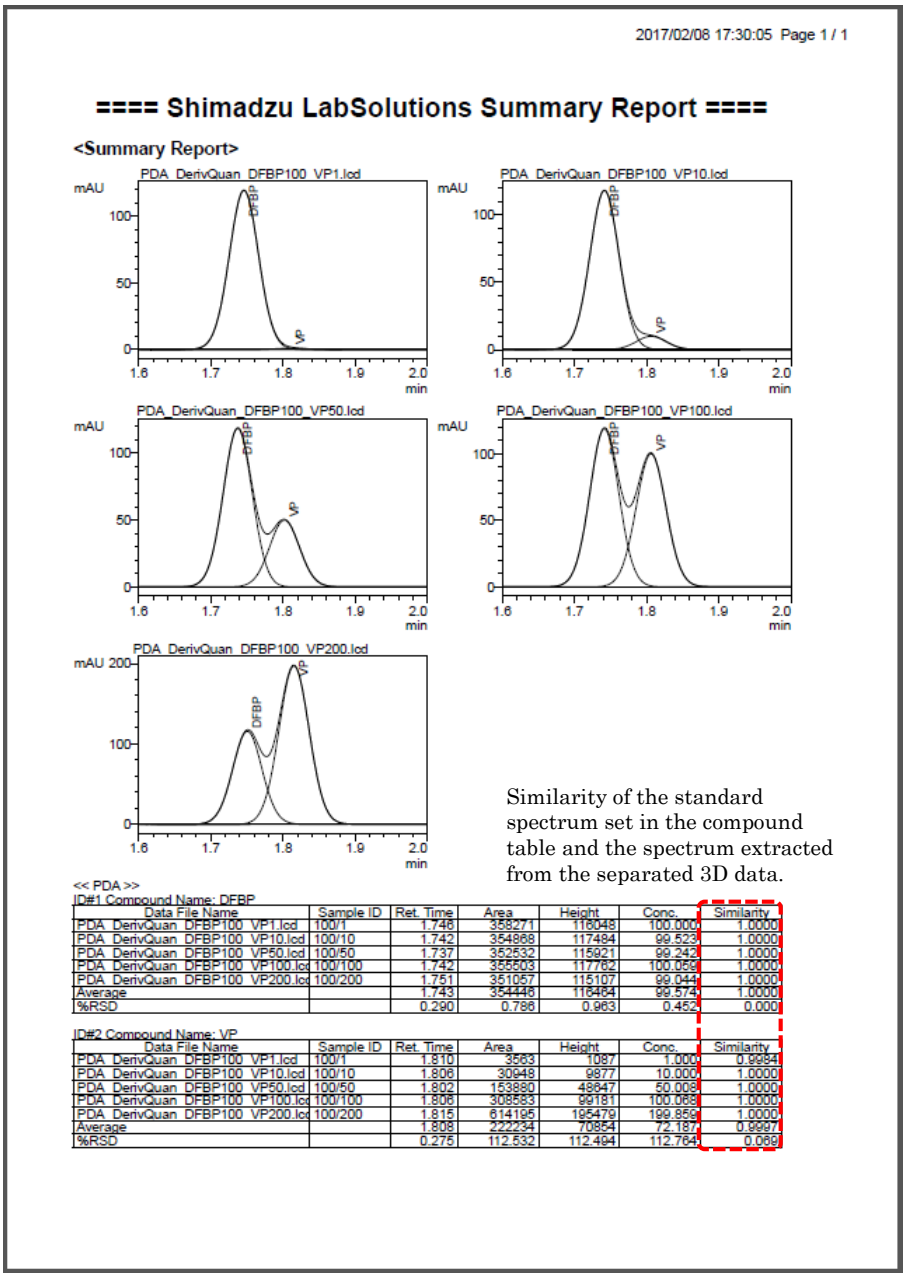
   PDA_DerivQuan_DFBP100_VP1.lcd

   PDA_DerivQuan_DFBP100_VP10.lcd

   PDA_DerivQuan_DFBP100_VP50.lcd

   PDA_DerivQuan_DFBP100_VP100.lcd

   PDA_DerivQuan_DFBP100_VP200.lcd
4) Click [Preview] on the assistant bar to display and confirm the summary report.



2017/02/08 17:30:05 Page 1 / 1

==== Shimadzu LabSolutions Summary Report ====

<Summary Report>

PDA_DerivQuan_DFBP100_VP1.lcd

PDA_DerivQuan_DFBP100_VP10.lcd

PDA_DerivQuan_DFBP100_VP50.lcd

PDA_DerivQuan_DFBP100_VP100.lcd

PDA_DerivQuan_DFBP100_VP200.lcd

Similarity of the standard spectrum set in the compound table and the spectrum extracted from the separated 3D data.

<< PDA >>

ID#1 Compound Name: DFBP

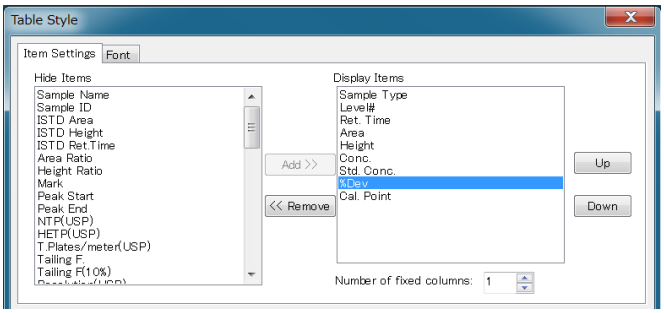
Data File Name	Sample ID	Ret. Time	Area	Height	Conc.	Similarity
PDA_DerivQuan_DFBP100_VP1.lcd	100/1	1.746	358271	116048	100.000	1.0000
PDA_DerivQuan_DFBP100_VP10.lcd	100/10	1.742	354868	117484	99.523	1.0000
PDA_DerivQuan_DFBP100_VP50.lcd	100/50	1.737	352532	115921	99.242	1.0000
PDA_DerivQuan_DFBP100_VP100.lcd	100/100	1.742	355503	117762	100.059	1.0000
PDA_DerivQuan_DFBP100_VP200.lcd	100/200	1.751	351057	115107	99.044	1.0000
Average		1.743	354446	116484	99.574	1.0000
%RSD		0.290	0.786	0.963	0.452	0.000

ID#2 Compound Name: VP

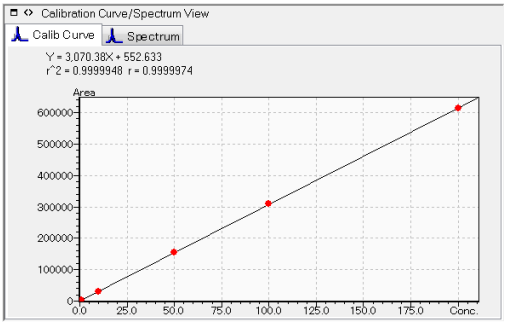
Data File Name	Sample ID	Ret. Time	Area	Height	Conc.	Similarity
PDA_DerivQuan_DFBP100_VP1.lcd	100/1	1.810	3563	1087	1.000	0.9984
PDA_DerivQuan_DFBP100_VP10.lcd	100/10	1.806	30948	9877	10.000	1.0000
PDA_DerivQuan_DFBP100_VP50.lcd	100/50	1.802	153980	48847	50.008	1.0000
PDA_DerivQuan_DFBP100_VP100.lcd	100/100	1.806	308583	99181	100.069	1.0000
PDA_DerivQuan_DFBP100_VP200.lcd	100/200	1.815	614195	195479	199.859	1.0000
Average		1.808	222234	70854	72.187	0.9997
%RSD		0.275	112.532	112.494	112.764	0.069

Step 9. Optimization of deconvolution parameters on Quant Browser

1) Run [Browser] from LabSolutions main menu and open the Quant Browser window.
2) Perform [Table Style…] in the right click menu of Quantitative Results View and move %Dev. to Display Items.



3) Select “PDA_Deconvolution_Batch.lcb” from Data Explorer.
(Double click the file or drag & drop it on the view.)
4) Confirm the separated chromatogram by selecting each row on Quantitative Results View.
5) Select ID#2 VP on Quantitative Results View. Then calibration curve and the values of r^2 and r are displayed on the Calib. Curve tab of Calibration Curve/Spectrum View.



6) Select Edit mode on Method View. Select “Type” on the Multi chromatogram table and change Deconvolution Wavelength from Auto to 200-340nm on the Wavelength Settings dialog window.
7) Chang Method View from Edit to View to perform reprocessing.
8) Confirm %Dev. of VP on Quantitative Results View.

Quantitative Results View ID# 2 VP

Data#	Data Filename	Height	Conc. (mg/	Std. Conc.	%Dev	Cal. Point
1	PDA_DerivQuan_DFBP100_VP1.lcd	1,085	0.996	1	-0.42	☑
2	PDA_DerivQuan_DFBP100_VP10.lcd	9,875	9.910	10	-0.90	☑
3	PDA_DerivQuan_DFBP100_VP50.lcd	48,554	49.849	50	-0.30	☑
4	PDA_DerivQuan_DFBP100_VP100.lcd	99,220	100.405	100	0.40	☑
5	PDA_DerivQuan_DFBP100_VP200.lcd	195,551	199.840	200	-0.08	☑

9) All %Div. values are within ±1.0%
10) After optimizing the deconvolution parameters, perform [File]-[Close All Data Files] and save method and data files by selecting Save current data file ? “Yes”.

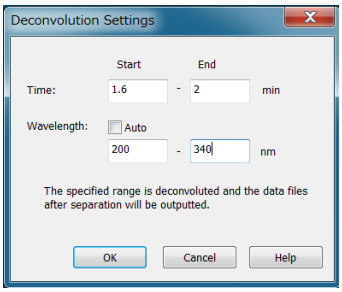
Step 10. Exporting deconvoluted data file

1) Following Step 9, operate as follows.
2) Select “PDA_DerivQuan_DFBP100_VP100.lcd” on Data Explorer.
3) Perform [File Conversion]-[Convert LabSolutions Data File to PDA Deconvolution Results....] in the right click menu and set as follows:

   Time range 1.6 - 2.0 min

   Wavelength range 200 - 340 nm

   Then click [OK].



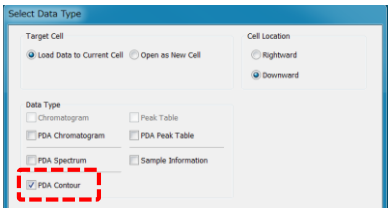
4) PDA 3D data is deconvoluted and the following 3D data files are exported.

   PDA_DerivQuan_DFBP100_VP100_Deconvoluted_1,lcd

   PDA_DerivQuan_DFBP100_VP100_Deconvoluted_2,lcd

   PDA_DerivQuan_DFBP100_VP100_Deconvoluted_3,lcd

5) Open Data Browser window from the main assistant bar.
6) From Data Explorer, select the three deconvoluted data files (Deconvoluted_1 through Deconvoluted_3) and then drag and drop on the cell of Data Browser.



7) Check “PDA contour” on the Select Data Type window and press [OK].

   Then three PDA contours are displayed on the cells.