References

1. Miller, R. D.; Michl, J. *J. Chem. Rev.* **1989**, *89*, 139.
2. Ishikawa, M.; Kumada, M. *Revs. Si, Sn and Pb Compounds* **1979**, *IV*, 7.
3. Gaspar, P. P.; Holten, D.; Konieczny, S.; Corey, J. Y. *Acc. Chem. Res.* **1987**, *20*, 329.
4. Gaspar, P. P.; Boo, B. H.; Chari, S.; Ghosh, A. K.; Holten, D.; Kirmaier, C.; Konieczny, S. *Chem. Phys. Lett.* **1984**, *105*, 153.
5. Ishikawa, M.; Kumada, M. *J. Organomet. Chem.* **1972**, *42*, 325.
6. Ishikawa, M.; Takaoka, T.; Kumada, M. *J. Organomet. Chem.* **1972**, *42*, 333.
7. Karatsu, T.; Miller, R. D.; Sooriyakumaran, R.; Michl, J. *J. Am. Chem. Soc.* **1989**, *111*, 1140.
8. McKinley, A. J.; Karatsu, T.; Wallraff, G. M.; Miller, R. D.; Sooriyakumararan, R.; Michl, J. *Organometallics* **1988**, *7*, 2567.
9. McKinley, A. J.; Karatsu, T.; Wallraff, G. M.; Thompson, D. P.; Miller, R. D.; Michl, J. *J. Am. Chem. Soc.* **1991**, *113*, 2003.
10. Ishikawa, M.; Kumada, M. *J. Chem. Soc. D.* **1971**, 489.
11. Davidson, I. M. T.; Michl, J.; Simpson, T. *Organometallics* **1991**, *10*, 842.
12. Stewart, J. J. P. *J. Comp. Chem.* **1989**, *10*, 221.
13. Coolidge, M. B.; Stewart, J. J. P. *OCPE Program No. 455.*
14. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902.
15. Dewar, M. J. S.; Jie, C. *Organometallics* **1987**, *6*, 1486.
16. The product mixtures in all these reaction systems were purified under the same gas chromatographic condition.
17. Barton, T. J. unpublished results. Also the synthetic procedure can be found in: Boo, B. H. The Chemistry of Silylsilylenes: Generation and Reactions of Trimethylsilylsilylene and Silylsilylene, Doctoral dissertation, Washington University, St. Louis, 1984.
18. Since one molecule of VI contains two subunits of silylene II, the yields of VI and VI-d_2 reported here is corrected to be twice larger than the measured yield based on the amount of the unrecovered starting material.
19. Davidson, I. M. T.; Howard, A. V. *J. Cheml. Soc., Faraday Trans. I*, **1975**, *71*, 69.
20. The photolysis of I in the presence of ethanol and 2-prop-2-en-1-ol is under study in our laboratory.
21. Only the mass spectrum is obtained since a tiny amount of the product is formed; mass: M^+ 98 (5.7), 84 (71), 83 (18), 69 (30), 56 (100), 55 (55), 44 (13), 43 (15), 42 (28), 41 (75), 39 (32), 32 (24).
22. It should be noted that the % decompositions do not exactly correspond to the scavenging efficiencies of the reactive silicon species since the UV irradiation time and the mole ratios of the reaction mixtures little differ each other.

A Quantitative Determination of Overlapped Chromatographic Peaks of Dysprosium and Yttrium Using Target Transformation Factor Analysis

Joon Myong Song, Chul Lee*, and Koo Soon Chung

Department of Chemistry, Sogang University, Seoul 121-742
**Department of Chemistry, Hanyang University, Seoul 133-791*

Received September 10, 1993

Rare earth elements (REE) were individual separated by applying the gradient elution *via* HPLC using α-hydroxy isobutyric acid (HIBA) as an eluent. However, the overlap of Y and Dy peaks was too severe to obtain the resolution of these two peaks. The target transformation factor analysis (TTFA) was applied to resolve the elution peaks of Y and Dy. $[A]_{raw}$ formed from the absorbances of mixed solution was factor analyzed. The abstract factor analysis (AFA) was used to determine the number of components that contributed to the poorly resolved peaks. The error theory in the AFA showed that the number of components was 2. The test vectors which correspond to pure component were selected from the standard solutions of Y and Dy. TTFA was accomplished by target testing. The results showed that the resolution of two peaks as well as the determination of Y and Dy were possible by the factor analysis.

Introduction

The utilization of high performance liquid chromatography (HPLC) procedures for the separation and determination of rare earth elements (REE) was well documented.[1] The REE were usually separated using a concentration gradient of eluting agent. However, because of close similarity in their chemical properties, the resolutions of HPLC for the separation of individual REE were difficult to achieve. For example, the gradient elution by HPLC was applied to separate individual REE using α-hydroxy isobutyric acid (HIBA) as an eluent. However, the overlap of the Y and Dy was too severe

to obtain the resolution of these two peaks. In order to improve the peak resolution, the elution conditions were adjusted by varying HIBA concentration and the pH of eluent. However, these attempts were unsatisfactory for rapid and complete separation of these two elements. The coelution of Y and Dy, as pointed out by Cassidy and Elchuk,[2] was mainly attributed to the similarity of Y and Dy in ionic size.

Accordingly, it was thought desirable to apply chemometric method with which Y and Dy could be determined *via* a purely mathematical route using their unresolved elution peaks. Factor analysis, which is one of the most powerful methods in chemometrics, has been successfully applied to resolve overlapped peaks in HPLC[3].

In the present work, factor analysis was carried out on the overlapped peaks of Y and Dy which were obtained when REE were eluted with HIBA. The results showed that the resolution of two peaks as well as the determination of Y and Dy were possible by factor analysis.

Theory

In HPLC, the raw data matrix, $[A]_{raw}$, is the absorbances of the eluate to be eluted through the column, which are measured at the regular wavelength interval while the overlapped peaks are observed. The rows in $[A]_{raw}$ are the time periods during which the elution peaks are observed and the columns in $[A]_{raw}$ are the chromatogram which are obtained at regular wavelength interval by the multichannel detector. The elements of $[A]_{raw}$, A_{ij}, measured in a cell of unit path length at time i and wavelength j, are analyzed by factor analysis. The A_{ij} can be expressed as Eq. (1)

$$A_{ij}=\sum_{k=1}^{n}(c_{ik}\cdot e_{kj}) \tag{1}$$

where c_{ik} is the concentration of the component k at time i and e_{kj} is the molar absorptivity of the component k at wavelength j. The above equation can be converted into another form by matrix notation.

$$\underset{i\times j}{[A]_{raw}}=\underset{i\times n}{[C]}\ \underset{n\times j}{[E]} \tag{2}$$

Where $[A]$ is an $i\times j$ absorbance matrix for n components, $[E]$ is an $n\times j$ pure absorptivity matrix, and $[C]$ is an $i\times n$ concentration matrix. In order to yield more accurate results statistically, the number of rows and columns should be 2^n or more. The raw data matrix is first converted into a covariance or correlation matrix. The correlation matrix is obtained by premultiplying the normalized data matrix by its transpose.

$$[Z]_N=[A]_{N,raw}[A]^T_{N,raw} \tag{3}$$

$[A]^T_{N,raw}$ is the transposed matrix of the normalized raw data matrix. Throughout this work, the correlation matrix about the origin has been used. The correlation matrix is decomposed by the method of principal component analysis,[4] in which this matrix is diagonalized by finding a matrix $[Q]$ such that

$$[Q]^{-1}[Z][Q]=[\lambda_i\cdot\delta_{ik}] \tag{4}$$

Here δ_{ik} is the well-known Kronecker delta and λ_i is an eigenvalue of the set of Eq. (6)

$$\delta_{ik}=\begin{matrix}0 \ \textit{if}\ i\neq k\\ 1\ \textit{if}\ i=k\end{matrix} \tag{5}$$

$$[Z]Q_i=\lambda_i Q_i \tag{6}$$

where Q_i is the ith column of $[Q]$. These columns called eigenvectors are usually normalized to form an orthonormal set. Hence

$$[Q]^{-1}=[Q]^T \tag{7}$$

The absorptivity matrix is expressed by Eq. (8) and the concentration matrix is also expressed by Eq. (9)

$$[E]_{abst}=[Q]^T \tag{8}$$

$$[C]_{abst}=[A]_{raw}[Q] \tag{9}$$

These two matrixes are abstract absorptivity and concentration matrixes, which are used to reproduce the original data matrix.

Experimental error invariably produces a larger number of eigenvectors than is required by the pure factor space. Accordingly, the significant eigenvectors (the number of components) and the residual eigenvectors (random noise, or error in the spectra) have to be distinguished. In this step, three types of error are used, which are real error (RE), imbedded error (IE) and extracted error (XE). These errors are calculated by the following equations.

$$RE=\left[\frac{\sum_i^r\sum_k^c A^2_{ik}\ \sum_{j=n+1}^{c}\lambda_j}{r\ c\ (c-n)}\right]^{(1/2)} \tag{10}$$

$$XE=(RE)\sqrt{\frac{(c-n)}{c}} \tag{11}$$

$$IE=[(RE)^2-(XE)^2]^{1/2} \tag{12}$$

where c is the number of rows or columns, which is smaller while r is the number of rows or columns which is larger. n is the number of factors used to reproduce the data matrix. λ_i is the i-th eigenvalue in Eq. (4).

It is difficult to obtain the number of appropriate vectors by the RE method. Besides, the method is not usable in this work since it depends on an accurate estimate of the error. The imbedded error function is useful since it requires no knowledge of experimental error[5]. When the correct number of factors is employed, IE function reaches a minimum. The minimum of IND (indicator) function like that of IE function also appears in correct number of factors. The IND function is much more sensitive method than the IE in that the minimum of the IND function appears more obviously than that of the IE function. The indicator function is defined as

$$IND=\frac{RE}{(c-n)^2} \tag{13}$$

The transformation process is necessary to convert abstract eigenvectors into vectors that have physical and chemical meaning. $[C]_{abst}$ is used in target transformation. The suspected factors are represented by a test matrix $[C]_{test}$. The test matrix obtained from empirical knowledge and intuition consists of the known concentration of the standard

solution and has the dimension ($c \times n$). The transformation matrix $[T]$ is obtained by the following Eq. (14)

$$\underset{n_c \times n_c}{[T]} = [\lambda]^{-1} \underset{n_c \times c}{[C]^{T}_{abst}} \underset{c \times n_c}{[C]_{test}} \quad (14)$$

where $[C]^{T}_{abst}$ is the transposed matrix of $[C]_{abst}$, n_c is the number of the determined components by eigenvector analysis. The transformation matrix from Eq. (14) is used to obtain predicted concentration matrix according to Eq. (15).

$$\underset{c \times n_c}{[C]_{pred}} = \underset{c \times n_c}{[C]_{abst}} \underset{n_c \times n_c}{[T]} \quad (15)$$

Whether the test matrix consists of true factors can be judged by comparing the test matrix with the predicted matrix. If the test matrix corresponds to the true factor, the differences between the corresponding elements are less than expected value. The real column matrix is calculated by premultiplying the abstract column matrix with the transformation matrix. The validity of a test vector can be judged by some error functions. They are the apparent error in the target test vector (AET), the real error in the predicted vector (REP), the real error in the target test vector (RET), and the SPOIL function.

$$AET = \left[\frac{\sum_{i=1}^{r} (\bar{r}_i - \bar{\bar{r}}_i)^2}{r} \right]^{(1/2)} \quad (16)$$

$$REP = (RE)_{n=n_c} (T_1^T \cdot T_1)^{1/2} \quad (17)$$

$$RET = [(AET)^2 - (REP)^2]^{1/2} \quad (18)$$

Here $\bar{r}_i$ and $\bar{\bar{r}}_i$ are the ith elements of the predicted vector and the test vector, respectively. $(RE)_{n=n_c}$ is the real error obtained from AFA using n_c factors. T_1 is a transformation vector defined in Eq. (14). The SPOIL values are calculated from RET and REP values as follows.

$$SPOIL \cong \frac{RET}{REP} \quad (19)$$

The value of the SPOIL supply an excellent criterion to judge the overall validity of a suspected target[6]. The SPOIL values between 0.0 and 3.0 are proper to be accepted as a target.

Experimental

Various procedures are currently employed to separate the individual REE by HPLC. Of the procedures for the separation, IPC (ion pair chromatography) is widely used in HPLC. The reagents used as the mobile phase were 1 M α-hydroxyisobutyric acid (HIBA) that was used for the complexing agent and 0.1 M sodium n-octanesulfonate (O. S.) which was used for the ion pair reagent. The HIBA solution was buffered at pH 4.6 with NaOH. The stationary phase was 5 μm Superisorb C_{18} column and 1.5×10^{-4} M 3,6-bis[(o-arsenophenyl)azo]-4,5-dihydroxy-2,7-naphthalenedisulfonic acid (Arzenazo III) was used as post column reagent. In the present method, it was found that individual REE could be separated and determined within 30 mins. as shown in Figure 1. The individual REE was separated *via* gradient concentration from 0.05 to 0.09 M HIBA over the initial 10 min at 1 mL min^{-1} and from 0.09 to 0.4 M over the succeeding 21 min. and then held at 0.4 M for 5 min. during the final elution.[7]

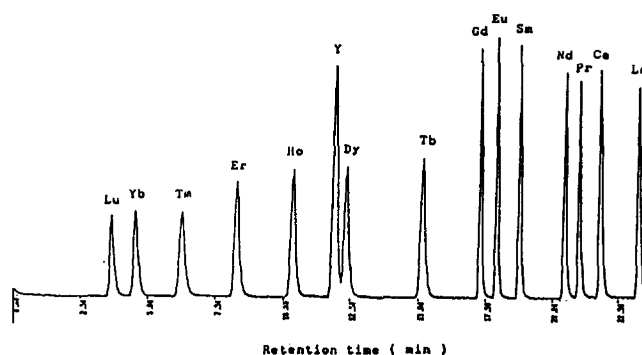

Figure 1. The separation of the rare earth elements using α-hydroxy isobutyric acid (HIBA). The individual REE was separated *via* gradient concentration from 0.05 to 0.09 M HIBA over the initial 10 min. at 1 mL min^{-1} and from 0.09 to 0.4 M over the succeeding 21 min. and then held at 0.4 M for 5 min. during the final elution.

Reagents. The O. S., ion pair reagent, (Kodak), and HIBA and Arsenazo III (Fluka) were used. The distilled water was filtered through the Mill-Q deionizing unit (Barnsted Sybron CO). The stock solutions of each REE were prepared by the corresponding oxides (99.9%, produced by the Cerac), as follows. Exact amount of REE corresponding to 2000 ppm was dissolved with 5 mL of concentrated HNO_3 by warming on the hot plate. The solution was transferred to the volumetric flask, and filled to the mark with 1% HNO_3. The standard solution of the REE was made by diluting the stock solution. All the mobile phase and the color developing agent were filtered through 0.45 μm Millipore filter before use.

HPLC System. Gradient elutions were performed using the Waters Associates pumps (M 510) equipped with a Waters M 660 gradient controller. Samples were injected *via* a Waters U6K sampling valve. The post-column reagent was added with the pump, at the same flow rate used for elution of the REE from the column. Pulse damper (Model LP-21 Lo-Pulse, Scientific Systems, Inc.) and restriction coil were used to reduce the pulsation from the pump. The eluates were monitored with a Waters photodiode array detector. The detector output was monitored with a Waters Data Module. The details of the procedure and the experimental condition of the HPLC system are given elsewhere.[7]

TTFA. The TTFA was applied to determine each component of poorly resolved peaks, *i.e.*, Y and Dy in the chromatogram of the REE. The compositions of the sample solutions are given in Table 1. The 1st solution was the mixed standard solution that consisted of 13 ppm each of Y and Dy. The 2nd and 3rd solution contained 10 ppm each of Y and Dy, respectively. Before injecting the sample solution, the HPLC column was previously equilibrated with the mobile phase containing 0.03 M O.S. and 0.01 M HIBA. The eluted metal-HIBA complexes were monitored by the photodiode array detector after a postcolumn reaction with Arsenazo III. The poorly resolved peaks of Y and Dy were observed from 11.5 min to 13.0 min. A 16-row by 10-column data mat-

Table 1. Composition of sample solution and test vector solutions

	Dy concentration	Y concentration
1st solution	13 ppm	13 ppm
2nd solution	10 ppm	0
3rd solution	0	10 ppm

Table 2. Results of abstract factor analysis on the chromatogram of the sample, which was formed from the absorbances measured by the photodiode array detector (retention time=11.5-13.0 min, wavelengths=648-657 nm)

Result n	Eigenvalue	RE ($\times10^{-4}$)	IE ($\times10^{-4}$)	IND ($\times10^{-5}$)
1	9.9998	11.19	3.539	1.382
2	1.559×10^{-4}	4.546	2.033	0.710
3	2.330×10^{-5}	1.676	0.034	0.342
4	1.049×10^{-6}	1.478	0.935	0.411
5	1.049×10^{-6}	1.145	0.810	0.458

rix, $[A]_{raw}$, was absorbances of 16 fractions (per 6 sec) measured at 1 nm interval from 648 nm to 657 nm using photodiode array detector, which was obtained from the 1st solution in Table 1. $[A]_{raw}$ was analyzed by the abstract factor analysis (AFA). The 2nd and 3rd solutions were used as the test vectors. Test vectors of two single solutions were constructed from absorbances of 16 fractions(per 6 sec) measured at 653 nm at which the maximum absorbance was observed.

Results and Discussion

The error theory in the AFA was applied to determine the number of components. Because experimental error in the raw data matrix invariably produces a larger number of eigenvectors than is required by the pure factor space, the method is necessary to distinguish the significant eigenvector from the residual eigenvector. This was accomplished by observing the values of the RE, the IE, and the IND that were obtained with $n=1$ to $n=5$ factors, as shown in Table 2. The extended values of RE, IE, and IND which correspond to n above 5 were attributed to noise and accordingly were neglected. The RE can be estimated from experimental errors. However, because such information is difficult to obtain in many real chemical problems, as in the case of the present work, the IE and IND were used. The IE is due to the fact that a fraction of the error from the data mixes into the reproduction process. According to the IE function, it can be deduced that the number of factors is 3 because the IE function decreases as we use the primary eigenvectors, *viz.*, $n=3$, in the data reproduction but increases when we begin to include the secondary eigenvectors, $n=4$, in the reproduction. The IND value for data matrix $[A]_{raw}$ also reached a minimum at $n=3$. However, we know that the number of components that contribute to the unresolved peaks is 2, which is Y and Dy. However, the data obtained from by the IND and the IE function indicated that IE function is much smaller at $n=3$ than at $n=2$ and also that the minimum of IND function reaches at one more than the proper number of factor as shown in Table 2. Such conclusion may be ascribable to the baseline effect which contributes to the total absorbance in the raw data matrix. This effect which contributes to the concentration error is not very much as shown in Table 3, *i.e.*, because the difference of the absorbance between 11.5 and 13.0 min., during which the unresolved peaks are observed, is not very much. Accordingly, the failure here to show the proper number of factors should not be considered as significant.

The test vectors were selected from the absorbances of Y and Dy measured at 653 nm from 11.5 to 13.0 min. because these vectors correspond to the components of the poorly resolved peaks. The elements of the test vectors correspond

Table 3. The elements of the raw data matrix, $[A]_{raw}$ (row: retention time=11.5-13.0 min, column: wavelengths=648-657 nm)

	648	649	650	651	652	653	654	655	656	657
11.5	0.00130	0.00139	0.00171	0.00123	0.00188	0.00346	0.00233	0.00094	0.00126	0.00136
11.6	0.01405	0.01473	0.01486	0.01406	0.01455	0.01708	0.01614	0.01451	0.01486	0.01491
11.7	0.05523	0.05594	0.05582	0.05629	0.05805	0.05975	0.05833	0.05670	0.05695	0.05700
11.8	0.1131	0.1144	0.1152	0.1160	0.1182	0.1212	0.1199	0.1180	0.1185	0.1184
11.9	0.1744	0.1773	0.1793	0.1797	0.1809	0.1839	0.1829	0.1810	0.1815	0.1814
12.0	0.1151	0.1163	0.1174	0.1189	0.1208	0.1223	0.1213	0.1201	0.1203	0.1202
12.1	0.06413	0.06518	0.06629	0.06669	0.06813	0.06991	0.06889	0.06762	0.06802	0.06803
12.2	0.07084	0.07208	0.07199	0.07233	0.07447	0.07710	0.07595	0.07435	0.07481	0.07486
12.3	0.1166	0.1193	0.1200	0.1202	0.1225	0.1262	0.1251	0.1231	0.1235	0.1236
12.4	0.1105	0.1140	0.1161	0.1169	0.1182	0.1205	0.1196	0.1182	0.1187	0.1187
12.5	0.05233	0.05338	0.05478	0.05528	0.05615	0.05817	0.05734	0.05606	0.05647	0.05663
12.6	0.01988	0.02078	0.02074	0.02023	0.02119	0.02367	0.02272	0.02111	0.02137	0.02136
12.7	0.01025	0.00971	0.00911	0.00891	0.01033	0.01282	0.01179	0.01013	0.01046	0.01070
12.8	0.00432	0.00513	0.00419	0.00359	0.00527	0.00797	0.00681	0.00503	0.00531	0.00529
12.9	0.00343	0.00349	0.00288	0.00255	0.00408	0.00672	0.00533	0.00374	0.00413	0.00432
13.0	0.00340	0.00347	0.00200	0.00154	0.00324	0.00527	0.00389	0.00226	0.00263	0.00262