

# Use of Evolutionary Factor Analysis in the Spectroelectrochemistry of *Escherichia coli* Sulfite Reductase Hemoprotein and a Mo/Fe/S Cluster

Robert L. Keeseey and Michael D. Ryan\*

Chemistry Department, Marquette University, P.O. Box 1881, Milwaukee, Wisconsin 53201

**The deconvolution of spectroelectrochemical data is often quite difficult if the spectra of intermediates are not known. Factor analysis, however, has been shown to be a powerful technique which can make it possible to deconvolute overlapping spectra. In this work, evolving factor analysis will be used to determine the number of intermediates and the spectra of those species for two typical spectroelectrochemical experiments: linear scan voltammetry and chronoabsorptometry in a thin-layer cell. The first system was the reduction of *E. coli* sulfite reductase hemoprotein (SiR-HP). Principal factor analysis indicated that three species were present. By using evolving factor analysis, the potential regions where each of the species were present were identified, and their concentrations and spectra were determined by the use of the mass balance equation. The spectra of the one-electron (SiR-HP<sup>1-</sup>) and two-electron (SiR-HP<sup>2-</sup>) reduced product were compared with previous work. The second experiment was the chronoabsorptometry of Cl<sub>2</sub>FeS<sub>2</sub>-MoS<sub>2</sub>FeCl<sub>2</sub><sup>2-</sup> in methylene chloride. This experiment indicated that five species were present during the experiment. The entire set of 61 spectra were fit by assuming that there were 4 species present during the electrolysis. The rate constant for the appearance of subsequent species fit quite well with the rate constant for the disappearance of previous species. The spectra of the intermediates and final product were obtained using evolving factor analysis and a mass balance equation. Identification of the fifth species, which was probably the initial reduction product, Cl<sub>2</sub>FeS<sub>2</sub>MoS<sub>2</sub>FeCl<sub>2</sub><sup>3-</sup>, was difficult due to its low concentration and the fact that it was present in the same time region as the starting material.**

Factor analysis has been shown to be a powerful technique for the analysis of complex analytical data.<sup>1</sup> However, only a few reports have been presented on the use of factor analysis in voltammetry, and most of these have focused on the quantitative aspects.<sup>2–4</sup> Most chemical applications of factor analysis have dealt

with chromatography,<sup>5–7</sup> injection analysis,<sup>8,9</sup> and chemical equilibrium.<sup>10–13</sup> Surprisingly, there have been no applications of factor analysis to spectroelectrochemical data. Spectroelectrochemistry has many of the same complexities that have been observed in the chromatographic and flow injection data, namely, an unspecified number of components whose spectral features may not be well characterized.

The spectra of the redox product(s) in spectroelectrochemical data can be readily analyzed if only one redox couple is present or if the redox potentials for multielectron transfers are well separated. In these cases, the spectral deconvolution can be done in a straightforward manner. On the other hand, it will be very difficult to determine the spectrum of the intermediate and deconvolute the spectra if the product of the electron transfer is not stable or if the *E*'s of consecutive electron transfers are close together.<sup>14,15</sup> In addition to the problem of deconvolution, it is difficult to unambiguously identify the number of components present in solution if more than two species are present. Both of these problems can be readily addressed by the use of factor analysis.

## THEORY

Factor analysis is a methodology to analyze large data sets by reducing the data to their lowest dimensionality.<sup>16</sup> This is achieved through abstract matrices which have no physical significance. Factor analysis normally begins with the determination of the number of principal factors in the data matrix

(1) Malinowski, E. R. *Factor Analysis in Chemistry*; John Wiley & Sons: New York, 1991.  
(2) Diazcruz, J. M.; Tauler, R.; Grabaric, B. S.; Esteban, M.; Casassas, E. J. *Electroanal. Chem.* **1995**, *393*, 7–16.

(3) Grabaric, B. S.; Grabaric, Z.; Tauler, R.; Esteban, M.; Casassas, E. *Anal. Chim. Acta* **1997**, *341*, 105–120.  
(4) Ni, Y. N.; Bai, J. L.; Jin, L. *Anal. Lett.* **1997**, *30*, 1761–1777.  
(5) Maeder, M. *Anal. Chem.* **1987**, *59*, 527–530.  
(6) Gargallo, R.; Tauler, R.; Cuestasanchez, F.; Massart, D. L. *TrAC, Trends Anal. Chem.* **1996**, *15*, 279–286.  
(7) Sanchez, F. C.; Vandeginste, B. G. M.; Hancewicz, T. M.; Massart, D. L. *Anal. Chem.* **1997**, *69*, 1477–1484.  
(8) Gemperline, P. J.; Hamilton, J. C. *J. Chemom.* **1989**, *3*, 455–461.  
(9) Malinowski, E. R. *J. Chemom.* **1992**, *6*, 29–40.  
(10) Kankare, J. J. *Anal. Chem.* **1970**, *42*, 1322–1326.  
(11) Gamp, H.; Maeder, M.; Meyer, C. J.; Zuberbühler, A. D. *Talanta* **1985**, *32*, 1133–1139.  
(12) Meinrath, G.; Schweinberger, M. *Radiochim. Acta* **1996**, *75*, 205–210.  
(13) Darj, M. M.; Malinowski, E. R. *Anal. Chem.* **1996**, *68*, 1593–1598.  
(14) Yap, W. T.; Marbury, G.; Blubaugh, E. A.; Durst, R. A. *J. Electroanal. Chem.* **1989**, *271*, 325–329.  
(15) Langhus, D. L.; Wilson, G. S. *Anal. Chem.* **1979**, *51*, 1139–1144.  
(16) Rogers, L. J.; Adams, M. *J. Pharm. Sci.* **1997**, *3*, 333–336.

(**D**). Using a series of spectra of the solution at each potential (or time) as columns and the absorbance readings at each wavelength as rows, the data matrix can be decomposed into two matrices,

$$\mathbf{D} = \mathbf{R}_{\text{abst}} \mathbf{C}_{\text{abst}} \quad (1)$$

where  $\mathbf{R}_{\text{abst}}$  and  $\mathbf{C}_{\text{abst}}$  are the abstract row and column matrices. Various procedures can be used to obtain the  $\mathbf{R}_{\text{abst}}$  and  $\mathbf{C}_{\text{abst}}$  matrices, the most common being the use of singular value decomposition (SVD),

$$\mathbf{D} = \mathbf{USV}' \quad (2)$$

where  $\mathbf{U}$  is an abstract orthonormal eigenvector that spans the row space,  $\mathbf{S}$  is a diagonal matrix whose elements are the square roots of the eigenvalues ( $\lambda$ ), and  $\mathbf{V}'$  is the transpose of  $\mathbf{V}$  which is an abstract, orthonormal matrix that spans the column space. From the SVD, the abstract row and column matrices can be calculated as follows:

$$\mathbf{R}_{\text{abst}} = \mathbf{US} \quad (3)$$

$$\mathbf{C}_{\text{abst}} = \mathbf{V} \quad (4)$$

Principal factor analysis estimates directly the number of factors (species) in a set of spectra. This is possible because the number of significant factors (chemical species) that are present equals the number of nonzero eigenvalues (the eigenvalues are the square of the diagonal elements of the  $\mathbf{S}$  matrix). The number of diagonal elements in  $\mathbf{S}$  is equal to the minimum of the number of rows or columns in  $\mathbf{D}$ . Therefore, to find two factors, you will need at least two wavelengths and two solutions (3 factors: 3 wavelengths and 3 solutions, etc.). In practice, though, there are two sets of eigenvalues: the principal factors in the spectra and the noise factors. Real data, of course, contain noise, and most of the eigenvalues will be attempts of the method to fit the noise to the spectra and do not contain useful information. The eigenvalues that are not associated with a real factor in the solution will roughly correspond to the noise level. Because the method attempts to fit both the data and the noise, one will need many more than the minimum number of spectra/wavelengths in order to obtain reliable data.

Methods that are available to distinguish the noise factors from the real factors are the real error (RE), the imbedded error (IE), the chi value ( $\chi$ ), or an empirical indicator function (IND).<sup>1</sup> The value of RE can be calculated from the eigenvalues,  $\lambda$ , the number of rows ( $r$ ), columns ( $c$ ), and principal factors ( $n$ ).

$$\text{RE} = \left( \sum_{j=n+1}^c \lambda_j^0 / r(c-n) \right)^{1/2} \quad (5)$$

The RE measures the difference between the pure data and the experimental data. The RE will fall below the expected noise in the data when all of the real factors have been included. To avoid estimating the noise level, Malinowski has discovered an IND which reaches a minimum value when the correct number of

principal factors are used.

$$\text{IND} = \text{RE} / (c-n)^2 \quad (6)$$

Equations 5 and 6 apply when  $r > c$ . If  $r < c$ , then  $r$  and  $c$  should be interchanged in these expressions.

The abstract row and column matrices are transformed into the real concentration ( $\mathbf{M}$ ) and molar absorptivity ( $\mathbf{E}$ ) matrices once the number of factors has been estimated. This is done by the transformation matrix ( $\mathbf{T}$ ), where,

$$\mathbf{M} = \mathbf{R}_{\text{abst}} \mathbf{T} \quad (7)$$

and,

$$\mathbf{E} = \mathbf{T}^{-1} \mathbf{C}_{\text{abst}} \quad (8)$$

where the number of columns in  $\mathbf{R}_{\text{abst}}$  and rows in  $\mathbf{C}_{\text{abst}}$  has been reduced to equal the number of principal factors ( $n$ ). There is an equivalence between the number of chemical species and the number of factors, so it might be convenient to think of each factor as being associated with a given chemical species. This is not true. It is only after applying the transformation matrix,  $\mathbf{T}$ , that the first row of  $\mathbf{E}$  will correspond to species 1. The first row of  $\mathbf{C}_{\text{abst}}$  corresponds to the first factor, which is itself a linear combination of all the principal factors (as is the second factor, third factor, etc.). Therefore, the appearance of a second factor means that there are two chemical species, but it does not mean that factor 1 is species 1 and factor 2 is species 2.

While eqs 1–6 can be readily solved, the determination of  $\mathbf{T}$ , which is needed for eq 7, presents the greatest challenge. An examination of eq 7 shows that, if there are  $n$  factors in the data, one can also write  $n$  equations of the following form:

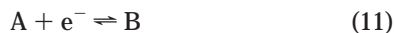
$$m_i = \mathbf{R}_{\text{abst}} t_i \quad (9)$$

where  $m_i$  is the concentration vector of species  $i$  and  $t_i$  is the transformation vector of species  $i$ . The vectors  $m_i$  and  $t_i$  are the  $i$ th column of  $\mathbf{M}$  and  $\mathbf{T}$ , respectively. Each row of  $m_i$  and  $\mathbf{R}_{\text{abst}}$  corresponds to a given spectrum in the spectroelectrochemical experiment. It is not necessary to know all the values of  $m_i$  in order to solve for  $t_i$ . In addition, we only need to know the concentration of species  $i$ , not all the species that are in the spectra. If we know some of the concentration values in  $m_i$ , we can generate a new expression where we use only those spectra (rows) for which the concentration of species  $i$  is known.

$$m_i^0 = \mathbf{R}_i^0 t_i \quad (10)$$

where  $m_i^0$  contains only the concentrations of species  $i$  that are known and  $\mathbf{R}_i^0$  contains the rows of  $\mathbf{R}_{\text{abst}}$  that correspond to the  $m_i^0$  vector. The concentration values in  $m_i$  can then be calculated from eq 9 once  $t_i$  is known from eq 10. This process can be repeated for all the species in order to generate the transformation matrix  $\mathbf{T}$ , as is illustrated below.

Let's consider the reduction of species A to B and then to C:



There are three species in the above example, and one expects to have three principal components ( $n = 3$ ). Therefore,  $\mathbf{R}_{\text{abst}}$  will have three columns and one row for each spectral scan.  $\mathbf{C}_{\text{abst}}$  will have three rows and one column for each spectral wavelength. At the beginning of the scan, [A] is finite, while [B] = [C] = 0. By the end of the scan, [A] = 0, while [C] will be nonzero.

The procedure that was used in this work to generate the vector  $m_i^0$  is called evolutionary factor analysis (EFA).<sup>17</sup> In an evolutionary process, each species appears only once and has a concentration equal to zero for some region(s) of potential (or time). As each new species appears (e.g., when the potential is scanned through the  $E^0$ ), an eigenvalue associated with that species will also appear. Therefore, if one calculates a series of  $\mathbf{S}$  matrices by successively adding spectra to the data set, the eigenvalues for each principal factor will begin to appear in the potential scan (forward eigenvalues) as they deviate from the background noise. The process can be repeated by starting from the opposite end of the series of spectra to produce eigenvalues (reverse eigenvalues) that indicate the potential (or time) where a species disappears. In this case, the last two spectra are used to obtain the first set of reverse eigenvalues, then the last three, and so forth until all the spectra (from last to first) are used to obtain the reverse eigenvalues. The potential (time) range where each species is present is arrived at by considering the forward and reverse eigenvalues.

From EFA, we can determine the potential (spectrum number) where the second and third eigenvalues are significantly larger than the noise, hence, the potential where their concentrations are nonzero. If we know the potentials (or times) where the concentrations are nonzero, we may infer that their concentrations are zero outside of that potential (time) region. By knowing the potentials (or times) where [A] = 0, it is possible to solve for  $t_A$  using eq 10. Let us consider an experiment with 8 spectra (rows) in the experiment, and [A] is equal to zero in spectra 6–8.  $\mathbf{R}_A^0$  will contain only the last three rows of  $\mathbf{R}_{\text{abst}}$  as shown below:

$$\mathbf{R}_{\text{abst}} = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \\ r_{41} & r_{42} & r_{43} \\ r_{51} & r_{52} & r_{53} \\ r_{61} & r_{62} & r_{63} \\ r_{71} & r_{72} & r_{73} \\ r_{81} & r_{82} & r_{83} \end{pmatrix} \Rightarrow \mathbf{R}^0 = \begin{pmatrix} r_{61} & r_{62} & r_{63} \\ r_{71} & r_{72} & r_{73} \\ r_{81} & r_{82} & r_{83} \end{pmatrix} \quad (13)$$

and eq 10 becomes

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} r_{61} & r_{62} & r_{63} \\ r_{71} & r_{72} & r_{73} \\ r_{81} & r_{82} & r_{83} \end{pmatrix} \begin{pmatrix} t_1 \\ t_2 \\ t_3 \end{pmatrix} \quad (14)$$

The trivial solution is that  $t_A$  is zero, but a nontrivial solution can be obtained if we assume a value (e.g., 1) for one of the elements of  $t_i$ . In this case, eq 14 can be rearranged to

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} r_{61} & r_{62} & r_{63} \\ r_{71} & r_{72} & r_{73} \\ r_{81} & r_{82} & r_{83} \end{pmatrix} \begin{pmatrix} 1 \\ t_2 \\ t_3 \end{pmatrix} \Rightarrow \begin{pmatrix} -r_{61} \\ -r_{71} \\ -r_{81} \end{pmatrix} = \begin{pmatrix} r_{62} & r_{63} \\ r_{72} & r_{73} \\ r_{82} & r_{83} \end{pmatrix} \begin{pmatrix} t_2 \\ t_3 \end{pmatrix} \quad (15)$$

and it is now possible to solve for the nontrivial solutions of  $t_A$ . In assuming that  $t_1 = 1$ , only the relative changes in [A] can be determined, not the actual concentrations. This is corrected for in the end using the mass balance equation.

This process can be repeated for B and C so that we will have the entire matrix  $\mathbf{T}$ . The relative concentrations of each species for each spectra can be calculated as follows once the  $\mathbf{T}$  matrix is known:

$$\mathbf{m} = \mathbf{R}_{\text{abst}} \mathbf{T} \quad (16)$$

where  $\mathbf{m}$  is the relative concentration array. This procedure is mathematically equivalent to rank annihilation evolving factor analysis (RAEFA)<sup>18</sup> and to window factor analysis (WFA).<sup>19</sup>

The actual concentration of the initial material (A) can be easily found because A is the only species present at the beginning of the scan. The normalization factor for A,  $nm_A$ , is equal to

$$nm_A = C_A / m_{11} \quad (17)$$

where  $C_A$  is the initial molar concentration of A and  $m_{11}$  is the relative concentration of A in the first spectrum (element in the first column and first row of  $\mathbf{m}$ ). Typically, the values of  $m_A$  in the first several spectra are averaged in order to calculate  $nm_A$ . The concentration of A in any spectra  $i$ ,  $C_{A,i}$ , is then equal to

$$C_{A,i} = nm_A m_{i1} \quad (18)$$

where  $m_{i1}$  is the relative concentration in the  $i$ th row and the first column. Similar procedures can be used to calculate the values of  $nm_B$  and  $nm_C$  if there is a region where only one of the species is present. Often there is an overlap, so the concentrations of B and C can be determined by a set of linear equations:

$$C_A = C_{A,i} + nm_B m_{i2} + nm_C m_{i3} \quad (19)$$

where  $m_{i2}$  and  $m_{i3}$  are defined analogously to  $m_{i1}$  in eq 18. There are only two unknowns in the equation above ( $nm_B$  and  $nm_C$ ) so

(17) Gampp, H.; Maeder, M.; Meyer, C. J.; Zuberbühler, A. D. *Chimia* **1985**, *39*, 315–317.

(18) Gampp, H.; Maeder, M.; Meyer, C. J.; Zuberbühler, A. D. *Anal. Chim. Acta* **1987**, *193*, 287–293.

(19) Den, W.; Malinowski, E. R. *J. Chemom.* **1993**, *7*, 89–98.

the problem will be overdetermined and one can use a least-squares procedure to determine the best values for  $nm_B$  and  $nm_C$ . Once the concentrations of A, B, and C are determined, the transformation matrix for their actual concentrations can be calculated by rearranging eq 7. The molar absorptivity matrix, **E**, can then be calculated by use of eq 8, where the rows of **E** correspond to the spectrum of each species.

## EXPERIMENTAL SECTION

**Equipment.** The optically transparent thin-layer electrochemical cell (OTTLE) that was used for the Mo/Fe/S cluster has been previously described.<sup>20</sup> The working electrode was a platinum mesh. The OTTLE cell that was used for the *Escherichia coli* sulfite reductase hemoprotein (SiR-HP) was a low-volume cell with a methyl viologen-modified gold minigrid electrode.<sup>21</sup> The cell compartment was isolated from the ambient atmosphere by a glovebag and the bag was purged with dinitrogen cooled to 9–10 °C. The light beam was sealed with microscope cover slips. Visible spectra were obtained using a Hewlett-Packard 8452 diode array spectrophotometer.

**Chemicals.** The complex  $(\text{Ph}_4\text{P})_2[\text{MoS}_4\text{Fe}_2\text{Cl}_4]$  was synthesized by a literature procedure.<sup>22</sup> The *E. coli* sulfite reductase hemoprotein was isolated according to the methods of Siegel et al.<sup>23,24</sup> The *E. coli* B contained a pBR322 plasmid, which had the ampicillin resistance gene and the *E. coli* B sulfite reductase genes for flavoprotein (*CysI*), hemoprotein (*CysJ*), and uroporphyrinogen III methyltransferase (*CysG*) (gift from N. Kredich, Duke University).

**Procedures.** The reference spectra for the SiR-HP experiments were obtained using the OTTLE cell with a standard buffer. After the reference spectrum was obtained, the standard buffer was removed and the cell dried with a water aspirator. The thawed SiR-HP was degassed for 10 min using a vacuum pump. The degassed SiR-HP was taken into a 1-mL tuberculin syringe and delivered into the dried OTTLE cell. The Ag/AgCl (saturated KCl) reference electrode was inserted into the OTTLE cell overflow chamber. The entire setup was transferred to the diode array spectrophotometer sample compartment. The cell holder was then screwed into the sample compartment, covered by a glovebag. The glovebag was purged with cooled dinitrogen (9–10 °C). To further reduce dissolved dioxygen, the potentiostat was set to –0.49 V until the current reading reduced to background levels. Before a potential scan was performed, the potentiostat was adjusted to –0.400 V.

The solution containing the Mo/Fe/S cluster in methylene chloride was prepared in a glovebox and placed into the OTTLE cell. The potential was stepped from –0.60 to –1.05 V versus Ag/AgNO<sub>3</sub> in acetonitrile, and the spectral acquisition was simultaneously initiated. The mathematical analysis of the spectra by factor analysis was carried out using MATLAB (The Math Works, Inc.).

(20) Lin, X. Q.; Kadish, K. M. *Anal. Chem.* **1985**, *57*, 1498–1501.

(21) Landrum, H. L.; Salmon, R. T.; Hawkrigde, F. M. *J. Am. Chem. Soc.* **1977**, *99*, 3154–3158.

(22) Coucouvanis, D.; Baenziger, N. C.; Simhom, E. D.; Stremple, P.; Swenson, D.; Simopoulos, A.; Kostikas, A.; Petrouleas, V.; Papaefthymiou, V. *J. Am. Chem. Soc.* **1980**, *102*, 1732–1734.

(23) Siegel, L. M.; Murphy, M. J.; Kamin, H. *J. Biol. Chem.* **1973**, *248*, 251–264.

(24) Young, L. J.; Siegel, L. M. *Biochemistry* **1988**, *27*, 5984–5990.

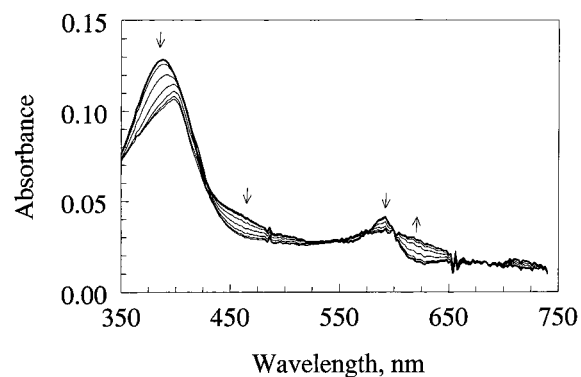
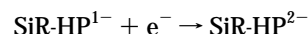
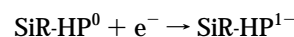


Figure 1. Thin-layer visible spectroelectrochemistry of *E. coli* sulfite reductase hemoprotein. Bold line is the initial spectrum. Other lines are repetitive scans at  $E = -0.60, -0.64, -0.68, -0.72, -0.76,$  and  $-0.80$  V versus Ag/AgCl. Scan rate, 0.30 mV/s;  $E_i = -0.40$  V;  $E_f = -0.80$  V; pH 7.2.

## RESULTS AND DISCUSSION

**Spectroelectrochemistry of *E. coli* Sulfite Reductase Hemoprotein.** *E. coli* sulfite reductase hemoprotein directly reduces at a methyl viologen-modified gold minigrid electrode. The visible spectral changes for a scan rate of 0.30 mV/s are shown in Figure 1. The spectra obtained are consistent with the photochemical reduction of SiR-HP,<sup>25</sup> indicating that the following redox reactions had occurred:



where the oxidation states of the prosthetic groups of SiR-HP are as follows: SiR-HP<sup>0</sup>, ferric siroheme, [4Fe-4S]<sup>2+</sup>; SiR-HP<sup>1-</sup>, ferrous siroheme, [4Fe-4S]<sup>2+</sup>; and SiR-HP<sup>2-</sup>, ferrous siroheme, [4Fe-4S]<sup>+</sup>.<sup>25</sup>

Reversal of the scan regenerated the original oxidized spectrum. A series of isosbestic points was observed during the initial stage of the reduction, and a second set was observed near the end of the scan. Between these two sets of isosbestic points, though, no isosbestic points were discernible, indicating the presence of all three oxidation states. It was impossible to obtain the spectrum of the intermediate (SiR-HP<sup>1-</sup>) because of the spectral overlap.

The first step in the principal factor analysis of the spectroelectrochemistry of SiR-HP was the determination of the number of principal factors. The eigenvalues were determined from the square of the diagonal elements in the **S** matrix, determined by SVD (eq 2). The results are summarized in Table 1. The value of the RE and the IND were calculated using eqs 5 and 6, respectively. The IND reached minimum at 3 factors. This is consistent with a two-electron reduction of SiR-HP, with no evidence of any other intermediates.

EFA can now be used to determine the potential regions where each of the species (SiR-HP<sup>0</sup>, SiR-HP<sup>1-</sup>, SiR-HP<sup>2-</sup>) are present. The forward EFA for the first five eigenvalues is shown in Figure 2. There are two important features of this figure. First, only one factor is needed to generate the initial spectrum. Factors 2 and 3, which are initially insignificant, begin to increase significantly

(25) Janick, P. A.; Siegel, L. M. *Biochemistry* **1982**, *21*, 3538–3547.

Table 1. Principal Factor Analysis of SiR-HP Spectroelectrochemistry

principal factor	eigenvalue	RE $\times 10^5$	IND $\times 10^8$
1	10.74	254	160
2	0.0242	52.5	34.5
3	0.0010	9.57	6.62
4	$4.33 \times 10^{-6}$	9.06	6.62
5	$2.95 \times 10^{-6}$	8.72	6.72
6	$2.70 \times 10^{-6}$	8.36	6.84
7	$2.26 \times 10^{-6}$	8.09	7.00
8	$2.14 \times 10^{-6}$	7.79	7.15
9	$1.88 \times 10^{-6}$	7.52	7.34
10	$1.79 \times 10^{-6}$	7.24	7.54

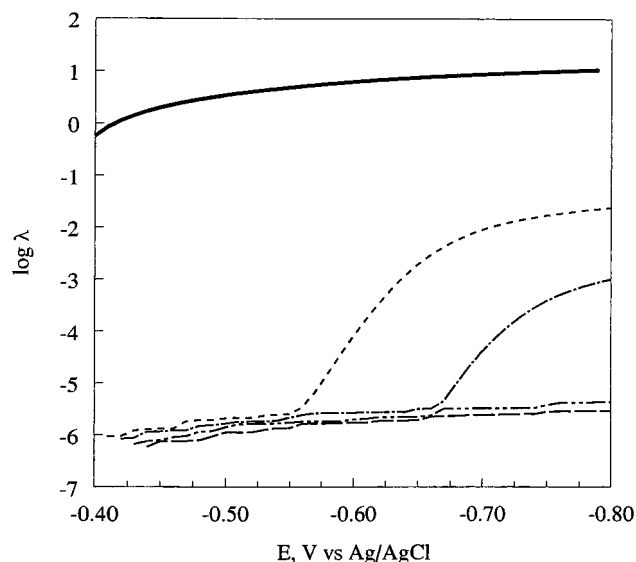


Figure 2. Forward EFA eigenvalues for the thin-layer spectroelectrochemistry of *E. coli* sulfite reductase hemoprotein: factors 1 (—), 2 (---), 3 (- · - ·), 4 (- · · -), and 5 (- - -).

above the background values as the potential reaches more negative values. Second, factors 4 and 5 (and all the larger factors) never differ significantly from the noise level. This further confirms the presence of three factors and gives us the potential windows where each of the species is initially formed. The forward and reverse EFA of the spectroelectrochemical data for the first three principal factors are shown in Figure 3.

On examining Figure 3, only one factor (factor 1) is present between the beginning of the scan and  $-0.55$  V (spectrum 16). At that point, a second factor (factor 2) appears. It is reasonable to assume that this corresponds to the appearance of the one-electron-reduced species, SiR-HP<sup>1-</sup>, but we must not equate factor 1 with SiR-HP<sup>0</sup> and factor 2 with SiR-HP<sup>1-</sup> (see Theory section). Similarly, at  $-0.66$  V (spectrum 27), a third factor appears, which corresponds to the appearance of SiR-HP<sup>2-</sup> (as above we must not equate species number and factor number). From the forward EFA, we can deduce that factor 1 appears at spectrum 1, factor 2 at spectrum 16, and factor 3 at spectrum 27.

To generate the window where each factor is present, the reverse EFA will be examined (Figure 3). Except for the last set of spectra, the data sets for the forward and reverse EFA are different. As a result, we should not equate factor 1 of the forward data set with factor 1 of the reverse data set. Instead, we want to

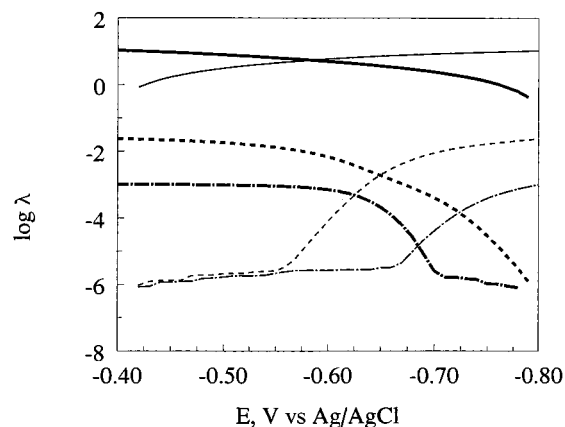


Figure 3. Forward and reverse EFA eigenvalues for the thin-layer spectroelectrochemistry of *E. coli* sulfite reductase hemoprotein: factors 1 (—), 2 (---), and 3 (- · - ·). Bolder lines are the reverse EFA eigenvalues.

determine when each factor appears in the reverse EFA (which actually corresponds to the disappearance of the factor in the scan). In the reverse EFA, a second factor appears immediately in the analysis (Figure 3). This means that there are still two factors (chemical species) at  $-0.80$  V (spectrum 41) or the second factor has just disappeared. The species that would make the most chemical sense would be the most reduced species, SiR-HP<sup>1-</sup> and SiR-HP<sup>2-</sup>. At  $-0.70$  V (spectrum 32), a third factor (SiR-HP<sup>0</sup>) appears (Figure 3). Combining the forward and reverse EFA, SiR-HP<sup>0</sup> was present between  $-0.40$  and  $-0.70$  V (spectra 1–31), SiR-HP<sup>1-</sup> was present between  $-0.55$  V and  $-0.80$  V (spectra 16–41), and SiR-HP<sup>2-</sup> between  $-0.64$  and  $-0.80$  V (spectra 27–41).

Each factor can now be transformed independently using eq 10. From the above analysis, the potential region from  $-0.7$  to  $-0.8$  V (spectra 32–41) contains only SiR-HP<sup>1-</sup> and SiR-HP<sup>2-</sup>. After the  $\mathbf{R}_{\text{abst}}$  matrix was reduced to three columns (for the three factors), the  $\mathbf{R}^0$  matrix for SiR-HP<sup>0</sup> was constructed using rows 32–41 of  $\mathbf{R}_{\text{abst}}$  (see eq 13). The transformation vector of SiR-HP<sup>0</sup> ( $t_{\lambda}$ ) was then calculated (eq 10). A similar procedure can be repeated for the third species (SiR-HP<sup>2-</sup> (C), spectra 1–26). The  $\mathbf{R}^0$  matrix for C was constructed using the first 26 rows of  $\mathbf{R}_{\text{abst}}$ . The intermediate species, SiR-HP<sup>1-</sup>, is not present in spectra 1–15 ( $-0.40$  to  $-0.55$  V). Using these spectra to generate the  $\mathbf{R}^0$  for SiR-HP<sup>1-</sup>, the concentration of SiR-HP<sup>1-</sup> does not decrease at more negative potentials as expected when SiR-HP<sup>2-</sup> appears. This result is not physically meaningful in that, when the concentrations are determined using eq 19, a negative concentration for SiR-HP<sup>1-</sup> is required. An examination of the reverse EFA for factor 2 in Figure 3, though, indicates that this factor begins near the noise level, suggesting that SiR-HP<sup>1-</sup> has nearly disappeared at  $-0.8$  V. In addition, a comparison of the spectrum at  $-0.80$  V shows that it is very close to SiR-HP<sup>2-</sup>. If we add row 41 ( $E = -0.80$  V) of  $\mathbf{R}_{\text{abst}}$  to  $\mathbf{R}^0$  for the SiR-HP<sup>1-</sup>, we obtain a new curve for SiR-HP<sup>1-</sup> where the concentration decreases to zero at the end of the scan. Equations 17–19 were then used to calculate the actual concentrations (normalization factors were 0.381 and 0.365), which are plotted in Figure 4.

Is the assumption that led to the inclusion of spectrum 41 correct? The first indication of its validity is the shape of the

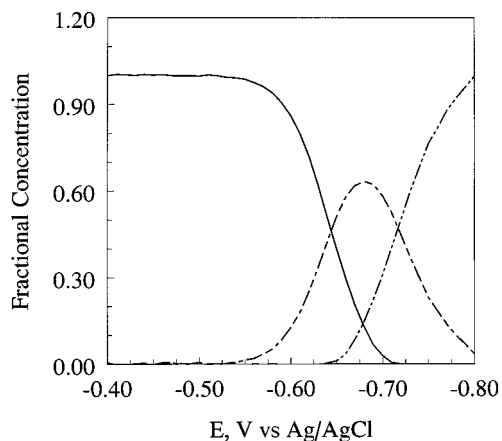


Figure 4. Concentration profiles for SiR-HP<sup>0</sup> (—), SiR-HP<sup>1-</sup> (---), and SiR-HP<sup>2-</sup> (- · - ·).

arbitrary concentration versus potential curve. When rows with significant concentration of species are included in  $\mathbf{R}^0$ , the concentration profile will oscillate around zero with significant positive and negative values. The “normal” concentration profile for SiR-HP<sup>1-</sup> indicates that [SiR-HP<sup>1-</sup>] is close to zero at -0.8 V. It is possible to obtain a better value for the [SiR-HP<sup>1-</sup>] by an iterative process. First, we estimate the arbitrary concentration of SiR-HP<sup>1-</sup> at -0.80 V, the value is put into eq 14, and the concentration profile and spectra are solved. From the three spectra and concentrations, all the intermediate spectra can be generated and compared to the experimental spectra. Using a series of values (between 0 and 10% remaining) for the concentration of SiR-HP<sup>1-</sup> at -0.80 V, the minimum error was obtained at 3.5%, which was used in this work. In any case, the overall concentration profile of SiR-HP<sup>1-</sup> does not depend significantly on the concentration at -0.80 V.

The actual concentrations were then used to calculate the  $\mathbf{T}$  matrix (eq 7), and the spectra for each species were obtained using eq 8. The spectra are shown in Figure 5, along with the literature spectra, which were obtained by photochemical reduction.<sup>25</sup> The spectrum of SiR-HP<sup>0</sup> agrees quite well with the literature spectrum, indicating that the starting material was spectroscopically the same as the one reported. The spectrum of SiR-HP<sup>2-</sup> also agrees reasonably well, especially with regard to the  $\lambda_{\max}$  in the Soret band (398 nm). The broad absorbance between 550 and 650 nm was observed in both our spectrum and the literature spectrum, though the peak absorbance appeared at shorter wavelengths in our work. The broadness of the bands and the noise in the thin-layer cell make these differences less significant. There are significant differences in the SiR-HP<sup>1-</sup> obtained in this work versus the literature spectrum. In particular, we observe a sharper Soret band at 396 nm versus a much broader Soret band with 389 nm with a shoulder at about 375 nm. The longer wavelength band agrees well with the literature with a  $\lambda_{\max}$  at 585 nm. It is important to note that previous workers were no more successful than we were in obtaining a pure SiR-HP<sup>1-</sup> spectrum because of the thermodynamic constraints. In both cases, a deconvolution process was necessary. Our deconvolution process was able to use all the spectra during the entire reduction, while their work considered only those spectra where only two

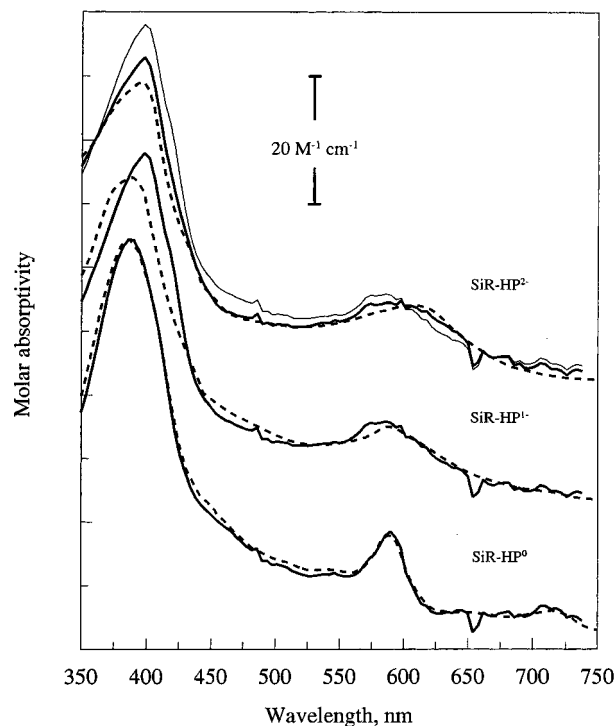
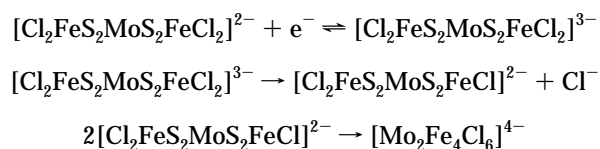


Figure 5. Spectra of SiR-HP<sup>0</sup>, SiR-HP<sup>1-</sup>, and SiR-HP<sup>2-</sup> as calculated by EFA: bold solid lines, spectra calculated by EFA; dashed lines, spectra obtained from ref 25; light solid line, SiR-HP<sup>1-</sup> spectrum.

species were present at any time. The lack of systematic error as well as the small values for the residual errors (less than  $5 \times 10^{-4}$  absorbance unit) indicates that the spectra are effectively deconvoluted.

**Spectroelectrochemistry of (Ph<sub>4</sub>P)<sub>2</sub>(Cl<sub>2</sub>FeS<sub>2</sub>MoS<sub>2</sub>FeCl<sub>2</sub>), (I).** The electrochemistry and spectroelectrochemistry of **I** have been recently reported.<sup>26</sup> The Mo/Fe/S cluster is reduced reversibly at -1.00 V versus Ag/AgNO<sub>3</sub> to a trianion, which loses Cl<sup>-</sup> and dimerizes to a suspected cubane structure.



If the potential is stepped from -0.40 to -1.05 V, reduction of the complex occurs. The spectral changes during the first 24 s are shown in Figure 4 of ref 26. Figure 6 shows the spectral changes over the entire experiment. A total of 61 spectra was obtained. Principal factor analysis of the data is summarized in Table 2. At least four to six species are present based on the RE and IND functions. Forward and reverse EFA are shown in Figure 7.

An analysis of Figure 7 shows that the first four eigenvalues are present early in the experiment (mathematically, it takes at least four spectra or 15 s to determine the first  $\lambda_4$ ). The value of  $\lambda_4$ , though, is close to the background noise at 20 s, indicating that it was probably just beginning to appear at this time. The

(26) Liu, Y. M.; Chen, J. H.; Ryan, M. D. *Inorg. Chem.* **1998**, *37*, 425–431.

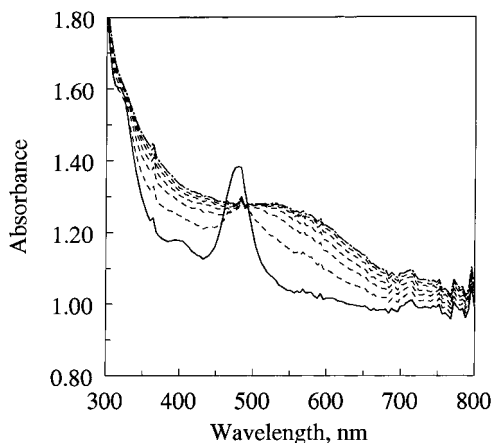


Figure 6. Thin-layer chronoabsorptometry of  $[\text{Cl}_2\text{FeS}_2\text{MoS}_2\text{FeCl}_2]^{2-}$  in methylene chloride and 0.10 M TBAP. Solid line is the initial spectrum; dashed lines are intermediate spectra at 10, 25, 50, 100, and 175 s; dash-dot line is the final spectrum at 300 s.

Table 2. Principal Factor Analysis of Mo/Fe/S Spectroelectrochemistry

principal factor	eigenvalue	RE $\times 10^4$	IND $\times 10^7$
1	1200	231.	64.2
2	3.48	86.0	24.7
3	0.466	8.04	10.1
4	0.0797	4.25	2.47
5	0.00337	3.76	1.35
6	$2.92 \times 10^{-4}$	3.65	1.24
7	$7.62 \times 10^{-5}$	3.55	1.25
8	$6.55 \times 10^{-5}$	3.47	1.26
9	$4.87 \times 10^{-5}$	3.41	1.28
10	$4.58 \times 10^{-5}$	3.34	1.31

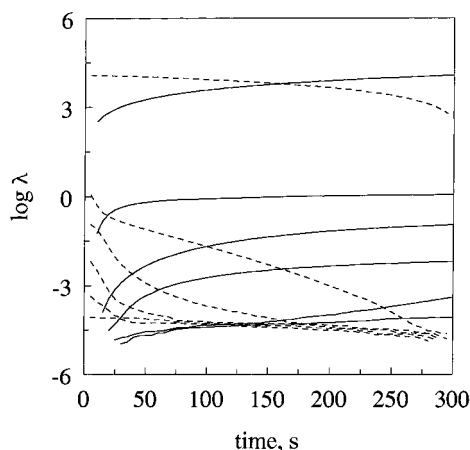


Figure 7. Forward and reverse EFA eigenvalues for the spectroelectrochemistry of  $[\text{Cl}_2\text{FeS}_2\text{MoS}_2\text{FeCl}_2]^{2-}$ . Forward EFA are solid lines; reverse EFA are dashed lines.

reverse eigenvalues were much more informative. Only one eigenvalue was observed above the noise level between 275 and 300 s. This indicates that the final product had been formed by that time. The second eigenvalue deviates from the noise level at 275 s, while the third appears at 150 s. The fourth eigenvalue was significant only below 75 s. From these results, four species were observed within the following time windows: species 1, 0–75 s; species 2, 0–150 s; species 3, 0–275 s and species 4, 20–300 s. We can make one more refinement. Only the starting species is

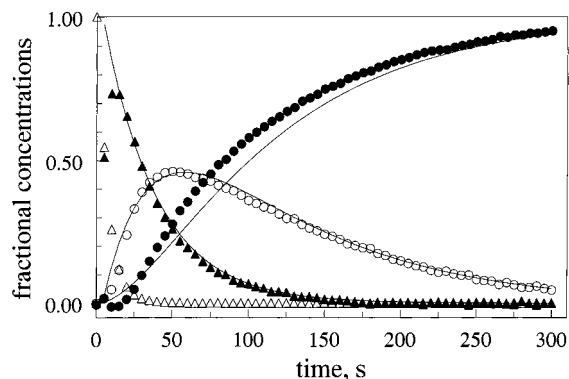
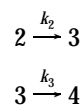


Figure 8. Fractional concentration profiles for the spectroelectrochemistry of  $[\text{Cl}_2\text{FeS}_2\text{MoS}_2\text{FeCl}_2]^{2-}$  generated by factor analysis: species 1, open triangles; species 2, filled triangles; species 3, open circles; Species 4, filled circles. Solid lines are the best fit to the kinetic expressions given by eqs 20–22.

present at 0 s, and this species is most likely species 1 (we can confirm this once the spectra have been obtained). Therefore, species 2 and 3 are not present at  $t = 0$  s. The final windows where each species is present are therefore the following: 0–75, 5–150, 5–275, and 20–300 s.

The regions of zero concentrations for each species can be readily determined from the preceding time windows. The value of each transformation vector,  $t$ , was determined using eq 10. The relative concentration matrix,  $\mathbf{m}$ , was determined using eq 16. The actual concentrations were then calculated using eqs 17–19. The results are shown in Figure 8.

As will be verified by the spectra below, species 1 is the starting material,  $\mathbf{I}$ . This material was electrolyzed rapidly and was completely reduced at about 30 s, the time constant for the electrolysis. Species 2, 3, and 4 appeared in sequential order, with only species 4 remaining after 300 s. One could then postulate the following sequential reactions:



The concentration profiles for species 2, 3, and 4 can be easily calculated for this sequential mechanism:

$$[2] = [2]_i e^{-k_2 t} \quad (20)$$

$$[3] = (k_2 [2]_i / (k_3 - k_2)) (e^{-k_2 t} - e^{-k_3 t}) \quad (21)$$

$$[4] = ([2]_i / (k_3 - k_2)) (k_2 e^{-k_3 t} - k_3 e^{-k_2 t}) + [2]_i \quad (22)$$

where we have ignored the initial rise of species 2 as the starting material is reduced. The solid lines in Figure 8 are the best-fit lines using the kinetic equations above and  $k_2 = 0.029 \text{ s}^{-1}$  and  $k_3 = 0.011 \text{ s}^{-1}$ . The excellent fit of the concentration profiles to the single or double decays, and the fact that consistent rate constants were obtained for the three profiles, indicate that the concentration profiles obtained by factor analysis are quite reasonable.

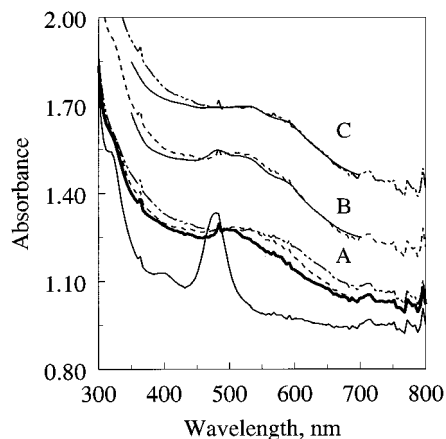
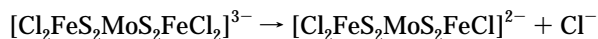


Figure 9. (A) Visible spectra of Mo/Fe/S species obtained by factor analysis: solid line,  $[\text{Cl}_2\text{FeS}_2\text{MoS}_2\text{FeCl}_2]^{2-}$  species 1; bold line, species 2; dashed line, species 3 ( $[\text{Mo}_2\text{S}_8\text{Fe}_4\text{Cl}_6]^{4-}$ ); dash-dot line, species 4. (B) Visible spectra of species 3 and previously reported spectra obtained spectroelectrochemically at  $-1.05$  V. (C) Visible spectra of species 4 and spectra obtained spectroelectrochemically at  $-1.12$  V. Solvent, methylene chloride;  $0.10$  M TBAP.

An analysis of the rate constants indicates that both reactions are much slower than the dissociation of  $\text{Cl}^-$ , as previously measured by cyclic voltammetry.<sup>27</sup>



The rate constant for this reaction is  $0.5 \text{ s}^{-1}$ , which gives a  $t_{1/2}$  of about  $1.2$  s. Little of the initial reduction product was seen because the data were taken in  $5$ -s intervals. On the other hand, a fifth factor was observed in the reverse EFA, which was only seen in the first several spectra. This species may be the trianion, but attempts to obtain a concentration profile and spectrum by EFA were unsuccessful probably because of its low concentration and the fact that it was present in only a few spectra.

The spectra for the first four principal factors were obtained using eq 8, and the results are shown in Figure 9, curves A. The spectra for species 1 were indeed the starting material. The latter three spectra were similar. Previous work<sup>26</sup> has shown that the longer wavelength region increased in absorbance as the complex is subjected to further reduction. Curves B of Figure 9 compare the spectrum of species 3 with a previously reported spectroelectrochemical spectrum obtained at  $-1.05$  V.<sup>26</sup> The spectrum of species 4 was similar to species 3, but the species absorbed more strongly in the longer wavelength region. Curves C of Figure 9 compare the spectrum of species 4 with a spectrum obtained from the Mo/Fe/S cluster at  $-1.12$  V. This is the potential of the second (irreversible) wave that was seen for **I** at slow scan rates and is due to the reduction of the product formed at the first wave. This spectrum was also the same as that observed in the borohydride reduction of **I**, which has been shown to be reduced by more than one electron.<sup>26</sup> Therefore, further reduction will occur if sufficient time is allowed, and the formation of species 4 is

probably due to the slow electrolysis of species 3 rather than a homogeneous reaction. It is not clear which species is responsible for species 2, but a likely candidate based on the reaction mechanism is  $[\text{Cl}_2\text{FeS}_2\text{MoS}_2\text{FeCl}]^{2-}$ , which is the initial complex formed by the loss of chloride. Further work is being continued in our laboratory to obtain vibrational spectra of these intermediates in order to verify this mechanism.

## CONCLUSIONS

This work shows that factor analysis is a powerful tool for the rapid analysis of spectroelectrochemical data especially where there is substantial overlap of the various redox species. The first application utilized a linear scan method while the second involved a potential step technique. The most challenging aspect of the method is the determination of the transformation matrix, **T**. For the EFA method to be successful, each of the species must be absent in some of the spectra. As a result, this approach will fail for the case of semi-infinite diffusion where the starting complex is colored (not transparent at the wavelength range chosen). In the preceding case, other methods of factor analysis must then be used to determine the transformation matrix. The deconvolution may also be difficult if the concentrations of the two species present are disproportionately far apart. As a result, even though there is evidence for the trianion of the Mo/Fe/S cluster (factor 5), its spectrum could not be determined because it is present at low concentrations in the same window as the starting material. Hence, it was impossible for the method to separate these two factors.

There are other situations where this problem may occur. In general, if a reaction sequence forms two stable, but spectrally distinct species in a fixed stoichiometric ratio, it will be impossible to separate the species because they will appear in the same time window. The spectrum that will be obtained will be a composite of the two species, and only one chemical species will be predicted by factor analysis (if the two species appeared and disappeared at the same rate). Finally, for those species that appear and then disappear during the experiment, it is necessary to use data from before and after their appearance in order to accurately transform the data. It is easiest to generate the transformation matrix if the experiment is continued until only one species is present at the end of the experiment. Otherwise, the separation of the remaining components may be difficult, but not insurmountable.

The large quantity of data that are utilized in this methodology helps to ensure the validity of the results. Significant errors in choosing which spectra to use in calculating the transformation matrix can generally become apparent through inconsistencies in the result, such as negative concentrations or molar absorptivities. In addition, the shape of the concentration profile can be used to check the consistency of the results, as shown in this work. Because the deconvolution is based on a model-free methodology, consistency between a particular model and the experimental data is not required in the deconvolution process and is instead an additional check on the work.

Finally, considerable care must be taken to distinguish between artifacts and actual species. Systematic variations in the data may

(27) Coucouvanis, D.; Simhon, E. D.; Stremple, P.; Ryan, M.; Swenson, D.; Baenziger, N. C.; Simopoulos, A.; Papaefthymiou, V.; Kostikas, A.; Petrouleas, V. *Inorg. Chem.* **1984**, *23*, 741–749.

appear as additional factors. For example, other data taken in our laboratory (not reported here) have shown that the appearance of a strong band that deviates from Beer's law will give rise to an additional factor. In fact, any solution changes that affect the spectrum such as solvation<sup>28</sup> and ionic strength will give rise to additional factors. Therefore, even though EFA is a model-free method, it is important to eventually assign physical significance to the observed factors in order to eliminate the possibility that the observed factor may be an artifact.

---

(28) Edward, J. T.; Wong, S. C. *J. Am. Chem. Soc.* **1977**, *99*, 4229.

#### ACKNOWLEDGMENT

The authors acknowledge support from the NIH-AREA Grant R15 GM/OD55917. The authors also thank N. Kredich for the pBR322 plasmid containing SiR.

Received for review September 29, 1998. Accepted February 18, 1999.

AC981079H