ACADEMIC PRESS

Available online at www.sciencedirect.com
SCIENCE DIRECT®

Journal of Catalysis 213 (2003) 126–134

JOURNAL OF CATALYSIS

www.elsevier.com/locate/jcat

Spectral reconstruction of in situ FTIR spectroscopic reaction data using band-target entropy minimization (BTEM): application to the homogeneous rhodium catalyzed hydroformylation of 3,3-dimethylbut-1-ene using $Rh_4(CO)_{12}$

Chuanzhao Li, Effendi Widjaja, and Marc Garland *

Department of Chemical and Environmental Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 119260

Received 5 March 2002; revised 3 August 2002; accepted 23 August 2002

Abstract

The homogeneous catalytic hydroformylation of 3,3-dimethylbut-1-ene was studied, starting with $Rh_4(\sigma\text{-}CO)_9(\mu\text{-}CO)_3$ as catalyst. The multiple experiment in situ spectroscopic data were preconditioned to subtract the absorbance due to background moisture, carbon dioxide, and solvent. The *preconditioned* data were then subjected to band-target entropy minimization (BTEM) in order to recover the pure component spectra of the species present—*using no libraries and no a priori information*. The pure component spectra of the main species present, namely, the organic reactant 3,3-dimethylbut-1-ene, the organic product 4,4-dimethylpentanal, the catalyst precursor $Rh_4(\sigma\text{-}CO)_9(\mu\text{-}CO)_3$, and the observable organometallic intermediate $RCORh(CO)_4$, were all readily recovered. In addition, it was possible to recover the expected minor species, namely 2-methyl-3,3-dimethylbutanal, $Rh_6(CO)_{16}$, and a recently identified cluster $Rh_4(\sigma\text{-}CO)_{12}$. The latter two species exist at ppm levels. The new BTEM algorithm shows that successful and detailed in situ spectroscopic system identification for catalytic studies is possible given no prior information.

Keywords: Homogeneous catalysis; Spectral reconstruction; BTEM; Entropy minimization; Un-modified rhodium; Hydroformylation; In situ FTIR

1. Introduction

The hydroformylation of alkenes using soluble cobalt and rhodium complexes is one of the largest homogeneously catalyzed reactions worldwide [1]. Numerous detailed in situ spectroscopic measurements and kinetic analyses of the unmodified rhodium catalyzed hydroformylation of alkenes have been reported [2–7]. These studies have clearly shown that homogeneous catalytic research, driven by copious in situ spectroscopic data, can have an enormous impact on chemical, kinetic, and mechanistic insight. Ultimately, however, such insight relies on our ability to interpret very complex, multicomponent, in situ spectroscopic reaction data.

Recently, we introduced an advanced algorithm for pure component spectral reconstruction from reaction mixtures. This new algorithm, called band-target entropy minimization (BTEM), has now been successfully applied to FTIR spectroscopic data of nonreactive mixtures, organometallic reactions, and homogeneous catalytic reactions [8,9]. The foundations of this algorithm lie in the concept of the information content or entropy of a signal [10] and as such the BTEM algorithm goes far beyond normal deconvolution. The search for the minimum entropy constituents of a signal (spectrum) forms the basis for reconstructing the pure component spectra. Entropy minimization is closely associated with pattern recognition [11]. A few forerunners of pure component spectral reconstruction, far less robust than the present BTEM algorithm, exist [12–16].

The utility of the BTEM algorithm in catalysis arises from its rather straightforward application by the experimentalist. For example, consider an experiment where solvent and reagents are sequentially added to a batch reactor, and in situ spectroscopic measurements are performed. Let n be the number of preliminary spectroscopic measurements, k be the number of spectra measured during the catalysis,

* Corresponding author.
E-mail address: chemvg@nus.edu.sg (M. Garland).

0021-9517/02/$ – see front matter
PII: S0021-9517(02)00009-X

and v be the number of channels of spectroscopic data. Then for e experiments, an absorbance matrix $A^{\exp}$ is measured.

The experimental absorbance matrix can be preconditioned to subtract background and solvent. The preconditioned data A^{pre} can be decomposed (vector-space decomposition) and a disentangled matrix of right singular vectors V^{T} obtained. The final crucial step is the reassembly of the disentangled vectors into the pure component spectra using minimum entropy criteria. The term band-target arises from the fact that the experimentalist can visually identify and subsequently target disentangled spectral features in the vector-space decomposition V^{T}. The identified feature is retained and the entire associated whole function or pure component spectrum is generated. *Thus no a priori information is used in the BTEM algorithm*:

$$A^{\exp} \rightarrow A^{\text{pre}} \rightarrow V^{\text{T}} \rightarrow a_{s\times v}. \tag{1}$$

In the following, we examine multiple experiment in situ spectroscopic FTIR data from the unmodified rhodium-catalyzed hydroformylation of 3,3-dimethylbut-1-ene (33DMB) to 4,4-dimethylpentanal (44DMP) using the BTEM algorithm. The reasons are twofold. First, this reaction is one of the most studied homogeneous catalyzed reactions, and second, we wish to see if new and previously unrecoverable information can be retrieved.

2. Experimental

2.1. General information

All solution preparations were carried out under argon (99.9995%, Soxal, Singapore) using standard Schlenk techniques [17]. The argon was further purified prior to use by passage through deoxy and zeolite columns. All reactions were carried out under carbon monoxide (99.97%, Soxal, Singapore) and hydrogen (99.999%, Soxal, Singapore) after further purification through deoxy and zeolite columns.

The precious metal complex $Rh_4(CO)_{12}$, with stated purity of 98% min, was obtained from Strem Chemicals (Newport, MA) and was used without further purification, although trace quantities of the high-nuclearity cluster $Rh_6(CO)_{16}$ are virtually always present. The n-hexane solvent (stated purity $>$ 99.6%, Fluka AG) was refluxed over sodium potassium alloy under argon. 3,3-Dimethylbut-1-ene (99%, Fluka AG, Switzerland) was used as obtained.

2.2. Equipment

In situ spectroscopic studies were performed in a 1.5-L stainless steel (SS316) autoclave (Büchi-Uster, Switzerland) which was connected to a high-pressure infrared cell. The autoclave ($P_{\max} = 22.5$ MPa) was equipped with a packed magnetic stirrer with six-bladed turbines in both the gas and liquid phases (Autoclave Engineers, Erie, PA) and was constructed with a heating/cooling mantle. A high-pressure membrane pump (Model DMK 30, Orlita AG, Geissen Germany) with a maximum rating of 32.5-MPa and a 3-L/h flow rate was used to circulate the n-hexane solutions from the autoclave to the high-pressure IR cell and back to the autoclave via jacketed 1/8-in. (SS316) high-pressure tubing (Autoclave Engineers). The entire system, autoclave, transfer lines, and infrared cell, was cooled using a Polyscience Model 9505 cryostat and could be kept isothermal ($\Delta T \leqslant 0.5$ °C) at 298–313 °C. Temperature measurements were made at the cryostat, autoclave, and IR cell with PT-100 thermoresistors. The necessary connections to vacuum and gases were made with 1/4-in (SS316) high-pressure tubing (Autoclave Engineers) and 1.0, 5.0, and 10.0 piezocrystals were used for pressure measurements (Keller AG, Winterthur, Switzerland). The entire system was gas-tight under vacuum as well as at 20.0 MPa, the maximum operating pressure.

The high-pressure infrared cell was constructed at the ETH Zürich of SS316 steel and could be heated and cooled. The CaF_2 single crystal windows (Korth Monokristalle, Kiel, Germany) had dimensions of diameter 40 mm by thickness 15 mm. Two sets of Viton and silicone gaskets provided sealing, and Teflon spacers were used between the windows. The construction of the flow-through cell [18] is a variation on a design due to Noack [19] and differs in some respects from other high-pressure infrared cells described in the literature (for a review, see Whyman [20]). The high-pressure cell was situated in a Perkin–Elmer 2000 FTIR infrared spectrometer. The cell chamber was purged with purified nitrogen (99.999%, Soxal, Singapore). The resolution was set to 4 cm^{-1} for all spectroscopic measurements. A schematic diagram of the experimental setup can be found in Ref. [6].

2.3. In situ spectroscopic and kinetic studies

All the experiments were performed in a similar manner. First, single-beam background spectra of the IR sample chamber were recorded. Then 200 ml n-hexane was transferred to the autoclave under argon. Under 0.2-MPa CO pressure, infrared spectra of the n-hexane in the high-pressure cell were recorded. The total system pressure was raised to 2.0-MPa CO, and the stirrer and high-pressure membrane pump were started. After equilibration, infrared spectra of the CO/n-hexane solution in the high-pressure cell were recorded. A solution of 5 ml 3,3-dimethylbut-1-ene (33DMB) dissolved in 50 ml n-hexane was prepared, transferred to the high-pressure reservoir under argon, pressured with CO, and then added to the autoclave. After equilibration, infrared spectra of the 33DMB/CO/n-hexane solution in the high-pressure cell were recorded. A solution of 50 mg $Rh_4(CO)_{12}$ dissolved in 50 ml n-hexane was prepared, transferred to the high-pressure reservoir under argon, pressured with CO, and then added to the autoclave. After equilibration, infrared spectra of the $Rh_4(CO)_{12}$/33DMB/CO/n-

hexane/solution in the high-pressure cell were recorded. Then 1.0 MPa hydrogen was added to initiate the syntheses. Spectra were recorded at 10-min intervals in the range 1000–2500 cm^{-1}. In every 6-h experiment, 30–36 spectra were taken.

A total of six experiments were performed in this study. Five experiments were conducted under the same conditions, namely 5 ml 33DMB, 50 mg $Rh_4(CO)_{12}$, 2.0 MPa CO, and 1.0 MPa H_2 in 300 ml *n*-hexane. The other one was conducted under different conditions, namely 5 ml 33DMB, 50 mg $Rh_4(CO)_{12}$, 1.0 MPa CO, and 1.0 MPa H_2 in 300 ml *n*-hexane.

The postreaction mixture was first irradiated with an ultraviolet lamp (ACE 7825 UV, 450 W) to destroy the metal complexes and then analyzed by GC (HP6890; HP-1 methyl siloxane capillary column, 150 °C; FID, 250 °C) to determine the organic compositions.

3. Computational aspects

3.1. Computation

All algorithms were implemented in MATLAB. Calculations were performed on a WinNT Pentium III Xeon Work Station having two 450-MHz processors with 2-GB RAM.

3.2. Spectral preconditioning

Each experimental solution spectrum $\mathbf{A}^{\text{exp}}_{1\times\nu}$ was preconditioned by subtraction of (i) the background reference, $\mathbf{A}^{\text{ref}}_{\text{bac}}$, (ii) the CaF_2 cell with *n*-hexane reference, $\mathbf{A}^{\text{ref}}_{\text{cell+hex}}$, and (iii) the dissolved CO reference, $\mathbf{A}^{\text{ref}}_{\text{CO}}$, according to the following equation, where the coefficients x_i are scalar subtraction factors [21]:

$$\mathbf{A}^{\text{pre}}_{1\times\nu} = \mathbf{A}^{\text{exp}}_{1\times\nu} - \sum_i x_i \mathbf{A}^{\text{ref}}_i. \tag{2}$$

Entropy minimization was the basis for determining the optimal values of x_i. The discrete form of the Shannon entropy function for any arbitrary spectrum forms the objective function, and the golden section search was used for optimization of x_i values [22]. The entropy function was based on the first derivative of the resultant subtracted spectrum $\mathbf{A}^{\text{subtract}}_{1\times\nu}$. The final spectrum, after optimal subtraction of $\mathbf{A}^{\text{ref}}_{\text{bac}}$, $\mathbf{A}^{\text{ref}}_{\text{cell+hex}}$, and $\mathbf{A}^{\text{ref}}_{\text{CO}}$ is the desired preconditioned spectrum $\mathbf{A}^{\text{pre}}_{1\times\nu}$. The algorithm is robust and stable, and produces smooth preconditioned spectrum with extremely good baselines:

$$\min g = -\sum_\nu h_\nu \ln h_\nu, \tag{3}$$

where

$$h_\nu = \left|\frac{\mathrm{d}A^{\text{subtract}}_\nu}{\mathrm{d}\nu}\right|. \tag{4}$$

3.3. Spectral analysis

The consolidated matrix of preconditioned reaction absorbance data is denoted $\mathbf{A}_{ke\times\nu}$, where k denotes the number of spectra in one experiment, e denotes the number of experiments, and ν is the number of data channels associated with the experimental FTIR wavenumber range and interval taken.

According to the Lambert–Beer–Bougeur law, the absorbance matrix is a linear product of a concentration matrix $\mathbf{C}_{ke\times s}$ (which incorporates the path length l) and an absorptivity matrix $\mathbf{a}_{s\times\nu}$. This absorbance data matrix $\mathbf{A}_{ke\times\nu}$ can be subjected to singular value decomposition (SVD) [23], to obtain its abstract orthonormal matrices $\mathbf{U}_{ke\times ke}$ and $\mathbf{V}^{\mathrm{T}}_{\nu\times\nu}$ with its diagonal singular matrix $\boldsymbol{\Sigma}_{ke\times\nu}$.

Furthermore, $\mathbf{A}_{ke\times\nu}$ can be

$$\mathbf{A}_{ke\times\nu} = \mathbf{C}_{ke\times s}\mathbf{a}_{s\times\nu} + \varepsilon_{ke\times\nu} = \mathbf{U}_{ke\times ke}\boldsymbol{\Sigma}_{ke\times\nu}\mathbf{V}^{\mathrm{T}}_{\nu\times\nu}, \tag{5}$$

approximated by

$$\mathbf{A}_{ke\times\nu} \approx \widehat{\mathbf{C}}_{ke\times s}\hat{\mathbf{a}}_{s\times\nu} = \mathbf{U}_{ke\times s}\boldsymbol{\Sigma}_{s\times z}\mathbf{T}^{-1}_{s\times z}\mathbf{T}_{s\times z}\mathbf{V}^{\mathrm{T}}_{z\times\nu}, \quad ke \geqslant z \geqslant s, \tag{6}$$

where s is the number of species recovered and j is the number of right singular vectors used for spectral reconstruction. Note that $\mathbf{T}^{-1}_{s\times z}$ is the generalized inverse for a rectangular matrix, $\hat{\mathbf{a}}_{s\times\nu}$ is the matrix of averaged pure component expectations for s species, $\widehat{\mathbf{C}}_{ke\times s}$ is its corresponding expectation for concentration calculated from

$$\widehat{\mathbf{C}}_{ke\times s} = \mathbf{U}_{ke\times s}\boldsymbol{\Sigma}_{s\times z}\mathbf{T}^{-1}_{s\times z} = \mathbf{A}_{ke\times\nu}\hat{\mathbf{a}}^{\mathrm{T}}_{s\times\nu}\left(\hat{\mathbf{a}}_{s\times\nu}\hat{\mathbf{a}}^{\mathrm{T}}_{s\times\nu}\right)^{-1}, \tag{7}$$

and $\varepsilon_{k\times\nu}$ is a combination of experimental error and spectral nonlinearities [24].

Instead of solving all rows of the transformation matrix $\mathbf{T}_{s\times z}$ at once, the BTEM algorithm solves the problem one transformation row at a time [8,9], and hence one spectrum at a time is resolved. The expectation for each spectrum $\hat{\mathbf{a}}_{1\times\nu}$ is then given by

$$\hat{\mathbf{a}}_{1\times\nu} = \mathbf{T}_{1\times z}\mathbf{V}^{\mathrm{T}}_{z\times\nu}, \quad z \geqslant s, \tag{8}$$

with the corresponding expectation for concentration $\widehat{\mathbf{C}}_{ke\times 1}$ given by

$$\widehat{\mathbf{C}}_{ke\times 1} = \mathbf{A}_{ke\times\nu}\hat{\mathbf{a}}^{\mathrm{T}}_{\nu\times 1}\left(\hat{\mathbf{a}}_{1\times\nu}\hat{\mathbf{a}}^{\mathrm{T}}_{1\times\nu}\right)^{-1}. \tag{9}$$

3.4. Band-target entropy minimization (BTEM) algorithm

In this method, the vectors in the $\mathbf{V}^{\mathrm{T}}$ matrix were first inspected for significant spectral features. Such features usually appear only in the first few vectors of $\mathbf{V}^{\mathrm{T}}$, as they represent most of the variance in the data matrix $\mathbf{A}_{ke\times\nu}$. Since BTEM targets these features one at a time, through the use of Eq. (8), narrow intervals corresponding to these local extrema are assigned.

For example, the characteristic bridging carbonyl band for $Rh_4(CO)_{12}$ occurs at circa 1886 cm^{-1}. Therefore, upon visual inspection, clearly identifiable spectral features/extrema can be seen at circa 1886 cm^{-1} in the initial five $\mathbf{V}^T$ vectors. Accordingly, one sets the BTEM algorithm to target the interval 1885–1887 cm^{-1}. The algorithm then reconstructs the simplest function associated with this spectral feature.

The reason for taking a region of wavenumbers for the targeted band rather than an exact singular band peak wavenumber is that *nonlinearities* due to band-shifting and shape changes exist in real spectra. In the process of transforming the right singular vectors, the targeted band peak absorbance $\hat{\mathbf{a}}_{1\times\nu}^{\text{target}}$ is normalized to 1.0. It is worth reiterating that no *a priori* information as such is required by this BTEM approach, only visual inspection of the right singular vectors.

3.5. Corana's simulated annealing

The Corana simulated annealing algorithm was employed to perform the global optimizations of the highly nonlinear BTEM objective function. Corana's simulated annealing is a random optimization method, modified from the original simulated annealing algorithm [25], with dynamic stepsize generation and strict convergence criterion enhancements [26]. This algorithm has been extensively used and has proven capable of obtaining global solutions for highly nonlinear optimization problems [27–29]. A key feature of Corona's algorithm is its ability to escape local minima. This feature is based on the Metropolis selection criteria.

4. Results

4.1. Preconditioned spectra

Spectra of the hydroformylation reaction conducted with $Rh_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$ and 33DMB for six experiments were preconditioned as described in the computational section in order to subtract the background, CaF_2 cell with *n*-hexane and dissolved CO spectra. A sample of the optimal subtraction result is shown in Fig. 1, where background, the CaF_2 cell with *n*-hexane, the dissolved CO spectrum, a reaction spectrum, and a preconditioned spectrum are presented. After preconditioning, all reaction spectra from the six sets of experiments were collected into one preconditioned absorbance data matrix $\mathbf{A}_{188\times4751}$, with 188 reaction spectra (rows) and 4751 channels of data (columns). As an example, the preconditioned reaction spectra from one of the experimental runs are shown in Fig. 2 (15 spectra are taken). It is immediately seen that in the wavenumber range 2140–2500 cm^{-1}, there is only the baseline spectrum without any other spectral contribution from organic or organometallic components. Therefore, it is possible to delete the channels of data correlated to this wavenumber range. After subtraction, the preconditioned absorbance data matrix is reduced to $\mathbf{A}_{188\times2951}$. For subsequent discussion, it is useful to note the exceptionally small magnitude of the organometallic signals in Figs. 1 and 2.

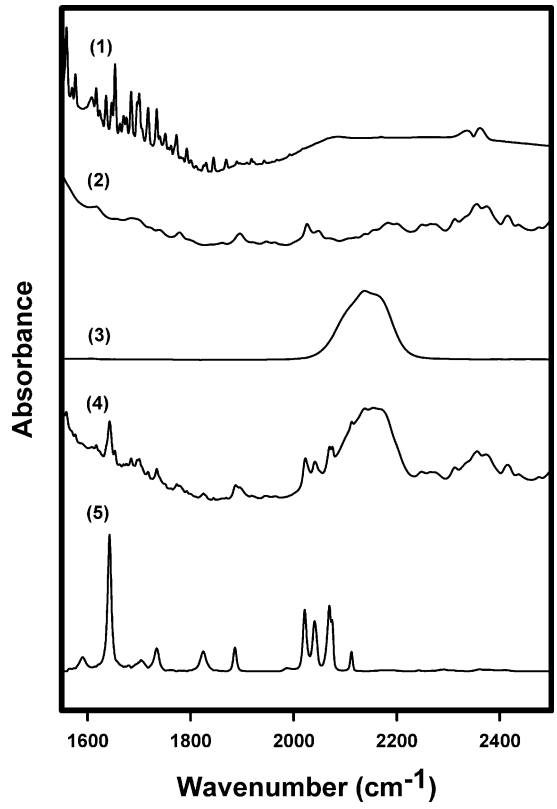

Fig. 1. Preconditioning of 33DMB/$Rh_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$ reaction spectra: (1) background spectrum; (2) CaF_2 cell with *n*-hexane spectrum; (3) dissolved CO spectrum; (4) reaction spectrum; (5) preconditioned spectrum.

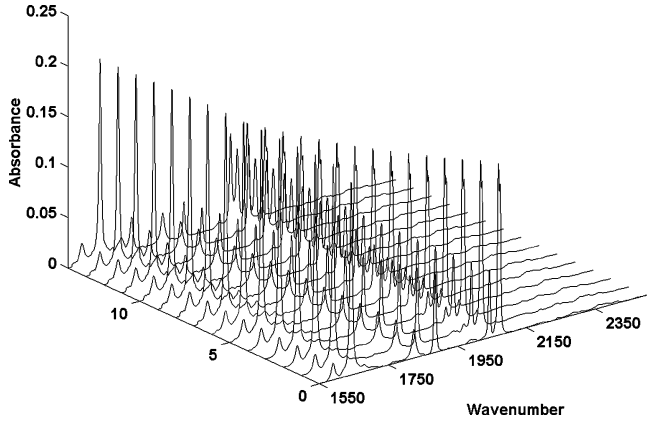

Fig. 2. Consolidated preconditioned absorbance data matrix taken from the first 15 reaction spectra of the 5th experiment.

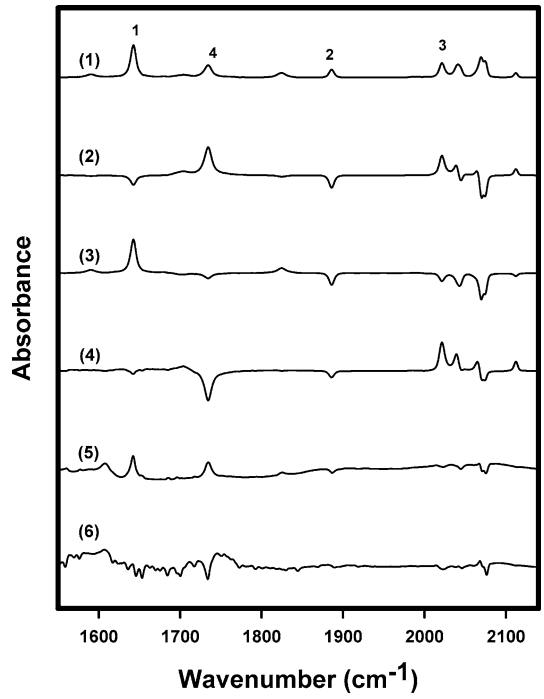

Fig. 3. First six right singular vectors of the $\mathbf{V}^{\mathrm{T}}$ matrix: (1) 1st vector; (2) 2nd vector; (3) 3rd vector; (4) 4th vector; (5) 5th vector; (6) 6th vector.

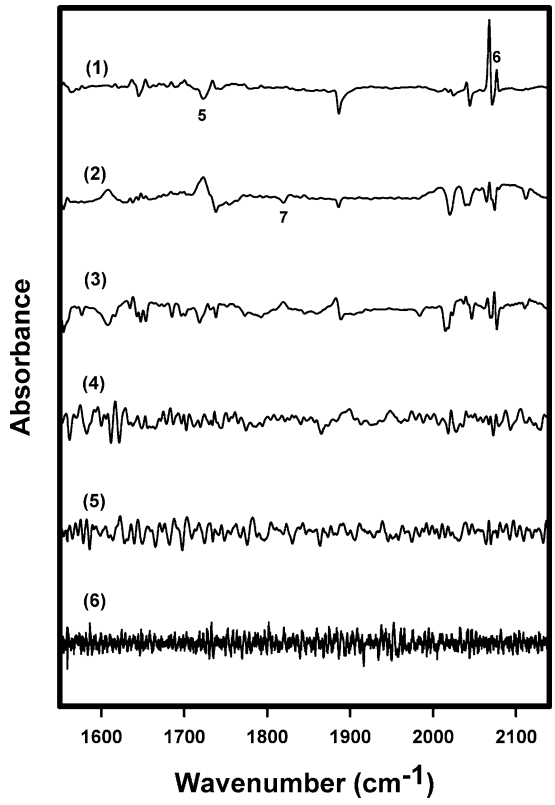

Fig. 4. Six other right singular vectors of the $\mathbf{V}^{\mathrm{T}}$ matrix: (1) 9th vector; (2) 10th vector; (3) 20th vector; (4) 40th vector; (5) 80th vector; (6) 188th vector.

4.2. Singular value decomposition

Singular value decomposition was used to decompose the preconditioned absorbance data matrix $\mathbf{A}_{188\times2951}$, yielding the orthonormal matrices $\mathbf{U}_{188\times188}$ and $\mathbf{V}^{\mathrm{T}}_{2951\times2951}$ and the diagonal singular value matrix $\boldsymbol{\Sigma}_{188\times2951}$. The row vectors in the $\mathbf{V}^{\mathrm{T}}$ matrix are orthogonal basis vectors that contain the abstract information on the pure component spectra. Although a total of 2951 right singular vectors of $\mathbf{V}^{\mathrm{T}}$ were obtained after decomposition, only the first 188 $\mathbf{V}^{\mathrm{T}}$ vectors are physically meaningful since only 188 reaction spectra are in the original data array $\mathbf{A}_{188\times2951}$. The first six $\mathbf{V}^{\mathrm{T}}$ vectors are presented in Fig. 3, while Fig. 4 presents the 9th, 10th, 20th, 40th, 80th, and 188th $\mathbf{V}^{\mathrm{T}}$ vectors.

It is important to note that the right singular vectors in the $\mathbf{V}^{\mathrm{T}}$ matrix are ordered according to their contribution to the total variance in the observations. Therefore, the first few vectors are associated with real chemically important signals in the system, and the remainder are associated primarily with the random instrumental and experimental noise. Accordingly, inspection of the first 10 right singular vectors strongly suggests the existence of chemically important spectral features. Indeed, there are regions of localized signal intensity, and these correspond to the regions where absorbance from organic and metal-carbonyl vibrations can be expected. The 20th and 40th vectors have a minimum of localized signal and are predominantly random noise. The final 80th and 188th right singular vectors seem to have only randomly distributed noise. Essentially no localized signal intensity can be identified.

4.3. Pure component spectra reconstruction results via BTEM (major components)

After decomposition, the right singular vectors of the $\mathbf{V}^{\mathrm{T}}$ matrix were inspected for significant spectral features/extrema to be used as targets in the BTEM algorithm. From the first $\mathbf{V}^{\mathrm{T}}$ vector, four significant bands (indicated by numbers in Fig. 3) were chosen for the reconstruction. These spectral features also appear in the 2nd, 3rd, 4th, and 5th right singular vectors. The four-targeted band regions were chosen as 1640–1644, 1885–1887, 2020–2023, and 1733–1735 cm^{-1} according to the number in Fig. 3. Ultimately four major components were resolved using first derivative entropy minimization as shown in Fig. 5.

These reconstructions correspond to the organic reagent 3,3-dimethylbut-1-ene, the catalyst precursor $\mathrm{Rh}_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$, the observable organometallic intermediate RCO $\mathrm{Rh(CO)}_4$, and the organic product 4,4-dimethylpentanal. The number of vectors in $\mathbf{V}^{\mathrm{T}}$ used for the reconstructions were; $z = 15$ for 33DMB, $\mathrm{Rh}_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$, 44DMP and $z = 18$ for $\mathrm{RCORh(CO)}_4$. These represent excellent estimates of the pure component spectra. Indeed, all primary absorbance maxima are in the correct positions (Table 1), and the similarity of the estimates of 33DMB and $\mathrm{Rh}_4(\sigma$-

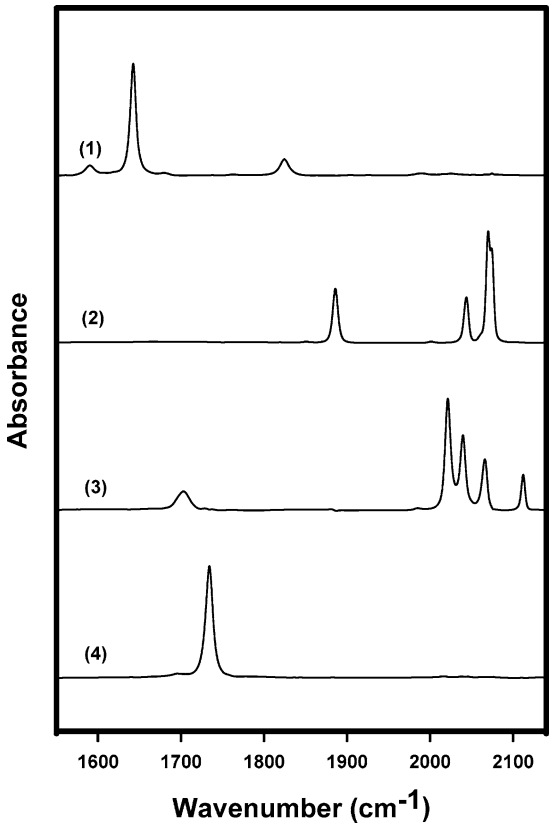

Fig. 5. The reconstructed pure component spectra of major components: (1) 33DMB; (2) $Rh_4(\sigma\text{-}CO)_9(\mu\text{-}CO)_3$; (3) $RCORh(CO)_4$; (4) 44DMP.

$CO)_9(\mu\text{-}CO)_3$ to real experimental samples is in excess of 0.999 in both cases.[1]

In addition, the pure spectra are very smooth. The values of the first derivative entropies were 5.82, 5.66, 6.11, and 5.44 for 33DMB, $Rh_4(\sigma\text{-}CO)_9(\mu\text{-}CO)_3$, $RCORh(CO)_4$, and 44DMP, respectively. An independent experimental measurement of a dilute standard of 33DMB in hexane provided a value of 5.83. An independent measurement of a dilute standard of $Rh_4(CO)_{12}$ in hexane provided a value of 5.62. A previous pure component reconstruction using entropy minimization also provided a value of 5.62.

4.4. Pure component spectra reconstruction results via BTEM (minor components)

Besides resolving the major components with significant signals in the reaction spectra, the BTEM algorithm was also used to reconstruct the minor components. These minor components are identifiable by local spectral intensity in the later right singular vectors. For example, in the 9th right singular vector, two regions denoted as 5 and 6 can be identified and in the 10th right singular vector another region denoted by 7 can be identified. Three minor components were resolved from these regions, namely an unknown species X, a newly identified complex $Rh_4(\sigma\text{-}CO)_{12}$, and $Rh_6(CO)_{16}$.

The minor components could be resolved because many more right singular vectors were used compared to the previous section. To reconstruct the unknown X, the targeted region was chosen as 1720–1726 and 100 right singular vectors were used. Instead of first derivative entropy minimization, the first derivative absorbance summation and integrated area minimization were utilized. The first derivative entropy value was 6.65 and the band maximum was 1723 cm^{-1}. To reconstruct $Rh_4(\sigma\text{-}CO)_{12}$, the targeted region was chosen as 2075.5–2077 cm^{-1} and 100 right singular vectors were taken. The transformation was done by first derivative entropy minimization. The primary band maxima for $Rh_4(\sigma\text{-}CO)_{12}$ occur at 2068 and 2076.2 and an entropy value $= 5.33$ was obtained. To reconstruct $Rh_6(CO)_{16}$, the targeted region was chosen as 1818–1820 and 50 right singular vectors were used. Instead of first derivative entropy minimization, the second derivative absorbance summation and integrated area minimization was utilised. The reconstructed pure component spectrum is rather noisy due to the very low signal intensity,[2] the similarity to an authentic experimental standard is 0.908, and the first derivative entropy value is 6.72.[3] The estimates of these spectra are shown in Fig. 6.

Though the two primary bands of $Rh_6(CO)_{16}$ at circa 1819.2 and 2075 cm^{-1} could be resolved, the estimated spectrum clearly shows some distortion. This can be seen from the uncorrelated signal at circa 1640 cm^{-1}, which

[1] A well-established measure for similarity is defined by the inner product of the estimated spectrum and experimental spectrum in an L^2 norm. A value of identically 1 means that there is exact agreement.

[2] The band maxima occur at 1639.2, 1819.2, and 2075 cm^{-1}. This can be compared to independent experimental values of 1819 and 2075.4 cm^{-1}.

[3] An independent measurement of a dilute standard of $Rh_6(CO)_{16}$ in hexane provided a value of 5.25.

Table 1
Peak position comparison between results obtained from present study and from previous studies

	Band maxima (cm^{-1})	
	This study	Other studies
33DMB	1590.2 1642.4 1679.4 1824.6	1590.4 1642.6 1679.4 1824.6[a]
$Rh_4(\sigma\text{-}CO)_9(\mu\text{-}CO)_3$	1886 2043.8 2069.8 2074.4	2043.8 2069.8 2074.4 1885.8[a]
$RCORh(CO)_4$	1703.2 2021.4 2039.6 2066 2112.2	1698 ± 3 2020 2039 2065 2111 (Ref. [2])
44DMP	1734.2	1734 (Ref. [4])

[a] Obtained from independent experimental study.

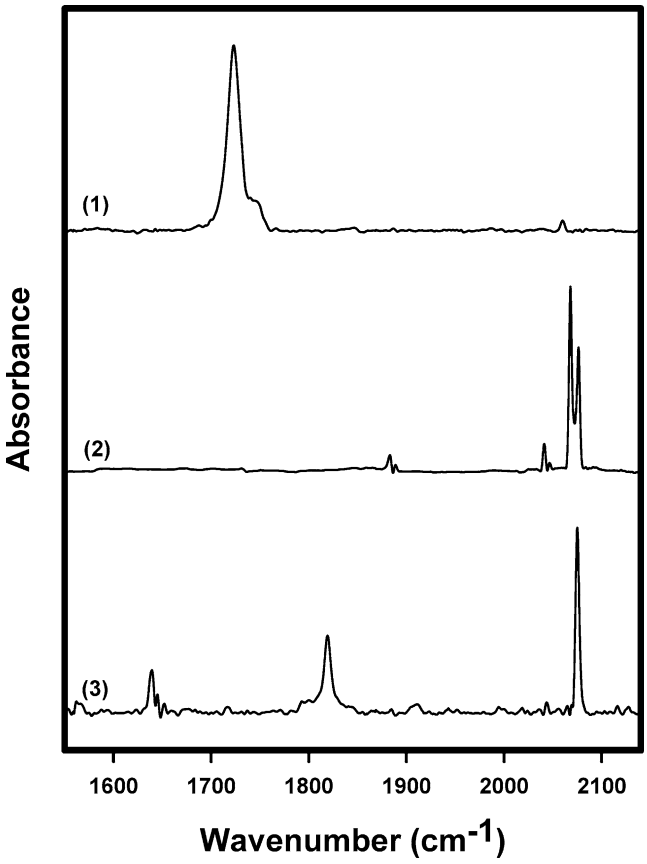

Fig. 6. The reconstructed pure component spectra of minor components: (1) unknown species X; (2) $Rh_4(\sigma\text{-}CO)_{12}$; (3) $Rh_6(CO)_{16}$.

belongs to 33DMB instead of $Rh_6(CO)_{16}$. This type of problem can particularly occur for components with very small signals. To solve this problem, one of the possible techniques is to maximize the variation of the experimental conditions—i.e., add additional experiments under new reaction conditions.

We briefly note that the mean concentration of $Rh_6(CO)_{16}$ is estimated to be 1 ppm. At such low signal/noise ratio, additional signals can get mixed into the pure component reconstruction.

4.5. Relative concentration profile

After obtaining all estimates of pure component spectra, it is possible to calculate the relative concentrations of these components. These are relative concentrations since the pure component spectra are normalized. The exact magnitude of absorptivities and therefore concentrations cannot be obtained without further numerical techniques. By solving the following inverse problem using a least squares approach, the relative concentration profiles of each component for all six experiments can be obtained:

$$\widehat{\mathbf{C}}_{ke\times s} = \mathbf{A}_{ke\times\nu}\hat{\mathbf{a}}_{s\times\nu}^{\mathrm{T}}\left(\hat{\mathbf{a}}_{s\times\nu}\hat{\mathbf{a}}_{s\times\nu}^{\mathrm{T}}\right)^{-1}. \tag{10}$$

Figure 7 shows these relative concentrations. It is worth noting that due to the distortion seen in the $Rh_6(CO)_{16}$ spectrum, as well as its very small intensity, this estimated concentration profile was not included (see discussion for further details).

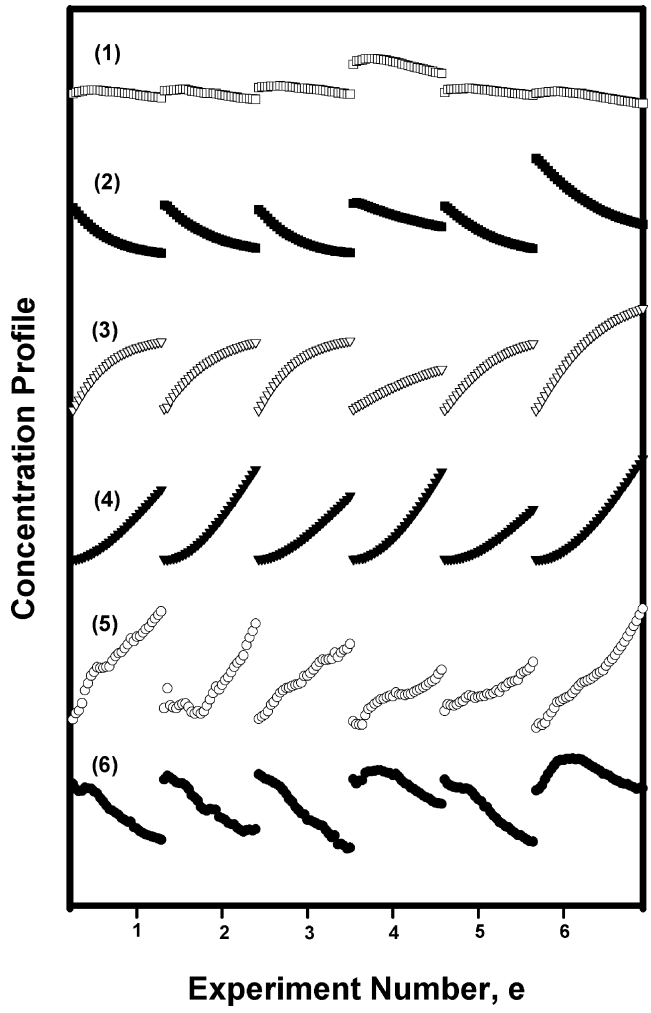

Fig. 7. Relative concentration profile of reconstructed components for all six experiments: (1) 33DMB; (2) $Rh_4(\sigma\text{-}CO)_9(\mu\text{-}CO)_3$; (3) $RCORh(CO)_4$; (4) 44DMP; (5) unknown species X; (6) $Rh_4(\sigma\text{-}CO)_{12}$.

5. Discussion

5.1. Spectral reconstruction, relative concentration profiles, integrated intensities

The results clearly show that it was possible to reconstruct very good estimates of pure component spectra using the BTEM algorithm in conjunction with multiple experiments. The major reconstructed components are in excellent agreement with our previous results. It was even possible to recover minor components existing in trace amounts.

The relative concentration profiles for the four major components (Fig. 7) are exactly as we expect. The profiles of the catalyst precursor $Rh_4(\sigma\text{-}CO)_9(\mu\text{-}CO)_3$ [9], its isomer $Rh_4(\sigma\text{-}CO)_{12}$, and the organic reagent are monotonically de-

Table 2
Percentage of integrated absorbance of each component in the preconditioned data and in the original experimental data

Component	Integrated absorbance in the preconditioned data (%)	Integrated absorbance in the original experimental data (%)
33DMB	36.612	1.957
$Rh_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$	22.522	1.204
$RCORh(CO)_4$	25.937	1.387
44DMP	12.387	0.662
2M33DMB	1.001	0.054
$Rh_4(\sigma\text{-CO})_{12}$	1.356	0.072
Total	99.815	5.336

creasing and profiles of the organometallic intermediate and the organic product are monotonically increasing. The profile of the unknown species X is monotonically increasing and mirrors that of the aldehyde 4,4-dimethylpentanal.

The total integrated intensities for the major and minor species, over all the experiments, are shown in Table 2. It is seen that circa 95% of the experimental intensity is due to the background and solvent. All organic and organometallic reactants contribute only 5.3% of the total signal.

Closer inspection of the pure component spectrum for X indicates that it is organic. Indeed, the band is fairly broad, and it is positioned in an organic carbonyl region. The wavenumber position is 1723. This strongly suggests that the unknown X is the minor aldehyde isomer 2-methyl-3,3-dimethylbutanal (2M33DMB). Furthermore, the concentration profile mirrors the profile of 44DMP. The integrated intensities suggest that the ratio of the minor to major aldehyde is circa 0.08. GC analysis of the postreaction mixture indicates a ratio of 0.11.

The change in the partial pressure of CO in these experiments, and hence the change in its liquid phase mole fraction, resulted in the selectivity difference for aldehydes. This in turn lifts any linear dependence in the spectra (collinearity) and allows the two isomers to be resolved mathematically.

The difficulties in recovering the concentration profile of $Rh_6(CO)_{16}$ are associated with its very low concentration and hence the accuracy of solving the inverse problem in Eq. (10). As Table 2 indicates, the total integrated intensity of $Rh_6(CO)_{16}$ must be less than 0.2% of the total preconditioned signal (the missing intensity). This indicates that its time-averaged concentration is circa 1% of the nominal rhodium concentration or less.

5.2. Maximizing signal recovery—spectral nonlinearities

There are only seven observable species in the reactive system (four major and three minor). Therefore, if the signals were stationary, only seven vectors in $\mathbf{V}^{\mathrm{T}}$ should have localized signals. This is clearly not the case. Localized features can certainly be found in the first 20 or more vectors. Therefore, serious spectral nonlinearities, such as changing band positions and band shapes, exist. The use of $z > s$ right singular vectors in the BTEM algorithm provided a means of maximizing signal recovery. In particular, it provided the means of recovering the minor pure component spectra, whose spectral features are almost at the noise level. The use of more vectors allowed smoother and simpler pure component estimates. The integrated intensities of the minor components 2M33DMB and $Rh_4(\sigma\text{-CO})_{12}$ are only circa 1.0 and 1.4% of the preconditioned signal, or 0.054 and 0.073% of the original experimental data.

5.3. Further explanation of the global search

As a final note, a brief explanation of the global search is useful. For example, during the BTEM recovery of the pure component spectrum of $Rh_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$, many intermediate estimates were generated. These estimates are to some extent random; indeed, the global search needs a random element. However, the long-term trend of the search provides a progressively better estimate of $Rh_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$ in time. This progression is recorded in Fig. 8. The initial estimate corresponds to the randomly chosen starting point for the search. The later estimates correspond to the 10th, 20th, and last (coldest) annealing temperature. The

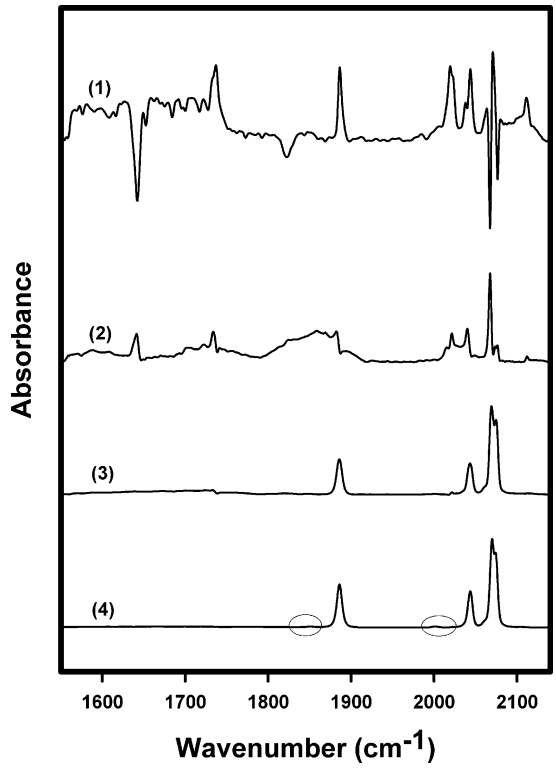

Fig. 8. Spectral reconstruction progression of $Rh_4(\sigma\text{-CO})_9(\mu\text{-CO})_3$ during entropy minimization process by simulated annealing optimization: (1) 1st temperature with entropy value of resolved spectrum = 6.90; (2) 10th temperature with entropy value = 6.64; (3) 20th temperature with entropy value = 5.86; (4) final temperature with entropy value = 5.66.

final spectrum is accurate enough that the vibrations from the 1.1% ^{13}CO isotopomers can be seen (see circled regions).

6. Conclusion

Successful spectral recovery from multiple experimental runs of a very dilute homogeneous organometallic catalyzed reaction was achieved, given no libraries of other a priori information. All anticipated major component spectra were reconstructed. In addition, spectra of unexpected minor components were also obtained. The present analysis of the rhodium-catalyzed hydroformylation of alkenes, using the new band-target entropy minimization algorithm, marks an important point in the investigation of homogeneous catalytic reactions.

Although detailed in situ spectroscopic studies of the unmodified rhodium-catalyzed hydroformylation of alkenes have appeared previously, the present study shows the considerable utility of an advanced spectral recovery algorithm. Band-target entropy minimization clearly holds considerable promise for other detailed analyses of in situ spectroscopic homogeneous as well as heterogeneous catalytic studies.

References

[1] R.S. Dickson, in: Homogeneous Catalysis with Compounds of Rhodium and Iridium, Academic Press, San Diego, CA, 1985, pp. 139–143.
[2] M. Garland, G. Bor, Inorg. Chem. 28 (1989) 410.
[3] M. Garland, P. Pino, Organometallics 10 (1991) 1693.
[4] M. Garland, Organometallics 12 (1993) 535.
[5] C. Fyhr, M. Garland, Organometallics 12 (1993) 1753.
[6] J. Feng, M. Garland, Organometallics 18 (1999) 417.
[7] G. Liu, R. Volken, M. Garland, Organometallics 18 (1999) 3429.
[8] W. Chew, E. Widjaja, M. Garland, Organometallics 21 (2002) 1882.
[9] E. Widjaja, C. Li, M. Garland, Organometallics 21 (2002) 1991.
[10] C.E. Shannon, Bell System Tech. J. 27 (1948) 379.
[11] S. Watanabe, Pattern Recognition 13 (1981) 381.
[12] K. Sasaki, S. Kawata, S. Minami, Appl. Opt. 22 (1983) 3599.
[13] K. Sasaki, S. Kawata, S. Minami, Appl. Opt. 23 (1984) 1955.
[14] S. Kawata, H. Komeda, K. Sasaki, S. Minami, Appl. Spectrosc. 39 (1985) 610.
[15] Y. Zeng, M. Garland, Anal. Chim. Acta 359 (1998) 303.
[16] Y. Pan, L. Susithra, M. Garland, J. Chemometrics 14 (2000) 63.
[17] D.F. Shriver, M.A. Drezdzon, The Manipulation of Air-Sensitive Compounds, Wiley, New York, 1986.
[18] U.K. Dietler, Dissertation, 5428, ETH Zurich, 1974.
[19] K. Noack, Spectrochim. Acta Part A 24 (1968) 1917.
[20] R. Whyman, in: H.A. Willis, J.H. van der Maas, R.G.J. Miller (Eds.), Laboratory Methods in Vibrational Spectroscopy, 3rd ed., Wiley, New York, 1987, Chapt. 12.
[21] L. Chen, PhD thesis, National University of Singapore, 2002.
[22] G.V. Reklaitis, A. Ravindran, K.M. Ragsdell, Engineering Optimization: Methods and Applications, Wiley, New York, 1983.
[23] G.H. Golub, C.F. Van Loan, Matrix Computations, Johns Hopkins Univ. Press, Baltimore, 1996.
[24] M. Garland, E. Visser, P. Terwiesch, D.W.T. Rippin, Anal. Chim. Acta 351 (1997) 337.
[25] S. Kirkpatrick, C. Gelatt Jr., M. Vecchi, Science 220 (1983) 671.
[26] A. Corana, M. Marchesi, C. Martini, S. Ridella, ACM Trans. Math. Software 13 (1987) 263.
[27] W.L. Goffe, G.D. Ferrier, J. Rogers, J. Econometrics 60 (1994) 65.
[28] S.P. Asprey, R. Batres, T. Fuchino, Y. Naka, Ind. Eng. Chem. Res. 38 (1999) 2364.
[29] S.P. Asprey, Y. Naka, J. Chem. Eng. Jpn. 32 (1999) 328.