

Article

**Combination of Single Bead FTIR and Chemometrics in  
Combinatorial Chemistry: Application of the Multivariate  
Calibration Method in Monitoring Solid-Phase Organic Synthesis**

Bing Yan, and Hongbin Yan

*J. Comb. Chem.*, **2001**, 3 (1), 78-84 • DOI: 10.1021/cc000063d • Publication Date (Web): 28 November 2000

Downloaded from <http://pubs.acs.org> on March 20, 2009

**More About This Article**

---

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 1 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

[View the Full Text HTML](#)



**ACS Publications**  
High quality. High impact.

# Combination of Single Bead FTIR and Chemometrics in Combinatorial Chemistry: Application of the Multivariate Calibration Method in Monitoring Solid-Phase Organic Synthesis

Bing Yan<sup>\*,†</sup> and Hongbin Yan<sup>‡</sup>

*ChemRx Advanced Technologies, Inc., 385 Oyster Point Boulevard,  
South San Francisco, California 94080, and Analytical Division, Barr Laboratories, Inc.,  
Pomona, New York 10970*

*Received July 20, 2000*

The single bead FTIR method has been used in quantitative analyses of solid-phase organic synthesis (SPOS) such as the determination of reaction kinetics and conversion yield. These studies rely on data analysis methods to extract quantitative information from IR spectra. However, the IR spectrum of a solid-phase sample contains vibrational bands from the starting material, product, and the polymer support itself. The coexistence of multiple chemical components causes severe spectral overlaps and sometimes makes quantitative analysis extremely difficult. In some cases, it is impossible to extract qualitative and quantitative information from overlapped IR spectra. In this paper, we use partial least squares (PLS), a chemometrics method, to achieve qualitative and quantitative analysis of samples that generate severely overlapped IR spectra. The primary loading factor obtained from a PLS calculation only displays those spectral features that have undergone changes during a SPOS reaction. Disappearing and emerging organic functional groups generate negative and positive signals, respectively, in the primary loading factor, thus allowing qualitative analysis of the reaction with improved precision. The scores of the primary loading factors of spectra taken at various times during a reaction provide quantitative information allowing the study of the reaction kinetics directly on solid support. On the basis of the analysis of three diverse SPOS reactions, the PLS method has demonstrated the unique capability of extracting quantitative and qualitative information from the overlapped IR spectra. It is a powerful data analysis tool for the monitoring of SPOS reactions in combinatorial chemistry.

## Introduction

Combinatorial synthesis<sup>1</sup> is becoming a powerful tool for many disciplines to discover novel leads and further optimize these leads to generate new chemical entities (such as new drugs). Compared to solution-phase and liquid-phase (soluble polymers) synthetic methods, solid-phase organic synthesis (SPOS)<sup>2</sup> has advantages in several key aspects such as the realization of true combinatorial synthesis (such as split synthesis), the ease of reaction workup and automation, and the use of a large excess of reagent to drive a reaction nearly to completion without the requirement of purification. Because of these advantages, SPOS has become the most useful synthetic method in combinatorial chemistry. However, reactions on solid phase also have a major disadvantage, i.e., solution-phase reactions must be first transferred to the dissimilar solid matrix. This reaction optimization process takes most of the time in solid-phase combinatorial synthesis.

Although various analytical methods can assist the optimization of the reaction procedures, most of the methods require the cleavage of the products from the support before analysis. This not only consumes excess amounts of time,

sample, and labor, but it is also detrimental to unstable intermediates and products. Methods capable of direct on-support analysis are ideal for monitoring reaction processes in combinatorial chemistry. One method used routinely for this purpose is single bead FTIR microspectroscopy.<sup>3</sup> Using this method, a drop of resin suspension is taken at any time during a reaction and washed with a solvent, and a few beads are compressed between two IR transparent plates (NaCl, KBr, diamond), and a spectrum is then taken by focusing on a single flattened bead on one side of the plate in transmission mode.

The monitoring of reactions by FTIR is based on the facile detection of the organic functional group interconversions via chemical reaction or by appearance or disappearance of functional groups carried by building blocks or protecting groups introduced or removed during a reaction. The functional group to be monitored need not be directly involved in the reaction that is to be optimized. Therefore, for the rehearsal of a library synthesis on solid support, building blocks can always be selected to contain an IR detectable group at a remote site.

Besides the qualitative monitoring of the solid-phase reaction, single bead FTIR also provides quantitative information such as on-resin conversion yield and reaction kinetics.<sup>4</sup> In fact, owing to its sensitivity and speed, the single bead FTIR method has provided most of the kinetic data for

\* Corresponding author: Dr. Bing Yan, ChemRx Advanced Technologies, Inc., 385 Oyster Point Blvd., South San Francisco, CA 94080; Tel. 650-829-1163; Fax 650-829-1266; E-mail bing\_yan@axystech.com.

† ChemRx Advanced Technologies, Inc.

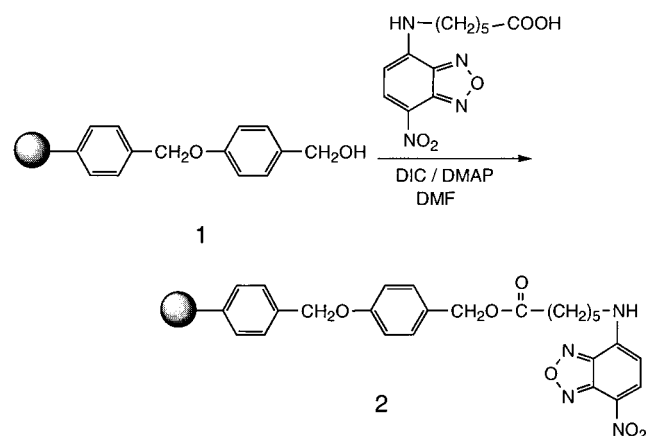
‡ Barr Laboratories, Inc.

solid-phase organic reactions.<sup>4,5</sup> In these studies, the spectral changes of individual reaction beads were used to monitor the time-dependent functional group transformations. By using IR bands from polystyrene as internal standard, quantitative analysis can be carried out by analyzing spectra taken at various times from different beads with various diameters and thickness. Because of the general lack of kinetics information, solid-phase reactions are commonly carried out for longer than necessary times (e.g., 12–24 h) before cleaving compounds for analysis. From single bead FTIR studies, we have found that most solid-phase organic reactions proceed at a much higher rate than generally expected. Slow, incomplete reactions at intermediate steps will cause the accumulation of impurities in the final library. However, extending reaction times beyond completion may also increase the amount of impurities by promoting side reactions or causing product decomposition. In multistep syntheses, the timely completion of each reaction is crucial to guarantee the quality of the final library. Knowledge of the on-support conversion yield and reaction kinetics can guide the selection of optimal reaction conditions.

IR is a spectroscopic technique based on the detection of molecular vibrations. The vibrational modes observed for a given molecule are determined by its atomic composition, the type of chemical bonds present, molecular symmetry, and effects from the surrounding environments. Any changes in these properties can result in perturbation of the spectral properties of the sample. For an  $n$  atom nonlinear molecule, there can be a maximum of  $3n - 6$  degrees of vibrational freedom, and therefore  $3n - 6$  fundamental vibrational modes. Therefore, a vibrational spectrum is extremely rich in information content. In principle, comparison of the IR spectra from a starting material and a product can detect changes in any chemical reaction. In reality, the rich spectral features in resin IR spectra often lead to severe spectral overlap, making it difficult to quantify changes in starting material and product. Because of this, a tremendous need exists for the quantitative evaluation of complex spectra arising from two or more components.

If an IR band from the product is well separated from those of the starting material and the polymer, the relative amount of each resin-bound component can be estimated by simply comparing integrated peak areas. In other cases, IR bands of the product are overlapped with the starting material, the resin, or most likely, both. In this case, peak fitting programs have occasionally been used to extract quantitative information from slightly overlapped spectra.<sup>5</sup> However, analyses performed using either single peak integration or peak fitting methods are subject to errors in measurement, integration, and peak fitting procedure and to baseline drift. In some reactions, the formation of a product only generates overlapped spectral features. Solid-phase samples from these reactions show no distinguishable change in the IR spectrum other than in the fingerprint region in which the spectral overlapping problem is more serious because resin backbones and the attached organic molecule all have extensive vibrational signals in this region. This problem has prevented the analysis of a portion of SPOS reactions by FTIR methods. Therefore, a more reliable and versatile data analysis method

### Scheme 1



is needed to accurately characterize and quantitate organic molecules generated in combinatorial chemistry.

In this investigation, we introduce a novel application of the chemometrics method of partial least squares (PLS) to analyze complex single bead IR spectra from three diverse SPOS reactions. Our results demonstrate that this chemometrics method can be applied to a broad range of SPOS reactions. A compelling advantage of this method is that it is capable of extracting both qualitative and quantitative information from severely overlapped IR spectra. While FTIR has been one of the most practical and useful methods for monitoring SPOS reactions, the analysis of spectral data by chemometrics should greatly enhance the scope of the applications available to this spectroscopic technique.

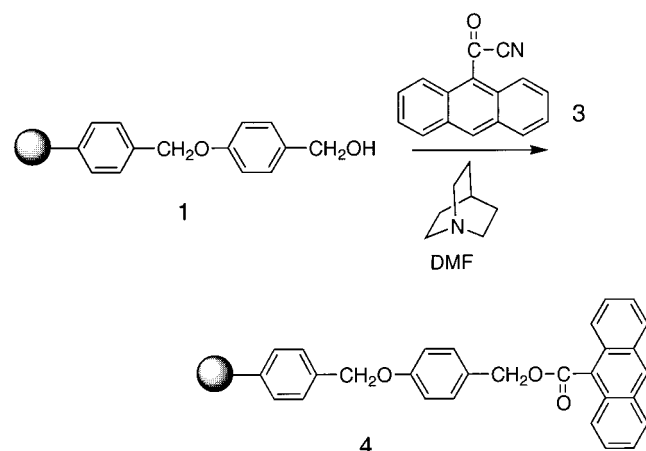
### Materials and Methods

**Materials.** All the resins used in this study were purchased from NovaBiochem (San Diego, CA). Resins are based on 1% cross-linked divinylbenzene-styrene copolymer. They are 100–200 mesh with a loading of 0.6–1.2 mmol/g. All reagents, if not specified, were purchased from Aldrich (Milwaukee, WI).

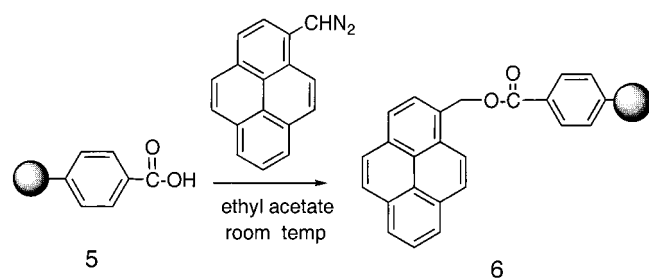
**Single Bead FTIR Microspectroscopy.** All spectra were collected on a Nicolet Magna 550 IR spectrophotometer coupled with a Nic-Plan IR microscope, using OMNIC software. The microscope is equipped with a 15X cassegrain objective and a liquid-nitrogen-cooled mercury–cadmium–telluride (MCT) detector. The view mode aided in locating a single bead. The transmission mode was used for the whole bead measurement. The general procedure for IR measurement is as in ref 3. A flattened bead with a diamond window (SpectraTech, Shelton, CT) was used throughout experiments in transmission mode. A clean diamond window (SpectraTech, Shelton, CT) was used to collect the background spectrum. Data were collected at 4  $\text{cm}^{-1}$  resolution, and 64 scans were averaged.

**Solid-Phase Synthesis. (1) Synthesis of 2.** The reaction is shown in Scheme 1. Wang resin **1** (60 mg, 1.0 mmol/g) was swollen with 4 mL of DMF for 15 min and then drained. Diisopropylcarbodiimide (89 mg, 0.3 mmol) was added to a solution of succinimidy 6-(*N*-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)hexanoate (38 mg, 0.3 mmol) in 1 mL of dry DMF. The mixture was added to the resin after being stirred for 5 min. Then 4-(*N,N*-dimethylamino)pyridine (DMAP)

## Scheme 2



## Scheme 3



(7.3 mg, 0.06 mmol) was added to the resin suspension. The reaction mixture was then mixed at room temperature for specified times.

**(2). Synthesis of 4.** The reaction is shown in Scheme 2. In a 1 mL Supelco filtration tube was added 10 mg (11.7  $\mu\text{mol}$ ) of resin 1. The sealed tube was shaken with 0.5 mL of DMF on an orbital shaker for 15 min and then drained on a visiprep vacuum manifold. 9-Anthrolylnitrile **3** (approximately 2-fold molar ratio relative to the loading on resin) in 200  $\mu\text{L}$  of DMF and 20  $\mu\text{L}$  quinuclidine (0.23 mg/ $\mu\text{L}$ ) were added to the resin. The mixture was mixed on a Glas-Col Laboratory rotator (32 rpm) at room temperature for specified times. Resin beads were taken at various times, rinsed with THF (1 mL  $\times$  10) and DCM (1 mL  $\times$  10), and drained. The beads were analyzed by single bead FTIR.

**(3). Synthesis of 6.** The reaction is shown in Scheme 3. A total of 10 mg (2.6  $\mu\text{mol}$  of carboxyl groups) of resin **5** was added to a 1.0 mL Supelco filtration tube. The sealed tube was shaken with 0.5 mL of ethyl acetate on an orbital shaker for 15 min. The solvent was drained on visiprep vacuum manifold. PDAM (approximately 2-fold excess relative to the loading of the resin) in 500  $\mu\text{L}$  of ethyl acetate was added. The tube was sealed with a polystyrene plug vented with a small needle to avoid pressure buildup from nitrogen gas produced during the reaction. The reaction mixture was rotated on a Glas-Gol laboratory rotator at (32 rpm) at room temperature for specified times.

**Partial Least Square (PLS) Method.** Chemometrics<sup>6</sup> is a suite of mathematical and statistical methods for solving chemical problems. It can be employed to obtain maximum information from experimental data. Traditional methods of information recovery from a spectroscopic experiment generally involve the measurement of a signal at a single selected

frequency. Recent advances in instrumentation and computer technology have facilitated the use of multivariate calibration methods,<sup>7</sup> which rely on the concerted changes in many frequencies to extract information from the process of study.

Multivariate calibration provides an estimate of physical and chemical properties of materials and the component concentration on the basis of their infrared spectra.<sup>7</sup> It is no longer necessary to identify isolated analyte bands or to perform subjective baseline corrections for analysis. Many frequencies or even the entire spectral region can be used to perform a qualitative or quantitative analysis on spectral data. These methods can also achieve increased precision from the redundant information in the spectra, account for baseline variations, and provide outlier detection (outliers are those minor component spectra or data which are different from the major component of detection or not associated with the process of concern).

One of the multivariate calibration methods is the PLS<sup>8</sup> method. The PLS method has been designed to extract factors that explain the covariance between a set of "Y-block" dependent variables (e.g., absorbance values across a spectral region) and one or more "X-block" dependent variables (e.g., an analyte concentration that is closely related to reaction time in this study). The tremendous data compression of PLS also leads to a signal-to-noise enhancement over the original spectral data and a significant improvement in precision.

The mathematical algorithm for PLS analysis can be simplified as

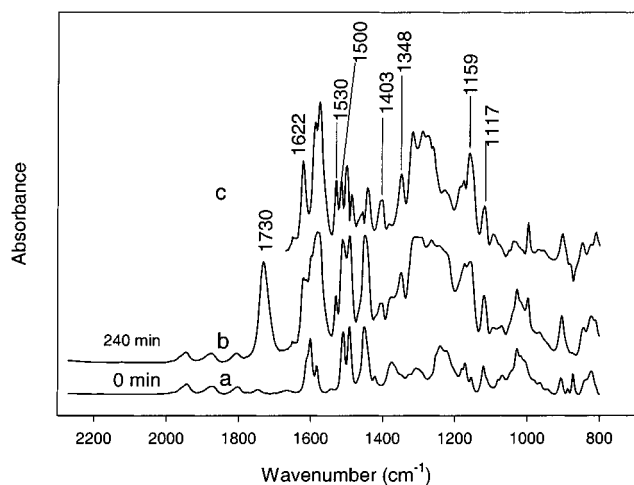
$$\mathbf{A}_{np} = \mathbf{S}_f \mathbf{F}_{fp} + \mathbf{E}_{np}$$

where  $\mathbf{A}$  is the spectral data matrix;  $\mathbf{S}$  is the score matrix;  $F$  is the loading factor,  $E$  is the spectral residuals,  $n$  is the number of spectra,  $p$  is the number of data point, and  $f$  is the number of the rank, which is chosen to produce the optimal prediction. All the data were mean centered before the PLS calculation. The spectral data matrix is transformed into a new full-spectrum representation that requires only a few loading factors. A loading factor results from changes in the coordinate system of the spectral data. These factors form a set of full-spectrum basis factors which will reproduce each of the calibration spectra to within the error represented by  $E$ . Qualitative and chemically interpretable information are included in loading factors. The scores can be considered as spectral intensities in the new full-spectrum coordinate system of the loading factors. Scores are related to the analyte concentrations and can be used to perform quantitative analysis on spectral data.

All precalculation data processing in this study was performed using Win-IR software (Bio-Rad, Cambridge, MS). Multivariate calculations were carried out with the Win-IR/PLS Plus program.

## Results and Discussions

To establish the generality of the PLS method in the qualitative and quantitative analysis of SPOS reactions, three different types of reactions involving diverse compound features and covering a wide range of spectral characteristics in the IR spectra were studied in this investigation.



**Figure 1.** IR spectrum of (a) **1** (starting material); (b) **2** (product); (c) the primary loading factor for the spectral changes during the synthesis of **2**.

**Single Bead FTIR Microspectroscopy.** Single bead analysis is required for one-bead-one-compound library analysis. However, the choice of single bead IR analysis in this work is not due to sample limitation, but rather due to the better spectral quality in comparison with that obtained from other macro sampling methods.<sup>9</sup>

Solid-phase organic reactions, when properly agitated, can occur rather uniformly. Bead-to-bead variation in reaction yield has been found to be very small for most mixing methods.<sup>10</sup> Reactions studied in this work were carried out using the optimal mixing method on the basis of previous experiments. Although a single bead is a good representation of the entire population of beads,<sup>3a</sup> quantitative analysis is still questionable because bead to bead variation in size and thickness (path length) renders their spectra in a wide absorbance range. To perform quantitative analysis, the polystyrene band near 1946 or 1874  $\text{cm}^{-1}$  was selected as the internal standard peak. Reactive linkers are distributed evenly throughout beads with various sizes. The ratio of the product signal to that of the polystyrene is constant for beads with various diameters. Therefore, it is not necessary to analyze kinetics on the same bead throughout a reaction. Comparing beads with different diameters or beads flattened to different thicknesses does not constitute a problem as long as polystyrene bands are used as internal standard peaks.

Finally, the need for an internal standard to correct effects from variable bead sizes or thicknesses is only required for single bead integration and peak-fit analysis methods. Variability in bead sizes and thicknesses is no longer a problem for the chemometric analysis method.

**Analysis of Reaction 1.** The reaction in Scheme 1 was analyzed first in order to validate the analysis method. Spectra of the starting resin **1**, the product **2**, and the spectra taken at various times during the course of the reaction were recorded. The spectra of **1** and **2** are shown in Figure 1 (spectra a and b).

When an IR band of the starting material or the product does not overlap with other bands, the kinetics of the reaction can be analyzed based on the time-dependent changes in its peak area. When bands overlap, however, clear-cut kinetics

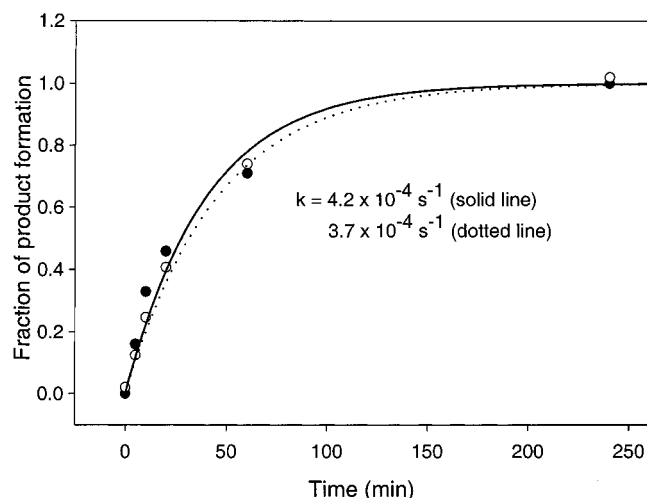
**Table 1.** Peak Assignments for the Starting Material and Product in Reactions 1–3

compound	frequency ( $\text{cm}^{-1}$ )	assignment
product <b>2</b>	1730	$\text{C}=\text{O}$
	1622	$\text{N}=\text{C}-$
	1500	$\text{NO}_2$ asymmetric
	1530	$\text{CN}-\text{H}$
	1348	$\text{NO}_2$ symmetric
	1117	$\text{C}-\text{O}-\text{C}-$
	1722	$\text{C}=\text{O}$
product <b>4</b>	1235	$\text{C}-\text{O}-\text{C}-$ asymmetric
	1149	$\text{C}-\text{O}-\text{C}-$
	890	anthryl, $\text{C}-\text{H}$ out-of-plane bending
reactant <b>5</b>	1720, 1689	$\text{C}=\text{O}$
	1281	$\text{C}-\text{O}$
product <b>6</b>	1721	$\text{C}=\text{O}$
	1269	$\text{C}-\text{O}-\text{C}-$
	1096	$\text{C}-\text{O}-\text{C}-$
	846	pyrene out-of-plane $\text{C}-\text{H}$ bending

information is not obtainable by simple integration. Therefore, the first task of our investigation is to prove that the PLS analysis method can obtain accurate kinetics information in a spectral overlapping situation. In this reaction, a totally isolated band at 1730  $\text{cm}^{-1}$  from **2** is also available for single peak integration analysis. Therefore, this reaction can be monitored by both the PLS analysis method and the single peak integration method. Comparison of these two methods provides a way of validating the PLS analysis method in the analysis of overlapping IR spectra.

As depicted in Figure 1, many spectral changes are involved in this reaction. An intense and well-isolated band at 1730  $\text{cm}^{-1}$  is due to the  $\text{C}=\text{O}$  stretching vibration of the ester carbonyl in **2**. Many other bands in the fingerprint region below 1670  $\text{cm}^{-1}$  are significantly overlapped. The integration of the band area at 1730  $\text{cm}^{-1}$  is used as a reference to compare with the PLS calculation of spectral changes in the fingerprint region (800–1670  $\text{cm}^{-1}$ ) where spectral features are overlapped.

The PLS calculation of a series of spectra taken at 0, 5, 10, 20, 60, and 240 min. for the region from 800 to 1670  $\text{cm}^{-1}$  yielded a primary loading factor with a weight of 90% (Figure 1c). This loading factor reflects all the correlated spectral changes involved in this reaction. The PLS calculation also generated two other minor loading factors (5% each). These minor loading factors showed no correlation with the primary loading factor. The origin of these features was not considered in this work. Some of the assignable spectral features of **2** in this loading factor are listed in Table 1. Some changes of spectral features due to the formation of **2** are marked in Figure 1c. A number of key IR bands, which are previously buried due to overlapping, now become unambiguously resolved and identified in the loading factor (Figure 1c): the peak at 1622  $\text{cm}^{-1}$  assigned to the  $\text{C}=\text{N}$  double bond stretching vibration of **2** and the  $\text{NO}_2$  asymmetric stretching mode at 1500  $\text{cm}^{-1}$ . Most vibrational bands are much better resolved in the loading factor. These include the  $\text{NO}_2$  symmetric stretching vibration at 1348  $\text{cm}^{-1}$ , and the  $\text{C}-\text{N}-\text{H}$  bending vibration at 1530  $\text{cm}^{-1}$  (Figure 1c). On the other hand, vibrational bands that do not undergo changes during the reaction did not show up in the loading factor

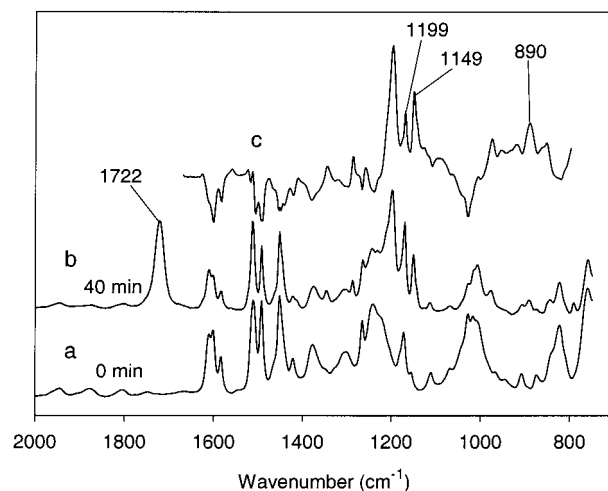


**Figure 2.** Time course for the synthesis of **2** based on the analysis of spectra taken at 0, 5, 10, 20, 60, and 240 min during the reaction: (a) from PLS calibration method (filled symbols and the best fit as a solid line); (b) from integration of the single peak area in the range from 1670 to 1780  $\text{cm}^{-1}$  (open symbols and the best fit as a dotted line).

such as the strong resin signal at 1600  $\text{cm}^{-1}$  and signals at 1375 and 1028  $\text{cm}^{-1}$ . A few negative bands are also observed. They represent the species where concentration is decreasing during the reaction. As demonstrated in this example, the loading factor provides augmented and enriched information about the changes of all spectral features in a SPOS reaction. Therefore, an important application of the PLS, as shown in this reaction, is that the loading factor provides enhanced information about the nature of the chemical species involved in the reaction.

The scores obtained from PLS analysis of IR spectra of samples taken at various times during the reaction can be considered as intensities of the loading factor and can be used for quantitative analysis of the spectral data. On the basis of IR changes, this reaction is complete after  $\sim 2$  h. The scores are therefore normalized by assigning the score at 240 min to 1. The normalized scores of the primary loading factor are shown in Figure 2 as solid circles. The best fit to a pseudo first-order reaction rate equation was plotted as a solid line. Curve-fitting analysis yielded a rate constant of  $4.2 \times 10^{-4} \text{ s}^{-1}$  for the PLS calibration method. These scores were obtained from the analysis of all the spectral intensities in the overlapped fingerprint region (800–1670  $\text{cm}^{-1}$ ). As a comparison, the time course that was calculated from the integration of peak areas of the isolated carbonyl band at 1730  $\text{cm}^{-1}$  for **2** is also shown in Figure 2 as open symbols and the best fit as a dotted line. The rate constant derived from the single peak area integration is  $3.7 \times 10^{-4} \text{ s}^{-1}$ .

As shown in Figure 2, time courses calculated from both spectral regions are very similar. The analysis of this reaction also demonstrates another valuable feature of the PLS method, i.e., it can extract quantitative information for a chemical reaction even though severe spectral overlapping is encountered. When spectral overlapping severely blurs spectral changes, PLS is the method of choice to extract quantitative information during a SPOS reaction. The valida-



**Figure 3.** IR spectrum of (a) **3** (starting material); (b) **4** (product); (c) the primary loading factor for the synthesis of **4**.

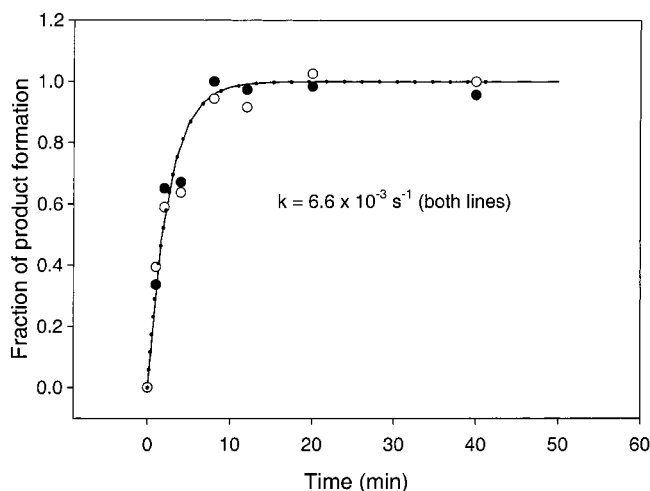
tion of the PLS method in the analysis of reaction 1 is further corroborated by the analysis of the following reaction.

**Analysis of Reaction 2.** The synthesis of **4** is shown in Scheme 2. IR spectra taken at 0, 1, 2, 4, 8, 12, 20 and 40 min were used in PLS analysis. This reaction was studied to further validate the ability of PLS to analyze severely overlapping IR spectra. As in the previous example, the isolated ester carbonyl band at 1722  $\text{cm}^{-1}$  allowed single peak analysis to be performed for comparison. The spectra of **1** and **4** are shown in Figure 3a and b, respectively. The primary loading factor (70% weight) is shown in Figure 3c. There are two other loading factors (15% each). These variations were totally random as the reaction progresses. These were not considered.

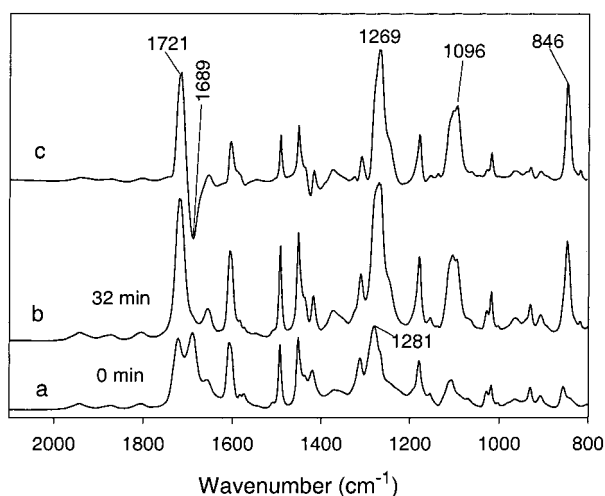
As shown in Figure 3, an intense, well-isolated ester carbonyl band at 1722  $\text{cm}^{-1}$  from **4** was also observed during this chemical reaction. The integration of the band at 1722  $\text{cm}^{-1}$  is used as a reference for a further validation of the PLS method by comparing with results from the analysis of the overlapped fingerprint region.

The PLS calculation was performed on a series of spectra taken at 0, 1, 2, 4, 8, 12, 20, and 40 min during the synthesis using the fingerprint region (800–1670  $\text{cm}^{-1}$ ). In the primary loading factor, a few distinguishable spectral features are marked (Figure 3c). The positive bands from **4** are C–O–C asymmetric and symmetric vibration at 1235, 1149  $\text{cm}^{-1}$  for the ester functional group, and a band at 890  $\text{cm}^{-1}$  for the C–H out-of-plane bending vibration of anthracene. The latter band is blurred in Figure 3b, but it is clearly observed in Figure 3c. The detailed peak assignments are included in Table 1. Negative bands are observed. They represent the decreasing concentration of **3** or some perturbation on resin during the reaction.

The normalized scores of the primary loading factor for a series of spectra taken at various times during the synthesis of **4** are shown in Figure 4 as filled circles. They are derived from the analysis of the fingerprint region (800–1670  $\text{cm}^{-1}$ ) and represent the time course for the synthesis of **4**. The best fit to a pseudo first-order reaction rate equation is plotted as a solid line. The rate constant was determined to be  $6.6 \times 10^{-3} \text{ s}^{-1}$ . The time course obtained from the integration



**Figure 4.** Time course for synthesis of **4** based on the analysis of spectra taken at 0, 1, 2, 4, 8, 12, 20, and 40 min during the reaction: (a) from PLS calibration method (filled symbols, the best fit as a solid line); (b) peak area integration in the range of 1670 ~ 1780  $\text{cm}^{-1}$  (open symbols, the best fit as a dotted line).



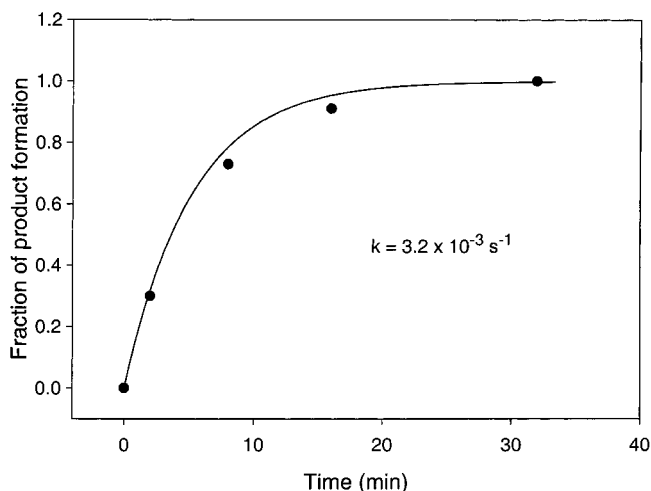
**Figure 5.** IR spectrum of (a) **5** (starting material); (b) **6** (product); (c) primary loading factor for the synthesis of **6**.

of the isolated peak at 1722  $\text{cm}^{-1}$  is also shown in Figure 4 as open circles and the best fit as a dotted line. The rate constant obtained from this analysis is identical with the PLS analysis.

PLS analysis of the reaction kinetics in this example indicates that the method can be applied to reactions that involve only minimal spectral features in the loading factor and demonstrates again that reliable reaction kinetics data can be obtained by the analysis of the fingerprint region.

**Analysis of Reaction 3.** The synthesis of **6** as shown in Scheme 4 is also studied. This reaction includes the disappearance of a carboxylic acid group and the formation of an ester group. The IR spectra of **5** and **6** are displayed in Figure 5 (a and b). All the primary bands from **5** and **6** suffer from certain interference due to overlapping bands.

PLS analysis was performed on a series of spectra taken at 0, 2, 8, 16, and 32 min using the spectral region from 800 to 2270  $\text{cm}^{-1}$ . The primary loading factor (71%) from the PLS calibration is shown in Figure 5c. As indicated, there are many changes during this reaction. The negative absorp-



**Figure 6.** Time course for the synthesis of **6** based on the PLS analysis of spectra taken at 0, 2, 8, 16, and 32 min. Experimental data are shown as filled symbols and the best fit as a solid line.

tion band appears at 1689  $\text{cm}^{-1}$  and is assigned to the carbonyl stretching vibration of the carboxylic acid from **5**. It underwent a decrease in intensity because it is the starting material for the reaction. The positive absorption band at 1721  $\text{cm}^{-1}$  is assigned to the carbonyl stretching vibration of the ester group from **6**. This band has an increased intensity, and it is coincidentally overlapped with the one of the carbonyl bands from carboxylic acid. These changes in band intensity indicate the conversion of an acid to an ester. The C–O stretching vibration at 1281  $\text{cm}^{-1}$  is shown in Figure 5a and b, but does not appear in Figure 5c because it is not changed in this reaction process. The asymmetric stretching vibration of C–O–C at 1269  $\text{cm}^{-1}$  is overlapped with the band at 1281  $\text{cm}^{-1}$  in Figure 5b, but it can easily be identified in Figure 5c. The out-of-plane bending vibration of pyrene at 846  $\text{cm}^{-1}$  is overlapped with a band nearby in Figure 5b, but it is better defined in Figure 5c. All the peak assignments for **5** and **6** are listed in Table 1.

The normalized scores of the loading factors from the analysis of spectra of various samples taken at different times during the synthesis of **6** are plotted in Figure 6. Curve-fitting analysis yields a rate constant of  $3.2 \times 10^{-3} \text{ s}^{-1}$ .

Due to the spectral overlap, the conventional peak integration method cannot be used to give accurate quantitative results. The quantitative information can only be obtained using PLS analysis. The analysis of this reaction demonstrates again the unique capability of the multivariate method in obtaining quantitative information on SPOS reactions when there is no isolated spectral feature.

**Validation of the PLS Method.** The compelling advantage of PLS is the capability to extract qualitative and quantitative information from severely overlapped IR spectra. This has been demonstrated repeatedly by the analysis of three reactions in this investigation. First, the analysis of reactions 1 and 2 has provided a head-to-head comparison of the PLS method with the single peak integration method. The nearly identical results obtained by these two different methods showed that the PLS method can be as accurate as the analysis of an isolated signal. The analysis of reaction 3 further provided direct evidence that totally overlapped

spectra can be analyzed and quantitative information can be obtained using PLS analysis. Quantitative analyses using the PLS method have been instrumental for the kinetic characterization of all reactions in this investigation.

**Enhanced Information Content.** Compared with conventional isolated peak analysis, the PLS method can provide greatly enhanced information content. First, one of the advantages in PLS analysis is that only vibrational modes that undergo changes during a chemical reaction show up in the primary loading factor. The vibrational modes that did not participate in the reaction do not show up in the loading factors. This is totally different from the regular IR spectrum. This also differs from difference spectroscopy because there will be no interference from the subtraction artifact in PLS analysis. Second, vibrational modes from the starting material can show up as negative bands. This information is not available from regular IR analysis and it is very important for the purpose of monitoring the reaction based on changes from both the starting material and the product. Third, previously buried bands can be revealed in the primary loading factor. This has been demonstrated throughout our analysis. Fourth, quantitative information is easily obtained by calculating scores of the primary loading factor for different samples. This application has been repeatedly demonstrated by kinetic studies of all reactions analyzed in this paper. PLS method also has limitations. For example, it can only reveal changes due to reaction or other changes. It cannot provide spectra for a pure component.

### Conclusion

IR spectrum of a SPOS sample contains signals from starting material, product, and the polymer support. The major difficulty in monitoring a SPOS reaction is attributable to the severe spectral overlap caused by the coexistence of multiple chemical components. The FTIR method has been a very useful method for the monitoring of the SPOS directly on resin support and for kinetics studies of SPOS reactions. However, the extraction of quantitative and qualitative information from overlapped IR spectra remains a major challenge. Because of this difficulty, a portion of SPOS reactions has been classified as reactions that cannot be monitored by FTIR. Up to now, the quantitative analysis of single bead FTIR data has been performed using the conventional single peak area method and, in some cases, the peak deconvolution method. These methods are only applicable to spectra with an isolated band or bands with only slight overlap.

In this investigation, we have introduced a chemometrics method, namely PLS, for quantitative and qualitative analysis of samples with severe overlapped spectra. The primary loading factor obtained from PLS analysis only displays those

spectral features that have undergone changes during a reaction. Disappearing and emerging organic functional groups generate negative and positive signals in the primary loading factor, allowing enhanced qualitative analysis of the reaction. The scores of the primary factors of spectra taken at various times during a reaction provide quantitative information, allowing the study of the reaction kinetics directly on solid support. On the basis of analysis of three diverse SPOS reactions, the PLS method has demonstrated the unique capability of extracting quantitative and qualitative information from complex spectra. Chemometrics has been shown to be an essential analytical tool for the monitoring of SPOS reactions.

**Acknowledgment.** We acknowledge the assistance of Lina Liu and Dr. Qun Sun for providing us with the solid-phase synthesis samples.

### References and Notes

- (1) (a) Gordon, E. M.; Barrett, R. W.; Dower, W. J.; Fodor, S. P. A.; Gallop, M. A. *J. Med. Chem.* **1994**, *37*, 1385. (b) Thompson, L. A.; Ellman, J. A. *Chem. Rev.* **1996**, *96*, 555. (c) DeWitt, D. H.; Czarnik, A. W. *Acc. Chem. Res.* **1996**, *29*, 114. (d) Balkenhohl, F.; Bussche-Hunnefeld, C.; Lansky, A.; Zechel, C. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2288. (e) Lam, K. S.; Lebl, M.; Krchňák, V. *Chem. Rev.* **1997**, *97*, 411.
- (2) Leznoff, C. C. *Acc. Chem. Res.* **1978**, *11*, 327. (b) Akelah, A.; Sherrington, D. C. *Chem. Rev.* **1981**, *81*, 557. (c) Frechet, J. M. J. *Tetrahedron* **1981**, *37*, 663. (d) Hodge, P. In *Synthesis and separations using functional polymers*; Sherrington, D. D., Hodge, P., Eds.; Wiley: Chichester, 1988; Chapter 2. Chemometrics text book.
- (3) (a) Yan, B.; Kumaravel, G.; Anjaria, H.; Wu, A.; Petter, R.; Jewell, C. F., Jr.; Wareing, J. R. *J. Org. Chem.* **1995**, *60*, 5736–5738. (b) Yan, B.; Kumaravel, G. *Tetrahedron* **1996**, *52*, 843–848.
- (4) (a) Yan, B.; Fell, J. B.; Kumaravel, G. *J. Org. Chem.* **1996**, *61*, 7467. (b) Yan, B.; Sun, Q.; Wareing, J. R.; Jewell, C. F. *J. Org. Chem.* **1996**, *61*, 8765.
- (5) (a) Li, W.; Yan, B. *J. Org. Chem.* **1998**, *63*, 4092. (b) Sun, Q.; Yan, B. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 361.
- (6) Kramer, R. *Chemometric Techniques for Quantitative Analysis*; Marcel Dekker: New York, 1998.
- (7) (a) Haaland, D. M. *Computer-enhanced analytical spectroscopy*; Jurs, P. C., Ed.; Plenum Press: New York, Vol. 3. (b) Haaland, D. M. *Practical Fourier Transformation Infrared Spectroscopy*; Academic Press: San Diego, 1990. (c) Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley & Sons: New York, 1989.
- (8) (a) Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193–1202. (b) Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1208–1217.
- (9) Yan, B.; Gremlich, H.-U.; Moss, S.; Coppola, G. M.; Sun, Q.; Liu, L. *J. Comb. Chem.* **1999**, *1*, 46–54.
- (10) Li, W.; Yan, B. *Tetrahedron Lett.* **1997**, *38*, 6485–6488.