On-Line Fermentation Monitoring by Mid-infrared Spectroscopy

GUNTA MAZAREVICA, JOSEF DIEWOK, JOSEFA R. BAENA, ERWIN ROSENBERG, and BERNHARD LENDL*

Institute of Chemical Technologies and Analytics, Vienna University of Technology, Getreidemarkt 9/164-AC, A-1060 Vienna, Austria (J.D., J.R.B., E.R., B.L.); and Department of Mathematics and Physics, University of Latvia, Zellu str. 8., LV-1002 Riga, Latvia (G.M.)

A new method for on-line monitoring of fermentations using mid-infrared (MIR) spectroscopy has been developed. The method has been used to predict the concentrations of glucose and ethanol during a baker's yeast fermentations. A completely automated flow system was employed as an interface between the bioprocess under study and the Fourier transform infrared (FT-IR) spectrometer, which was equipped with a flow cell housing a diamond attenuated total reflection (ATR) element. By using the automated flow system, experimental problems related to adherence of CO_2 bubbles to the ATR surface, as well as formation of biofilms on the ATR surface, could be efficiently eliminated. Gas bubbles were removed during sampling, and by using rinsing steps any biofilm could be removed from the ATR surface. In this way, constant measuring conditions could be guaranteed throughout prolonged fermentation times (~8 h). As a reference method, high-performance liquid chromatography (HPLC) with refractive index detection was used. The recorded data from different fermentations were modeled by partial least-squares (PLS) regression comparing two different strategies for the calibration. On the one hand, calibration sets were constructed from spectra recorded from either synthetic standards or from samples drawn during fermentation. On the other hand, spectra from fermentation samples and synthetic standards were combined to form a calibration set. Differences in the kinetics of the studied fermentation processes used for calibration and prediction, as well as the precision of the HPLC reference method, were identified as the main chemometric sources of error. The optimal PLS regression method was obtained using the mixed calibration set of samples from fermentations and synthetic standards. The root mean square errors of prediction in this case were 0.267 and 0.336 g/L for glucose and ethanol concentration, respectively.

Index Headings: **On-line fermentation monitoring; Fourier transform infrared spectrometry; FT-IR spectrometry; Biofilm formation; Process analysis.**

INTRODUCTION

Cultivation of microorganisms is an increasingly used manufacturing route towards several products ranging from bulk antibiotics to high-value therapeutic proteins.[1] However, the control of these bioprocesses is still suboptimal. Only a few parameters (pH, O_2, T, . . .) can be monitored *in situ*. Additional information required for process control and decision making is usually based on data produced by infrequent sampling, which in addition often has a significant time delay from sampling to obtaining the final result.[2] Generally, analyses of target compounds are performed off line by wet chemical assays often involving enzymatic reactions or separation techniques, such as high-performance liquid chromatography (HPLC) or gas chromatography (GC).[3,4] These assays require a tedious sample preparation that is usually time consuming.

The ideal method for biochemical process control should enable direct rapid, precise, and accurate determination of several target compounds, with minimal or no sample preparation and reagent consumption. These requisites are currently fulfilled by spectroscopic methods, most commonly based on infrared[4–9] and Raman[8–10] spectroscopy.

Infrared spectroscopy has been employed for years as an alternative technique for fermentation monitoring. The relative merits of both near-infrared (NIR) and mid-infrared (MIR) spectroscopies have been considered for various systems,[4–9] and it has been established that because MIR spectroscopy measures the fundamental vibration modes of molecules, it has the advantage over NIR that unknown species can also be identified. In addition, the use of such a technique for on-line or *in situ* monitoring enables the detection of transient signals that cannot be captured otherwise (with e.g., off-line and at-line methods). However, the high absorption of water poses a problem for the direct application of IR spectroscopy for monitoring aqueous systems. Several approaches have been developed to achieve the on-line measurement of bioprocesses minimizing this water interference. For this purpose, attenuated total reflection (ATR) based systems have been employed, placing the ATR element in different interfaces between the instrument and the process under study. Doak et al.[11] used a multireflection diamond ATR insertion probe specifically modified to be inserted through a port of the reaction vessel to achieve *in situ* monitoring of *Escherichia coli* fermentation. An extensive study on the stability of the probe using aqueous standards has shown a relative variability up to 20%, having only very limited data from one real fermentation. A different approach was described using an ATR probe in which the transfer of infrared light from the detector to the dip sensor is achieved through an optical fiber to take *in situ* measurements. This system was employed to determine the concentrations of major components in alcoholic and lactic fermentations.[12] However, potential difficulties when working with probe-based systems are biofilm formation on the ATR surface as well as gas bubble adherence.

Partial least-squares (PLS) regression[13–15] has become a very popular multivariate data analysis method to extract information on the concentration of main compo-

Received 2 December 2003; accepted 3 March 2004.

* Author to whom correspondence should be sent. E-mail: blendl@mail.zserv.tuwien.ac.at.

0003-7028/04/5807-0804$2.00/0

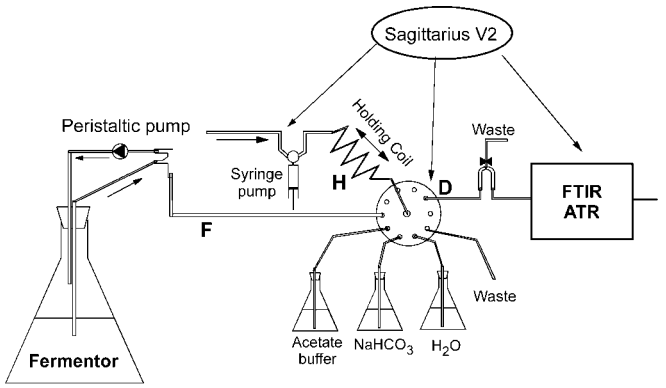

FIG. 1. System for the on-line FT-IR monitoring of yeast fermentation.

nents from the recorded spectra.[4–6,16] PLS regression performs a calibration on the basis of data obtained from one fermentation, and the PLS method obtained is later applied to the prediction of concentration values of future fermentation samples not included in the initial calibration set.

The objective of the present paper was to develop a robust on-line monitoring method for fermentation monitoring based on Fourier transform infrared (FT-IR) spectroscopy with ATR. By way of example, the biotransformation of glucose into ethanol by yeast was selected because it is a well-known system with wide-spread applications in industry (baker's yeast, beer elaboration, antibiotics production, etc.). The experimental and chemometric problems encountered during method development were found to be:

- CO_2 adherence on the ATR surface,
- biofilm formation,
- glucose mutarotation,
- different kinetics in the recorded concentration profiles of glucose and ethanol during various fermentations used for calibration and prediction, and
- precision of the HPLC reference method.

The different problems are discussed and proper solutions presented. The developed method is characterized by full automation, short analysis times, simultaneous multi-analyte determination capability, and robustness. It is applicable to other bioprocesses and different target analytes without requiring experimental modifications. Therefore, this method should be of high interest especially for modern bioprocess development where multiple analytes must be measured with high time resolution and low dead time between sampling and measurement.

MATERIALS AND METHODS

Apparatus. ***Automated Flow System and Fourier Transform Infrared Detection.*** The complete experimental setup used is schematized in Fig. 1. A computer-controlled flow system acts as an interface between the fermentor and the FT-IR spectrometer containing the ATR cell for measurements. The flow system consists of Teflon tubing (inner diameter, 1 mm), one glass F-joint, one glass Y-joint, a Cavro (Sunnyvale, CA) XP 3000 syringe pump equipped with a 5 mL syringe, a peristaltic pump (Masterflex, Cole-Parmer, IL), and a selection valve (Vici Valco, USA). The characteristic geometry of the F- and Y-joints, together with a proper flow rate provided by the peristaltic pump (continuously working) or the syringe pump, allowed the removal of the CO_2 bubbles produced during the fermentation. For spectrum acquisition a horizontal diamond ATR cell with three internal reflections (DurasampleII, SensIR) mounted on a Bruker Equinox55 spectrometer equipped with a liquid nitrogen cooled mercury cadmium telluride (MCT) detector was used. The ATR cell was integrated in a homemade flow cell and connected to the automated flow system via a 70-cm-long Teflon tubing and the selection valve (Fig. 1). The flow system was controlled via the in-house-written MS Visual Basic 6.0 based program Sagittarius V2 (1.02.0003). Each FT-IR spectrum was recorded with the instrument software OPUS (Bruker, Germany) at a resolution of 4 cm^{-1} by coadding 128 scans. The spectral range recorded extended from 4000 to 700 cm^{-1} with the main spectral information contained within the 2000–800 cm^{-1} region.

High-Performance Liquid Chromatography Separation. High-performance liquid chromatography analyses for the determination of glucose and ethanol concentrations were performed using an HP series 1100 HPLC analyzer with a refraction index detector equipped with an 8% Ca^{2+} column (300 mm × 7.8 mm, Phenomenex, CA) and a temperature control system. The separation was achieved under isocratic conditions (flow rate: 0.6 mL/min) with distilled water as the mobile phase and at a temperature of 75 °C. The peak area was selected as the analytical signal.

Samples. ***Fermentation Samples.*** Baker's yeast (Mautner Markhof Hefe, Austria Hefe AG) fermentations in a 0.1 M sodium acetate buffer (Merck, pH = 5.0) were studied in a 1000 mL fermentor.

Three different fermentation experiments were performed under different initial concentrations, glucose supply, and fermentation time conditions (see Table I). In all cases, the pH of the solution decreased by about 0.05 units, but such small changes were found to not affect the activity of the yeast cells. Samples were taken from fermentors every 20 minutes, and glucose and ethanol concentrations were determined by HPLC as described in the next section. Changes in glucose and ethanol content throughout each fermentation are shown in Fig. 2. As expected, the concentration of glucose diminished in the medium as it was converted into ethanol during yeast fermentation.

Synthetic Samples. In addition to the samples collected from fermentation processes, 16 synthetic glucose and ethanol mixtures in 0.1 M acetate buffer were prepared. The glucose and ethanol content in the synthetic samples ranged from 0 to 30 and 0 to 15 g/L, respectively.

Reference Samples for Partial Least-Squares Calibration Set. For PLS regression, three different calibration sets can be distinguished (see Table I): including only the FT-IR spectra of the prepared synthetic samples with known ethanol and glucose concentration (S); including both spectra of synthetic samples with known concentration (S) and spectra of samples from the fermentation A subjected to reference analysis by HPLC; and including only the samples from the fermentation A.

In addition, to evaluate the influence of the HPLC ref-

TABLE I. Fermentation conditions and synthetic sample composition.

Samples	Sample origin	Starting concentration (g/L)			Additional glucose supply	Fermentation time (h)	Reference values by
		Dry yeast	Glucose	Ethanol			
		5	30	0	No	~7.5	HPLC
A	Fermentation	10	30	0	No	~4.5	HPLC
B	Fermentation	5	10	0	50 mL (100g/L) after 2.6, 4.6 and 6.3 h	~7.5	HPLC
C	Fermentation						
S	Synthetic mixtures	···	0–30	0–15	···	···	Weight
S_{HPLC}	Synthetic mixtures	···	0–30	0–15	···	···	HPLC

erence method on the results, the PLS calibration based on the synthetic samples was not only performed using the known concentration values for ethanol and glucose (S in Table I) but also using those values obtained after analyzing the synthetic samples with the HPLC method (S_{HPLC} in Table I).

Analytical Procedure. ***Automated Flow System and Fourier Transform Infrared Measurements.*** The flow system and the FT-IR detector were computer controlled through the OPUS and Sagittarius V2 software. For every measurement, a sequence of steps was automatically executed by the programs (see also Fig. 1):

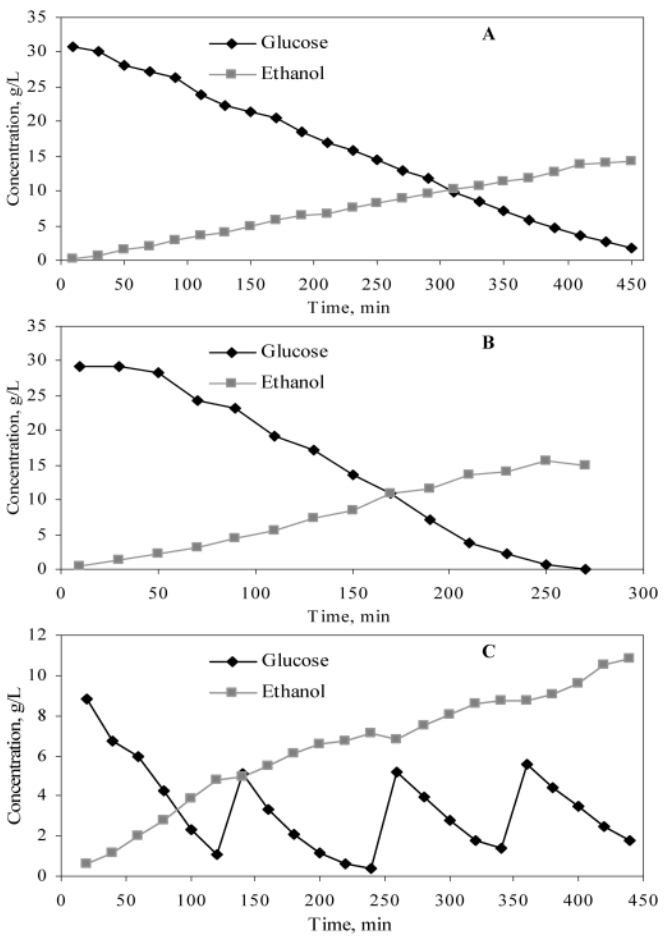

FIG. 2. Concentration profiles of ethanol and glucose obtained by HPLC reference analysis of samples collected from fermentations A, B, and C.

(1) Recording of the background spectrum (acetate buffer solution, pH = 5.0).
(2) Conditioning of sampling tubing (F in the figure) with fresh solution by pumping two fractions of 3 mL from the fermentor to the waste.
(3) Recording of the spectrum of the fermentation mixture by driving a third 3 mL fraction to the ATR cell. Simultaneously, a 2 mL volume sample is manually taken from the fermentor to perform the corresponding reference analysis.
(4) Cleaning of the flow system (tubing H, D) and ATR surface by flushing a 5% $NaHCO_3$ solution for 5 min, followed by a stream of distilled water for 10 min.

Every cycle (including background measurement, system conditioning, sample measurement, and cleaning steps) took 20 min to be completed. The next measurement cycle began immediately after the end of the previous one, allowing a sample throughput of 3 h^{-1}.

High-Performance Liquid Chromatography Procedure. Samples collected during the fermentations in parallel to the FT-IR measurements as well as synthetic samples (S_{HPLC}) were subjected to HPLC analysis to determine the concentration of ethanol and glucose. Whereas standards were directly injected, samples drawn from the fermentation were centrifuged for 10 min to separate the solution from the yeast cells, thus interrupting the fermentation process. The supernatant was then filtered through a Millipore 0.45 μm filter and stored at 4 °C until HPLC analysis was performed.

Under the selected chromatographic conditions, the peaks of glucose and ethanol, as well as the acetate buffer, could be satisfactorily resolved, as can be seen in Fig. 3. Using this method and standard solutions, calibration curves for both glucose and ethanol were constructed. The main analytical features of the method are listed in Table II.

Data Analysis: Partial Least-Squares Regression. The recorded FT-IR spectra together with the known concentration values or the results from HPLC reference analysis were analyzed using partial least-squares (PLS) regression, performed with the OPUS and UNSCRAMBLER (CAMO, Trondheim, Norway) software. The latter program also allows for analysis of the regression coefficient vectors. PLS regression has emerged as one of the most popular multivariate calibration methods. PLS is a modeling procedure that finds the latent variables in the spectral matrix that will best predict the latent variables in the sample concentrations matrix.

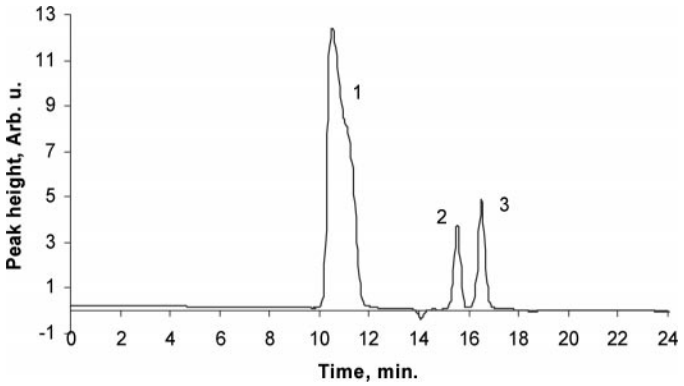

FIG. 3. HPLC spectra of a fermentation sample. Peaks: (*1*) glucose, (*2*) acetate, and (*3*) ethanol.

Partial least-squares regression yields different PLS models when different numbers of factors are used. These models are characterized by the determination coefficient, R^2, which evaluates the fit of the model to the calibration data, and different errors such as RMSEC = root mean square error of calibration, RMSECV = root mean square error of cross-validation, and RMSEP = root mean square error of prediction. All of these errors are calculated for different numbers of factors used in the PLS models and can be defined with the following formula:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}}$$

where n = number of calibration or prediction samples; y_i = true concentration; and $\hat{y}_i$ = predicted concentration. RMSEC is a measure of how good a PLS model fits the calibration data set. RMSECV indicates the predictive ability of the model within the calibration set and the optimal number of factors to be used. The RMSEP error evaluates the response of an established calibration model versus an independent test sample set.

Before and during establishing PLS models, spectra were checked for outliers by visual inspection of principal component score plots and by searching for spectra with unusually high spectral residuals in PLS modeling. The number of PLS factors to be included in the PLS calibration models was in all cases assessed by cross-validating the calibration set. Cross-validation (CV) was done as full leave-one-out CV. In this technique, one calibration sample at a time is taken out of the n sample calibration set, a PLS model is established with the remaining $n - 1$ samples, and the concentration of the taken out sample is predicted. This procedure is repeated until all calibration samples have been taken out and predicted. The n prediction residuals (predicted minus reference concentration) are used to calculate the RMSECV of the calibration model. Then, the number of PLS factors was assessed by plotting the RMSECV versus the number of PLS factors. The optimum number of PLS factors can be generally found at the minimum of the plot, but if RMSECV values at lower PLS factor numbers are not significantly bigger than the minimum RMSECV, the smaller number of PLS factors is chosen in order to avoid over-fitting and to obtain more robust PLS models. The choice of the number of PLS factors is supported by statistical decision criteria in both OPUS and Unscrambler. The RMSECV from full leave-one-out CV tends to give an optimistic estimation of the true prediction error, but all calibration models were also validated versus independent test sample sets to give the RMSEP (see Results and Discussion section).

TABLE II. Analytical figures of merit for glucose and ethanol determination by HPLC.

	Linear regression equation[a]	R	Standard deviation of the method
Ethanol	$y = 9.993x - 0.926$	0.997	0.3
Glucose	$y = 23.305x - 14.134$	0.998	1.5

[a] y = peak area; x = concentration (g/L).

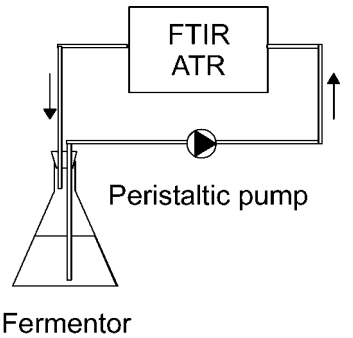

FIG. 4. Initial design for the on-line fermentation monitoring system.

RESULTS AND DISCUSSION

Initial Experiments. In a first attempt to monitor yeast fermentations, a simple flow system based on only a peristaltic pump (Fig. 4) was used to connect the fermentor with the ATR flow cell. In this setup, the flow was continuously circulated over the ATR surface. However, this system proved to be inadequate for robust bioprocess monitoring due to the different phenomena that took place. Illustrative examples of the spectra that represent these problems are presented in Fig. 5. As a reference,

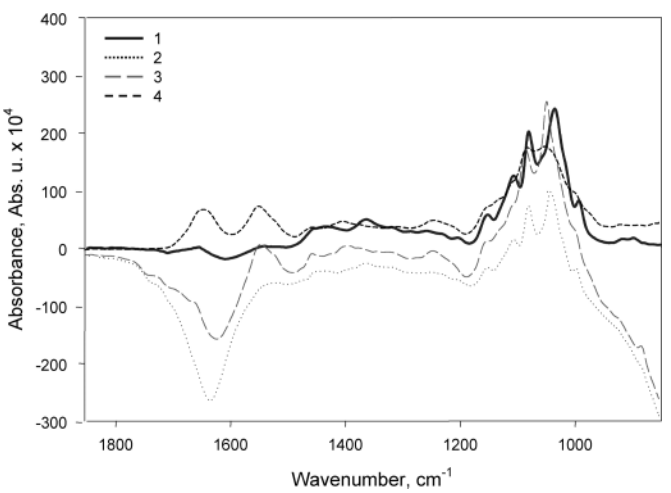

FIG. 5. Fermentation spectra showing the effect of the presence of CO_2 and biofilm formation: (*1*) optimal yeast fermentation, (*2*) presence of CO_2, (*3*) CO_2 and biofilm formation, and (*4*) original yeast.

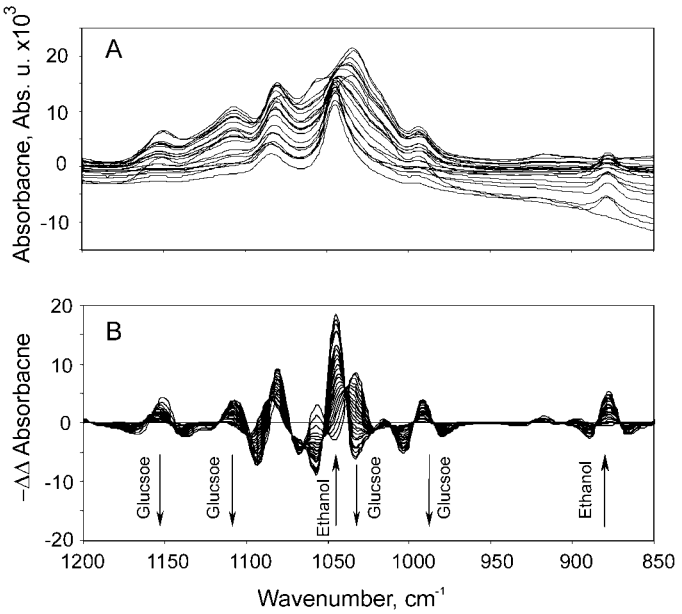

FIG. 6. FT-IR spectra from fermentation A recorded every 20 minutes (initial conditions: 30 g/L glucose, 5 g/L dry yeast). (**A**) Original spectra, (**B**) second-derivative spectra. Main glucose peaks: 1150, 1106, 1079, 1034, and 994 cm^{-1}; main ethanol peaks: 1085, 1045, and 878 cm^{-1}.

the spectrum of yeast fermentation in optimal conditions is depicted as spectrum 1. The first problem detected was a strong negative peak at 1640 cm^{-1}, as seen in spectrum 2. The fermentation process consumes glucose to produce ethanol and CO_2, resulting in the presence of gas bubbles in the solution. These bubbles displace water from the ATR surface and produce a negative contribution to the baseline when the aqueous background is subtracted.

The second problem can be observed in spectrum 3 of Fig. 5. Yeast cells can be deposited and adsorbed onto the ATR surface and form a biofilm that will go on growing during the fermentation. As a consequence, two characteristic peaks of yeast at 1644 cm^{-1} and 1547 cm^{-1} (amide I and II) are noticeable as the reaction proceeds. Since the yeast spectrum (spectrum 4 of Fig. 5) additionally presents a strong absorption in the 1200–900 cm^{-1} region, where the most significant signals for ethanol and glucose are also located, it is necessary to remove the biofilm for accurate determination of the components of interest by FT-IR.

Optimized Flow System. The two experimental problems exposed above required modification of the original flow system. The optimized flow system is shown in Fig. 1. To avoid the problem of the presence of CO_2 bubbles on the ATR flow cell, the system was modified to include two degassing mechanisms (Fig. 1). In the first, an extra loop consisting of an F-joint glass tube and a peristaltic pump was placed close to the fermentor. The solution was pumped at a low flow rate (0.5 mL/min) in such a way that the CO_2 produced by yeast cells as well as a part of the solution were sent back to the fermentor, while the rest of the solution was sent to the detector through tubing F. As the fermentation was relatively fast, a little amount of CO_2 was still produced in tubing F and D, leading to the ATR. In order to remove this additional contribution, one more glass element was added just

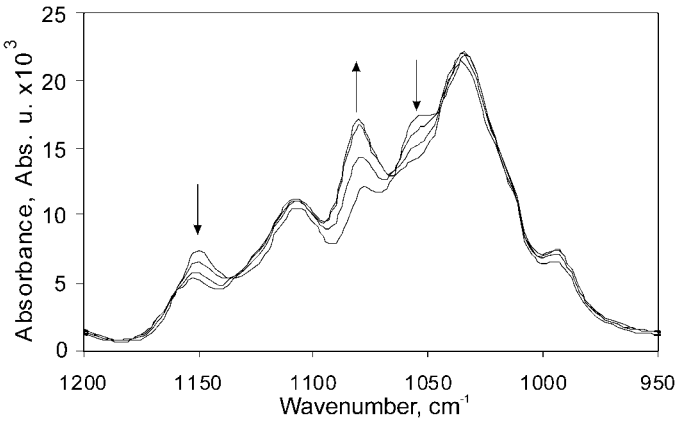

FIG. 7. Changes in α-glucose (30 g/L in buffer) spectra with time. Spectra were recorded immediately after dissolving α-glucose and after 20, 60, and 120 minutes.

ahead of the FT-IR spectrometer. By means of it, the fermentation solution was divided into two parts and only the degassed flow was sent to the ATR measurement cell.

The second problem, biofilm formation, can be easily avoided by including a cleaning step in the experimental protocol. In this step, a stream of 5% $NaHCO_3$ followed by distilled water was sent through both H and D Teflon tubing and circulated over the ATR surface. After this treatment, the ATR cell was ready for a new measurement. The versatility of the computer-controlled flow system allowed for the implementation of this on-line cleaning step without the need for disassembling the setup.

Fourier Transform Infrared Spectra of Fermentation. Figure 6 shows the original spectra of fermentation A and the corresponding second-derivative spectra, obtained using the optimized system described in the previous section. Spectral derivatives can remove baseline artifacts, such as spectral offsets and sloping baselines. They can also be considered as a pseudo resolution enhancement technique because of their ability to highlight slight variations in the slope and contours of bands. As can be observed in the figure, second-derivative spectra showed higher spectral band resolution than the original spectra and baseline shifts were eliminated. Band assignments for ethanol and glucose peaks are also shown (Fig. 6B).

Interesting features observed in the first spectrum of fermentation A are indicated with asterisks in Fig. 6. Concretely, unexpected changes at 1056 (decrease), 1080 (increase), and 1154 cm^{-1} (decrease) can be observed. This can be explained by looking at the reagents employed: at the beginning of the fermentation, α-glucose was added to the yeast solution and the isomerization into β-glucose was not taken into account. The spectrum of β-glucose shows an absorption maximum at 1080 cm^{-1} that can be assigned to the C–O stretching mode of the anomeric C–O–H group.[17] Figure 7 shows the spectra recorded from a freshly prepared solution of α-glucose during the first two hours. As can be seen, the sugar mutarotation is clearly visible in the recorded spectra after 1 hour: the maximum at 1080 cm^{-1} increases as β-glucose is formed and the maxima at 1056 and 1154 cm^{-1} (α-glucose C–O stretching modes) decrease. Once equilibrium is reached, the spectra remain constant. These spec-

TABLE III. Results from PLS regression (number of factors = 3).

Glucose							
All spectra	S → B	A → B	AS → B	All spectra	S → C	A → C	AS → C
RMSEP, g/L	2.290	0.963	1.470	Without first spectra[a]	0.488	2.600	0.409
Without first spectra[a]							
R^2	99.99	99.92	99.83		99.99	99.92	99.83
RMSEC, g/L	0.086	0.250	0.346		0.086	0.250	0.346
RMSECV, g/L	0.110	0.371	0.404		0.110	0.371	0.404
RMSEP, g/L	0.640	0.554	0.650		0.352	2.340	0.267
Ethanol							
All spectra	S → B	A → B	AS → B	All specra	S → C	A → C	AS → C
RMSEP, g/L	1.420	0.440	0.520	Without first spectra[a]	0.507	3.88	0.327
Without first spectra[a]							
R^2	99.99	99.93	99.93		99.99	99.93	99.93
RMSEC, g/L	0.055	0.114	0.122		0.055	0.114	0.122
RMSECV, g/L	0.074	0.176	0.152		0.074	0.176	0.152
RMSEP, g/L	0.563	0.455	0.499		0.515	3.720	0.336

[a] Excluding the first three spectra from fermentation A and B and the first two spectra from fermentation C. Notation: (AS) calibration set including data from fermentation A and synthetic samples S; (AS → B) calibration model AS applied to predict samples from fermentation B.

tral changes are also present in the first spectra of fermentation A (Fig. 6B).

Chemometric Analysis. When multivariate methods such as PLS regression are employed, the selection of the proper spectral range is an important issue. The spectral range should thus include information describing the main features of the analyte and matrix components while regions dominated by artifacts or noise should be excluded.[7] In the present study, two regions were used: from 1198 to 960 cm^{-1} and from 892 to 862 cm^{-1}.

In order to evaluate the origin of the errors that can be found when performing PLS regression, three different calibration sets were formed, based on the data of fermentation A with concentrations obtained by HPLC and the data of the synthetic samples S using the known concentration values from their preparation (see also Table I). The calibration sets consisted of data from fermentation A (denoted as A), data from synthetic samples S (denoted as S), and the combination of A and S (denoted as AS). The results obtained for each one of the different calibrations together with the cross-validation results are listed in Table III. In all cases, samples from other fermentations (B and C) were used for test set validation of the calibration models (A, S, AS).

Glucose Mutarotation. Initially, all spectra of fermentation samples were included in PLS calibration and prediction. As can be seen in Table III, the value of RMSEP obtained for the prediction of fermentations B or C using the calibration models A, S, or AS is unusually high, especially for glucose concentration. Specifically, when using AS and S calibration sets to predict the glucose concentration of fermentation B, the RMSEP values are 1.470 and 2.290 g/L, respectively. This error can be explained by keeping in mind that the first spectra of each fermentation also contained information about the conformational change undergone by glucose. A further comparison of the spectra showed that this information was contained in the first three spectra of fermentations A and B and just in the first two spectra of fermentation C. Thus, to check the influence of this extra-variability of the data, these spectra were excluded from the set of samples, and PLS calibrations were run again. The results are also shown in Table III. In this case, the values of RMSEP were acceptable, especially when compared to the errors obtained when all spectra were included. The exclusion of the first spectra of the fermentation samples resulted in a decrease of the RMSEP for glucose concentration by a factor between 1.11 and 3.6. From a chemometric point of view, the mutarotation of glucose is problematic as it includes an additional spectral variation in the fermentation process that is represented by only a very small number of spectra from the fermentation and that is not properly captured by the calibration. However, this problem can be easily avoided experimentally once it is identified by using the already mutarotated glucose solutions in the fermentation process.

TABLE IV. Results from PLS regression of synthetic samples (number of factors = 3).

	Glucose		Ethanol	
	S	S_{HPLC}	S	S_{HPLC}
R^2	99.9	99.89	99.99	99.94
RMSEC, g/L	0.086	0.337	0.055	0.128
RMSECV, g/L	0.110	0.413	0.074	0.149

Synthetic versus Real Fermentation Samples. From Table III, it can be clearly seen that RMSEC and RMSECV are much smaller for PLS models using only synthetic samples (S) compared to PLS models based on real fermentation samples (A, AS). The synthetic PLS models thus provide good results when predicting synthetic samples but lack the information about the matrix components in real fermentation samples. Therefore, as expected, RMSEP values are much higher than RMSEC and RMSECV values for the same PLS method. On the other hand, PLS calibrations based on fermentation samples include valuable information about real samples but are affected by experimental errors introduced by the reference analysis required for the determination of glucose and ethanol concentrations.

Influence of High-Performance Liquid Chromatography Reference Method. In order to investigate the origin of the error introduced in those PLS predictions of real samples based on calibration set S (S→B, S→C), the synthetic samples were analyzed using the HPLC reference method (S_{HPLC}), and the concentration values obtained were employed as inputs for the PLS calibration. The results are shown in Table IV, together with those

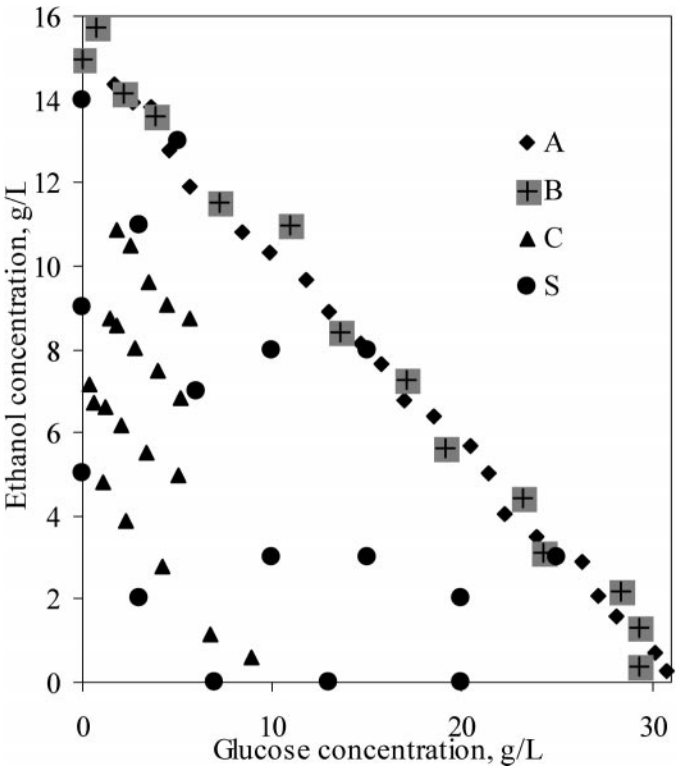

FIG. 8. Sample space spanned in PLS regressions. A, B, and C are data from fermentations; S is data from synthetic samples.

obtained for calibration set S. As can be seen, in this case both RMSEC and RMSECV values are in the same range as those obtained when using real fermentation samples for calibration, indicating that the sources of error are not only the lack of information about the real matrix composition but mostly the experimental errors imposed by the HPLC reference method itself.

Test-Set Validation. As can be seen in Table III, both calibrations S and A give a similar error when predicting the concentration values of B, but the situation is different when predicting fermentation C. In the latter case, the prediction error provided by calibration A is unusually high. This can be explained by having a look at the concentration ranges of the different calibration sets A and S and the test set samples B and C. In Fig. 8 the concentrations of ethanol are plotted versus the concentrations of glucose for the three fermentations and also for the synthetic samples. While fermentations A and B are located in the same region of the plot, fermentation C always has a lower content of glucose than A and B. On the other hand, the synthetic samples are distributed covering the whole range of concentrations that are possible. As a consequence, the PLS model calibrated with samples from a specific fermentation (such as A) will only be adequate to predict the concentration of a similarly behaving fermentation (as is B). Obviously, it will fail in the prediction of a different fermentation procedure generating a wider concentration range (case of C) since the model would have to extrapolate outside the range in which it was calibrated. Better prediction results are obtained when using a synthetic calibration set covering all of the possible concentration range, although the best result is provided by a mixed calibration set including both synthetic samples (distributed over the whole concentration range) and real samples (introducing the information of the real sample matrix), as can be seen in Table III (AS→C).

CONCLUSION

Robustness is a critical issue in the development of calibration models for on-line implementation to provide information intended to be used as the basis of reliable process control. Keeping this idea in mind, a method including an SIA system as interface between the fermentor under study and an MIR-ATR detector has been developed. Experimental and chemometric sources of error have been considered and minimized, providing a method that can be successfully applied to several bioprocesses carried out under different experimental conditions, with only the limitation of being performed in the same fermentation medium. The main errors have been found in the precision of the reference method (HPLC) for the target analytes and in the improper selection of the calibration samples. Thus, a calibration set including synthetic as well as experimental samples covering the whole range of concentrations in which the analytes can be present is mandatory to obtain fast and reliable results in real-time monitoring of bioprocesses.

ACKNOWLEDGMENTS

G.M. thankfully acknowledges the financial support received from a Marie Curie fellowship from the European Union (HPMT-CT-2000-00059). Furthermore B.L, G.M., J.D., and J.R.B. acknowledge financial support received from the Austrian Science Foundation within project No 13686. J.R.B. also acknowledges a postdoctoral grant held by the Spanish Secretaría de Estado de Educación y Universidades and cofinanced by the European Social Foundation.

1. H. C. Vogel and C. L. Todaro, *Fermentation and Biochemical Engineering Handbook: Principles, Process Design, and Equipment* (Noyes Publications, New Jersey, 1996).
2. S. Vaidyanathan, S. A. Arnold, L. Matheson, P. Mohan, B. McNeil, and M. Harvey, Biotechnol. Bioeng. **74,** 376 (2001).
3. K. S. Y. Yeung, M. Hoare, N. F. Thornhill, T. Williams, and J. D. Vaghjiani, Biotechnol. Bioeng. **63,** 684 (1999).
4. R. C. Willson, in *Manual of Industrial Microbiology and Biotechnology*, A. L. Demain and J. L. Davies, Eds. (ASM Press, Washington, D.C., 1999), 2nd ed., pp. 266–272.
5. D. Picque, D. Lefier, R. Grappin, and G. Corrieu, Anal. Chim. Acta **279,** 67 (1993).
6. Ph. Fayolle, D. Picque, and G. Corrieu, Vib. Spectrosc. **14,** 247 (1997).
7. J. B. Cooper, K. L. Wise, W. L. Welch, M. B. Sumner, B. K. Wilt, and R. R. Bledsoe, Appl. Spectrosc. **51,** 1613 (1997).
8. S. Sivakesava, J. Irudayaraj, and A. Demirci, J. Indust. Microbiol. Biotechnol. **26,** 185 (2001).
9. A. C. McGovern, D. Broadhurst, J. Taylor, N. Kaderbhai, M. K. Winson, D. A. Small, J. J. Rowland, D. B. Kell, and R. Goodacre, Biotechnol. Bioeng. **78,** 527 (2002).
10. A. D. Shaw, N. Kaderbhal, A. Jones, A. M. Woodward, R. Goodacre, J. J. Rowland, and D. B. Kell, Appl. Spectrosc. **53,** 1419 (1999).
11. D. L. Doak and J. A. Phillips, Biotechnol. Prog. **15,** 529 (1999).
12. Ph. Fayolle, D. Picque, and G. Corrieu, Food Control **11,** 291 (2000).
13. K. R. Beebe and B. R. Kowalski, Anal. Chem. **55,** 1007A (1987).
14. E. V. Thomas, Anal. Chem. **66,** 795A (1994).
15. H. Martens and T. Naes, *Multivariate Calibration* (John Wiley and Sons, Chichester, 1989).
16. F. Despagne, D. L. Massart, and P. Chabot, Anal. Chem. **72,** 1657 (2000).
17. D. M. Back, D. F. Michalska, and P. L. Polavarapu, Appl. Spectrosc. **38,** 173 (1984).