Front. Chem. Sci. Eng. 2015, 9(3): 386–406
DOI 10.1007/s11705-015-1533-3

RESEARCH ARTICLE

Real time monitoring of bioreactor mAb IgG3 cell culture process dynamics via Fourier transform infrared spectroscopy: Implications for enabling cell culture process analytical technology#

Huiquan Wu (✉)[1,2], Erik Read[3], Maury White[1], Brittany Chavez[3], Kurt Brorson[3], Cyrus Agarabi[1], Mansoor Khan[1]

1 Division of Product Quality Research (DPQR, HFD-940), Office of Testing and Research (OTR), Office of Pharmaceutical Quality (OPQ), Center for Drug Evaluation and Research (CDER), US Food and Drug Administration (FDA), Federal Research Center at White Oak, 10903 New Hampshire Ave, Silver Spring, MD 20993, USA

2 Process Assessment Branch II, Division of Process Assessment 1 (DPA 1), Office of Process and Facilities (OPF), Office of Pharmaceutical Quality (OPQ), Center for Drug Evaluation and Research (CDER), US Food and Drug Administration (FDA), Federal Research Center at White Oak, 10903 New Hampshire Ave, Silver Spring, MD 20993, USA

3 Division of Biotechnology Review and Research II (DBRRII), Office of Biotechnology Products (OBP), Office of Pharmaceutical Quality (OPQ), Center for Drug Evaluation and Research (CDER), US Food and Drug Administration (FDA), Federal Research Center at White Oak, 10903 New Hampshire Ave, Silver Spring, MD 20993, USA

Abstract Compared to small molecule process analytical technology (PAT) applications, biotechnology product PAT applications have certain unique challenges and opportunities. Understanding process dynamics of bioreactor cell culture process is essential to establish an appropriate process control strategy for biotechnology product PAT applications. Inline spectroscopic techniques for real time monitoring of bioreactor cell culture process have the distinct potential to develop PAT approaches in manufacturing biotechnology drug products. However, the use of inline Fourier transform infrared (FTIR) spectroscopic techniques for bioreactor cell culture process monitoring has not been reported. In this work, real time inline FTIR Spectroscopy was applied to a lab scale bioreactor mAb IgG3 cell culture fluid biomolecular dynamic model. The technical feasibility of using FTIR Spectroscopy for real time tracking and monitoring four key cell culture metabolites (including glucose, glutamine, lactate, and ammonia) and protein yield at increasing levels of complexity (simple binary system, fully formulated media, actual bioreactor cell culture process) was evaluated via a stepwise approach. The FTIR fingerprints of the key metabolites were identified. The multivariate partial least squares (PLS) calibration models were established to correlate the process FTIR spectra with the concentrations of key metabolites and protein yield of in-process samples, either individually for each metabolite and protein or globally for all four metabolites simultaneously. Applying the 2nd derivative pre-processing algorithm to the FTIR spectra helps to reduce the number of PLS latent variables needed significantly and thus simplify the interpretation of the PLS models. The validated PLS models show promise in predicting the concentration profiles of glucose, glutamine, lactate, and ammonia and protein yield over the course of the bioreactor cell culture process. Therefore, this work demonstrated the technical feasibility of real time monitoring of the bioreactor cell culture process via FTIR spectroscopy. Its implications for enabling cell culture PAT were discussed.

Keywords process analytical technology (PAT), Fourier-transform infrared (FTIR) spectroscopy, partial least squares (PLS) regression, mouse IgG3, bioreactor cell culture process, real time process monitoring

Received July 30, 2015; accepted August 31, 2015

E-mail: Huiquan.wu@fda.hhs.gov

Disclaimer: The article reflects the views of the authors and should not be construed to represent FDA views or policies

Dedicated to the 120th Anniversary of Tianjin University

1 Introduction

Antibody biopharmaceuticals are generally produced in mammalian cell culture. Potential sources of antibody variability [1] may arise from the type and mode of bioreactor (i.e., batch, fed batch, perfusion, etc.), media composition, culture duration, engineering parameters or inconsistent raw materials. Media formulation and raw material lot-to-lot variability, especially with biologically derived components, can occur and impact process performance. Typically variability in the starting material (i.e., harvested cell culture) can not only be passed through the manufacturing process ultimately to drug substance, but also challenge the downstream chromatography impurity clearance process. The reason for the possible challenge existing is that the downstream chromatograph impurity clearance process may not be designed to handle the unanticipated variability. Therefore, from a process development and regulatory science perspective, developing strategies for optimizing culture media compositions and adjusting bioreactor culture conditions in real time is highly desirable. Monitoring mAb manufacturing processes *via* modern process analytical technology (PAT) tools [2,3] is one innovative approach that might be adopted to achieve this goal. In general, there are a number of well perceived and/or demonstrated advantages of implementing PAT in the pharmaceutical sector [4–6] such as acquiring real time process information, enabling in-depth process and product understanding, achieving better product and process design, enabling better process control and real time release testing, etc. Compared to its counterpart of small molecule product area where PAT applications and implementations are wide spreading [7–15], the adoption of PAT in biotechnology product area has been more challenging as evidenced by its slow yet incremental progress. For example, developing and implementing non-invasive and inline spectroscopy-based PAT strategies for real time monitoring and control mammalian cell culture bioreactors for biological manufacturing has been challenging due to the stringent autoclave requirement for any device to be inserted to the bioreactor.

Near infrared spectroscopy (NIR) has gained popularity over the last two decades because of its convenience and relatively large volume sampling. However, due to the facts that the bands observed in NIR predominantly arose from stretching of O – H, C – H, and N – H bonds and the commonality of these substructures across organic molecules, the differences between NIR spectra of different compounds are often very subtle, thus resulting in a lower molecular selectivity than by using FTIR. In mammalian cell cultures, the analytes most commonly and easily measured include: (1) Carbon and energy sources, which is glucose in most cases; (2) Nitrogen sources, which often include glutamine and other amino acids; and (3) Waste products such as lactate, ammonia and glutamate. Simultaneous measurement of these in synthetic mixtures prepared in aqueous solutions has been reported using a FTIR spectrometer while samples were placed in rectangular cells composed of Infrasi quartz [16]. Strategies to measure those in bioprocess and other fluids have been proposed using *in situ* NIR [17–20] and inline real time Raman spectroscopy [21]. It was shown that the latter technique was able to effectively monitor glutamine, glutamate, glucose, lactate, ammonium, viable cell density, and total cell density in real time, due to its outstanding molecular selectivity as it interrogates fundamental vibrational modes.

FTIR spectroscopy [22,23] can produce exquisite molecular selectivity. In the mean while, its absorption signals are strong. However, due to the requirement for similar and short sample and reference light paths, the instrument configurations are limited. Despite these limitations, FTIR remains a sensitive and highly reproducible physicochemical analytical technique that can identify structural moieties of biomolecules on the basis of their IR absorption. Each biomolecule will exhibit a unique FTIR spectrum, representing the vibrations of its structural bonds. As a promising physicochemical technique, FTIR spectroscopy has been used for off-line monitoring changes in biological samples [24], detection of classes of compounds (e.g., proteins) rather than specific biochemical species, and characterization of protein secondary structure [25] because specific regions in the IR spectra are not only characteristic to protein secondary structure, but they also correspond to the amount of protein in the sample [26]. In addition, FTIR spectra can also be used to monitor changes in sugars due to the fact the C – O stretching in the polysaccharide region (1200–800 cm^{-1}) provides information on the amount of sugar present in the samples [27]. However, all of the aforementioned FTIR spectroscopic studies were conducted in off-line fashion. In a study [28] published in 2010, it was shown that off-line FTIR spectroscopy when combined with chemometrics based on supervised learning (viz., partial least squares regression (PLSR)) can be used to predict accurately antibody production, glucose utilization, and lactate accumulation in cultures of antibody producing murine myeloma NS0 and CHO cell lines. However, to the best of our knowledge, there is no existing literature that reports real time inline process monitoring of bioreactor cell culture process via FTIR spectroscopy.

Here, we report real time monitoring of bioreactor cell culture process via integration of measuring inline FTIR spectra from bioreactor cell cultures directly, parallel off-line determinations of nutrients and antibody concentrations in the media, and process chemometrics. The calibrated partial least squares (PLS) regression models were used to predict concentration profiles of four key cell culture metabolites (glutamine, glucose, lactate, and ammonia) and mAb yield over the course of bioreactor cell culture process. These results demonstrate the

technical feasibility of using FTIR spectroscopy for real time monitoring medium utilization, waste product accumulation, and protein production by bioreactor mammalian cell cultures. The implications for enabling cell culture real time PAT are discussed.

2 Materials and methods

2.1 Materials

Glucose (D-glucose anhydrous) and ammonium chloride was purchased from Fisher Scientific (Fair Lawn, NJ, USA). L-(+)-lactic acid (Free Acid), Fisher BioReagents was purchased from Fisher HealthCare (Pittsburgh, PA, USA). L-(+)-glutamine was purchased from Acros Organics (New Jersey, NJ, USA). Deionized (DI) water was obtained in-house from a Millipore Advantage A10 water purifier (18.2 MΩ resistivity). The base media consisted of Chemically Defined (CD) Hybridoma AGT culture media with 8 mmol/L glutamine (Gibco Life Technologies, Grand Island, NY, USA). For cell culture operations the base media was supplemented with individual amino acids including 0.15 mmol/L aspartate, 0.4 mmol/L cysteine, 0.3 mmol/L methionine, 0.7 mmol/L threonine, 0.15 mmol/L tryptophan, and 0.3 mmol/L tyrosine (Sigma-Aldrich, St. Louis, MO) as described in Read et al. [29]. For PLS modeling of the FTIR spectra, the FTIR spectra were subjected to the Savitzky-Golay polynomial smoothing with seven points, prior to the 2nd derivative preprocessing algorithm.

2.2 FTIR spectroscopy

The FTIR spectra were collected using a Mettler-Toledo ReactIR 45 m Fiber MultiplexIR with MCT Detector using HappGenzel apodization with a DiComp diamond probe connected via AgX 9.5 mm × 2 m silver halide fiber (Mettler-Toledo, Columbia, MD, USA). Sampling was acquired from 2000 to 650 cm^{-1} at 8 cm^{-1} wavenumber resolution, 1024 scans per average, and 1× *g*. The starting water spectrum was subtracted from subsequent spectra to amplify spectral changes. The relationships between the real time FTIR spectra and the concentrations of the components were determined via PLS algorithm implemented in the iCIR 4.1 software (Mettler-Toledo, Columbia, MD, USA).

2.3 Real time inline FTIR characterization of binary systems (individual components in water) at constant temperature

First, 500 mL of DI water in a 1000 mL beaker was brought to 37 °C in a water bath as measured by a glass thermometer. Once the water temperature had stabilized, the given dose of component was added to the beaker in powder form and allowed to dissolve completely to form a binary mixture of individual component and water. Spectra were then collected via the ReactIR after the temperature was established. The averaged spectrum was calculated to represent the spectrum corresponding to the 1st given concentration of the component in the binary mixture. The next dose of the same component was added to the beaker and was dissolved completely to increase the concentration of the component in the binary mixture to the 2nd given concentration. Similarly, the next set of FTIR spectra was collected and the averaged spectrum was obtained to represent the spectrum corresponding to the 2nd given concentration of the specific component in its binary mixture. This procedure was repeated until the expected concentration range of the particular component in its binary mixture was covered. The procedure was repeated again for the 2nd component, the 3rd component, and the 4th component with 500 mL of fresh DI water in a 1000 mL beaker, respectively, until all of the binary mixtures of component and water with expected component concentration ranges were covered and FTIR spectra were acquired. All of the individual components were added in fixed amounts at each dose. The concentration ranges covered in these experiments were chosen to go from zero to triple the concentration that could be expected in an actual cell culture batch. For details on the exact binary mixtures that were characterized with FTIR, please refer to the Appendix Tables A1−A4.

2.4 Real time inline FTIR monitoring of bioreactor media supplementation

A stepwise addition of base CD hybridoma media to the uninoculated bioreactor study was designed to examine how the FTIR spectroscopy responds to the pre-specified changes of media component concentration. First the bioreactor was charged with 4.0 L of base media, then individual components, dissolved in base media first, were added stepwise up to their maximum expected concentration in concentrated doses of 40 mL. After each addition, a 40 mL sample was taken for analysis on the NOVA 100 Bioanalyzer. All process controls were maintained at the same values as they would be during a typical cell culture process. The set points for temperatures, pH, dissolved oxygen (DO), and agitation were 37 °C, 7.2, 30%, and 125 r/min, respectively. The higher than normal impeller speed (75 r/min) was used to enhance the mixing of the individual component solution with the base media. During this stepwise addition of the media to the bioreactor, the media in the bioreactor was monitored in real time by the inline FTIR spectroscopy. For details regarding the exact concentrations of individual components in the media after each dose during the stepwise addition process, please refer to Table 1 for glutamine and Tables A5−A8 in the Appendix for glucose, lactate, and ammonium.

Table 1 Accumulated glutamine concentrations in the medium during the stepwise addition of glutamine—measurement and prediction

Dose #	Calculated glutamine concentration, C_{cal} /(mmol·L^{-1})	Measured glutamine concentration by Nova, C_{Nova} /(mmol·L^{-1})	Data preprocessing algorithms applied	PLS modeling results based on C_{cal}	PLS modeling results based on C_{Nova}
1	0	0	Spectral block: User specified region: 1900 to 900 cm^{-1}; single point B/L @ Zero; mean center.	Calibration model R^2 = 0.9652	Calibration model R^2 = 0.9225
2	1	2.605			
3	2	3.055		Leave-one-out cross validation model R^2 = 0.8849	Leave-one-out cross validation model R^2 = 0.8064
4	3	4	Component concentration block: Mean center		
5	4	4.56			
6	5	5.67		PLS latent variables based on RMSEC *vs.* factors : 3 factors	PLS latent variables based on RMSEC *vs.* factors : 3 factors
7	6	6.04			
8	7	6.37			
9	8	6.6			

2.5 Murine hybridoma culture

A murine IgG3:κ antibody-producing hybridoma adapted to suspension culture growth in this serum-free chemically defined media [29] was grown in 2 L spinner flasks prior to transfer to a 7.5 L capacity Bioflo 110 bioreactor (New Brunswick Scientific, Edison, NJ, USA). Each bioreactor contained 4.0 L working volume of supplemented media. Bioreactor parameters were set to maintain a gas sparge rate of 0.5 L/h, 30% dissolved oxygen, pH 7.2, 37 °C, and pitched blade impeller agitation at 75 r/min. Cultures were fed by bolus addition of 25 mL of 40 g/L glucose and 40 mmol/L glutamine at 72 and 96 h to increase final concentrations by 1 g/L and 1 mmol/L, respectively.

2.6 In-process sampling

Prior to inoculation, and at least three times within a 24 h period throughout the culture, 10 mL samples of cell culture fluid (CCF) were aseptically harvested from the culture for a total of 19 samples. The concentration of viable (VC/mL) and non-viable cells (NVC/mL) within each time-point sample was counted with a TC20 automated cell counter (BIORAD). These values were used to calculate the total cells concentration (TC/mL), viable fraction (Vf), and integral viable cell days (IVCD/mL) for each culture. Samples were clarified by centrifugation at 2000 $\times g$ for 20 min and the resulting supernatant was passed through a 0.22 μm PVDF filter to remove cell debris and large aggregates.

2.7 Analytical protein A affinity chromatography

Relative concentration of the IgG3 antibody in each in-process clarified CCF samples were assessed by analytical scale protein A chromatography across a POROS A/20 4.8 mm column (Life Technologies) by an Agilent 1200 HPLC (Agilent Technologies, Santa Clara, CA, USA) with UV absorbance detected at 280 nm, as described elsewhere [29].

2.8 Nutrient analysis

Glutamine (mmol/L), glucose (g/L), lactate (g/L), ammonium ion (mmol/L), were determined in triplicate for each in-process samples drawn from the bioreactors with an electrode based sensor-type nutrient analyzer (Nova Bioprofile 100 plus analyzer, Nova Biomedical, Waltham, MA, USA). The instrument was calibrated and verified against manufacturer supplied standards to ensure the reliability of membranes and the robustness of the values determined by the instrument.

2.9 Real time monitoring of the cell culture process via FTIR spectroscopy

Prior to cell inoculation of the bioreactor, the FTIR was allowed to equilibrate in the media at the operational set points and base line spectra were collected every 30 min. Upon inoculation of the bioreactor, real time FTIR spectra were taken every 30 min throughout the 120 h cell culture process. As stated previously, 10 mL aliquots of cell culture fluid (CCF) were aseptically harvested from the culture periodically (every 8 h) and were subjected to Nova nutrient analysis. The corresponding FTIR spectra were saved in the database of the iCIR 4.1 software (Mettler-Toledo, Columbia, MD, USA). Data from two batches of the bioreactor cell culture process were used to establish the PLS model.

3 Results and discussion

3.1 Real time inline FTIR characterization of binary systems and FTIR fingerprint identification of pure components

Calibration measurements were run for all four major commonly studied nutrients/waste products (glucose, glutamine, ammonia, lactate) at concentrations from 0 to 3 times the levels expected in a typical mammalian cell

culture bioreactor process. Fingerprint peaks were identified from the 1st derivative spectra at the highest concentration for each component. Calibration models for each individual component were built using the iCQuant package of the iCIR software (Mettler-Toledo, Columbia, MD, USA). Table 2 shows that even in this simple system of individual medium components, the number of PLS latent variables required for the calibration models of glucose, glutamine, ammonia, and lactate to achieve R^2 value greater than 0.99 are 9, 11, 6, and 7, respectively.

In multivariate regression analysis of spectral data, preprocessing (smoothing operation and derivatives) has become an integral part of chemometrics modeling so as to reduce unwanted background information (offsets, sloped baselines) or accentuate absorption features in intrinsically overlapping bands. It was found that such preprocessing algorithms are often useful in reducing the number of latent variables of the actual decomposition and lowering residual error [30]. In this work, the 2nd derivative preprocessing algorithm was applied to the FTIR spectra prior to the PLS modeling to examine if the number of PLS latent variables can be reduced to simplify the model interpretation. For comparison, the PLS modeling results of both with and without 2nd derivative algorithm were listed in Table 2. It was shown that for the binary system PLS models, the 2nd derivative preprocessing algorithm was able to reduce the number of PLS latent variables significantly and thus to improve the models' interpretability. However, this was achieved at the cost of lowered cross-validation R^2 values. Thus, the technical feasibility of using FTIR spectroscopy for characterization and quantification of individual media components in the simple binary system was demonstrated.

3.2 Real time monitoring of stepwise addition of individual components to base media in uninoculated bioreactor

A more complex system of stepwise individual component additions to basal media in an uninoculated bioreactor was monitored in real time via ReactIR 45m system. The resulted FTIR spectra were shown in Fig. 1 for the case of stepwise addition of glutamine into the medium. As expected, there were linear increases of certain peak intensities during the course of stepwise addition of individual components. PLS models were built to correlate the FTIR spectral intensities with both the calculated (by weight of component) and measured (by NOVA) concentration values for each component, as summarized in Table 3 for all four components and in Table 1 for glutamine. The ReactIR system is able to follow the addition of three components (glucose, glutamine, and lactate) sensitively and precisely. Evidently, there is a linear relationship between the peak intensity of presumed component and its concentration. With 3 PLS latent variables, the PLS models have R^2 values of 0.96−0.99. For ammonium, its PLS model has R^2 value of 0.767 with 1 PLS latent variable. For the measurement and prediction results of glucose, lactates, and ammonium during the stepwise addition processes, please refer to Appendix Tables A5−A7 for details.

It is important to note that the data available for the PLS modeling was limited. The selection of PLS factors was based on the plot of RMSEC *vs.* factors where a local or global minimum was observed. The models could be further improved if more data points were available within the concentration ranges interested.

3.3 mAb IgG3 cell culture process media dynamics model

To evaluate FTIR performance in a model bioprocess, a bench-scale murine IgG3 producing bioreactor model culture was developed that allows measurement of in-process samples by analytical protein A affinity chromatography, cell counts and nutrient analysis. A process trajectory for this model bioreactor hybridoma cell culture process is shown in Figs. 2(a)−(b). The starting nutrient concentrations for glucose and glutamine were 4 g/L glucose and 5 mmol/L glutamine, respectively. Fig. 2(a) shows that the viable cell density (VCD) follows the typical pattern of lag (starting at 1.20E + 5 cells/mL), growth, plateau, and then finally death phase as nutrients have been depleted. The IgG3 concentration in the culture increases slowly during the period of 48 h to 114 h, indicating a slow initial IgG3 secretion process. At 114 h, the IgG3 concentration increases very rapidly; coinciding with death phase period of antibody release typically seen with this hybridoma.

As measured by a NOVA bioanalyzer, the nutrient concentration data of four key cell culture metabolites behaved as expected during the course of the fed-batch model murine IgG3 hybridoma cell culture process, as shown in Fig. 2(b). As expected, the glucose concentration decreased due to its consumption within the initial 68 h. To maintain the energy supply to the cell culture process, additional glucose and glutamine stocks were introduced to the bioreactor at 69 h and 93 h; transient feed spikes were evident on the glucose concentration profile, but depletion occurred subsequently (Fig. 2(b)). The concentration profile of glutamine followed similar trends as those for glucose. On the other hand, the concentrations of murine IgG3 hydridoma cell culture waste products lactate and ammonia accumulated in the bioreactor cell culture media.

3.4 FTIR spectroscopy applied to bioreactor cell culture model

3.4.1 PLS model calibration for multiple nutrient components

Multivariate PLS models were developed to correlate the process FTIR spectra with the concentration data of the

Table 2 FTIR calibration model results for pure components of cell culture media in water at 37°C

Component & data preprocessing algorithm / FTIR spectra block	Concentration Range	Specific Wavenumber Range/cm^{-1}		Number of PLS latent variables		Calibration PLS model merits of figure						Characteristic Peaks / cm^{-1} Identified (without applying the 2nd derivative preprocessing)
						No 2nd derivative applied			2nd derivative applied			
		Baseline correction; Mean center	Baseline correction; 2nd derivative; Mean center	Baseline correction; Mean center; No 2nd derivative applied	Baseline correction; Mean center 2nd derivative applied	R^2	RMSEC	RMSECV	R^2	RMSEC	RMSECV	
Glucose	0.7−21 g·L−1	1200–1000	1800–1000	9	3	1			0.935	0.983	2.24	1394, 1625
Glutamine	0.7−21 mmol·L−1	1780–1000	1800–1300	11	4	1	9.79E-4	0.945	0.995	0.393	3.15	1401, 1573, 1658, 1692
Ammonia	0.7−21 mmol·L−1	1850–750	1850–750	6	4	0.993	0.69	4.34	0.984	1.07	6.07	1397, 1871, 1886
Lactate	1−30 g·L−1	1900–900	1800–1000	7	5	0.999	0.135	1.95	0.997	0.307	2.21	865, 1010, 1125, 1144, 1394, 1416, 1662

Component conc. block: mean center

Table 3 PLS model results for stepwise addition of individual components to base media in uninoculated bioreactor at constant temperature

Component	Calculated concentration range covered	Number of PLS latent variables	PLS calibration model R^2	
			concentration based on component weight	concentration based on NOVA data of in-process samples
Glucose /(g·L^{-1})	6–8.04	3	0.9954	0.9996
Glutamine /(mmol·L^{-1})	0–8	3	0.9652	0.9225
Lactate /(g·L^{-1})	0–20	3	0.9988	0.9982
Ammonia /(mmol·L^{-1})	0–10	3	0.7677	0.6956

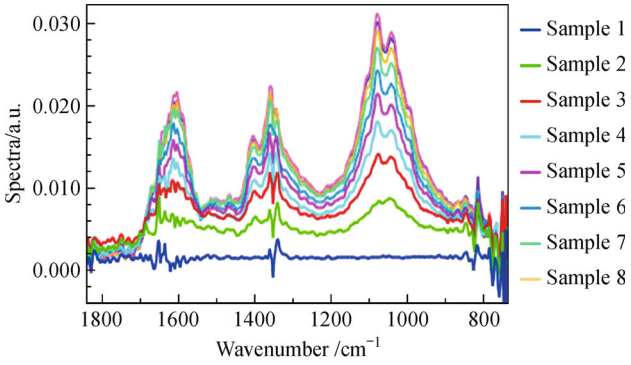

Fig. 1 FTIR spectroscopy responses to the stepwise additions of base CD hybridoma media with 40 mL of concentrated glutamine solution to the bioreactor. Legends in the Figure: Sample 1, water; Sample 2, after the 1st addition of 40 mL concentrated dose of glutamine; Sample 3, after the 2nd addition of 40 mL concentrated dose of glutamine; Sample 4, after the 3rd addition of 40 mL concentrated dose of glutamine; Sample 5, after the 4th addition of 40 mL concentrated dose of glutamine; Sample 6, after the 5th addition of 40 mL concentrated dose of glutamine; Sample 7, after the 6th addition of 40 mL concentrated dose of glutamine; Sample 8, after the 7th addition of 40 mL concentrated dose of glutamine

nutrients at various sampling time points. Two batches of bioreactor cell culture processes with different starting nutrient compositions and feeding strategies were initiated and monitored in real time via FTIR spectroscopy. A total of 50 FTIR spectra were acquired during the bioreactor cell culture process for batch A and batch B, as shown in Table 4. A total of 38 FTIR spectra acquired at various sampling time points over the entire courses of the two batches with corresponding nutrient compositions measured by Nova was taken as the calibration dataset for the PLS model calibration. The rest of 12 FTIR spectra acquired with corresponding nutrient compositions measured by Nova were taken as the testing dataset (also termed as prediction dataset). Two governing principles were followed during the process of selecting FTIR spectra for the calibration dataset and testing dataset: (1) The data points in each dataset cover the concentration ranges for each individual components as much as possible; and (2) For each individual components in the bioreactor cell culture process medium, the data points were distributed across the concentration range as uniformLy as possible. Both the calibration dataset and testing dataset were shown in Table 4. Two PLS modeling strategies were used, including global PLS models (where a common set of PLS latent variables is used for all components in the bioreactor media simultaneously) and individual PLS models (where a unique set of PLS latent variables is used for an individual component in the bioreactor media). The PLS

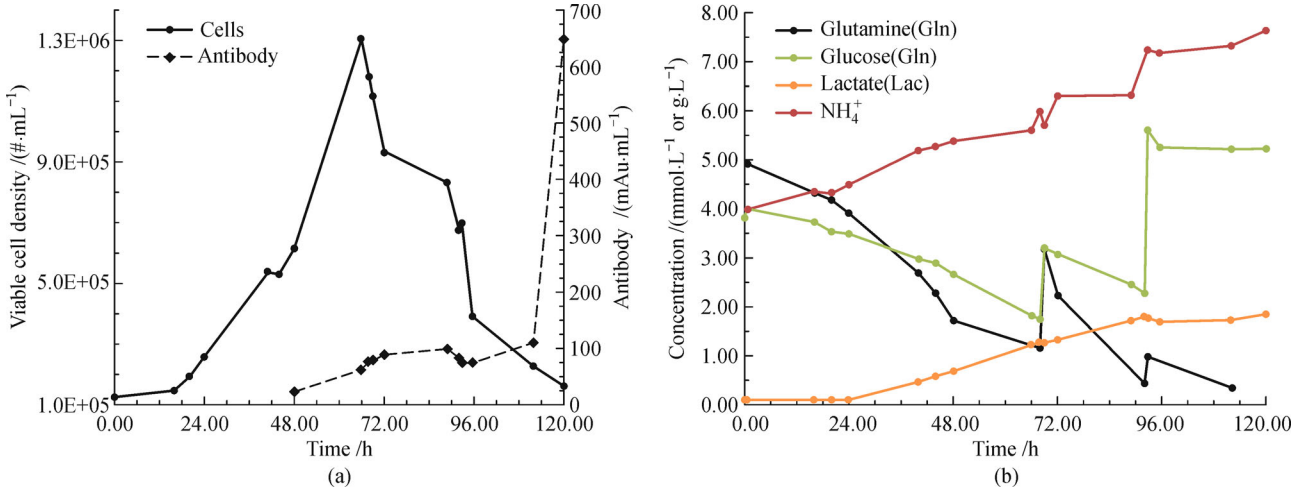

Fig. 2 (a) Process trajectory for a representative bioreactor cell culture process and IgG3 production; (b) Glutamine, glutamate, glucose, lactate, ammonia concentrations measured throughout the cell culture by Nova nutrient analyzer

Table 4 Calibration dataset and testing dataset for correlating process FTIR spectra with nutrients' concentrations at various time points during bioreactor mAb IgG3 cell culture process of two independent batches

Batch #	$T_{\text{spectra acquired}}$ /h	Spectra included in calibration?	Spectra included in test?	Glutamine /(mmol·L^{-1})	Glucose /(g·L^{-1})	Lactate/(g·L^{-1})	Ammonium /(mmol·L^{-1})
A	0.38	Yes		7.72	4.17	0.01	0.79
A	0.88	Yes		6.17	4.84	0.01	3.45
A	16.43	Yes		5.74	4.71	0.01	3.88
A	20.43		Yes	5.65	4.72	0.01	3.94
A	24.43	Yes		5.54	4.7	0.01	4.06
A	41.93	Yes		4.78	4.57	0.01	4.39
A	44.43		Yes	4.89	4.58	0.01	4.45
A	45.60	Yes		4.84	4.46	0.01	4.4
A	48.43		Yes	4.83	4.46	0.01	4.54
A	64.43	Yes		3.6	3.99	0.01	5.47
A	68.43	Yes		3.56	4.25	0.01	5.77
A	69.43		Yes	3.34	6.23	0.01	5.7
A	72.43	Yes		3.12	5.5	0.01	5.85
A	72.77	Yes		3.11	5.42	0.01	5.93
A	88.42	Yes		1.34	5.32	0.26	8.05
A	91.42		Yes	1.25	5.31	0.29	8.21
A	95.25	Yes		1.07	5.24	0.3	8.28
A	111.42		Yes	0.78	5.19	0.3	8.48
A	192.03	Yes		1.12	4.75	0.26	8.22
A	192.20	Yes		2.57	5	0.38	8.18
A	193.20	Yes		4	4.91	0.4	8.13
A	193.87	Yes		5.43	4.93	0.36	8.1
A	194.20	Yes		5.7	4.76	0.35	7.87
A	194.37	Yes		4	4.8	0.41	16.1
A	194.87		Yes	3.92	3.88	0.54	16.865
A	195.03	Yes		4.01	4.63	0.38	17.63
A	195.20	Yes		4.15	4.58	0.41	18.76
A	209.78	Yes		3.96	4.69	0.4	16.33
A	210.28	Yes		3.97	4.54	0.96	16.05
A	210.45		Yes	3.82	4.48	1.3	15.82
A	210.62	Yes		3.85	4.42	1.44	15.34
A	214.95	Yes		3.96	4.36	1.83	15.41
A	215.28	Yes		3.74	4.35	1.97	15.4
A	215.62		Yes	2.96	4.14	3.63	14.94
A	215.95	Yes		3.37	4.07	2.93	14.91
A	217.45	Yes		3.32	4.01	3.08	14.76
B	0.05	Yes		6.79	3.9	0	0.63
B	1.00	Yes		3.7	3.86	0	4.22
B	16.50		Yes	3.15	3.62	0	4.6
B	20.38	Yes		2.97	3.48	0	4.64
B	24.55	Yes		2.74	3.35	0.21	4.77
B	41.38	Yes		1.72	2.86	0	5.18

(*Continued*)

Batch #	$T_{spectra\ acquired}$ /h	Spectra included in calibration?	Spectra included in test?	Glutamine /(mmol·L^{-1})	Glucose /(g·L^{-1})	Lactate/(g·L^{-1})	Ammonium /(mmol·L^{-1})
B	44.72	Yes		1.35	2.78	0.61	5.49
B	66.71	Yes		0	1.9	1.26	6.21
B	69.22		Yes	0	1.81	1.32	6.18
B	69.38	Yes		2.06	3.08	1.33	6.24
B	72.38	Yes		1.39	2.95	1.41	6.75
B	89.88	Yes		0.27	2.36	1.78	7.34
B	92.71		Yes	0	2.27	1.86	7.16
B	93.38	Yes		0.41	5.28	1.81	7.96

calibration models were validated using the leave-one-out cross validation algorithm and were further tested by the testing datasets.

For the global PLS modeling results generated from testing datasets, the plots of predicted concentrations *vs.* actual measured concentrations and the linear regression equations for glucose, glutamine, lactate, and ammonium were shown in Figs. 3 (a)–(d), respectively. Similarly for the individual PLS modeling results generated from testing datasets, the plots of predicted concentrations *vs.* actual

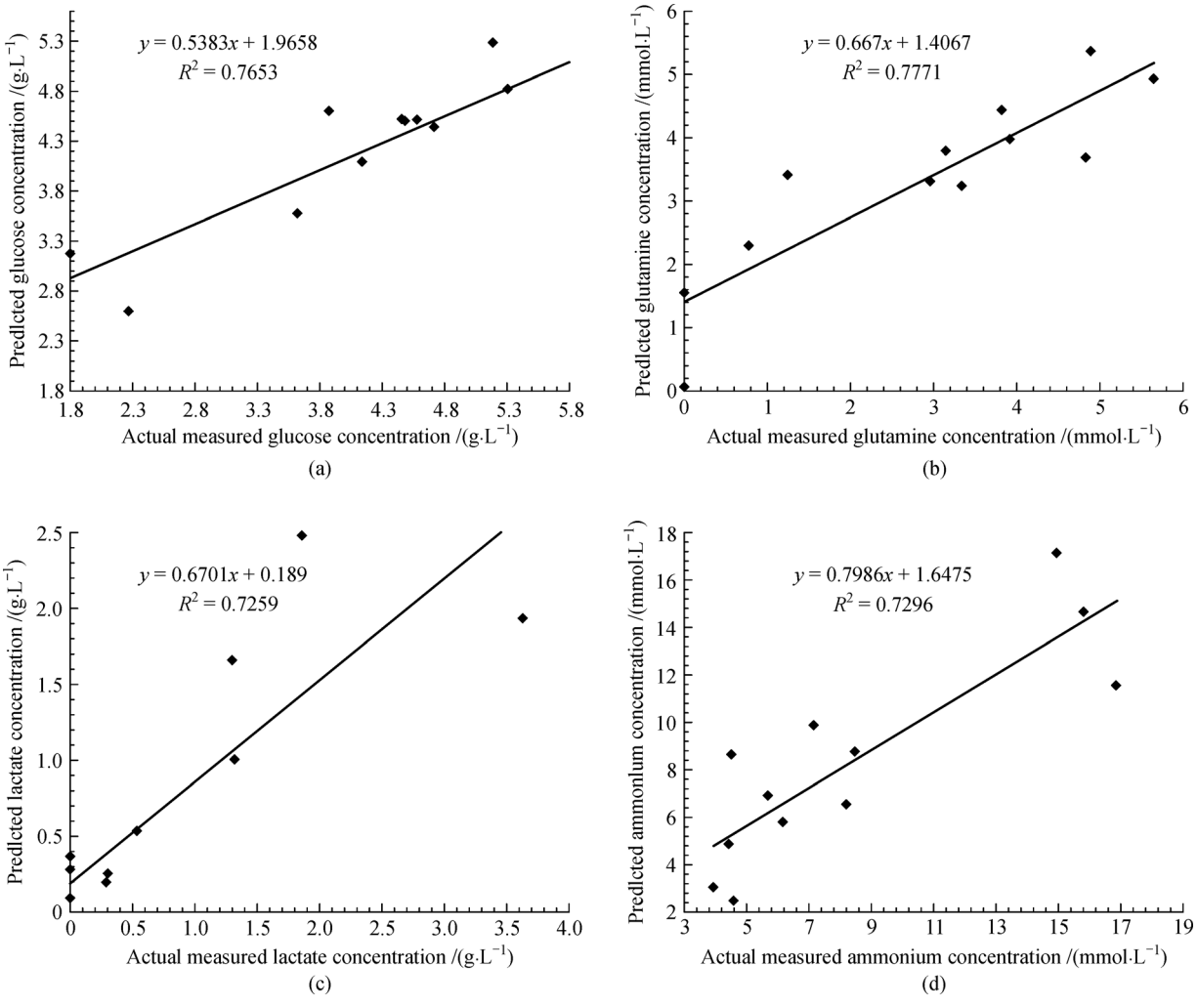

Fig. 3 Key metabolite concentration prediction results from the global PLS calibration models using 7 PLS factors for the bioreactor cell culture process of batch A and batch B. (a) Glucose; (b) Glutamine; (c) Lactate; (d) Ammonium

measured concentrations and the linear regression equations for glucose, glutamine, lactate, and ammonium were shown in Figs. 4(a)−(d), respectively. For all of the four metabolites studied, there is a linear relationship between the predicted concentrations and the actual measured concentrations, regardless the PLS modeling strategy used. The R^2 values for glucose, glutamine, lactate, and ammonium are 0.7653, 0.7771, 0.7259, 0.7296 for the global PLS models, and 0.9926, 0.9546, 0.9965, 0.9996 for the individual PLS models, respectively.

A comparison of figures of merits of the global and individual PLS models is listed in Table 5. Cumulative R^2 is cumulative explained variance by each factor for each response variable in the PLS model. Apparently for all of the four components studied, the individual PLS models have better figures of merits (such as Root-Mean-Square Error of Calibration (RMSEC), Root Mean Square Error of Cross-Validation (RMSECV), Cumulative R^2, $R^2_{\text{calibration}}$, $R^2_{\text{crossvalidation}}$) than the global PLS models. This is expected, given that individual PLS model construction uses specific wave length regions based on the characteristic peaks of individual components that were defined previously during the PLS calibration process from the binary system. It was noted that for bioreactor cell culture waste products lactate and ammonium, their individual PLS models' figures of merits are getting better when the number of PLS latent variables is increased.

The 2nd derivative preprocessing algorithm was applied to the FTIR spectra. It was found (data not shown) that for the real time bioreactor cell culture global PLS models, the 2nd derivative preprocessing algorithm was able to reduce the number of PLS latent variables needed and improve the models' interpretability, yet with comparable and reasonable model performance matrix as characterized by RMSEC, RMSECV, RMSEP, and cumulated R^2 value, etc. In addition, a comparison of model performance matrix when different numbers of PLS latent variables were selected was summarized in Table 6.

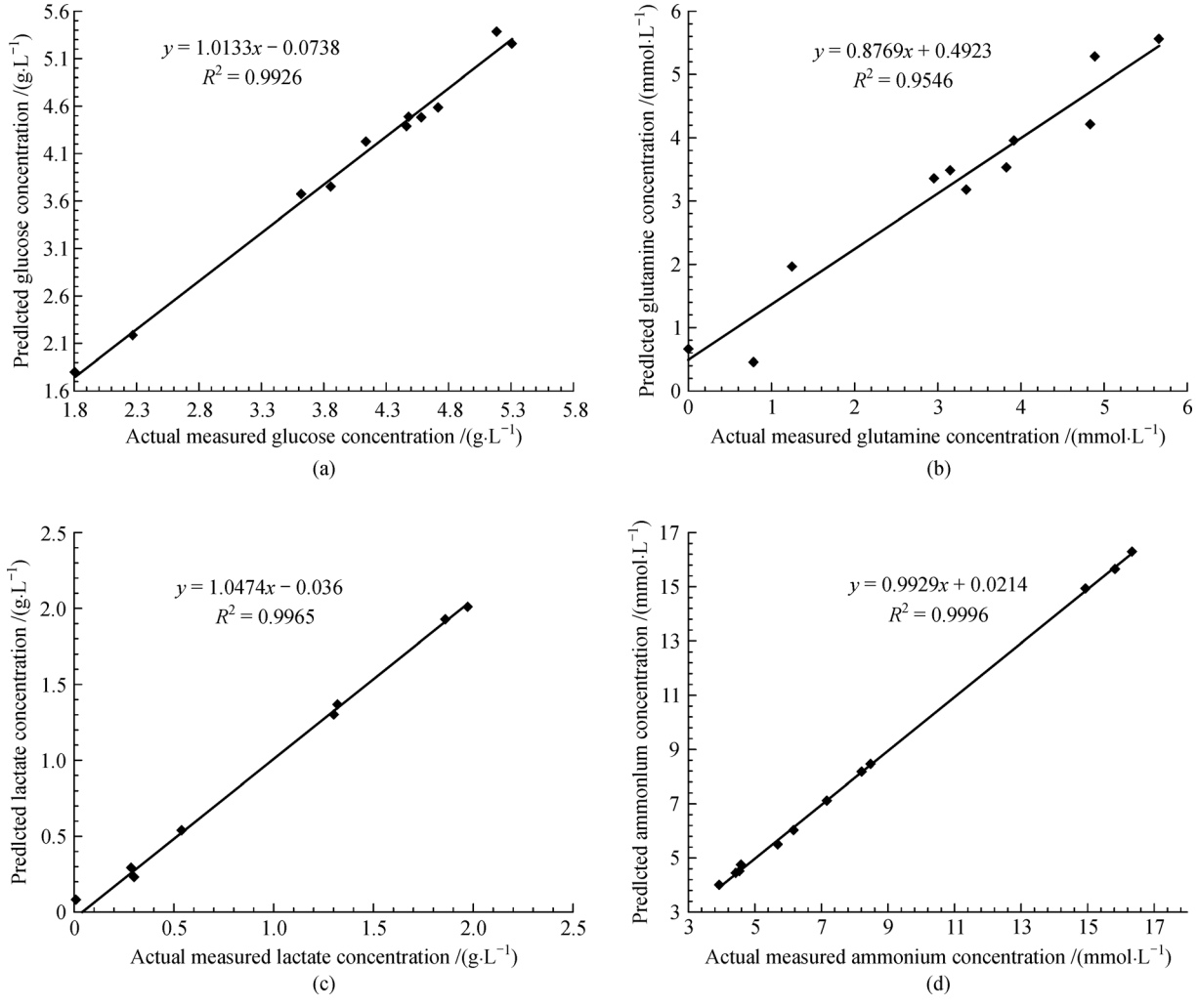

Fig. 4 Key metabolite concentration prediction results from the individual PLS calibration models for the bioreactor cell culture process for batch A and batch B. (a) Glucose, 4 PLS factors; (b) Glutamine, 7 PLS factors; (c) Lactate, 14 PLS factors; (d) Ammonium, 20 PLS factors

Table 5 Comparison of the global PLS model and the individual PLS model for concentration predictions of four key cell culture metabolites: PLS model calibration, cross validation, and testing results [a)]

Model type	Global PLS model for simultaneously prediction of four components				Individual PLS model for prediction of each component separately			
Nutrient component	Glucose /(g·L^{-1})	Glutamine /(mmol·L^{-1})	Lactate/(g·L^{-1})	Ammonium /(mmol·L^{-1})	Glucose /(g·L^{-1})	Glutamine /(mmol·L^{-1})	Lactate /(mmol·L^{-1})	Ammonium /(mmol·L^{-1})
Spectral data preprocessing algorithm	Mean center; 2nd derivative; variance scale				Mean center; 2nd derivative	Mean center; 2nd derivative; variance scale	Mean center; 2nd derivative	Mean center; 2nd derivative
User specified wave length regions	2000 to 600, single point baseline correction (B/L) @ 1300				1200 to 1000, B/L @Zero	1780 to 1000, B/L@ Zero	1850 to 1000, B/L @ Zero	1400 to 750, B/L@ Zero
Nutrient data treatment	Mean center; variance scale				Mean center	Mean center; 2nd derivative; variance scale	Mean center	Mean center
Number of latent variables	7				4	7	14	20
RMSEC	0.278	0.912	0.324	1.43	0.134	0.58	0.0559	0.0973
RMSECV	0.615	1.56	0.445	2.46	0.168	0.972	0.282	1.41
RMSEP	0.679	1.06	0.586	2.5	N/A	N/A	N/A	N/A
R^2 (Cumulative)	0.898	0.751	0.851	0.918	0.979	0.894	0.995	0.999
$R^2_{\text{calibration}}$	0.8979	0.7512	0.8506	0.9184	0.9789	0.8943	0.9950	0.9995
$R^2_{\text{cross validation}}$	0.7179	0.5047	0.8428	0.8669	0.9830	0.8155	0.9009	0.9447
R^2_{testing}	0.6978	0.7037	0.7177	0.7230	N/A	N/A	N/A	N/A

a) RMSEC: root-mean-square error of calibration; RMSECV: root mean square error of cross-validation; RMSEP: root-mean-square error of prediction. R_2 cumulative is cumulative explained variance by each factor for each response variable in the PLS model

Table 6 Effect of number of PLS latent variables on the global PLS model performance matrix after FTIR spectra had been subjected to the 2nd derivative preprocessing algorithm. (Data treatments for the FTIR spectra block including mean center, variance scale, 2nd derivative, user specified regions: 2000 to 660 cm^{-1}, single point baseline @ 1300 cm^{-1}; data treatments for the component concentration block including mean center and variance scale; quantitative analysis: leave-one-out (38 samples) cross-validation; F test statistic: 95% limit = 1.13, warn if > 1.13)

Number of PLS latent variables selected	Model performance matrix	Glucose /(g·L^{-1})	Glutamine /(mmol·L^{-1})	Lactate /(g·L^{-1})	Ammonium /(mmol·L^{-1})
4	RMSEC	1.12	0.537	0.367	2.28
	RMSECV	1.56	0.651	0.452	2.77
	RMSEP	1.27	0.683	0.587	1.91
	R^2 (Cumulative)	0.627	0.619	0.808	0.791
	Mahalanobis distance: 95% limit			0.71	
5	RMSEC	0.918	0.533	0.325	2.27
	RMSECV	1.59	0.699	0.448	2.8
	RMSEP	1.06	0.71	0.591	1.9
	R^2 (Cumulative)	0.748	0.625	0.849	0.792
	Mahalanobis distance: 95% limit			1.15	
6	RMSEC	0.913	0.321	0.324	2.13
	RMSECV	1.53	0.606	0.448	2.73
	RMSEP	1.05	0.693	0.586	2.05
	R^2 (Cumulative)	0.751	0.863	0.851	0.819
	Mahalanobis distance: 95% limit			1.64	

3.4.2 PLS model predictions of concentration profiles for multiple cell culture metabolites during cell culture process

In this work, the bioreactor cell culture process was continuously monitored in real time via FTIR probe. The FTIR spectra were acquired every 10 min. However, due to practical limitation, we did not have an auto-sampler installed which would allow us to conduct more frequent sampling from the bioreactor in an automatic fashion. Instead, manual sampling was performed for two to three times a day during the course of bioreactor cell culture process. As such, only limited nutrient data were available for building the PLS calibration models to correlate the FTIR spectra with the nutrient data. Therefore a wealth of real time FTIR spectra were left unused during the PLS model calibration process due to limited availability of experimental data of metabolites concentrations. However, this situation can be useful from PLS model testing perspective, as we could insert those unused real time FTIR spectra into the PLS calibration models to predict the concentrations of the four components at the time points when the real time FTIR spectra were taken. Because of the availability of wealthy unused process real time FTIR spectra, the concentration profile of each and every individual component can be predicted with much dense data density than the sporadically available NOVA data. Comparing these predicted concentration profiles with experimentally measured NOVA data can be used as a surrogated diagnosis tool, which helps to determine if the PLS calibration models established are suitable for intended use or not.

The time interval for the PLS model prediction was set as 10 min for both batch A and batch B. Given the time interval difference between the key component concentration prediction (10 min apart) and the experimental measurement via NOVA (at least 8 h apart), obviously the predicted concentration profiles have much higher data density than the available experimental measurement via NOVA. The predicted concentration results were compared to the experimental measurement data, as shown in Fig. 5 for Batch A and Fig. 6 for batch B. It can be seen that for both batch A and batch B, most of the predicted concentration profiles for glucose, glutamine, lactate, and ammonium agree well with the measured concentration data, except for batch B lactate where lactate's concentration was too low.

Previously it was shown that metabolites were either consumed or produced by the batch and fed-batch mode IgG3 hybridoma cultures [29]. Figure 2(b) is representative of the experimentally measured concentrations of metabolites via nutrient analyzer. It should be noted that the specific consumption between batches was different due to variability in cellular growth such as different starting nutrient concentrations for different batches; the overall nutrient consumption trend was consistent. For example, the starting nutrient concentrations for various batches were: 4 g/L glucose and 5 mmol/L glutamine for Fig. 2; 4.5 g/L glucose and 7 mmol/L glutamine for Figure 5;

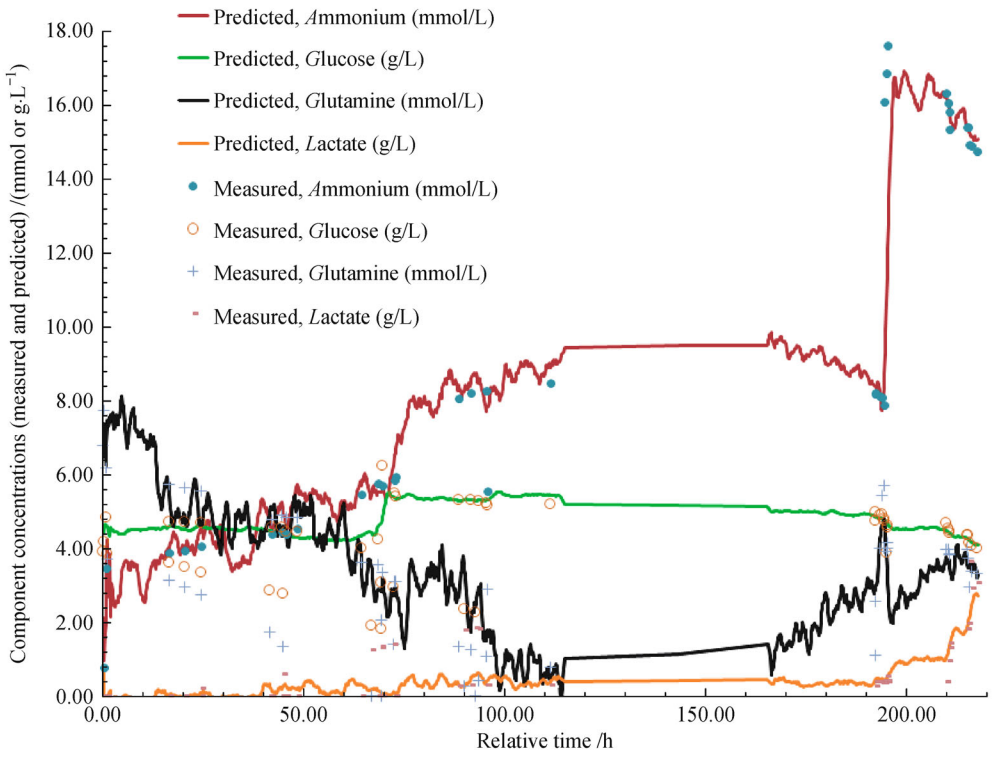

Fig. 5 The prediction results from individual PLS calibration models for each component in the bioreactor cell culture media for batch A in comparison with the experimental data. (a) Glucose; (d) Glutamine; (c) Lactate; (d) Ammonium (Note: batch monitoring disruption occurred due to unanticipated power supply disruption during the timeframe of 120−168 h. Extrapolated data were used to bridge the gap.)

4.25 g/L glucose and 4 mmol/L glutamine for Fig. 6. It is important to note that the goal of the experiments was not to generate identical cultures but to model nutrient concentrations based on FTIR data regardless of cell status. From multivariate model building perspective, including certain variability is essential. It can be seen that the global PLS calibration model can predict the key component concentrations of the in-process samples consistent with previous observations of their metabolism, with R^2 values of 0.828, 0.748, 0.807, and 0.886 for glucose, glutamine, lactate, and ammonium, respectively.

The global PLS model was based on simultaneously calibrating the same block of process FTIR spectra against the concentrations of glucose, glutamine, lactate, and ammonia. While this modeling method is relatively simple, the correlation coefficient for individual component could be limited due to lacking of focusing on component specific wave length range. To overcome this limitation, individual PLS models were established to correlate a block of process FTIR spectra within component specific wavelength regions with concentration data of a specific component. All FTIR spectra and nutrient data were subjected to mean center and 2nd derivative preprocessing. Additionally, glutamine concentration data was also subjected to variance scale preprocessing.

3.4.3 PLS model correlating the process FTIR spectra with the concentrations of the IgG3 antibody during the bioreactor cell culture process

A PLS model was constructed to correlate the process FTIR spectra with the concentrations of the IgG3 antibody of the in-process samples at various sampling points during the bioreactor cell culture process for the four scenarios: i) The 2nd derivative algorithm not applied, focused on wave length region of 1900 to 796 cm^{-1}; ii) The 2nd derivative algorithm not applied, focused on wave length region of 1277 to 796 cm^{-1}; iii) The 2nd derivative algorithm applied, focused on wave length region of 1900 to 796 cm^{-1}; and iv) The 2nd derivative algorithm applied, focused on wave length region of 1277 to 796 cm^{-1}. The datasets including FTIR spectra acquired at sampling time point and corresponding relative concentration of IgG3 antibody assessed by protein A chromatography for multivariate PLS model calibration and testing were provided in Table 7. The modeling results were summarized in Table 8. The prediction results of mAb concentrations at various time point during the bioreactor cell culture process for batch A and batch B were shown in Fig. 7(a) (for 7 PLS factors, specific wavelength regions from 1900 to 796 cm^{-1}) and Fig. 7(b)

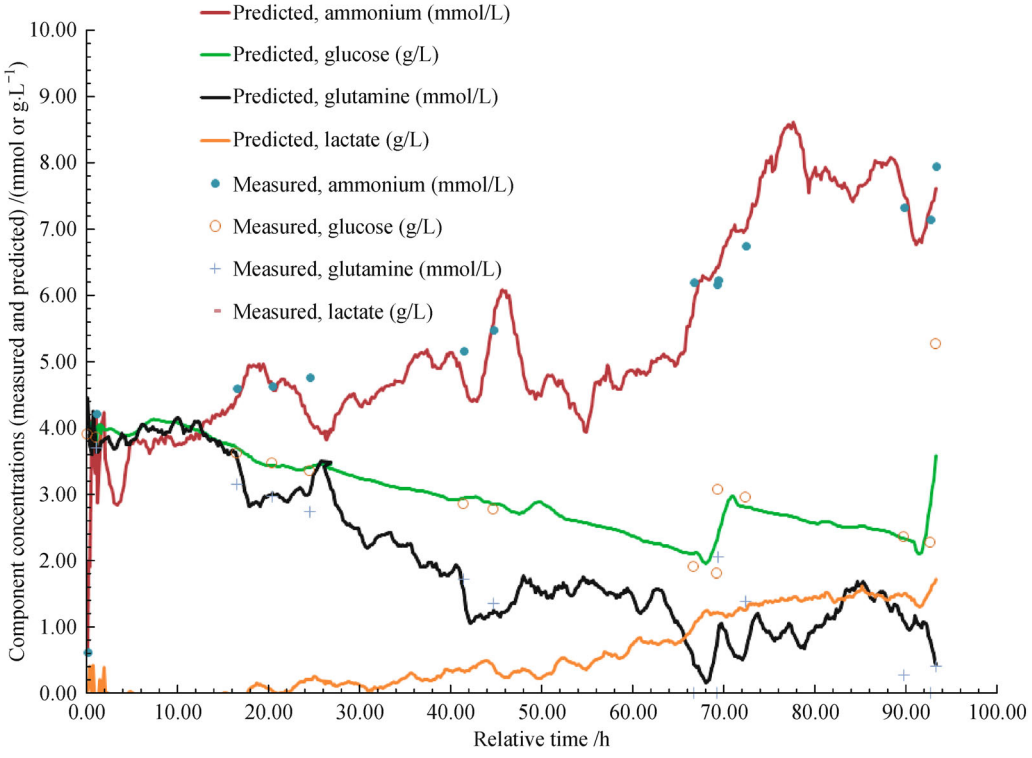

Fig. 6 The prediction results from individual PLS calibration models for each component in the bioreactor cell culture media for a batch B carried out at a different time in comparison with the experimental data. (a) Glucose; (d) Glutamine; (c) Lactate; (d) Ammonium

(for 9 PLS factors, specific wavelength regions from 1277 to 796 cm^{-1}).

It can see that the PLS modeling effort was quite successful as evidenced by the PLS model's figures of merits. In addition, Table 8 shows that the user specified wave length region can impact figures of merits of the FTIR-mAb PLS model substantially, especially for the scenarios where the 2nd derivative algorithm not applied. When the user specified wave length region is narrowed from (1900, 796) cm^{-1} to (1277, 796) cm^{-1}, the number of latent variables is increased from 7 to 9, which means the PLS model's complexity is increased a bit. However, most figures of merits for the PLS model (including RMSEC, RMSECV, and RMSEP) are improved substantially. The R^2 (Cum) value is increased from 0.995 to 0.999, too. This improvement is probably due to the fact that user specified range of (1277, 796) cm^{-1} covers the most characteristic spectral range. That is, based on the literature [23] of the FTIR spectra (1850–500 cm^{-1}) of IgG1–IgG4 in 20 g/L solutions, several major absorbance peaks which are unique to biomolecule IgG3, such as 1185, 1127, 1054 cm^{-1}, are covered and utilized.

When the 2nd derivative algorithm applied to the FTIR spectra prior to the PLS modeling, significant improvement was obtained as evidenced by the reduced number of PLS latent variables and improved model interpretability. The results were listed in the last two columns in Table 8 for comparison. Compared to their counterpart scenarios where the 2nd derivative algorithms not applied, apparently this improvement of model interpretability was achieved at some cost of other figures of merits, i.e., increased values for RMSEC/RMSECV/RMSEP and decreased $R^2_{\text{cumulative}}$ value.

3.5 Discussion

From a process control perspective [5], our observations on process trajectory and bioreactor cell culture media dynamics have certain important bioprocess implications. First, the real time process monitoring strategy can help to provide early indicator which enables us to make a go or no go process decision. For example, consumption of the energy source (glucose and glutamine) was predicted in real time during the cell culture process (Figs. 5 and 6). Under normal circumstances, the traditional feeding strategy based on in-process sampling plus off-line nutrient analysis is sufficient to feed the bioreactor and compensate for consumption in a timely manner. However, under aberrant situations such as microbial contamination where energy source consumption is unexpectedly high, it would be useful to get an early indicator to terminate the culture. The pace of the availability of the glucose and glutamine

Table 7 Calibration dataset and testing dataset for correlating process FTIR spectra with the mAb concentration at various time points during bioreactor mAb IgG3 cell culture process

Batch #	$T_{\text{spectra acquired}}$ /h	Spectra included in calibration?	Spectra included in testing?	mAb concentration /(mg·mL^{-1})
A	0.38	Yes		0
A	0.88	Yes		0
A	16.43	Yes		0
A	20.43	Yes		0
A	24.43	Yes		0.7843
A	41.93	Yes		2.22
A	44.43		Yes	3.16
A	48.43	Yes		4.34
A	69.43	Yes		9.07
A	72.43	Yes		10.3
A	88.42	Yes		12.3
A	91.42		Yes	12.4
A	95.25	Yes		11.9
A	111.42	Yes		12
B	0.05	Yes		0
B	1.05	Yes		0
B	16.55		Yes	0
B	20.38	Yes		0
B	24.55	Yes		0
B	41.38	Yes		0
B	66.72	Yes		0.62
B	69.22	Yes		0.67
B	69.38		Yes	0.66
B	72.38	Yes		0.74
B	89.88	Yes		2.3
B	92.72		Yes	2.32
B	93.38	Yes		2.31

concentrations in the bioreactor media provided via traditional sampling and analysis strategy may be too slow to direct timely feeding. Real time monitoring and prediction of the bioreactor cell culture nutrients can enable the direct timely feeding. As shown in Fig. 2(b), the modeled culture rapidly depleted glutamine well before the 96 h off-line sampling time point. The model-based prediction could have triggered a bolus of glutamine or feed it to a target concentration. On the other hand, by monitoring the glucose trend more effectively to a 1 g/L target concentration, we could have achieved more rational feed by skipping a 72 h feed and waited until 96 h to feed, or triggered an automated feed any time before or after 96 h. Therefore, there is no need to "break the seal" and draw a sample, no need to keep the nutrient analyzer calibrated or have a technician process samples. Eventually, we envision that a real time PAT data-loop would be a prediction populated from modeling software via OPC (object linking and embedding (OLE) for process control) to realize real time feedback/feed forward process control as highlighted in the FDA's PAT Guidance [2]. Furthermore, the PLS regression model which links process FTIR spectra with mAb yield can provide predicted mAb yield based on process FTIR spectra in real time. This information can be used to inform the process engineer and scientists regarding whether the bioreactor cell culture process is performed as expected or it may require certain process control action to remedy the situation.

As discussed in the introduction section, it has been recognized that it is challenging to develop an online real time monitoring strategy for a complicated yet dynamic bioreactor cell culture process. Normally, to handle the complexity of bioreactor culture and unique characteristics of each cell line, it is necessary to create process-specific multivariate real time monitoring models. In this work, PLS models were generated via a stepwise approach, i.e.,

Table 8 Comparison of the PLS models' figures of merits for correlating process FTIR spectra with the absolute concentrations of IgG3 antibody at various time points during the bioreactor cell culture process

Model type	PLS	PLS	PLS	PLS
Protein A measured result	mAb	mAb	mAb	mAb
FTIR spectral data preprocessing algorithm	Mean center	Mean center	Mean center; 2nd derivative	Mean center; 2nd derivative
User specified wave length regions	1900 to 796 cm^{-1}, Single point baseline correction @1900 cm^{-1}	1277 to 796 cm^{-1}, Single point baseline correction @1277 cm^{-1}	1900 to 796 cm^{-1}, Single point baseline correction @1900 cm^{-1}	1277 to 796 cm^{-1}, Single point baseline correction @1277 cm^{-1}
mAb data treatment	Mean center	Mean center	Mean center	Mean center
Method of determining PLS latent variables	Minimum PRESS	Minimum PRESS	Minimum PRESS	Minimum PRESS
Number of latent variables	7	9	4	4
Number of data points for calibration	22	22	22	22
Number of data points for testing	5	5	5	5
RMSEC /($g \cdot L^{-1}$)	0.335	0.12	1.28	0.842
RMSECV /($g \cdot L^{-1}$)	1.14	0.893	1.93	1.68
RMSEP /($g \cdot L^{-1}$)	1.2	0.698	3.48	1.44
$R^2_{Cumulative}$	0.995	0.999	0.922	0.966
$R^2_{calibration}$	0.995	0.999	0.922	0.966
$R^2_{testing}$	0.943	0.981	0.519	0.918
$R^2_{crossvalidation}$	0.968	0.980	0.903	0.930

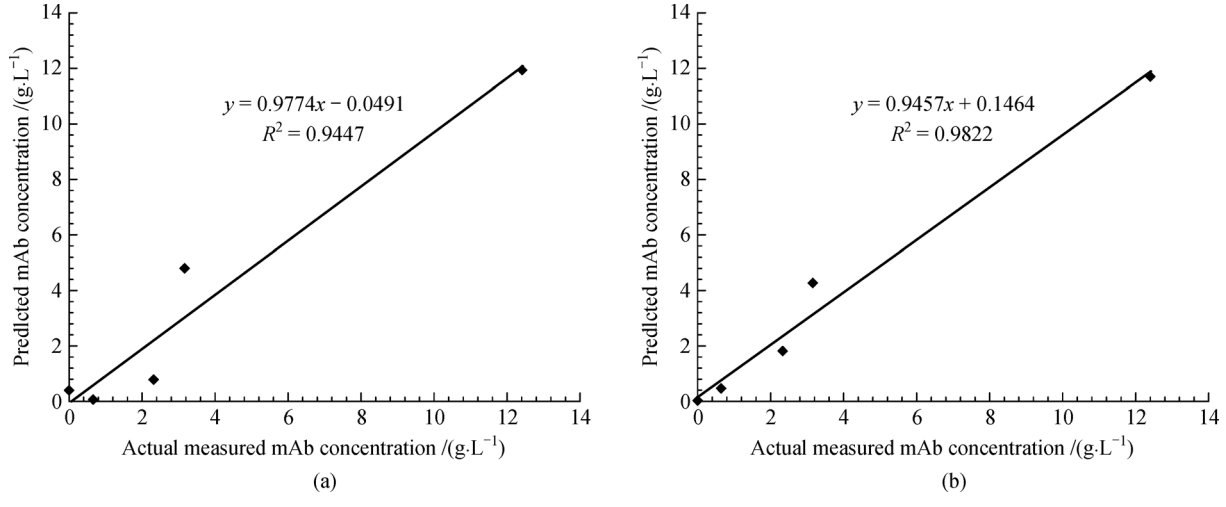

Fig. 7 The prediction results from PLS calibration model for mAb concentrations produced at various time points during the bioreactor cell culture process for batch A and batch B. (a) 7 PLS factors, model built based on specific regions from 1900 to 796 cm^{-1}, region to single point baseline @ 1900 cm^{-1}. (b) 9 PLS factors, model built based on specific regions from 1277 to 796 cm^{-1}, region to single point baseline @ 1277 cm^{-1}

first studying the simple binary system; then looking at the medium supplement system; afterwards examining the real bioreactor cell culture system. This stepwise approach allowed us to proof the concept of real time process monitoring via FTIR spectroscopy in a rational step-by-step fashion. Previously it was demonstrated that this stepwise approach is useful for developing an online real time monitoring strategy for a dynamic multicomponent pharmaceutical anti-solvent crystallization system [31]. In addition, this rational stepwise approach is proven useful from a process understanding perspective. Process understanding is an essential element highlighted in both FDA's PAT Guidance [2] and ICH Q8(R2) [32].

No attempt was made to develop a calibration model which includes the data from all of the three systems together. If the purpose of the modeling was to generate a model applicable to all three systems together, a different experimental design would be needed, in which case the

model would be more complicated than the ones presented in this work.

While this work exemplified the possibility of developing an integrated PAT approach for understanding the process dynamics of a lab scale bioreactor model mAb IgG3 cell culture process, some humility and caution is warranted. In terms of in-depth understanding of the complications of chemical, biological and bio-chemical characteristics and their interactions of a bioreactor cell culture process, there are plenty of space worthy for further investigation. Based on this work, it is important to note that: i) PLS regression has the desirable property that the precision of the model parameters can be improved with the increasing number of relevant variables and experimental measurements; ii) As the first step, this work illustrates an integrated PAT approach for real time monitoring of a lab scale bioreactor mAb IgG3 cell culture process. In the meanwhile, it opens up opportunities for further improvement. For example, we could certainly add more real time bioreactor cell culture process data from other batches along with more offline nutrient data for PLS model refinement. In addition, improving the ReactIR system including its FTIR probe design would be beneficial. Furthermore, developing innovative algorithms for real time process control could help to realize the ultimate goal of PAT, etc.

Some of the perceived benefits or potential advantages for the inline FTIR-based PAT strategy could include: i) Reducing nutrient sampling and analysis costs; ii) Reducing sampling associated risks of microbial contamination due to sampling; iii) Enhancing process control with higher density and more precise nutrient data; and iv) Facilitating process automation or terminating aberrant cultures. In summary, by adopting a rational stepwise approach, the technical feasibility of real time monitoring of the bioreactor cell culture process via FTIR spectroscopy has been demonstrated in this work.

4 Conclusions

The technical feasibility of using FTIR spectroscopy to track and monitor four key cell culture metabolites (including glucose, glutamine, lactate, and ammonia) and mAb yield of a bioreactor model mAb IgG3 cell culture process in real time was demonstrated via a rational stepwise approach. The level of complexity of the experiments was increased incrementally. The simple binary system of individual metabolite dissolved in water was taken for the 1st case where no dynamic consumption is expected. The stepwise addition of media to the uninoculated bioreactor was taken as the 2nd case to mimic the bioreactor charging process and test the sensitivity and specificity of the FTIR spectroscopy. As the 3rd yet most complicated case, the FTIR spectroscopy was applied to the actual bioreactor cell culture process for real time process monitoring, during which nutrient dynamics (such as energy source is depleted and waste product is accumulated during the course) is expected. Offline determinations of in-process nutrients and mAb yield were used to correlate with real time process FTIR spectra to establish multivariate PLS calibration models. The technical feasibility of using FTIR spectroscopy for characterization and quantification of individual media components was demonstrated in the 1st case. In addition, the capability of using FTIR spectroscopy to follow the addition of three components (glucose, glutamine, and lactate) to the un-inoculated bioreactor in real time sensitively and precisely was demonstrated via the 2nd case. For the 3rd case, the multivariate calibration models established show promise in predicting the concentration profiles of glucose, glutamine, lactate, ammonia, and mAb yield in real time during the entire course of the bioreactor cell culture process. Our data analysis shown that applying the 2nd derivative pre-processing algorithm to the FTIR spectra helps to reduce the number of PLS latent variables needed significantly and thus simplify the interpretation of the PLS models.

Finally, the important bioprocess implications of our work were discussed from process understanding and process control perspectives, such as providing early indicator which enables us to make a go or no go process decision, etc. In summary, this work illustrates that combining real time FTIR process monitoring, bioreactor in-process sample characterization, and process chemometrics can accelerate understanding cell culture process dynamics and provide vital information for enhanced process control.

Acknowledgements This work was financially supported by FDA Center for Drug Evaluation and Research (CDER) Regulatory Science and Review (RSR) funding RSR-12-42, RSR-13-32, FDA CDER Office of Testing and Research (OTR) Funding Wu-12-PAT, and FDA CDER OBP's PAT Critical Path project 1500. The unmatched technical support from Mettler-Toledo AutoChem team including but not limited to Vaso Vlachos, Paul Scholl, Simon Rea, and Jack Sue in the aspects of ReactIR probe autoclaving and ReactIR system optimization is greatly acknowledged. H. Wu wishes to thank Dr. Jennifer Maguire, Dr. Rapti Madurawe, Dr. Christine Moore, Dr. Vincent Vilker (retired), Dr. Patrick Faustino, Dr. Lucinda Buhse, and Dr. Lawrence Yu at the Office of Process and Facilities (OPF), Office of Pharmaceutical Quality (OPQ), Center for Drug Evaluation and Research (CDER), FDA for their management supports during the preparation and finalization of this manuscript. H. Wu is grateful to Prof. Jingkang Wang, Dr. Yingjin Yuan, Dr. Zhenhua Li at Tianjin University for the kind invitation, to Dr. Xiaowen Zhu for prompt communication and hard work, and to Yanni Li and Luli Cheng for their editorial assistance.

Appendix

For details on the exact binary mixtures that were characterized with FTIR, please refer to Appendix Tables A1–A4. For the measurement and prediction results of glucose, lactates, and ammonium during the stepwise addition processes, please refer to Appendix Tables A5–A7.

Table A1 Accumulated glucose concentrations in the binary system after each dose

Dose #	Amount of glucose added /mg	Calculated glucose concentration /(g·L^{-1})
1	350	0.7
2	350	1.4
3	350	2.1
4	350	2.8
5	350	3.5
6	350	4.2
7	350	4.9
8	350	5.6
9	350	6.3
10	350	7
11	1750	10.5
12	1750	14
13	3500	21

Table A2 Accumulated glutamine concentrations in the binary system after each dose

Dose #	Amount of glutamine added /mg	Calculated Concentration /(mmol·L^{-1})
1	51	0.872
2	51	1.396
3	51	2.094
4	51	2.792
5	51	3.490
6	51	4.188
7	51	4.886
8	51	5.584
9	51	6.282
10	51	6.932
11	255	10.469
12	255	13.959
13	511	20.953

Table A3 Accumulated lactate concentrations in the binary system after each dose

Dose #	Amount of Lactic acid added /mg	Calculated lactate concentration /(mmol·L^{-1})
1	31.9	0.708
2	31.9	1.417
3	30.5	2.094
4	32.1	2.806
5	31.9	2.837
6	30.4	4.190
7	30.8	4.873
8	32	5.584
9	30.6	6.263
10	32.7	6.989
11	159.4	10.528
12	156.2	14.067
13	315.2	21.066

Table A4 Accumulated ammonium concentrations in the binary system after each dose

Dose #	Amount of ammonium chloride added /mg	Calculated ammonium concentration /(mmol·L^{-1})
1	26.7	0.998
2	26.7	1.997
3	26.7	2.995
4	26.7	3.993
5	26.7	4.991
6	26.7	5.990
7	26.7	6.988
8	53.4	8.985
9	26.7	9.983
10	133.5	14.974
11	133.5	19.966
12	267	29.949

Table A5 Accumulated glucose concentrations in the medium during the stepwise addition of glucose—measurement and prediction

Dose #	Calculated glucose concentration, C_{cal}/(g·L^{-1})	Measured glucose concentration by Nova, C_{Nova} /(g·L^{-1})	Data preprocessing algorithms applied	PLS modeling results based on C_{cal}	PLS modeling results based on C_{Nova}
1	6	N/A	Spectral block: User specified region: 1900 to 900 cm^{-1}; single point B/L @ Zero; mean center. Component concentration block: Mean center	Calibration model R^2 = 0.9954 Leave-one-out cross validation model R^2 = 0.5785 PLS latent variables based on RMSEC *vs.* factors: 3 factors	Calibration model R^2 = 0.9996 Leave-one-out cross validation model R^2 = 0.3680 PLS latent variables based on RMSEC *vs.* factors: 3 factors
2	6	N/A			
3	6.04	4.12			
4	6.08	4.4			
5	6.58	5.37			
6	7.07	5.80			
7	7.56	6.38			
8	8.04	6.95			

Table A6 Accumulated lactate concentrations in the medium during the stepwise addition of lactic acid—measurement and prediction

Dose #	Calculated lactate concentration, C_{cal}/(g·L^{-1})	Measured lactate concentration by Nova, C_{Nova}/(g·L^{-1})	Data preprocessing algorithms applied	PLS modeling results based on C_{cal}	PLS modeling results based on C_{Nova}
1	0	0	Spectral block: User specified region: 1900 to 900 cm^{-1}; single point B/L @ Zero; mean center. Component concentration block: Mean center	Calibration model R^2 = 0.9988 Leave-one-out cross validation model R^2 = 0.8638 PLS latent variables based on RMSEC *vs.* factors: 3 factors	Calibration model R^2 = 0.9982 Leave-one-out cross validation model R^2 = 0.9230 PLS latent variables based on RMSEC *vs.* factors: 3 factors
2	4	0.59			
3	8	1.16			
4	12	1.53			
5	16	1.84			
6	20	2.16			

Table A7 Accumulated ammonium concentrations in the medium during the stepwise addition of ammonium chloride—measurement and prediction

Dose #	Calculated ammonium concentration, C_{cal}/(mmol·L^{-1})	Measured ammonium concentration by Nova, C_{Nova}/(mmol·L^{-1})	Data preprocessing algorithms applied	PLS modeling results based on C_{cal}	PLS modeling results based on C_{Nova}
1	0	0	Spectral block: User specified region: 1900 to 900 cm^{-1}; single point B/L @ Zero; mean center. Component concentration block: Mean center	Calibration model R^2 = 0.7677; Leave-one-out cross validation model R^2 = 0.6822; PLS latent variables based on RMSEC *vs.* factors: 1 factor	Calibration model R^2 = 0.6956; Leave-one-out cross validation model R^2 = 0.5196; PLS latent variables based on RMSEC *vs.* factors: 1 factor
2	2	2.99			
3	4	4.98			
4	6	8.24			
5	8	9.27			
6	10	12.95			

References

1. FDA/ICH. Guidance for Industry Q11 Development and Manufacture of Drug Substances. FDA, 2012. Available at: http://www.fda.gov/downloads/Drugs/Guidances/UCM261078.pdf (accessed on 07/28/2015)
2. FDA. Guidance for Industry PAT—A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance. FDA, 2004. Available at: http://www.fda.gov/downloads/Drugs/Guidances/ucm070305.pdf (accessed on 07/28/2015)
3. FDA. Advancing Regulatory Science at FDA: A Strategic Plan. FDA, 2011. Available at: http://www.fda.gov/downloads/scienceresearch/specialtopics/regulatoryscience/ucm268225.pdf (accessed on 07/28/2015)
4. Wu H Q, Khan M A, Hussain A S. Process control perspective for process analytical technology: Integration of chemical engineering practice into semiconductor and pharmaceutical industries. Chemical Engineering Communications, 2007, 194(6): 760–779

5. Read E K, Park J T, Shah R B, Riley B S, Brorson K A, Rathore A S. Process analytical technology (PAT) for biopharmaceutical products: Part I. Concepts and applications. Biotechnology and Bioengineering, 2010, 105(2): 276–284
6. Glassey J, Gernaey K V, Clemens C, Schulz T W, Oliveira R, Striedner G, Mandenius C F. Process analytical technology (PAT) for biopharmaceuticals. Biotechnology Journal, 2011, 6(4): 369–377
7. Simon L L, Pataki H, Marosi G, Meemken F, Hungerbühler K, Baiker A, Tummala S, Glennon B, Kuentz M, Steele G, Kramer H J M, Rydzak J W, Chen Z, Morris J, Kjell F, Singh R, Gani R, Gernaey K V, Louhi-Kultanen M, O'Reilly J, Sandler N, Antikainen O, Yliruusi J, Frohberg P, Ulrich J, Braatz R D, Leyssens T, von Stosch M, Oliveira R, Tan R B H, Wu H, Khan M, O'Grady D, Pandey A, Westra R, Delle-Case E, Pape D, Angelosante D, Maret Y, Steiger O, Lenner M, Abbou-Oucherif K, Nagy Z K, Litster J D, Kamaraju V K, Chiu M S. Assessment of recent process analytical technology (PAT) trends: A multiauthor review. Organic Process Research & Development, 2015, 19(1): 3–62
8. Wu H, Dong Z, Li H, Khan M A. An integrated process analytical technology (PAT) approach for pharmaceutical crystallization process understanding to ensure product quality and safety: FDA scientist's perspective. Organic Process Research & Development, 2015, 19(1): 89–101
9. Wu H, Heilweil E, Hussain A S, Khan M A. Process analytical technology (PAT): Quantification approaches in Terahertz spectroscopy for pharmaceutical application. Journal of Pharmaceutical Sciences, 2007, 97(2): 970–984
10. Wu H, Khan M A. THz spectroscopy: An emerging technology for pharmaceutical development and pharmaceutical process analytical technology (PAT) applications. Journal of Molecular Structure, 2012, 1020: 112–120
11. Wu H, Khan M A. Quality-by-design (QbD): An integrated approach for evaluation of powder blending process kinetics and determination of powder blending end-point. Journal of Pharmaceutical Sciences, 2009, 98(8): 2784–2798
12. Wu H, White M, Khan M A. An integrated process analytical technology (PAT) approach for process dynamics-related measurement error and process design space development for a pharmaceutical powder blending bed. Organic Process Research & Development, 2015, 19(1): 215–226
13. Wu H, White M, Khan M A. Quality-by-design (QbD): An integrated process analytical technology (PAT) approach for a dynamic pharmaceutical co-precipitation process characterization and process design space development. International Journal of Pharmaceutics, 2011, 405(1-2): 63–78
14. Wu H, White M, Berendt R, Foringer R, Khan M A. An integrated PAT approach for nucleation induction time measurement and nucleation mechanism assessment for a dynamic multi-component pharmaceutical antisolvent crystallization system. Industrial & Engineering Chemistry Research, 2014, 53(4): 1688–1701
15. Wu H, Lyon R C, Khan M A, Voytilla R, Drennen J III. Integration of near-infrared spectroscopy and mechanistic modeling for predicting film-coating and dissolution of modified release tablets. Industrial & Engineering Chemistry Research. Accepted for publication on May 19, 2015. http://pubs.acs.org/doi/abs/10.1021/ie504680m
16. Chung H, Arnold M A, Rhiel M, Murhammer D W. Simultaneous measurements of glucose, glutamine, ammonia, lactate, and glutamate in aqueous solutions by near-infrared spectroscopy. Applied Spectroscopy, 1996, 50(2): 270–276
17. Cervera A E, Petersen N, Lantz A E, Larsen A, Gernaey K V. Application of near-infrared spectroscopy for monitoring and control of cell culture and fermentation. Biotechnology Progress, 2009, 25(6): 1561–1581
18. Liu K Z, Mantsch H H. Simultaneous quantitation from infrared spectra of glucose concentrations, lactate concentrations, and lecithin/sphingomyelin ratios in amniotic fluid. American Journal of Obstetrics and Gynecology, 1999, 180(3): 696–702
19. Petersen N, Odman P, Padrell A E, Stocks S, Lantz A E, Gernaey K V. In situ near infrared spectroscopy for analyte-specific monitoring of glucose and ammonium in streptomyces coelicolor fermentations. Biotechnology Progress, 2010, 26(1): 263–271
20. Jose G E, Folque F, Menezes J C, Werz S, Strauss U, Hakemeyer C. Predicting mab product yields from cultivation media components, using near-infrared and 2D-fluorescence spectroscopies. Biotechnology Progress, 2011, 27(5): 1339–1346
21. Abu-Absi N R, Kenty B M, Cuellar M E, Borys M C, Sakhamuri S, Strachan D J, Hausladen M C, Li Z J. Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe. Biotechnology and Bioengineering, 2011, 108(5): 1215–1221
22. Budinova G, Salva J, Volka K. Application of molecular spectroscopy in the mid-infrared region to the determination of glucose and cholesterol in whole blood and in blood serum. Applied Spectroscopy, 1997, 51(5): 631–635
23. Petibois C, Cazorla G, Cassaigne A, Deleris G. Plasma protein contents determined by Fourier-transform infrared spectrometry. Clinical Chemistry, 2001, 47(4): 730–738
24. Ellis D I, Goodacre R. Metabolic fingerprinting in disease diagnosis: Biomedical applications of infrared and Raman spectroscopy. Analyst (London), 2006, 131(8): 875–885
25. Haris P I, Chapman D. Does Fourier-transform infrared-spectroscopy provide useful information on protein structures. Trends in Biochemical Sciences, 1992, 17(9): 328–333
26. Gross-Selbeck S, Margreiter G, Obinger C, Bayer K. Fast quantification of recombinant protein inclusion bodies within intact cells by FTIR spectroscopy. Biotechnology Progress, 2007, 23(3): 762–766
27. Stuart B H, Ando D J, eds. Biological Applications of Infrared Spectroscopy. New York: John Wiley & Sons, Inc., 1997, 191
28. Sellick C A, Hansen R, Jarvis R M, Maqsood A R, Stephens G M, Dickson A J, Goodacre R. Rapid monitoring of recombinant antibody production by mammalian cell cultures using fourier transform infrared spectroscopy and chemometrics. Biotechnology and Bioengineering, 2010, 106(3): 432–442
29. Read E K, Bradley S A, Smitka T A, Agarabi C D, Lute S C, Brorson K A. Fermentanomics informed amino acid supplementation of an antibody producing mammalian cell culture. Biotechnology Progress, 2013, 29(3): 745–753

30. Delwiche S R, Reeves J B III. A graphical method to evaluate spectral proprecessing in multivariate regression calibrations: example with Savtzky-Golay filters and partial least squares regression. Applied Spectroscopy, 2010, 64(1): 73–82
31. Wu H, Khan M A. Quality-by-design (QbD): An integrated process analytical technology (PAT) approach for real-time monitoring and mapping the state of a pharmaceutical co-precipitation process. Journal of Pharmaceutical Sciences, 2010, 99(3): 1516–1534
32. FDA/ICH. Guidance for Industry Q8(R2) Pharmaceutical Development. FDA, 2009. Available at: http://www.fda.gov/downloads/Drugs/Guidances/ucm073507.pdf (accessed on 07/28/2015)