APPLICATION NOTE

No. AN51449

Rapid, simultaneous monitoring of multiple components in mammalian cell culture with NIR spectroscopy

Cory Card, Brent Hunsaker, Tom Smith, Thermo Fisher Scientific (formerly HyClone), Logan, UT, USA
Jeffrey Hirsch, Thermo Fisher Scientific, Madison, WI, USA

The use of cellular physiology for production of target molecules has been practiced for centuries with early examples being the production of wine and beer via fermentation of yeast. Single-celled organisms (e.g., bacteria and yeast) and multi-celled organisms (plant or animal) can be harnessed to successfully produce otherwise chemically complex or low-yield materials. An example of this is the lactam ring-closure step in penicillin synthesis. Synthetically, this step is hard to achieve in the lab but is accomplished readily using fungi or bacteria containing isopenicillin N synthase (IPNS), a naturally occurring enzyme. Other examples are “lock-and-key” receptor complexes with perfect stereochemical specificity and large-scale protein scaffolds, all of which are accomplished readily using natural systems.

In the twentieth century, designer biological synthesis entered the age of mass production. One of the first efforts was the production of acetone, butanol and ethanol from starch using *Clostridium acetobutylicum*, a rod-shaped bacterium. This *Clostridium* was harnessed to produce, in a controlled fermentation, acetone, butanol and ethanol which are all important raw materials in the process for making Cordite (gunpowder) and Trinitrotoluene (TNT). The first large-scale manufacture of penicillin succeeded in time to treat soldiers wounded on D-Day in 1945. Large-scale industrial biological production also includes the million-ton annual production of citric acid from a fermentation with *Aspergillis Niger*. Biological manufacturing continues to evolve with large scale production of monoclonal antibodies, vaccines, biopolymers as well as techniques for post-translational modification (glycosylation).

Online analysis is an essential part of modern manufacturing protocols. Currently, biological manufacturers use multiple analytical tools to gauge critical process information such as osmolality, temperature and pH. Biological systems are extremely sensitive, complex matrices and require constant data feedback, analysis and control to achieve a successful reaction. Small variations in temperature, for example, will lead to cell death-scrapping an entire culture. Loss of nutrients in a batch-fed process or accumulation of waste will result in suboptimal reaction performance. Accumulation of excess ammonia, for example, has been found to interfere with protein glycosylation in CHO cells.

The use and acceptance of Fourier transform Near-infrared (FT-NIR) analysis as an online multicomponent analytical tool for the bioprocessing industry has grown substantially since the early 1990’s when some of the first studies were published. One of the main reasons for the popularity of FT-NIR as a measurement technique in cultures is the ability to provide rapid, accurate, high-resolution chemical analysis for major components of interest such as glucose or ammonia, even at low concentrations. While FT-NIR has been commonly used in other major industries like chemical, polymer and pharmaceutical for a number of years, adoption has been slower in biopharmaceuticals until very recently.

Near-infrared is a type of vibrational spectroscopy that uses overtones and combination bands to show the concentration of a particular analyte or analytes. Once light from an FT-NIR analyzer impinges on a sample, characteristic vibrational frequencies are absorbed by various molecular species like glucose or water, resulting in a unique spectrum. Collecting multiple spectra over time and combining them with known analyte concentrations allows the user to create a calibration using multivariate algorithms like Partial Least Squares (PLS). This type of analytical technology allows for accurate real-time predictions of multiple components within a single matrix, in this case, a broth.

The current study details experimental data from a cell culture of Human Embryonic Kidney (HEK 293) cells where a Thermo Scientific™ Nicolet™ Antaris™ FT-NIR Analyzer was used successfully to predict the four critical in-process components of the cell culture: glucose, lactate, ammonia and glutamine. This study, however, includes component predictions for pH and cell density in addition to the critical nutrient (glucose and glutamine) and waste (ammonia and lactate) calibrations.

Experimental

Human embryonic kidney-derived cells transformed with Ad5 DNA (HEK 293) were cultured in a chemically defined, serum-free culture medium (HyClone CDM4HEK 293) using a 10L working volume stirred-tank bioreactor at Thermo Fisher Scientific (formerly HyClone), Logan, UT. A recirculation loop from the bioreactor was used to continuously feed culture contents (cells and spent media) through an optical flow-cell. A temperature-controlled NIR transmission compartment housed the flow-cell throughout the entire process.

The culture was initiated with a cell seeding density of 3.0×10^5 viable cells/mL. The total process time was approximately 11 days. Cells reached a maximum population density of 8.5 to 9.0×10^6 viable cells/mL by day seven of the non-fed batch culture. Viability remained above 90% throughout most of the process until approximately day 10 as the cells reached senescence.

The Antaris FT-NIR analyzer (Method Development sampling system) was used to analyze the culture of HEK 293 cells. The analyzer was equipped with a liquid transmission sampling compartment with a heated-cell option and the cell temperature was kept at a constant 37° C. A 1.0 mm fixed-pathlength NIR flow-thru cell was purchased from Hellma®. The NIR spectral analysis range was from 4000 cm^{-1} to 10000 cm^{-1} at a resolution of 4 cm^{-1}. NIR data was collected every hour while simultaneously drawing aliquots for electrochemical analysis. The electrochemical analysis was performed on a Nova Biomedical BioProfile® 100 Plus analyzer. The values for pH, glutamine, glucose, lactate and ammonia from the BioProfile analyzer were then entered into a spreadsheet which was used in concert with the spectral data, to construct the near-infrared chemometric model. Cell density was also measured and fit in the model. A Beckman Coulter Vi-CELL® was used as the primary analysis technique for measuring cell density.

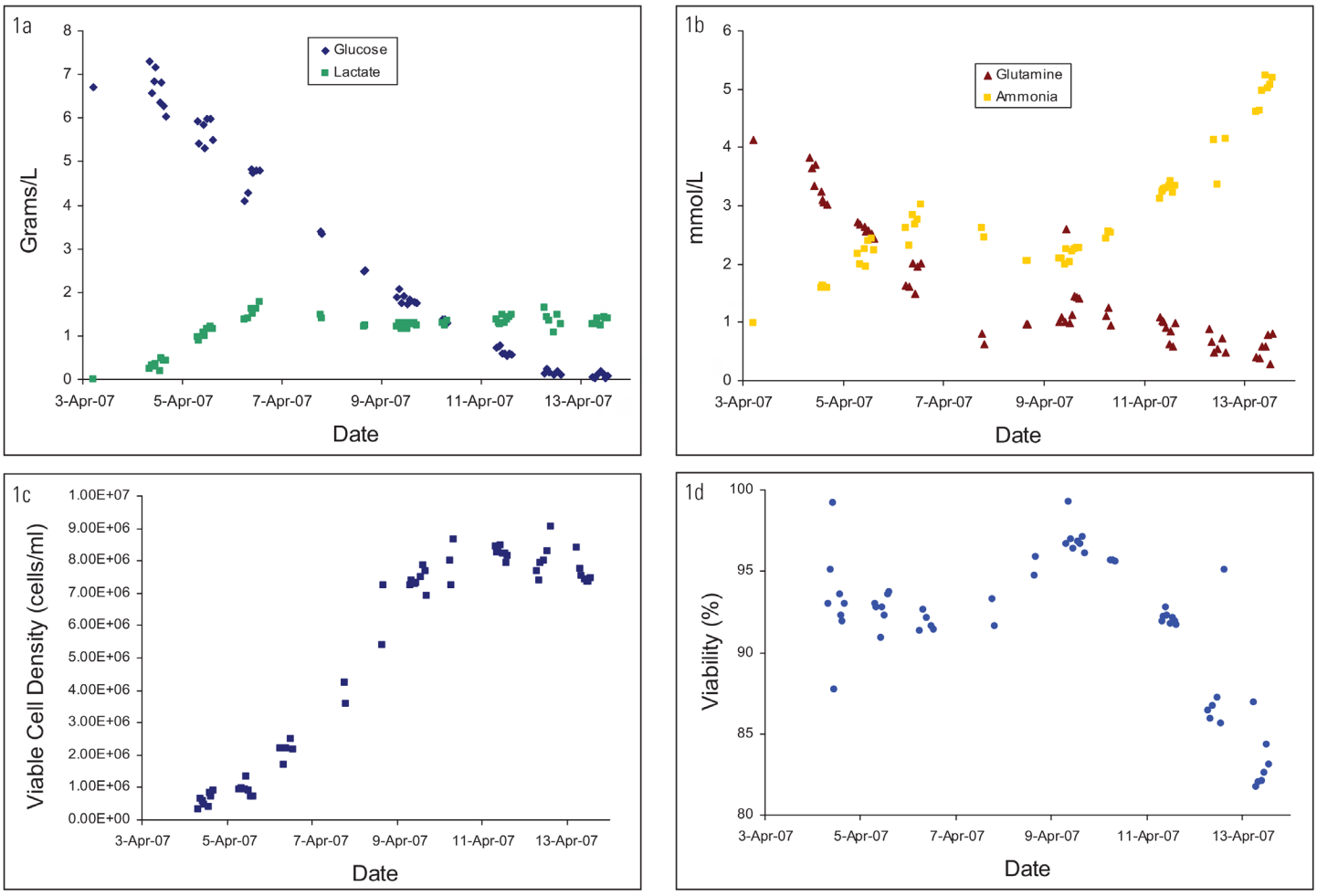
The chemometric model was built using Thermo Scientific™ TQ Analyst™ Software, the Thermo Scientific chemometrics package. The algorithm was Partial Least Squares (PLS1). Of the six components analyzed in this experiment, four (glucose, glutamine, pH and lactate) were treated together in a single chemometric model, while cell density and lactate were treated separately. First derivative with a Norris smooth (5,5) and mean-centering was the preferred pretreatment for glucose, glutamine, pH and ammonia. Lactate gave the best results with a slightly more aggressive smooth (11,5) as well as a combination of variance-scaling and mean-centering. Cell density was optimized with no spectral pretreatment except for a mild Savitsky-Golay 5,3 smooth and mean-centering. This implies that a good deal of the information on cell density is contained in baseline information, either offset or slope, or a combination of both. Other chemometric information including regions is contained in Table 1 where MC is mean centering and VS is variance scaling.

Results and discussion

In the current experiment, the major constituents of the HEK 293 culture medium were analyzed using potentiometric and amperometric techniques. Glucose and glutamine (nutrients), ammonia and lactate (wastes), and pH (a critical in-process variable) were all monitored using this methodology. The component concentration data versus time for the above components are shown in Figures 1a and 1b. Common y-axis scaling necessitated plotting glucose together with lactate and glutamine with ammonia.

Figures 1c and 1d show the changes in, respectively, viable cell population density in units of cells/ml and viability in units of per cent.

Trends in this data are commonly observed in batch-process cell culture as opposed to fed-batch cultures.



Figures 1a through 1d: Concentrations of glucose & lactate (1a), glutamine & ammonia (1b), cell density (1c) & cell viability (1d) throughout the course of the experiment.

Glucose and glutamine, the primary sources of carbon and nitrogen, respectively, are consumed as time progresses and ammonia, lactate and cell density increase. Glucose is consumed rapidly throughout the process until nearly completely consumed by day nine of the culture. Lactate, a product of glucose metabolism, accumulates early in the culture until reaching a peak level of approximately 1.8 g/L on day three. Thereafter, the cells utilize lactate as a carbon source so lactate level is decreased or maintained throughout the remainder of the process. An inverse correlation between pH and lactate level is also observed during the culture due to the increased presence of lactic acid.

Glutamine concentration decreases quickly during early culture. At 4.5 days post-seeding the level is maintained or even increased for several days. This increase in glutamine is further supported by the apparent consumption of glutamate (not shown) and ammonium beginning on day four. These metabolites are used in glutamine anabolism via the enzyme glutamine synthetase.

NIR spectra were collected throughout the culture process. The spectra are shown in Figure 2 on a common y-axis scale. The first salient feature of the spectral data is the major absorbances by water at approximately 5150 cm^{-1} and 7000 cm^{-1}. This behavior is expected from aqueous samples, which is why proper pathlength specification is critical to accurate measurements. In this case, the pathlength was fixed at 1.0 mm in order to keep the 7000 cm^{-1} within the photometric range of the NIR analyzer's InGaAs detector. The water combination band region around 5150 cm^{-1} was not used in this analysis as it is in a non-linear absorbance regime even at a small pathlength. The second significant spectral feature is a slight slope (tilt) and bias (offset) in the baseline. This can be caused by varying solids or cell density. The higher the cell density, the more NIR light scatters giving rise to the above-mentioned effects. This will come into play as the calibration for cell density is treated separately from the remaining components.

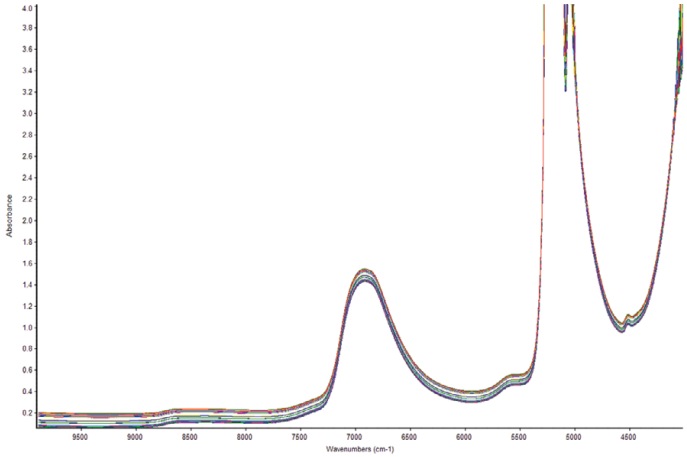


Figure 2: Common-scale FT-NIR spectra from the calibration set.

Component	Derivative	Smooth	Regions (cm-1)	MC/VS
Cell Density	None	SG 5,3	4250-4600	MC
pH	First	NO 5,5	4400-4650 5500-10000	MC
Lactate	First	NO 11,5	4200-4800 5477-7500	MC/VS
Glucose	First	NO 5,5	4400-4650 5500-10000	MC
Ammonia	First	NO 5,5	4400-4650 5700-6200	MC
Glutamine	First	NO 5,5	5700-6200	MC

Table 1: Data pre-processing parameters for the 6 components in the PLS regression model.

NIRS is a secondary technique, meaning NIR spectra must be correlated to primary or "known" data to be able to predict unknown samples. This correlation is embodied in a near-infrared calibration curve, a plot of sample data for each component with values from the primary method plotted on the x-axis and values from the NIR predictions on the y-axis. If the NIR analyzer is successful in predicting the analytes of interest, a linear relationship will exist between these data sets with a slope approaching 1 and a y-intercept approaching 0. FT-NIR spectra were correlated to data from the BioProfile analyzer to yield calibrations for lactate, glucose, glutamine, pH and ammonia. Cell density was correlated to a Vi-CELL automated cell counter.

The study of statistical correlations in chemical data sets is called chemometrics, a critical part of the science of NIR analytics. Chemometric analysis parameters can be broken down into two main categories, pretreatments and diagnostics. Data pretreatments are a way to increase the chances that a calibration will succeed. Common pretreatments include derivitization, smoothing and pathlength treatments. Diagnostics include the tools necessary to demonstrate or quantify the success of a calibration. The chemometric data treatments for the 6 components in the HEK 293 cell culture experiment are shown in Table 1. MC is mean centering and VS is variance scaling. NO is Norris smoothing and SG is Savitsky-Golay smoothing. Analysis regions are shown which were designed to avoid the highly absorbing water band around 5150 cm^{-1}.

The calibration model for the nutrients, waste and pH all used a first derivative pretreatment with mild smoothing. Cell density data was used without the aid of derivitization implying that most of the spectral response that contributes to the calibration model is contained in baseline effects. This is backed by loading spectra showing first loadings with little to no peak information. The only component that required more stringent smoothing was lactate. All component data was mean-centered. The pathlength treatment for all calibration models was set to constant as the path was physically controlled by a flow-thru cuvette set at 1.0 mm.

Once a calibration curve has been constructed, statistical methods are employed in order to quantify the correlation between the NIR spectra and the reference data. The main statistical parameters in this study are the correlation coefficient (R), root mean square error of calibration (RMSEC), root mean square error of cross validation (RMSECV) and factor analysis. Table 2 shows the calibration model output for these metrics. In addition, Table 2 contains component range information.

The calibration results and chemometric diagnostics show a strong correlation between the FT-NIR data and the primary results from both the BioProfile analyzer and the Vi-CELL cell counter. R values for all six components are between 0.926 and 0.995. The RMSECV calculation was performed as a "jack-knife" or a one-at-a-time cross validation showing a robust data model across components.

Component	Units	Range (Low)	Range (High)	Correlation Coefficient	Factors	RMSEC	RMSECV
Cell Density	Cells/ml	3.1 x 10^5	9.06 x 10^6	0.982	4	6.15 x 10^5	6.35 x 10^5
pH	pH	6.95	7.43	0.979	7	0.033	0.122
Lactate	g/L	0.19	1.77	0.926	4	0.140	0.201
Glucose	g/L	0.03	7.30	0.995	4	0.232	0.391
Ammonia	mmol/L	1.59	5.24	0.938	3	0.352	0.559
Glutamine	mmol/L	0.28	3.84	0.963	3	0.266	0.451

Table 2: Data from the TQ Analyst PLS1 calibration for all 6 cell culture components.

Calibration curves for two components, pH and glucose are shown, respectively, in Figures 3a and 3b.

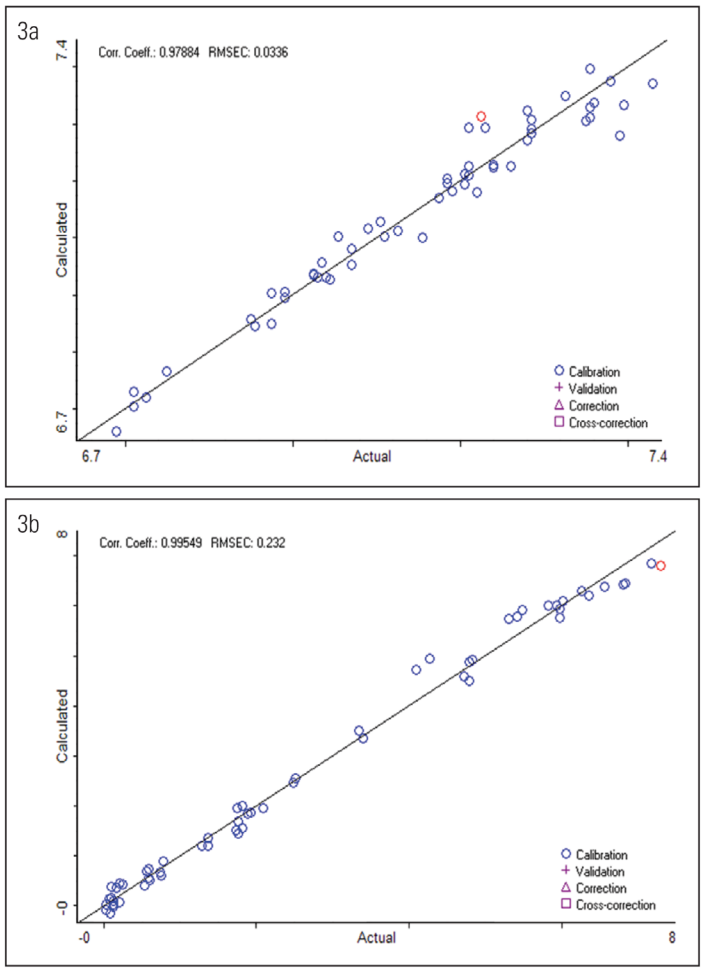
Another indicator of method performance in multivariate analysis is the Predicted Error Sum of Squares or PRESS plot. This is data that comes from the relationship between the error of cross validation, plotted on the y axis, versus the complexity of the model. In our case, model complexity is denoted by the number of factors. Too many factors can overfit a model while too few will underfit. The shape of the PRESS plot should be a struggle between competing contributions to the model, one from interference and one from estimation. The result should be plots that begin high, reach a minimum and then rise slightly from there. The two PRESS plots shown in Figures 4a and 4b are examples of this trend and indicate reasonable model performance. Figure 4a is the PRESS plot for glutamine and Figure 4b is for glucose.

Cost analysis

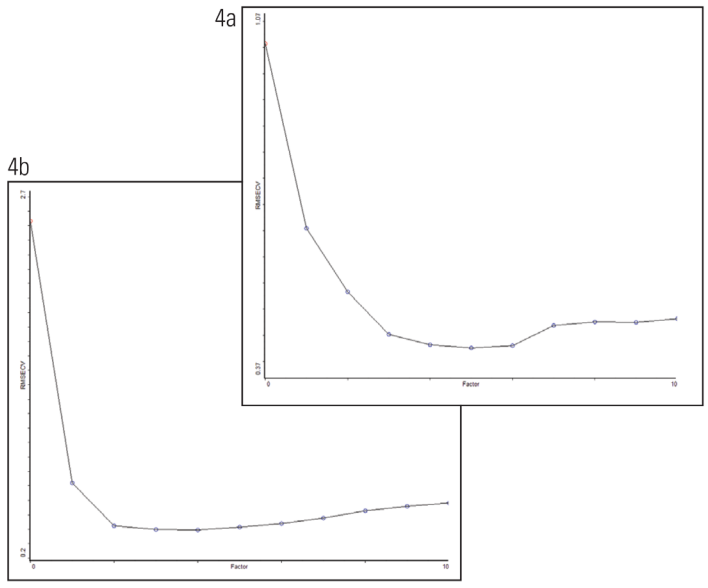
Table 3 shows the annualized cost of running and calibrating the BioProfile analyzer to acquire primary data on the listed components. Cell density and cell viability data were left out of this analysis. One of the main advantages of doing FT-NIR analysis is that it is fast, accurate, precise and creates no waste. In addition, once the method is calibrated, it takes only minutes a month to check the performance of the model, ensuring accurate NIR predictions. Whereas other techniques require not only constant maintenance, but a large cost investment for disposable materials such as calibration materials (reagent packs), membranes and electrodes.

Consumable	Cost Per Unit	Units/Yr	Annual Cost
Ammonium Electrode	$296.00	1	$296.00
Glutamine Electrode	$699.00	1	$699.00
Lactate Electrode	$699.00	1	$699.00
Glucose Electrode	$699.00	1	$699.00
pH Electrode	$509.00	1	$509.00
Glutamine Membranes	$51.00	12	$612.00
Glucose Membranes	$71.00	12	$852.00
Lactate Membranes	$71.00	12	$852.00
Reagent Packs	$291.00	52	$15,132.00
		Total	**$20,350.00**

Table 3: Annualized expenditures for primary analysis methodology for glucose, glutamine, pH and ammonia.



Figures 3a and 3b: Calibration curves for pH (3a) and glucose (3b).



Figures 4a and 4b: Predicted Error Sum of Squares, or PRESS, plots for glutamine (4a) and glucose (4b).

thermo scientific

Conclusion

The Antaris FT-NIR analyzer shows excellent correlations to the primary datasets from the BioProfile analyzer for glucose, glutamine, lactate, ammonia and pH and from the Vi-CELL for cell density. As these are all critical in-process components for judging the health of a mammalian cell culture, these data strongly suggest that near-infrared analysis of online cultures can readily provide rapid, accurate analysis in real-time.

Find out more at **thermofisher.com/nir**

ThermoFisher SCIENTIFIC

AN51449 0821 M