Research article

Journal of RAMAN SPECTROSCOPY

Received: 15 January 2015 Revised: 3 March 2015 Accepted: 5 March 2015 Published online in Wiley Online Library

(wileyonlinelibrary.com) DOI 10.1002/jrs.4686

Raman spectroscopy of complex defined media: biopharmaceutical applications

Gajendra P. Singh,[a]* Shireen Goh,[b] Michelangelo Canzoneri[c] and Rajeev J. Ram[a]

We demonstrate the detection of glucose and lactate concentrations with high accuracy in the supernatants of Chinese hamster ovary (CHO) cell culture, grown in shake flasks in batch fermentation mode, using Raman spectroscopy and explicit model-based classical least squares (CLS) algorithm. A deterministic Raman spectral library of pure components was created by acquiring Raman spectra from credible nutrient media constituents and CHO cell culture metabolites. Only analytes present with concentration above the instrument detection limit were included in this library. Residuals obtained after CLS analyses were used to identify missing components and to generate a revised library. An algorithmic sieve was thus construed to obtain an appropriate Raman spectral library from a complex chemical mixture that is well-defined but an industrial secret. High performance liquid chromatography (HPLC) was used to provide reference glucose and lactate concentrations. We demonstrate the detection of glucose and lactate concentrations in the supernatants of CHO cell culture using Raman spectroscopy and explicit model-based CLS analysis. An algorithmic sieve was construed to obtain an appropriate Raman spectral library from a complex chemical mixture that is well-defined but an industrial secret. HPLC was used to provide reference glucose and lactate concentrations. Copyright © 2015 John Wiley & Sons, Ltd.

Keywords: Raman spectroscopy; bioreactors; classical least squares; glucose detection; lactate detection

Introduction

Real time monitoring and control of bioprocesses are essential in improving their efficiency and reducing the cost of final product. With the advancement of biotechnology, bioprocesses are no longer restricted to the production of alcohols, organic and amino acids, and small biomolecules such as insulin, enzymes, and some antibiotics. Now, they are also a method of choice for using modified cell culture expression systems and the production of large biomolecules such as monoclonal antibodies that include some vaccines for human use.

Among the mammalian cell culture expression systems, the relatively high robustness of Chinese hamster ovary (CHO) cells in different bioreactor systems is well established. Furthermore, CHO cells are very efficient in post-translational glycosylation of proteins, making them closely resemble human glycosylation patterns in order to avoid or to reduce the immune responses after drug administration. Some examples of such proteins are follicle stimulating hormone, haemophilic FVIII, interferon-beta, tissue-plasminogen activator, and erythropoietin.[1]

To enhance the technologies available for monitoring and control of environmental parameters in bioreactors, several optical and spectroscopic modalities have been investigated as they are non-invasive and provide information in real time. Raman spectroscopy is one of the most promising techniques in this regard. It is based on the inelastic scattering of light and provides a biochemical fingerprint for the sample under investigation. High chemical specificity of Raman spectroscopy is the result of its measurement of vibrational energies in chemical bonds. Near infrared (NIR) laser excitation helps to minimize luminescence background generated either from the substrate or from the biological sample being probed. It also has higher penetration depth for biological materials and less tissue damage at higher excitation power compared with visible excitation. Miniaturization and integration of Raman technology is highly desirable, and in recent years, there has been an increase in efforts in this direction.[2]

Raman spectroscopy coupled with chemometrics has recently been used for real time monitoring and simultaneous prediction of multiple culture parameters including glutamine, glutamate, glucose, lactate, ammonium, viable cell density, and total cell density in CHO cell culture bioreactors.[3,4] It has also been used for the rapid identification, characterization, and quality assessment of complex cell culture media components used for industrial mammalian cell culture.[5] But, all of the studies have used implicit models, where reference measurements on a training sample set are performed to acquire information about the system to be investigated. A model is thus constructed, based on the training set, which implicitly accounts for any physical effects that influence the measured Raman spectra, allowing concentration estimation of unprocessed samples. However, the training samples are system specific and only model the particular system under investigation.

* *Correspondence to: Gajendra P. Singh, Physical Optics and Electronics Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.*
E-mail: gpsingh@mit.edu

a *Physical Optics and Electronics Group, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA*

b *Bioprocessing Technology Institute (BTI), 20 Biopolis Way, 06-01 Centros, Singapore 138668, Singapore*

c *Sanofi-Aventis Deutschland GmbH, Industrial Affairs, Biologics Industriepark Hoechst, Building D 706, Frankfurt am Main 65926, Germany*

They are not generally applicable to other systems. Implicit models do have utility and validity in bioprocess supervision applications, where the same bioreaction is to be repeatedly monitored for production purposes, but they have limited utility in bioprocess development applications where variations in operating conditions, growth rates, and medium composition are a necessary requirement. Explicit methods, based on physically modeling the system to be analyzed, are preferred for bioprocess development applications.[6] Explicit models have previously been used to estimate ethanol concentration in Baker's yeast fermentation,[7] and classical least-squares (CLS) fitting was used to estimate glucose and spiked quantities of glutamine, lactate, and ammonia, in filtered samples of a bioreactor.[8]

The industrial mammalian cell culture requires complex nutrient media components. Chemically defined media components are preferred to avoid unknown parameters and subsequently facilitate the regulatory approval of the product. The exact composition of the nutrient media in most cases is kept an industrial secret. Hence, it is not straightforward to use explicit methods such as CLS analysis for the prediction of metabolite concentrations.

In this article, we have shown that explicit model-based CLS algorithm can be used to detect mammalian (CHO) cell culture metabolites (glucose and lactate) with high accuracy using NIR Raman spectroscopy. An algorithmic sieve was construed and an appropriate Raman spectral library obtained from a complex chemical mixture that is well-defined but kept confidential by respective pharmaceutical industries. High performance liquid chromatography (HPLC) was used to obtain reference concentrations of glucose and lactate.

Materials and methods

Mammalian cells

Chinese hamster ovary cells obtained from Invitrogen Inc. were grown in a batch culture flask for 10 days in an incubator maintained with 5% CO_2 and 37 °C temperature. Similarly, CHO cells obtained from Sanofi-Aventis Deutschland GmbH Inc. were batch cultured for 13 days. The 2 ml supernatant was collected each day, in duplicate, and was stored in a freezer at −20 °C. Concentrations of glucose and lactate were estimated, in the supernatant, on an HPLC machine (Agilent Inc.), and the same sample was used to obtain Raman measurements immediately.

Confocal near infrared Raman spectroscopy system

A confocal Raman spectroscopy system was developed, consisting of fiber coupled external cavity diode laser at 830 nm (Ondax Inc.) for Raman excitation. A bandpass filter (Semrock Inc.) was used to clean the laser wavelength. A 100 μm thin fused silica window was used in the sample holder to minimize fluorescence background from optics. The excitation microscope objective (Nikon Inc.) used was 40 X in magnification and had 0.75 NA. Micrometer equipped x-y-z-translation stage (Thorlabs Inc.) helped to focus the sample and to maximize the sampling volume during Raman data acquisition. The 20 mM urea solution, 2 ml volume, was used as a calibration sample to find the sample holder position providing maximum Raman signal. The sample holder position was then locked firmly for the total course of the experiment. A 200 μm core multimode optical fiber (Thorlabs Inc) was used for the collection of Raman signal and also was served as a pinhole for confocal applications. A notch filter (Semrock Inc.) helped to block the excessive laser-back scattering light from optical elements and the sample. Collected Raman light was delivered to liquid nitrogen cooled charged coupled device camera (Princeton Instruments Inc.) through a dispersive spectrograph (Acton Inc.) containing 1200 grooves/mm grating. The resolution of the NIR Raman system was calculated as 8 cm^{-1} using the Raman peak of silicon at 520 cm^{-1}. The WINSPEC software (Princeton Instruments Inc.) was used to acquire Raman spectra and to convert into ASCII format. MATLAB (Mathworks Inc.) was used for further Raman data processing.

CLS analysis

Because the exact composition of CHO cell culture nutrient medium is kept confidential by respective pharmaceutical industries, we decided to select nutrient media constituents and CHO cell culture metabolites from a reliable literature. For example, Table 1 by Schröder *et al.*,[9] provides information in this regard and is used to develop the Raman spectral library. This table is shown in Fig. 1. We initially selected only those constituents from Fig. 1, which were known to generate Raman spectra (for example, the inorganic salts do not show Raman spectra as they dissociate into ions in an aqueous solution) and were present in concentration above the Raman instrument detection limit of ~1 mM. Because several of the amino acids mentioned in Table 1 have concentration reaching 1 mM, we selected the ones with high Raman cross section. Zhu *et al.*[10] provide in detail the Raman spectra of amino acids in solution. A deterministic Raman spectral library of pure components was thus obtained from credible nutrient media constituents and CHO cell culture metabolites and was used as a calibration routine for the CLS analysis. Phosphate buffered saline (PBS) was used to prepare all pure component solutions. Raman spectra were acquired from each day's supernatant. Polynomial fitting was not used to estimate and remove fluorescence background from Raman spectra before using the CLS analysis and performing concentration prediction for glucose and lactate. Cosmic rays were removed from the Raman spectra, and fifth order smoothing using Savitzky–Golay algorithm was performed using MATLAB (Mathworks Inc.) to suppress noise. The Raman spectra were normalized with respect to the CLS coefficient of water. It was also taken care that the concentration prediction of all basis spectra constituents generates non-negative numbers. A final Raman spectral library was thus conceived. Using CLS calibration routine developed with this Raman spectral library, concentrations of glucose and lactate were predicted. HPLC was used to obtain reference concentrations of glucose and lactate. Standard error was

Table 1. Comparison of explicit and implicit calibration models

Whelan *et al.*[3]			This work			
Partial least squares analysis (implicit)			Partial least squares (PLS) [leave-one-out cross validation (LOOCV)]		Classical least squares (CLS)	
	R^2	Standard error	R^2	Standard error	R^2	Standard error
Glucose	0.91	2.09	0.99	2.4	0.98	3.5
Lactate	0.99	11.49	0.99	3.1	0.99	5.2

Journal of RAMAN SPECTROSCOPY

Inorganic salts	Molar concentration (mM)	Carbohydrates	Molar concentration (mM)
CaCl2	1.9	d-Glucose	34
CuSO4·5H2O	4.81 × 10−6	l-Amino acids	
FeSO4·7H2O	1.20 × 10−3	l-Alanine	0.377
KNO3	9.02 × 10−4	l-Arginine·HCl	0.879
KCl	6.51	l-Asparagine·H2O	0.267
MgCl2	0.24	l-Aspartic acid	0.31
MgSO4	0.576	l-Cysteine·HCl·H2O	0.08
NaCl	52	l-Cystine·2HCl	0.349
NaHCO3	48.8	l-Glutamic acid	0.652
Na2HPO4	0.4	l-Glutamine	7.2
NaH2PO4·H2O	1.25	Glycine	0.52
Na2SeO3·5H2O	9.89 × 10−5	l-Histidine·HCl·H2O	0.28
ZnSO4·7H2O	1.20 × 10−3	l-Isoleucine	0.973
Vitamins and miscellaneous compounds		l-Leucine	1
d-Biotin	7.58 × 10−5	l-Lysine·HCl	1.04
dl-Pantothenic acid, calcium salt	1.09 × 10−2	l-Methionine	0.253
Choline chloride	7.44 × 10−2	l-Phenylalanine	0.491
Ethanolamine	0.02	l-Proline	0.537
Folic acid	1.21 × 10−2	l-Serine	0.52
Hypoxanthinea	0.1	l-Threonine	0.997
i-Inositol	8.79 × 10−2	l-Tryptophan	0.098
Linoleic acid	1.20 × 10−4	l-Tyrosine<comma> disc	0.569
Lipoic acid	4.07 × 10−4	l-Valine	1
Methotrexate (MTX)b	1 × 10−4 to 0.1	Peptones and proteins	
Nicotinamide	3.94 × 10−2	Fetuind	–
Phenol red	5.21 × 10−2	Insulin	–
Pluronic F-68c	–	holo-Transferrin	–
		Casein peptone soybean flour peptone	
Pyridoxal·HCl	2.36 × 10−2	brothe	–
Pyridoxine·HCl	1.21 × 10−4		
Riboflavin	1.32 × 10−3		
Sodium pyruvate	2.51		
Thiamine·HCl	1.46 × 10−2		
Thymidinea	0.016		
Vitamin B12	4.13 × 10−4		

Figure 1. List of constituents present in Chinese hamster ovary (CHO) cell culture nutrient media. Adapted from Schröder *et al.*.[9] Underlined constituents were part of the Raman spectral library used for classical least squares (CLS) calibration routine development.

calculated as the root mean-squared error of prediction, using Raman-predicted concentration values and HPLC-generated reference concentration values.

Results and discussion

The Raman spectral library, which was obtained after implementing the algorithmic sieve, is shown in Fig. 2. In the process of obtaining this library, some constituents that were mentioned as ~1 mM concentration in Fig. 1 such as L-leucine and L-valine were removed. Their presence made the concentration predictions worse and also the CLS fit to the supernatant Raman spectrum-produced residuals with structure. Removing any of the Raman spectra present in spectral library of Fig. 2 made the concentration predictions worse as did the addition of spectra from additional constituents. Thus, a final Raman spectral library for CLS analysis was conceived.

Raman spectra of L-arginine HCl, glucose, L-threonine, L-valine, sodium pyruvate, sodium bicarbonate, glycine, and sodium lactate were acquired at 20 mM concentration each in PBS. Raman spectra

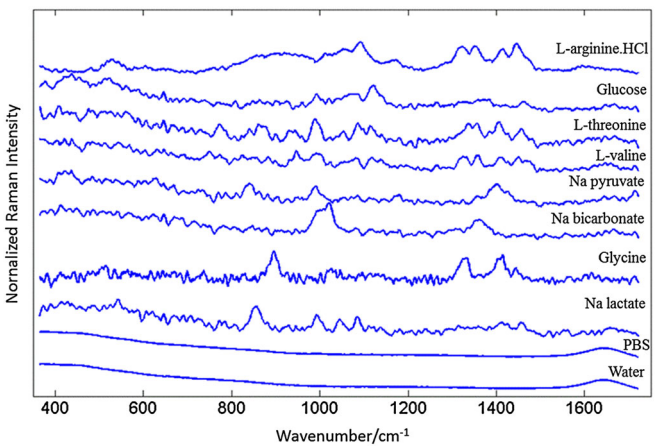

Figure 2. Raman spectral library consisting of pure components; some selected from Chinese hamster ovary (CHO) cell culture constituents mentioned in Fig. 1. PBS, phosphate buffered saline.

of PBS and water were also acquired and included in the Raman spectral library. In Fig. 2, the Raman spectrum of water has been subtracted from each of the solute Raman spectrum for better visualization. The main Raman peaks and bands observed in the spectra of pure components mentioned in Fig. 2 are listed in Table 2. A broad Raman peak at 1641 cm^{-1} was observed in the Raman spectrum of water, and a weak Raman peak at 992 cm^{-1} was observed in the Raman spectrum of PBS.

The detection limit of sodium lactate and D-glucose aqueous solutions is shown in Fig. 3, using the NIR Raman spectroscopy system. Here, we define the detection limit as noise equivalent concentration, i.e. the concentration at which $SNR = 1$. It was found that the detection limit of sodium lactate is better than 3 mM (obtained after extrapolation of curve), while that of D-glucose is better than 2 mM. The data in this figure were an average of three measurements, and the error bars were smaller than the shown size

Table 2. Main Raman peaks (Raman shift cm^{-1}) observed in the Raman spectra of pure components mentioned in Fig. 2. Water exhibited a broad Raman peak at 1641 cm^{-1} while phosphate buffered saline (PBS) at 992 cm^{-1}

Pure components from Fig. 2							
L-arginine HCl	Glucose	L-threonine	L-valine	Na pyruvate	Na bicarbonate	Glycine	Na lactate
—	435	—	—	—	—	—	—
—	516	—	—	—	—	—	—
530	—	—	—	—	—	—	—
—	—	—	750	—	—	—	—
—	—	771	—	—	—	—	—
—	—	—	—	840	—	—	—
—	—	—	—	—	—	—	855
—	—	861	—	—	—	—	—
—	—	—	—	—	—	897	—
919 (broad)	—	—	—	—	—	—	—
—	—	—	946	—	—	—	—
—	990	990	991	991	—	—	992
—	—	—	—	—	1027	—	—
—	—	—	—	—	—	—	1045
—	—	1055	—	—	—	—	—
—	1076	—	—	—	—	—	—
—	—	1085	1084	—	—	—	1085
1091	—	—	—	—	—	—	—
—	—	1112	—	—	—	—	—
—	1121	—	—	—	—	—	—
1172	—	—	—	—	—	—	—
1324	—	—	1324	—	—	—	—
—	—	1335	—	—	—	1333	—
1356	—	—	1358	—	1361	—	—
—	1374	—	—	—	—	—	—
—	—	—	—	1401	—	—	—
1415	—	1409	1412	—	—	1413	—
1453	1460	1456	1455	—	—	—	1456

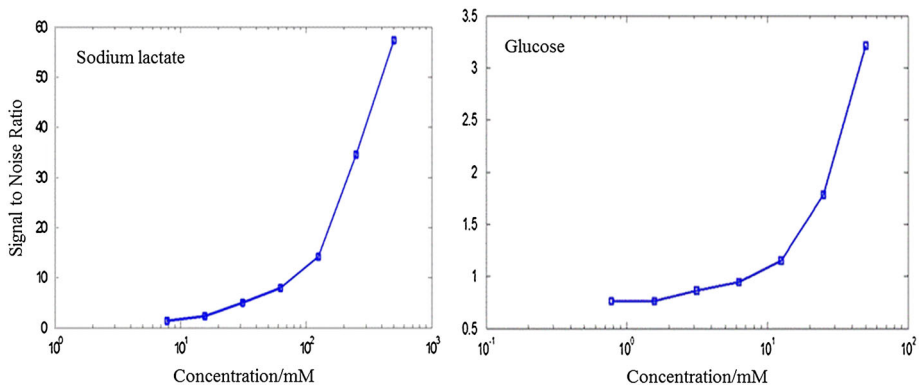

Figure 3. Detection limit of sodium lactate and D-glucose solutions measured using near infrared (NIR) Raman spectroscopy system. Error bars were smaller than the shown size of data points. For D-glucose, the mean of area under Raman peak at 1121 cm^{-1} was used for signal calculation, while for sodium lactate, the mean of area under Raman peak at 855 cm^{-1} was used.

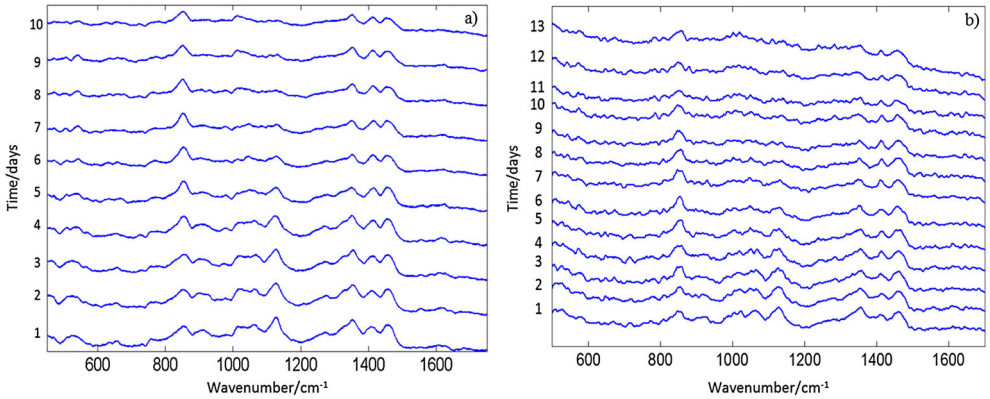

Figure 4. Raman spectra obtained from Chinese hamster ovary (CHO) cell culture supernatants using (a) Invitrogen cell line and (b) Sanofi cell line during 10 and 13 days of shake flask culture, respectively, in batch mode.

of data points. Each measurement consisted of ten spectral acquisitions. For D-glucose, the mean of area under Raman peak at 1121 cm^{-1} was used for signal calculation, while for sodium lactate, the mean of area under Raman peak at 855 cm^{-1} was used. Median standard deviation across the whole Raman spectrum of ten Raman spectra was used for noise estimation.

Figure 4(a) shows the Raman spectra obtained from Invitrogen CHO cell culture supernatants on each of the 10 days, while Fig. 4(b) shows the same for Sanofi CHO cell culture supernatants for 13 days. PBS Raman spectrum has been subtracted from each day's supernatant Raman spectrum for better visualization. The difference in Raman spectra on each day can clearly be observed. Main Raman peaks at 853 cm^{-1}, 900 cm^{-1} (broad), 1021 cm^{-1}, 1064 cm^{-1}, 1128 cm^{-1}, 1276 cm^{-1}, 1352 cm^{-1}, 1412 cm^{-1}, and 1458 cm^{-1} were observed.

The growth curves for Invitrogen and Sanofi CHO cell cultures respectively are shown in Fig. 5. The Sanofi CHO cell culture has a longer stationary phase because it has been optimized for antibody production.

Figure 6 shows the glucose and lactate concentration prediction in Invitrogen CHO cell culture supernatants using Raman spectroscopy and explicit model-based CLS algorithm. Comparison with reference concentration values obtained using HPLC is also shown. The experiment was performed three times, and average values of measurement are plotted. The error bars were equivalent to 2 mM concentration for Raman measurements and 0.5 mM concentration for HPLC measurements. The time required for glucose and lactate concentration measurements using HPLC was about 45 min per sample, after the HPLC was calibrated with the correct measurement column. Calibration of HPLC with a new column required almost 2 h. Because the same column could not be used for glucose/lactate and amino acids, for example, HPLC was found to be labor intensive. On the other hand, time required for Raman measurements was about 8 min per sample after the charged coupled device camera of NIR Raman system had been cooled to −90 °C. Cooling of the camera required about an hour. But once ready, the Raman system could be used to predict concentrations of multiple analytes from a single spectral measurement.

We also used an implicit calibration model and performed partial least squares (PLS) analysis using the Invitrogen CHO cell culture supernatant Raman spectra from 10 days and the reference concentrations provided by HPLC. The results of leave-one-out cross validation are shown in Table 1. In comparison with Whelan *et al.*,[3] it is observed that the correlation in glucose and lactate concentration predictions obtained using implicit calibration model is comparable, but the standard error for lactate concentration prediction is significantly less in our case. The correlation in concentration prediction using CLS is also comparable with slightly higher standard errors of prediction compared with PLS.

In Fig. 7, we show the glucose and lactate concentration prediction in Sanofi CHO cell culture supernatants using Raman

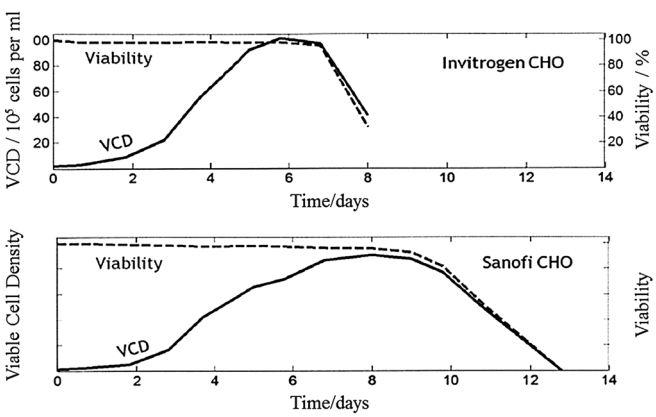

Figure 5. Growth curves for Invitrogen and Sanofi Chinese hamster ovary (CHO) cell cultures.[11] Sanofi cell line is optimized for antibody production and hence has a longer stationary phase. VCD, viable cell density.

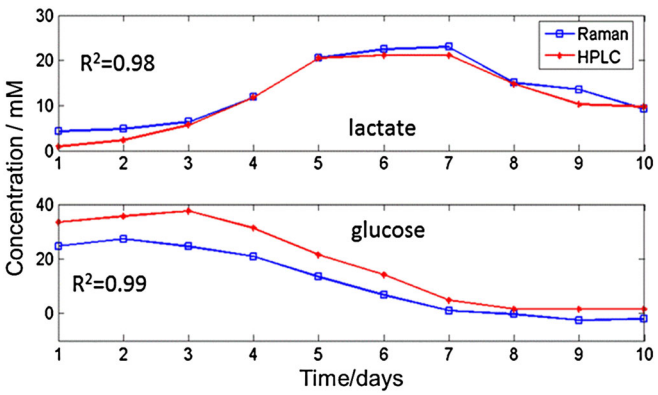

Figure 6. Prediction of glucose and lactate concentrations in Invitrogen Chinese hamster ovary (CHO) cell culture supernatants using Raman spectroscopy and explicit model-based classical least squares (CLS) analysis. The error bars were equivalent to 2 mM concentration for Raman measurements and 0.5 mM concentration for high performance liquid chromatography (HPLC) measurements.

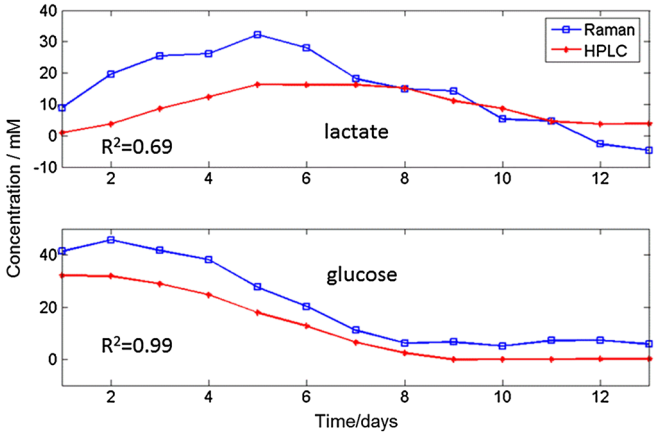

Figure 7. Prediction of glucose and lactate concentrations in Sanofi Chinese hamster ovary (CHO) cell culture supernatants using Raman spectroscopy and explicit model-based classical least squares (CLS) analysis. HPLC, high performance liquid chromatography.

spectroscopy and explicit model-based CLS algorithm. Comparison with reference concentration values obtained using HPLC is also shown. The slightly different prediction results for lactate in Figs 6 and 7 are because of the fact that we had used the same smoothing and basis spectra-based CLS model for both Invitrogen and Sanofi cell lines. These cell lines have different metabolism, as shown in Fig. 5, and the CLS algorithm can be sensitive to high frequency system noise or experimental noise because of the pseudo-inversion of the measurement matrix. One way to improve this in the future would be to introduce techniques such as regularization.[12]

Conclusions

We have shown for the first time that the detection of glucose and lactate concentrations with high accuracy is possible in CHO cell culture supernatants using Raman spectroscopy and explicit model-based CLS algorithm. When exact composition of the nutrient media is unknown, a deterministic Raman spectral library of pure components can be created by acquiring Raman spectra from credible nutrient media constituents and CHO cell culture metabolites. Residuals obtained after CLS analyses can be used to identify missing components or those present in excess. An algorithmic sieve thus generated provides a revised Raman spectral library that can be used for metabolite concentration predictions. Other CHO cell culture metabolites can also be monitored after improving the detection limit of the system, and it will require the use of respective HPLC columns to provide reference data values. Once validated, Raman spectroscopy-based concentration measurements require about an order of magnitude less time comparatively are not labor intensive and have the advantage of ease of automation. We believe that our algorithm and approach can be used for quality by design and process analytical technology. It also helps in the field of process development/understanding and process intensification.

Acknowledgements

We would like to thank Dr Jacqueline Wolfrum for helpful discussions and Professor Anthony Sinskey for providing access to his biology laboratory facilities.

References

[1] J. Zhu, *Biotechnol. Adv.* **2012**, *30*, 1158.
[2] Y. S. Yamamoto, H. Shinzawa, Y. Matsuura, Y. Ozaki, H. Sato, *Appl. Spectrosc.* **2011**, *65*, 844.
[3] J. Whelan, S. Craven, B. Glennon, *Biotechnol. Progr.* **2012**, *28*, 1355.
[4] N. R. Abu-Absi, B. M. Kenty, M. E. Cuellar, M. C. Borys, S. Sakhamuri, D. J. Strachan, M. C. Hausladen, Z. J. Li, *Biotechnol. Bioeng.* **2011**, *108*, 1215.
[5] B. Li, P. W. Ryan, B. H. Ray, K. J. Leister, N. M. S. Sirimuthu, A. G. Ryder *Biotechnol. Bioeng.* **2010**, *107*, 290.
[6] H. L. T. Lee, P. Boccazzi, N. Gorretb, R. J. Ram, A. J. Sinskey, *Vib. Spectrosc.* **2004**, *35*, 131.
[7] T. B. Shope, T. J. Vickers, C. K. Mann, *Appl. Spectrosc.* **1987**, *41*, 908.
[8] Y. Xu, J. F. Ford, C. K. Mann, T. J. Vickers, J. M. Brackett, K. L. Cousineau, W. G. Robey, *Proc. SPIE* **1997**, *2976*, 10.
[9] M. Schröder, K. Matischak, P. Friedl, *J. Biotechnol.* **2004**, *108*, 279.
[10] G. Zhu, X. Zhu, Q. Fan, X. Wan, *Spectrochim. Acta A* **2011**, *78*, 1187.
[11] S. Goh, Micro-bioreactor design for Chinese Hamster Ovary cells, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, **2013**.
[12] C. R. Vogel, *Computational Methods for Inverse Problems (Frontiers in Applied Mathematics)*, Soc for Industrial & Applied Math, Philadelphia, USA, **2002**.