Journal of Biotechnology 188 (2014) 148–157

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

In situ near infrared spectroscopy monitoring of cyprosin production by recombinant *Saccharomyces cerevisiae* strains

Pedro N. Sampaio[a,*], Kevin C. Sales[a], Filipa O. Rosa[a], Marta B. Lopes[a,b], Cecília R. Calado[a]

[a] Faculty of Engineering, Catholic University of Portugal, Estrada Octávio Pato, 2635-631 Rio de Mouro, Portugal
[b] Institute of Telecommunications, Instituto Superior Técnico, University of Lisbon, Av. Rovisco Pais, 1049-001 Lisbon, Portugal

ARTICLE INFO

Article history:
Received 26 April 2014
Received in revised form 17 July 2014
Accepted 23 July 2014
Available online 10 August 2014

Keywords:
In situ monitoring
Near infrared
Partial least squares regression
Recombinant cyprosin B
Saccharomyces cerevisiae

ABSTRACT

Near infrared (NIR) spectroscopy was used to *in situ* monitoring the cultivation of two recombinant *Saccharomyces cerevisiae* strains producing heterologous cyprosin B. NIR spectroscopy is a fast and non-destructive technique, that by being based on overtones and combinations of molecular vibrations requires chemometrics tools, such as partial least squares (PLS) regression models, to extract quantitative information concerning the variables of interest from the spectral data. In the present work, good PLS calibration models based on specific regions of the NIR spectral data were built for estimating the critical variables of the cyprosin production process: biomass concentration, cyprosin activity, cyprosin specific activity, the carbon sources glucose and galactose concentration and the by-products acetic acid and ethanol concentration. The PLS models developed are valid for both recombinant *S. cerevisiae* strains, presenting distinct cyprosin production capacities, and therefore can be used, not only for the real-time control of both processes, but also in optimization protocols. The PLS model for biomass yielded a $R^2 = 0.98$ and a RMSEP = 0.46 g dcw l^{-1}, representing an error of 4% for a calibration range between 0.44 and 13.75 g dcw l^{-1}. A $R^2 = 0.94$ and a RMSEP = 167 U ml^{-1} were obtained for the cyprosin activity, corresponding to an error of 6.7% of the experimental data range (0–2509 U ml^{-1}), whereas a $R^2 = 0.93$ and RMSEP = 672 U mg^{-1} were obtained for the cyprosin specific activity, corresponding to an error of 7% of the experimental data range (0–11,690 U mg^{-1}). For the carbon sources glucose and galactose, a $R^2 = 0.96$ and a RMSECV of 1.26 and 0.55 g l^{-1}, respectively, were obtained, showing high predictive capabilities within the range of 0–20 g l^{-1}. For the metabolites resulting from the cell growth, the PLS model for acetate was characterized by a $R^2 = 0.92$ and a RMSEP = 0.06 g l^{-1}, which corresponds to a 6.1% error within the range of 0.41–1.23 g l^{-1}; for the ethanol, a high accuracy PLS model with a $R^2 = 0.97$ and a RMSEP = 1.08 g l^{-1} was obtained, representing an error of 9% within the range of 0.18–21.76 g l^{-1}. The present study shows that it is possible the *in situ* monitoring and prediction of the critical variables of the recombinant cyprosin B production process by NIR spectroscopy, which can be applied in process control in real-time and in optimization protocols. From the above, NIR spectroscopy appears as a valuable analytical tool for online monitoring of cultivation processes, in a fast, accurate and reproducible operation mode.

1. Introduction

Biotechnology companies follow high standards on economic, technical, quality and regulatory procedures and are, therefore, encouraged to control and optimize their processes. To act accordingly, and to further automation large-scale microbial bioprocess, a thorough real-time estimation of the state of the process must be achieved (Harthun et al., 1997). In this view, the Food and Drugs Administration (FDA) introduced in 2004 the Process Analytical Technologies (PAT) Initiative with the goal of ensuring the quality of the final product, while achieving a high-knowledge of the process and preserving its reproducibility (Hakemeyer et al., 2012). PAT is used to design, analyze and control the manufacturing process based on monitoring the critical quality attributes of raw materials and in process. PAT is currently a key tool for the modern approach of the regulatory framework based on Quality by Design (QbD), that encloses the scientific understanding of the whole product life cycle, including the manufacturing process, together with risk-based approaches, to build quality into the process (http://www.ich.org/, 2014). Implementation of QbD implies the definition of the target product quality profile and consequently the production process that runs within a design space and uses PAT to monitor and control the process towards the target product quality. Therefore, the implementation of new

* Corresponding author. Tel.: +351 214 269 770; fax: +351 214 269 800.
E-mail address: pnsampaio@gmail.com (P.N. Sampaio).

http://dx.doi.org/10.1016/j.jbiotec.2014.07.454
0168-1656/

highly sensitive analytical techniques that enable the monitoring of the critical quality attributes of the process is highly relevant to evaluate the process during running and enabling its feed-back correction towards the design quality profile. The ideal approach of the process monitoring should be in-line and hopefully *in situ* to obtain information of the critical process attributes in real-time, while minimizing the perturbation of the process.

A relevant biotechnology process needing the development of *in situ* real-time monitoring is the production of recombinant cyprosin B by *Saccharomyces cerevisiae*. Cyprosin B, an aspartic protease present in *Cynara cardunculus* flowers, is largely used in cheese production due to its high clotting activity (Heimgartner et al., 1990). The flowers of this species are traditionally used in Portugal and Spain for cow, goat and ewe cheese making, such as Serpa and Serra cheeses, which are unique in Portugal. Clotting enzymes have been found in animals, plants, and microorganisms. Genetically engineered chymosin has been produced and was first recognized as safe (GRAS) by the Food and Drugs Administration in 1989. Currently, it is estimated that 50% of the marketed chymosin is produced by biotechnological processes and about 70% of domestic cheese in United States is produced using bioengineered chymosin.

Concerning the host cell to produce recombinant cyprosin, there are diverse advantages of using *S. cerevisiae*: as a microorganism, it presents high specific growth rates and, consequently, is generally associated to high productivities; it requires relative cheap media cultivations; the procedures for its stable genetic modification are well established and it enables the recombinant protein secretion facilitating the purification process. From all these reasons, this microorganism has being widely applied for proteins recombinant expression from research labs to industrial scale (Romanos et al., 1992).

Monitoring of industrial *S. cerevisiae* cultivation processes is, however, mainly based on off-line analysis of critical variables in the culture broth, such as biomass, product and carbon sources. This off-line analysis, which implies that a broth sample is taken from the bioreactor vessel, increases the cultivation contamination risk, besides the time delay for determining the target analyte concentration. Generally, the on-line information acquired in real-time during the cultivation process is the pH, dissolved oxygen concentration and the bioreactor stirring speed. However, due to complex relationships between the cell host general metabolism and growth, plasmid stability and product formation, it is highly difficult to infer from this *in situ* information the critical variables of the bioprocess. An ideal method for the bioprocesses control should consider a real-time monitoring that enables the accurate determination of several target culture variables, with minimal or no sample preparation and any reagent consumption. Process control is one of the most important tasks in the biotechnology processes the integrated process analysis and control of the bioprocesses can bring more advantages such as a constant update of process status, rapid feedback at failure, process quality proof, risk minimizing and analysis error reduction. In terms of economic point of view, the NIR control allows a faster action at process failure, automated sample preparation, increased production rate and product yield, and decreased by-product formation.

The ideal monitoring device should also be located directly in the culture broth and capable of coping with the high agitation/aeration levels. In this context, infrared (IR) spectroscopy, as a fast and non-destructive vibrational spectroscopy technique, has been explored. IR spectroscopy can be operated in the mid-infrared (MIR) and near-infrared (NIR) regions of the electromagnetic spectrum. While MIR spectroscopy reflects the fundamental vibrations of molecular bonds, NIR spectroscopy reflects overtones and combinations of vibrations, which makes MIR spectra theoretically more informative concerning the samples' biomolecular composition. However, due to the high absorption of water in the MIR region, it is usually necessary to take the samples from the bioreactor and subsequently dehydrate the samples, which increases the risk of bioreactor contamination and inputs a time delay in the analysis (Scholz et al., 2012). NIR spectroscopy is not so affected by the water present, and combined with chemometric techniques, also allows the construction of calibration models for the prediction of the critical variables of the bioprocess. Moreover, NIR fiber optic probes, that can be immersed directly in the culture broth and steam sterilized with it, enable the acquisition *in situ* of information in real-time (Lopes et al., 2014). NIR spectroscopy has been applied in diverse measurement approaches such as off-line, at-line (*i.e.* rapid on-line) and on-line (as reviewed in Cervera et al., 2009). As on-line the NIR probe has been used *ex situ*, where the probe is physically outside the bioreactor, and *in situ*, characterized by the NIR probe directly submersed in the bioreactor vessel. From these two approaches the *in situ* analysis is in accordance with QbD goals and therefore the one that must be achieved. Furthermore, the NIR-probe characteristics should enable the probe to resist to temperature and pressures of the heat sterilization process and should cope with the highly dynamic ranges of vibrations and gas phases effects due to the high variations in agitation speed and aeration levels normally occurring during microorganisms culture, besides coping with the possible probe fouling due to high biomass concentrations (Arnold et al., 2002; Tamburini et al., 2003).

There are very few works applying an *in situ* NIR probe along the whole culture process of microorganisms, that consequently cope with very high dynamic ranges of agitations, air-flow rate and biomass concentrations. Most of these works were conducted with prokaryotic cells like *Escherichia coli* (Lopes et al., 2014; Cimander and Mandenius, 2002; Arnold et al., 2002), *Vibrio cholera* (Navrátil et al., 2005), *Staphyloccocus* and *Lactobacillus* (Tosi et al., 2003; Tamburini et al., 2003).

The aim of the present work is to evaluate the *in situ* monitoring by NIR spectroscopy of the recombinant cyprosin B production by two *S. cerevisiae* strains (BJ1991 and W303-1A) transformed with the same expression vector, containing the *CYPRO11* gene (Sampaio et al., 2008). As the two recombinant expression systems of cyprosin present highly different behaviors it is also intended to develop predictive PLS models simultaneously valid for both systems and, consequently, that may be used to potentiate a real-time monitoring and control of the bioprocess and on optimization protocols. For this purpose, two cultivations conducted with these two expression systems were *in situ* monitored by NIR spectroscopy using an NIR fiber optic probe directly submersed in the culture broth, previously steam sterilized with the bioreactor vessel. The spectral information was correlated by PLS regression models with the off-line data obtained from enzymatic, chromatographic and spectrophotometric analysis of the main variables of the bioprocess: biomass, cyprosin enzymatic activity and specific activity, the carbon sources glucose and galactose, and the by-products acetic acid and ethanol.

2. Materials and methods

2.1. Cell culture conditions

2.1.1. Inoculum and enzyme production media

YNB medium [0.67% (w/v) yeast nitrogen base without amino acid and 2% (w/v) glucose] supplemented with auxotrophic amino acid and base (20 mg l^{-1} uracil and 70 mg l^{-1} tryptophan) for BJ1991 strain and (50 mg l^{-1} L-histidine; 20 mg l^{-1} L-adenine; 70 mg l^{-1} L-tryptophan and, 20 mg l^{-1} uracil) for W303-1A strain were previously sterilized by microfiltration and used to start *S. cerevisiae* culture. Afterwards, the 75 ml of YNB medium into