

Optical Spectroscopy Methods to Monitor Cells and Bacteria Concentrations and to Detect Contamination during Cell Culture: Application to the Fabrication of ATMPs

Bruno Wacogne^{1,2}, Déborah Legrand¹, Charles-Louis Azzopardi¹[0000-0003-2147-2042],
Christian Pieralli¹ and Annie Frelet-Barrand¹[0000-0002-0849-9040]

¹ FEMTO-ST institute, Univ. Bourgogne Franche-Comte, CNRS, 15B avenue des Montboucons, 25030 Besançon, cedex, France

² Centre Hospitalier Universitaire de Besançon, Centre d'Investigation Clinique, INSERM CIC 1431, 25000, Besançon, France
bruno.wacogne@univ-fcomte.fr

Abstract. Currently, the production of Advanced Therapy Medicinal Products is highly sensitive to any contamination sources and therefore takes place in clean and sterile environments. Several days are required for each production, making these products extremely expensive. Throughout the process, numerous quality controls must be performed. This is especially true during the expansion phase in order to monitor cell growth and to detect any contamination. Bioreactor's content must periodically be sampled to perform these controls. Two major drawbacks can be identified: a delayed knowledge of the quality control result and an additional risk of new contaminations due to sampling. In this work, we present optical spectroscopy methods which can be used to drastically reduce the risk of contamination. They provide a real time control of what happens in the bioreactor in a closed system manner. Cell concentrations are measured with an accuracy below 5% and contamination can be detected about 3 hours after it occurred. The real time operation leads to several tens of thousand dollars' savings because it allows stopping the production as soon as a problem arises. Consequently, the price of these products should be greatly reduced and they may be proposed to more patients.

Keywords: Optical Spectroscopy, Advanced Therapy Medicinal Product, Cell Growth Monitoring, Contamination Detection.

1 Introduction

New treatment solutions for patients with no further therapeutic options have recently emerged. They are called ATMPs (Advanced Therapy Medicinal Products). Some of them are based on the use of "drug" cells derived from genetic modification or tissue and cell engineering. Cells acquire new physiological functions, biological characteristics or reconstruction properties to the expense of substantial manipulations. Indeed,

natural processes of the body such as the use of stem cells for tissue regeneration, lymphocytes for cancer immunotherapy or apoptotic cells for anti-inflammatory purposes have inspired studies concerning these new biological drugs.

However, the production of these drugs requires the implementation of complex technologies of cell sorting, amplification, genetic transduction, amplification-division, activation, and this at several stages of production in sterile clean room type environment and in complex facilities. Also, the time needed to complete the production and the complex quality control processes further increase the fabrication costs. A schematic description of the fabrication process of CAR-T cells is given in figure 1 (adapted from Wang, 2016). It also applies to other ATMP productions.

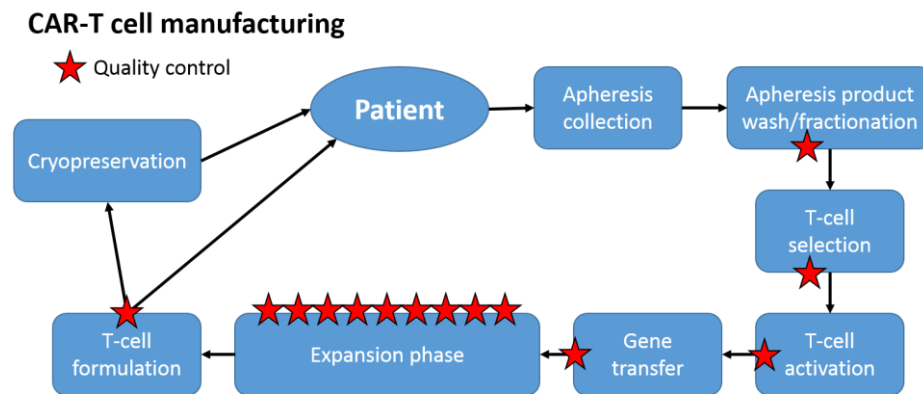


Fig. 1. Fabrication process of CAR-T cells. Adapted from (Wang, 2016)

Basically, the patient's blood is sampled and cells of interest (T cells for figure 1) are extracted/isolated and transduced to acquire the desired therapeutic properties. At this step, only well-transduced T cells are conserved. Now, these genetically modified cells are amplified/expanded in a bioreactor for a period that can extend up to 1 week. Eventually, the ATMP can be either injected to the patient or cryo-preserved before injection. Among others, one of the main constraints that must be addressed is the following.

Working in a controlled environment and preserving the closed system as necessary is crucial to meet the requirement of no contamination of the products. Frequently, the absence of containers, reagents or materials adapted to the protocol makes it difficult. Quality controls (red stars in figure 1) imply repeated samplings during the production, especially to control cell multiplication (expansion phase) and to detect any possible bacterial contamination (throughout the process). This increases the risk of contamination, the time of completion and requires increased traceability. Because these evaluations are time consuming, the production process still continues in parallel. In some cases, the production is stopped after several days when a contamination is detected. Obviously, this actually increases the cost and delay or even stop the delivery of the drug to the patient.

The work presented in this paper addresses the constraint of closed system and real time control of what happens in the bioreactor during the expansion phase. The goal is

to provide an easygoing/soft/simple method to monitor the cell growth and to detect contaminations as early as possible. In-line or real time measurement techniques have been widely studied either for cell culture or contaminant development monitoring, very rarely for both.

Various spectroscopic coupled to chemometric techniques were presented in Teixeira review (Teixiera, 2009) concerning cell culture. In addition, other methods based on ultrasonic measurements (Melchor, 2018) or capacitive techniques (Lee, 2016) have been proposed. Impedance monitoring, either in a bulk system (Cacopardo, 2019) or in a microfluidic chip (Fong Lei, 2014) has also been presented. But, in these references, indications whether or not the proposed method can be adapted in a closed system configuration were not discussed.

Concerning bacteria detection or monitoring, different sensors have been proposed to detect *Escherichia coli* (hereafter *E. coli*) by Ikonen (Ikonen, 2017). Modified Field Effect Transistors have been experimented to detect the same bacteria (Thakur, 2018). Detecting several contaminants with a single device is challenging. However, it has been demonstrated by using optical absorption spectroscopy (Theint, 2019), fiber optic Fourier Transform Infra-Red spectroscopy (Hassan, 2016), quartz crystal sensors (Chang, 2006) and electrochemistry (Safavieh, 2014). More generally, recent reviews concerning electrochemical biosensors (Cesewski, 2020) and impedimetric immunosensors (Leva-Bueno, 2020) for pathogen detection have been published.

Simultaneous cell monitoring and contaminant detection has only been the subject of very few papers. For example, Liu proposed advanced signal processing applied to Raman spectroscopy (Liu, 2017). Together with normal condition monitoring, authors demonstrated the detection of growth problems 5 hours after they stopped feeding the cells. They also detected effects of contamination with their monitoring algorithm. However, the nature of the contamination and the time required to detect it was not specified.

In this paper, we propose a proof of concept based on different optical spectroscopy methods to continuously monitor the evolution of cell concentration in a bioreactor and to issue an alarm signal shortly after a contamination occurred. The next section describes the experimental set-up, biological samples used in this study and spectroscopy methods employed. Section 3 presents the results obtained using several ways of exploiting information contained in transmission or absorption spectra of solution containing cells or bacteria (lymphocyte B cell and *E. coli*). Optical characterizations of the concentrations are based on colorimetric estimations of lymphocyte and *E. coli* solutions. These methods, together with the measurement of the maximum of the absorption spectra, can be used to monitor cell growth in real time. Other methods can be used for early detection of contamination. They are based on spectra shape analysis and Principal Component Analysis (PCA) respectively. Short discussions of these early results and aspects concerning socio-economic impacts will be given in section 4.

2 Materials and Methods

For this proof of concept, measurements are not performed in a closed system configuration. Adaptation of the method in this particular environment is shortly discussed in section 4.

2.1 Lymphocytes and *E. coli* Preparation.

Lymphocyte cell lines (Ramos, ATCC, USA) were cultured in X-Vivo (Lonza, Switzerland) with 5% FBS (Gibco™ 10270106) and 10% streptomycin/penicillin (100 µg/mL+100 UI/mL, CABPES01-0U, Eurobio) in a humidified 37°C, 5% CO₂ incubator. Cells were recovered after 2-3 days culture by centrifugation at 700 g, 10 min, 25°C. Different cell concentrations (10⁴×[1, 2, 4, 6, 8, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100] cells/mL) were prepared after dilution in autoclaved PBS 1x pH7.4 (Sigma, USA).

Escherichia coli DH5a (NEB, USA) were cultured in Luria Bertani overnight at 37°C, 180 rpm in a MaxQ incubator. They were recovered by centrifugation at 5000 g, 15 min, 20°C and re-suspended in autoclaved PBS 1x pH7.4 (Sigma, USA). Optical density of the re-suspension was measured in a spectrophotometer Shimadzu at 595 nm. Afterwards, different bacteria concentrations (10⁶×[1, 2, 4, 6, 8, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100] bacteria/mL) were prepared for experiments.

2.2 Experimental Set-up

The extremely simple experimental set-up is schematically presented in figure 2. The set-up was composed of a white light source (Ocean Optics HL 2000) connected to a cuvette holder (Avantes CUV-UV/VIS) via conventional step index optical fibers (Thorlabs M25L01). After propagation through the cuvette, light was launched into a spectrometer for transmission/absorption spectra acquisition (Ocean Optics QE-Pro). Fluorimeter polymethacrylate cuvettes were filled up to 3 mL with solutions of cells or bacteria (Sigma-Aldrich C0793-100EA). Spectra were measured using the specific feature available in the SpectraSuite software from Ocean Optics. Reference was obtained with a cuvette filled with PBS only. After transfer to PC, data processing was performed using MATLAB™ R2014b version. Spectra were truncated between 450 nm and 1120 nm in order to remove noisy data due to the transmission calculation in the SpectraSuite software.

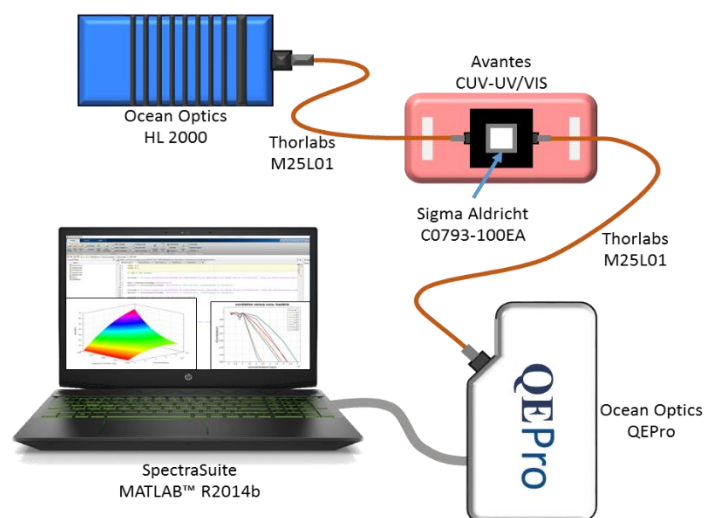


Fig. 2. Description of the experimental set-up.

2.3 Color Based Analysis of Transmission Spectra

The associated colors of the solutions were computed from transmission measurements. They do not actually correspond to the true colors of the solutions as discussed in section 4. However, they are expressed in different color spaces in order to study the evolution of colorimetric parameters with species concentrations.

Tristimulus values (CIE XYZ) were calculated from spectral distribution with a CIE 1931 standard colorimetric standard observer 2° and a CIE standard illuminant D65. Color of the sample expressed in tristimulus color space was then converted in other color spaces: CIELAB (CIE Lab 1976), sRGB (standard RGB) and HSV. HSV coordinates were determined from sRGB. Color space conversion were performed using the open source Python package “Colour” dedicated to color science (Mansencal, 2020).

2.4 Analysis of absorption spectra

Absorption spectra used in section 3.3 were slightly smoothed using a cubic spline algorithm in order to maximize the R^2 of the spectra fittings. Principal Component Analysis was performed with smoothed and normalized spectra.

In this work, so-called “contaminated spectra” presented below are artificial and made by adding spectra of lymphocytes and *E. coli*. This aspect will be discussed in section 4.

3 Experimental Results

3.1 Measuring Concentrations of Both Species with Colorimetric Description of Transmission Spectra

Examples of transmission spectra recorded with lymphocytes and *E. coli* for different concentrations are given in figure 3. Transmission data are used because they are more consistent with a colorimetric description of the spectra than absorption data.

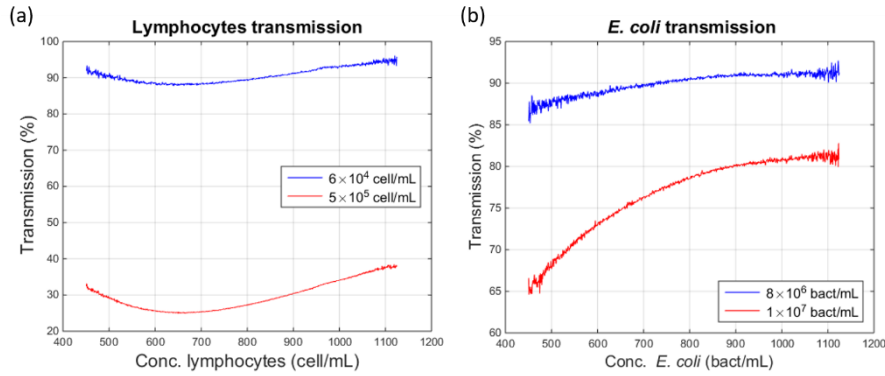


Fig. 3. Examples of transmission spectra of lymphocytes and *E. coli*. (a) Lymphocytes at 6×10^4 and 5×10^5 cell/mL, (b) *E. coli* at 8×10^6 and 1×10^7 bact/mL.

As previously mentioned, colors of the solution are described in several color spaces. Among them, we arbitrarily chose the following color spaces: XYZ, Lab, sRGB and HSV. The goal is to describe each parameter of these color spaces as a function of species concentration. Figure 4 shows the evolution of these colorimetric components as a function of the lymphocyte concentrations. In this figure, “star markers” correspond to experimental values. When the experimental data can be used to determine species concentrations, mathematical fitting of the experimental data is performed and displayed as continuous lines in the figure. In other cases, fitting is not performed.

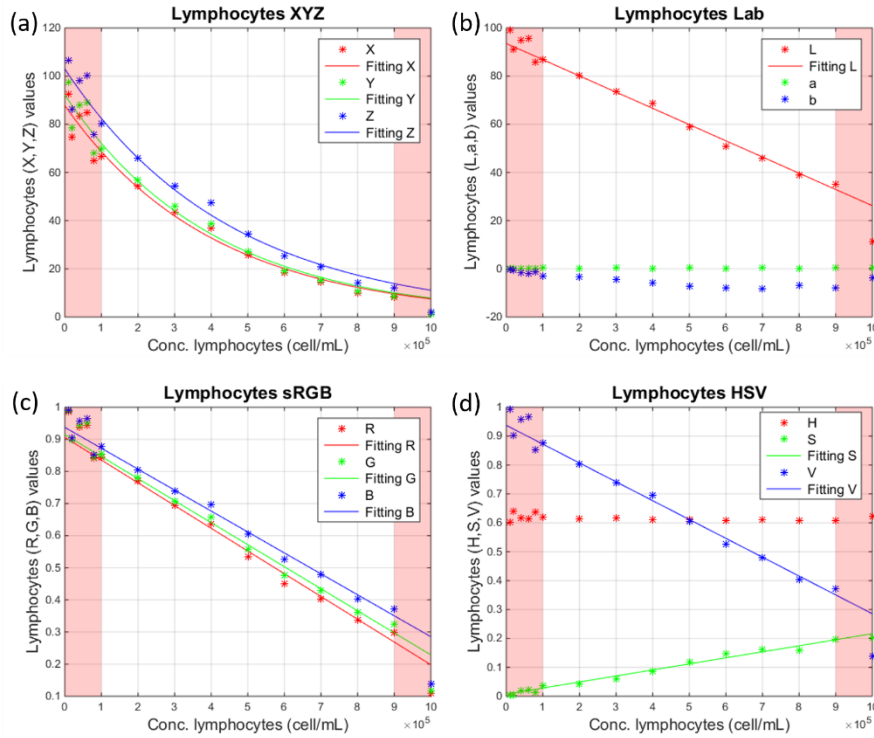


Fig. 4. Evolution of the colorimetric components with lymphocyte concentrations. (a) XYZ space, (b) Lab space, (c) sRGB space and (d) HSV space.

XYZ components of the transmission spectra can easily be described by an exponential function. However, and as it is depicted on the figure 4 with the light red areas, fitting XYZ only works in a reduced range: $[1 \times 10^5 - 9 \times 10^5]$ cell/mL. For low concentrations, transmission is of the order of 100% and no real difference can be observed in the transmissions spectra corresponding to these concentrations. Conversely, for high concentrations, the transmission approaches 0% and no coherent information can be extracted from the spectra.

The situation is slightly different for the Lab components. Here, only the “L” component is useful. It can be described with a linear regression. The “a” component is almost zero for any lymphocyte concentration and contains no exploitable information. The “b” component shows a non bijective evolution with the concentration. Therefore, it cannot be used to compute lymphocyte concentrations. For these two components, no attempts were made to describe their behavior. For the Lab space again, the useful measurement range is $[1 \times 10^5 - 9 \times 10^5]$ cell/mL.

The 3 components of the sRGB space can be used to compute the concentration using linear regressions in the above mentioned useful range. The HSV space only shows 2 exploitable components: “S” and “V”. They are described with linear regression in the same useful range.

The same analysis can be conducted with transmission spectra of *E. coli*. The result is given in figure 5. In this case, much less components can be used to compute bacteria concentrations. Indeed, for the XYZ and sRGB spaces, the evolutions of the components with the concentrations are non-bijective. This is also the case for the “L” and “V” components in corresponding color spaces. Concentrations can only be computed from the behavior of “b” and “S”. Evolutions of the components with the concentration are linear. Again, the useful range of the usable colorimetric components is $[1 \times 10^5 - 9 \times 10^5]$ bact/mL.

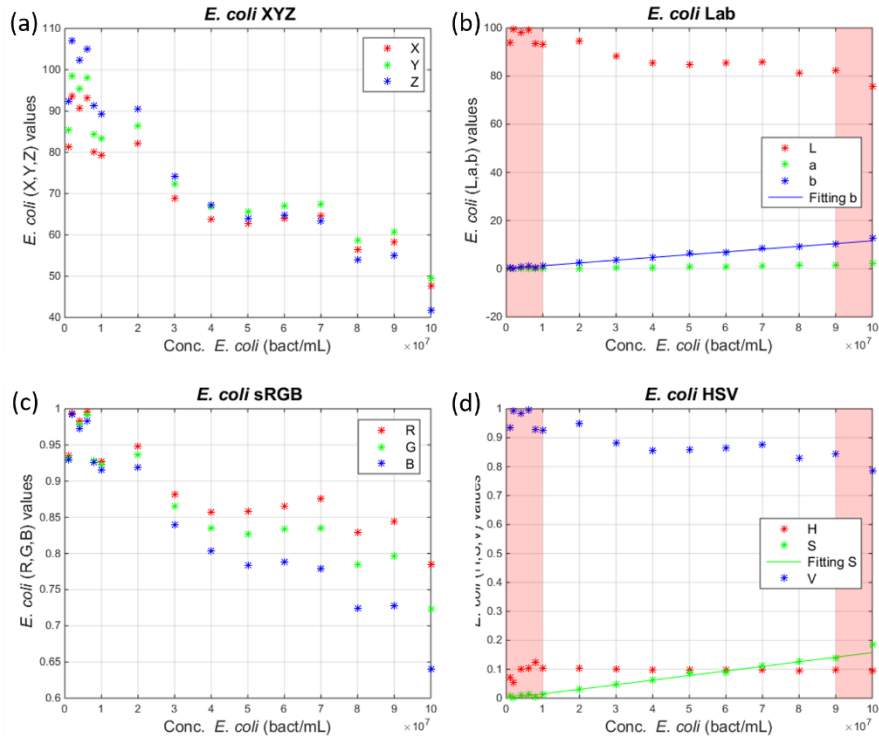


Fig. 5. Evolution of the colorimetric components with *E. coli* concentrations. (a) XYZ space, (b) Lab space, (c) sRGB space and (d) HSV space.

In order to define components offering the best accuracy when computing the concentration, we calculated the R^2 of each possible fitting. This is summarized in table 1.

Table 1. Summary of the methods used to describe the evolution of the lymphocyte and *E. coli* concentrations, type of function used to fit the experimental data, values of the fitting R^2 and useful ranges. “pol deg 2”: second order polynomial function, “exp”: exponential function, “linear”: linear regression. NA: not applicable. Light red highlighted rows correspond to non-exploitable components.

Lymphocytes				<i>E. coli</i>			
Method	Fit type	R^2	Range (cell/mL)	Method	Fit type	R^2	Range (bact/mL)
X	exp	0,9881	$1 \times 10^5 - 9 \times 10^5$	X	NA	NA	NA
Y	exp	0,988	$1 \times 10^5 - 9 \times 10^5$	Y	NA	NA	NA
Z	exp	0,9873	$1 \times 10^5 - 9 \times 10^5$	Z	NA	NA	NA
L	linear	0,9939	$1 \times 10^5 - 9 \times 10^5$	L	NA	NA	NA
a	NA	NA	NA	a	NA	NA	NA
b	NA	NA	NA	b	linear	0,9904	$1 \times 10^7 - 9 \times 10^7$
R	linear	0,9903	$1 \times 10^5 - 9 \times 10^5$	R	NA	NA	NA
G	linear	0,992	$1 \times 10^5 - 9 \times 10^5$	G	NA	NA	NA
B	linear	0,9936	$1 \times 10^5 - 9 \times 10^5$	B	NA	NA	NA
H	NA	NA	NA	H	NA	NA	NA
S	linear	0,9662	$1 \times 10^5 - 9 \times 10^5$	S	linear	0,9901	$1 \times 10^7 - 9 \times 10^7$
V	linear	0,9936	$1 \times 10^5 - 9 \times 10^5$	V	NA	NA	NA

R^2 values are all relatively high. The highest values are obtained with the “L” component of the lymphocyte transmission spectra ($R^2=0.9939$) and with the “b” component of the *E. coli* transmission spectra ($R^2=0.9904$). For these two examples, fitting functions $L_{lym}(C)$ for the lymphocyte L component and $b_{E.coli}(C)$ for the *E. coli* “b” components are given by:

$$L_{lym}(C) = m1 \cdot C + m2 \quad (1)$$

$$b_{E.coli}(C) = n1 \cdot C + n2 \quad (2)$$

Where $m1 = -6.735 \times 10^{-5}$, $m2 = 93.56$, $n1 = 1.147 \times 10^{-7}$ and $n2 = 0.1295$

Using equations (1) and (2) allows estimating the species concentrations as presented in figure 6. Blue diamonds correspond to the real concentrations and red crosses correspond to the concentrations calculated using equations (1) and (2). Standard deviations reported in the figure correspond to 1.89×10^4 cell/mL for the lymphocytes (3.8% at the center of the concentration range) and 2.37×10^6 bact/mL for *E. coli* (4.7 % at the center of the concentration range).

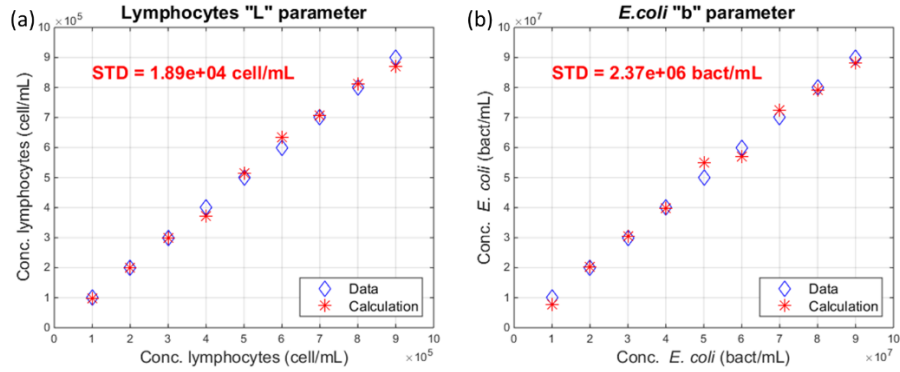


Fig. 6. Calculating the species concentrations from colorimetric parameters. (a) Lymphocytes concentrations using the “L” component in the Lab color space. (b) *E. coli* concentrations using the “b” component in the same color space.

STDs obtained using fitting equations corresponding to exploitable color spaces shown in table 1 are reported in table 2.

Table 2. Values of the Standard Deviation for the different fitting possibilities.

Lymphocytes			<i>E. coli</i>		
Method	Fit type	STD (cell/mL)	Method	Fit type	STD (bact/mL)
X	exp	$4,19 \times 10^4$	X	NA	NA
Y	exp	$3,84 \times 10^4$	Y	NA	NA
Z	exp	$4,27 \times 10^4$	Z	NA	NA
L	linear	$1,89 \times 10^4$	L	NA	NA
a	NA	NA	a	NA	NA
b	NA	NA	b	linear	$2,37 \times 10^6$
R	linear	$2,39 \times 10^4$	R	NA	NA
G	linear	$2,17 \times 10^4$	G	NA	NA
B	linear	$1,94 \times 10^4$	B	NA	NA
H	NA	NA	H	NA	NA
S	linear	$4,51 \times 10^4$	S	linear	$2,41 \times 10^6$
V	linear	$1,94 \times 10^4$	V	NA	NA

To summarize this section, concentrations of lymphocytes and *E. coli* can be measured over a large concentration range using the color associated to transmission spectra ($[1 \times 10^5 - 9 \times 10^5]$ cell/mL). Table 2 shows that the “L” component of the Lab color space leads to the lowest STD for cell concentration measurements (1.89×10^4 cell/mL, i.e. 3.78 % at the center of measurement range). For *E. coli* the lowest STD is obtained with the “b” component of the same Lab color space (2.37×10^6 bact/mL, i.e. 4.74 % at the center of measurement range). Aspects concerning the use of colorimetric analysis of transmission spectra will be shortly discussed in section 4.

Often in biology laboratories, concentration is measured using the absorption of a solution at a fixed wavelength. Measured solutions are diluted in order to obtain optical densities in a range which ensures a reliable measurement. In the next section, we show how the information contained in the whole absorption spectrum can be used to measure concentrations. This will also introduce section 3.3 where the possibility to exploit the shapes of absorption spectra can be used to detect culture contaminations.

3.2 Measuring Concentrations with the Shapes of the Absorption Spectra.

Examples of absorption spectra recorded with lymphocytes and *E. coli* for different concentrations are given in figure 7.

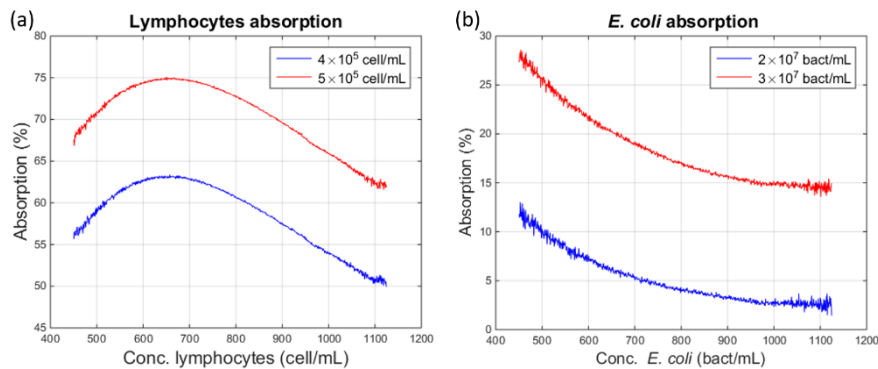


Fig. 7. Examples of absorption spectra of lymphocytes and *E. coli*. (a) Lymphocytes at 4×10^5 and 5×10^5 cell/mL, (b) *E. coli* at 2×10^7 and 3×10^7 bact/mL.

The shapes of spectral absorptions of the two species are completely different. The idea is, for each species, to model the evolution of the shapes of absorption spectra as a function of concentration and to use this mathematical description to calculate the concentration in species from any measured spectrum.

Indeed, the ulterior motive is to measure the absorption spectrum of the contents of the bioreactor continuously. For each recorded spectrum, the idea is to analyze the shape, to separate the part due to the contribution of lymphocytes from that due to *E. coli* and to calculate their respective concentrations. This would allow detecting a possible contamination in real time and stopping the fabrication of the ATMPs extremely early, i.e. before a large amount of money is lost and possibly reducing the final cost of the drug.

First, it is necessary to determine the evolution of the spectra of the two species as a function of the concentration. This is achieved using spectra mathematical fitting. We found a general equations which can be used for each of the species. We found that an exponential based function can be used to fit absorption spectra of lymphocytes while those of *E. coli* can be fitted with a 3rd order polynomial function. We then have:

$$Spec_{lym}(\lambda, C) = p1 \cdot \exp\left(-\left(\frac{\lambda-p2}{p3}\right)^2\right) + p4 \quad (3)$$

$$Spec_{E.Coli}(\lambda, C) = q1\lambda^3 + q2\lambda^2 + q3\lambda + q4 \quad (4)$$

Here, λ is the wavelength, C is the concentration. The concentration dependence exists through the “pi” and “qi” parameters which can be expressed as follows.

$$p1(C) = a1 \cdot \exp\left(-\left(\frac{C-a2}{a3}\right)^2\right) \quad (5)$$

$$p2(C) = b1 \cdot C + b2 \quad (6)$$

$$p3(C) = c1 \cdot C^2 + c2 \cdot C + c3 \quad (7)$$

$$p4(C) = d1 \cdot C^2 + d2 \cdot C + d3 \quad (8)$$

$$q1(C) = e1 \cdot C^2 + e2 \cdot C + e3 \quad (9)$$

$$q2(C) = f1 \cdot C^2 + f2 \cdot C + f3 \quad (10)$$

$$q3(C) = g1 \cdot C^2 + g2 \cdot C + g3 \quad (11)$$

$$q4(C) = h1 \cdot C^2 + h2 \cdot C + h3 \quad (12)$$

Coefficients used in equations (5) to (12) are given in table 3.

Table 3. Coefficients used in equations (5) to (12).

$p1(C)$	$a1=14.32$	$a2=4.1 \times 10^5$	$a3=6.145 \times 10^5$
$p2(C)$	$b1=7.21 \times 10^{-5}$	$b2=647$	
$p3(C)$	$c1=1 \times 10^{-10}$	$c2=-1.9 \times 10^{-5}$	$c3=325.7$
$p4(C)$	$d1=-5 \times 10^{-11}$	$d2=1.3 \times 10^{-3}$	$d3=6.791$
$q1(C)$	$e1=1.2 \times 10^{-23}$	$e2=-1.6 \times 10^{-15}$	$e3=1.2 \times 10^{-9}$
$q2(C)$	$f1=-3.3 \times 10^{-20}$	$f2=5.1 \times 10^{-12}$	$f3=-1.7 \times 10^{-6}$
$q3(C)$	$g1=2.8 \times 10^{-17}$	$g2=-5.5 \times 10^{-09}$	$g3=-8.3 \times 10^{-4}$
$q4(C)$	$h1=-1 \times 10^{-14}$	$h2=2.4 \times 10^{-6}$	$h3=1.234$

Figures 8 and 9 show the evolutions of the pi and qi functions and equations (5) to (12) are recalled in the figures.

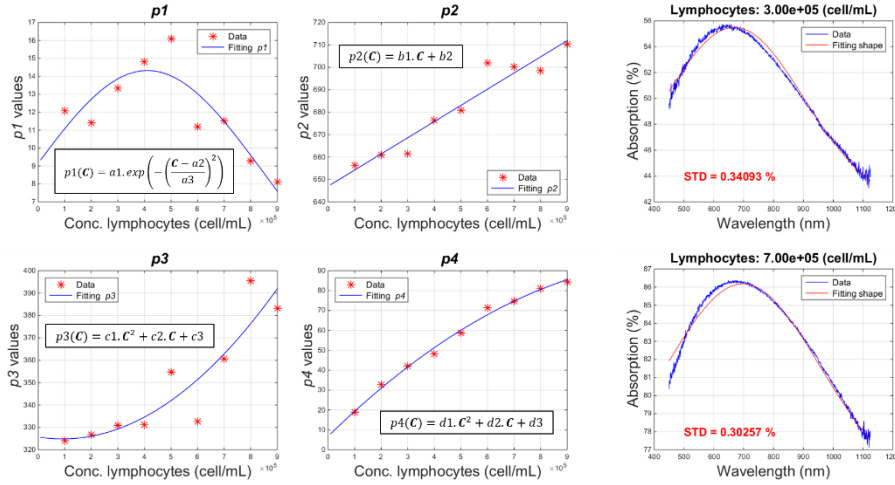


Fig. 8. Evolution of the p_i functions with the lymphocyte concentration (left hand side). Examples of experimental absorption spectra fitted with the $\text{Spec}_{\text{lym}}(\lambda, C)$ function (right hand side).

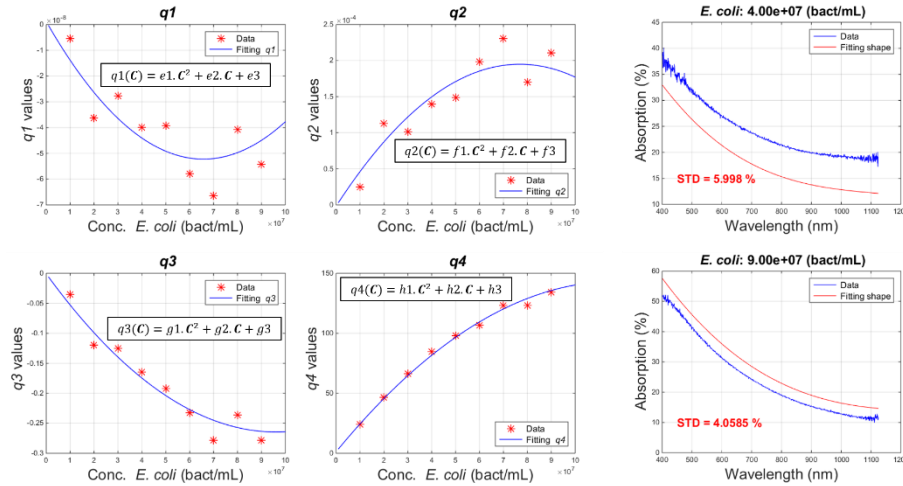


Fig. 9. Evolution of the q_i functions with the *E. coli* concentration (left hand side). Examples of experimental absorption spectra fitted with the $\text{Spec}_{E.Coli}(\lambda, C)$ function (right hand side).

Obviously, p_i and q_i data can hardly be fitted with mathematical function. Only p_4 and q_4 data can efficiently be modelled (this aspect will be discussed in section 4). Despite this, shapes of the absorption spectra can be modelled with a relatively good accuracy. For the lymphocytes, the STD is of the order of 0.3%. For the *E. coli* however, the STD is larger: between 4% to 6%.

Functions $\text{Spec}_{\text{lym}}(\lambda, C)$ and $\text{Spec}_{E.Coli}(\lambda, C)$ are represented in figure 10.

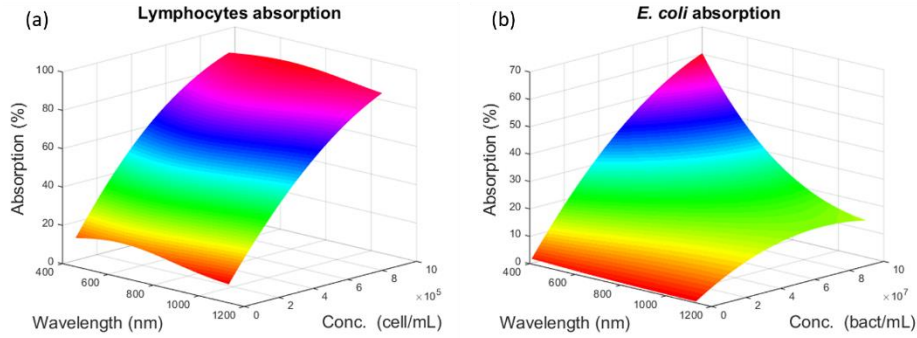


Fig. 10. Theoretical evolutions of the absorption spectra with species concentrations. (a) Lymphocytes, (b) *E. coli*. Adapted from (Wacogne, 2020).

These functions are used to calculate species concentrations from the shapes of the experimental absorption spectra. The result is shown in figure 11.

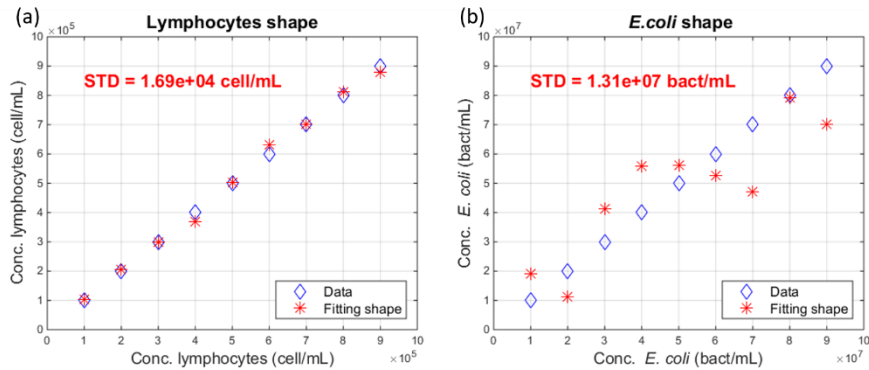


Fig. 11. Calculating the species concentrations from the shapes of the absorption spectra. (a) Lymphocytes concentrations. (b) *E. coli* concentrations. Adapted from (Wacogne, 2020).

The standard deviation for the lymphocytes is 1.69×10^4 cell/mL. This is lower than what was obtained using the “L” components in the Lab color space (1.89×10^4 cell/mL). For *E. coli* obviously, we failed measuring concentrations from the shape of the absorption spectra. Indeed, the standard deviation (1.31×10^7 bact/mL) is more than 5 times larger than what was obtained using the “b” component in the Lab color space (2.37×10^6 bact/mL). Therefore, the initial idea consisting in studying the shape of an absorption spectrum in real time in order to early detect any contamination seems impossible when trying to compute the actual concentrations values.

We now consider the evolution of the maxima of the lymphocyte absorption spectra with concentration. This evolution can be modelled with a second order polynomial function as depicted in figure 12(a). Here again, the useful range is $[1 \times 10^5 - 9 \times 10^5]$ cell/mL. Maxima of the absorption spectra can be modelled as follows.

$$Max_{lym}(C) = s1.C^2 + s2.C + s3 \quad (13)$$

Where: $s1 = -7.557 \times 10^{-11}$, $s2 = 1.527 \times 10^{-4}$ and $s3 = 16.74$

Calculation of lymphocyte concentrations from the evolution of the spectra maxima is shown in figure 12(b). Here, a standard deviation as low as 1.39×10^4 cell/mL (2.78 % at the center of the range) is obtained. This is the lowest values obtained so far in this study.

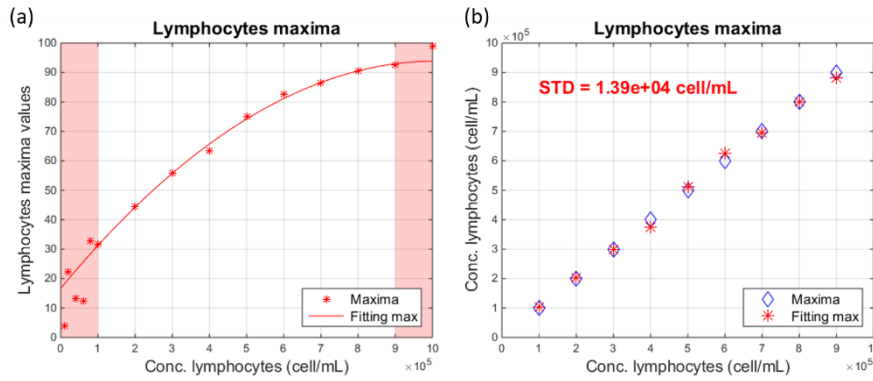


Fig. 12. Calculating the lymphocyte concentrations from the absorption maxima. (a) Evolution of the maxima with the lymphocyte concentrations. (b) Fitting concentrations using a second order polynomial function.

To summarize sections 3.1 and 3.2, concentrations of either lymphocyte or *E. coli* can be measured using white light spectroscopy. The best way to monitor lymphocyte concentration is to consider the evolution of the maxima of the absorption spectra. Concerning *E. coli*, the best method consists in measuring the “b” component in the Lab color space from the transmission spectra. Surprisingly, the evolution of the shape of the lymphocyte absorption spectra does not lead to a better accuracy than the simple measure of the absorption maximum.

However, estimating the shape of the absorption spectra can be used for another purpose. Indeed, the above described methods can be used to monitor the evolution of the concentrations of both species when only one species is present. When considering spectra of lymphocyte and *E. coli* mixtures, it is extremely difficult to separate their respective contributions. The problem becomes insoluble when several types of pathogens are considered.

We recall that the main goal is to measure the lymphocytes concentrations in a close –system environment and in real time. The goal is also to be able to detect any possible contamination in the same experimental environment. This can be achieved by considering things differently. This is the subject of the next section.

3.3 Monitoring Cell Concentrations and Detecting Contaminations: a New Approach

Fitting the Shape of the Absorption Spectra.

As long as everything is normal (no contamination) during the expansion phase, cell concentration can be monitored using methods described above. Also, we know the equation which describes the shape of the absorption spectra (equation (5)). In case of contamination, the shape of the spectrum resulting from the contribution of the lymphocytes and the contaminant differs from the ones corresponding to lymphocyte alone. In this situation, equation (5) should no longer be usable to fit the shape of absorption spectrum.

Practically, during the expansion phase, absorption spectra are recorded and fitted with a function representing the shapes of the lymphocyte spectra when they are alone. At this stage, only the shape of the spectra is analyzed. Therefore, equation (5) can be used in a simpler way than what is presented in section 3.2 where coefficients are expressed as a function of the concentration. This function is called $Abs_{lym}(\lambda)$. Here, the function only depends on the wavelength and coefficients are adjusted only to obtain the highest R^2 . Therefore, an accurate fitting (high R^2) using $Abs_{lym}(\lambda)$ means no contamination, lymphocyte concentration is calculated using equation (13), content of the bioreactor is adjusted according to this concentration and fabrication continues. Conversely, an R^2 value less than a threshold to be determined is the sign of a contaminated culture. Fabrication must then be stopped. Figure 13 explain this process.

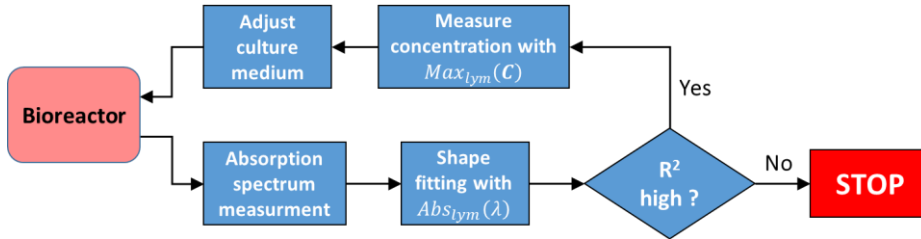


Fig. 13. In-line process to detect contaminations and to adjust the bioreactor composition as a function of the lymphocyte concentration.

As concentrations are not calculated, only the shape of the recorded spectrum is fitted. The following equation can efficiently be used.

$$Abs_{lym}(\lambda) = a \cdot \exp\left\{-\left(\frac{\lambda-b}{c}\right)^2\right\} + d \quad (14)$$

Here, constraints were put to coefficients a , b , c and d . Otherwise, the fitting algorithm (“trust region” in Matlab™ Curfitting toolbox) always finds a set of parameters to describe even contaminated spectra. The fitting bounds and starting points are summarized in table 4.

Table 4. Lower bounds, upper bounds and starting points used to fit absorption spectra with equation (14).

Coefficient	Lower bound	Upper bound	Starting point
a (%)	0	200	100
b (nm)	600	750	675
c (nm)	0	500	250
d (%)	0	200	100

Figure 14 shows examples of absorption spectra fittings using equation (14). Spectra were slightly smoothed as mentioned above.

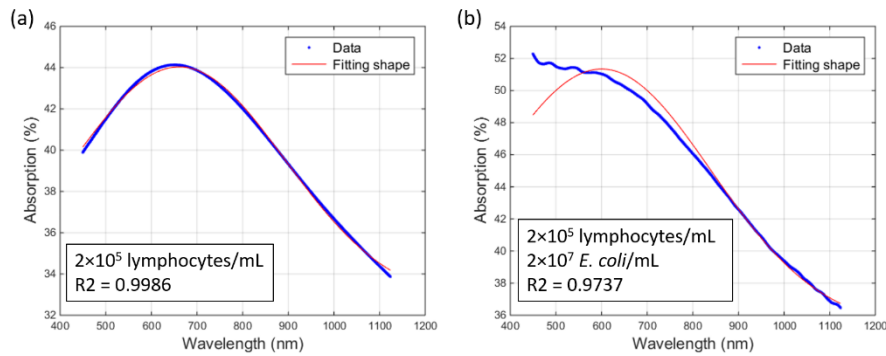


Fig. 14. Examples of spectra fitting using equation (14). (a) Lymphocyte concentration = 2×10^5 cell/mL, $R^2 = 0.9986$. (b) Lymphocyte concentration = 2×10^5 cell/mL, *E. coli* concentration = 2×10^7 bact/mL, $R^2 = 0.9737$.

In figure 14(a) fitting is performed with pure lymphocytes at 2×10^5 cell/mL. The R^2 value is quite high: $R^2 = 0.9986$. Indeed, with pure lymphocyte spectra, R^2 coefficients are always greater than 0.99 in the above mentioned useful range [$1 \times 10^5 - 9 \times 10^5$].

In figure 14(b) fitting is performed with the same lymphocyte concentration contaminated with 2×10^7 *E. coli*/mL. In this case and because the contamination modifies the shape of the recorded spectrum, R^2 decreases to 0.9737. Note that the so-called “contaminated spectra” are artificial and obtained by adding absorption spectra of both species as described in section 2.4.

For different lymphocyte concentrations in the useful range, figure 15 shows the evolution of R^2 with increasing concentrations of *E. coli*. Each curve corresponds to one lymphocyte concentration. It is observed that, for any lymphocyte concentrations, the contamination detection limit is of the order of $2.5 \cdot 10^7$ bact/mL. For this, we set a positivity threshold $R^2 = 0.988$ (arrows in the figure). Knowing that *E. coli* divides every 20 min and considering that the contamination is due to 1000 bact/mL, the warning signal can be issued 4h52 min post contamination.

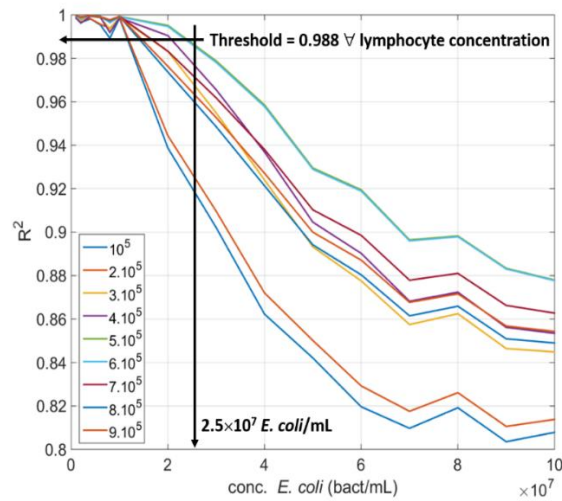


Fig. 15. Evolution of the R^2 coefficient with the concentration in *E. coli* for different lymphocyte concentrations. The legend corresponds to lymphocyte concentrations in cell/mL (Wacogne, 2020).

Using Principal Component Analysis.

Analyzing the shape of the absorption spectra allows detecting a contamination when $2.5 \cdot 10^7$ bact/mL are present in the bioreactor. But it is possible to further reduce the time required to issue the alert signal using Principle Component Analysis (PCA). In what follows, spectra are smoothed and normalized as mentioned above. PCA is conducted in two steps.

The first step consists in verifying that populations of both lymphocytes and *E. coli* can be distinguished. Figure 16 (a) shows that plotting PC2 vs. PC1 allows separating biological populations. The black line represents the frontier between these populations. It is defined by either the analysis of the mean and variance of the two distributions or the minima of the coordinates of lymphocytes (circles) and maxima of *E. coli* (crosses). Taking into account these data, a straight line which separates the PC1-PC2 domain in two zones is defined (Janné, 2001).

The second step consists in checking where a so-called “contaminated spectrum” is located in the PC1-PC2 space. Indeed, the 30 data used to generate figure 16(a) form a base for the pure lymphocyte and pure *E. coli* populations. Contaminated spectra correspond to all possible combinations of lymphocyte spectra added to *E. coli* spectra. They are processed one by one. Each contaminated spectrum is considered as a 31st data in the above mentioned base before a new PCA is performed. This is iterated for the 225 possible combinations. Indeed, in this PCA study, all *E. coli* spectra are considered and not only those corresponding to concentrations above 1×10^7 bact/mL as it was the case up to now. The result is shown in figure 16(b) with the contaminated spectra marked with red squares.

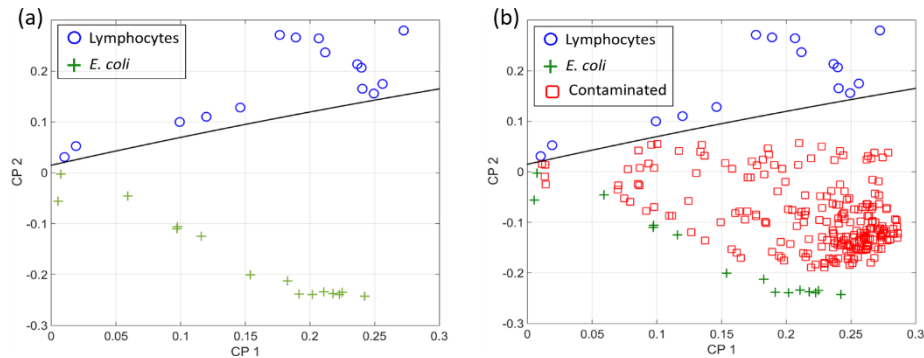


Fig. 16. Principal Component Analysis. Blue circles: lymphocytes, green crosses: *E. coli*, red squares: contaminated cultures and black line: separation between lymphocytes and *E. coli*. (a) ACP for pure lymphocyte and *E. coli*. (b) The same as (a) with contaminated cultures. Adapted from (Wacogne, 2020).

Contaminated spectra of contaminated culture are all situated in the *E. coli* region. This means that the bacteria detection limit is at least 1×10^6 bact/mL (the minimum *E. coli* concentration considered in this study). The warning signal can now be issued only 3h19min after a contamination with 1000 bact/mL occurred. Therefore, the in-line process proposed in Figure 13 can now be modified as presented in figure 17.

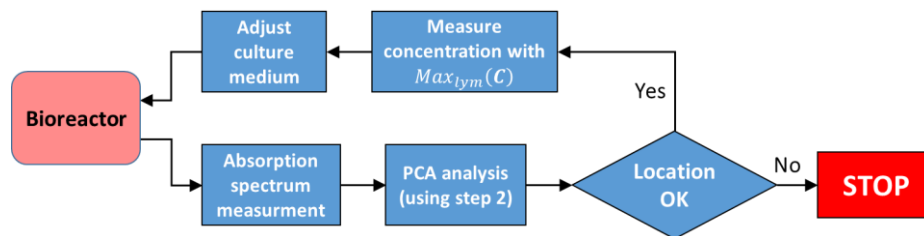


Fig. 17. In-line process to detect contaminations and to adjust the bioreactor composition based on Principal Component Analysis.

4 Discussion

4.1 Technical Aspects

Figures 4 to 6 show that analyzing transmission spectra using the associated color proves to be efficient to measure both species concentrations. Indeed, according to the results obtained in this study, this is the best way to measure *E. coli* concentration in real time and possibly in a closed system environment. In a general manner (for both lymphocyte and *E. coli*), we note that the “H” (Hue) and “a” (scale between green and red) parameters remain constant while the “b” (from blue to yellow) evolves very little. This means that the color of the solutions is not changing (or very little) with the con-

centration. The “L” (Luminance) and “V” (Value) parameters decrease with the concentration. This indicates that the intensity of the transmitted light decreases with the concentration. Although the color is not changing the associated color is reinforced with the concentration. This is observed with the increase of the “S” (Saturation parameter. “L” decreases but the color remains almost constant: this is confirmed by the fact that the “RGB” parameters decrease almost the same way with the concentration. Also, the relative values of the “RGB” parameters are constant which is consistent with a constant “H” parameter.

However, it is somehow inappropriate to talk of color when working with transmission spectra of biological solutions. Indeed, measuring the true color of the solutions would imply using a calibrated light source and studying the spectrum of the light reflected by the sample solution. Determining the true color of the solutions is not our purpose. Decomposition of the measured spectra in different color spaces is only a mathematical means of determining concentrations. For example, we also studied the entropy of the spectra. Estimating the quantity of information contained in the spectra using their entropy makes possible to measure concentration but in a reduced useful range which will not be presented in this paper.

Figures 8 to 12 show that, during ATMPs production, analyzing the shape absorption spectra of lymphocytes allows monitoring the expansion phase. However, looking closer at figures 8 and 9 shows that most of the coefficients used to describe the shape of the absorption spectra cannot be efficiently fitted with mathematical functions. Indeed, for lymphocytes, only the p4 function can be efficiently described (this is the same for *E. coli* considering q4). This means that a better concentration determination could be obtained using p4 only instead of the whole spectrum shape. This was not investigated in this study but looking at equation (4), we understand that this p4 function somehow represents the baseline of the recorded spectra. Functions p2 and p3 describe the shape of the spectra while p1, together with p4 contribute to the amplitude of the spectra. This is why indeed, measuring the concentration using the maxima of the absorption spectra (equation (13)) is particularly accurate (figure (12)).

This does not mean that the shape itself contains no useful information. It also provides powerful a tool to issue an alert signal less than 5 hours after a contamination with 1000 bact/mL occurred as it is shown in figure (15). Principal Component Analysis can be used to reduce the time required to issue a warning signal (figures (16)) the drawback being that it cannot be used to measure lymphocyte concentrations.

So-called “contaminated spectra” are artificial. They do not correspond to real lymphocyte and *E. coli* mixtures. These spectra consist of the addition of lymphocytes and *E. coli* spectra. Therefore, a bias could be introduced in the results presented here because adding absorption spectra may lead to artificial absorptions greater than 100%. The method described in section 3.3 is still valid because only the shape is considered and not the actual value of the maximum absorptions. We recall that the R^2 coefficient is only used to issue an alert signal. Equation depicted in figure 10(a) remains valid to monitor the expansion phase as long as no contamination occurs. Because it is performed using normalized spectra, results obtained using Principal Component Analysis are not affected. However, an ongoing and more realistic study involves real spectra recorded with real mixture of lymphocytes and *E. coli*.

Although they are not the only ones, we mentioned methods based on spectro(photo)metry in the introduction. One, based on Fourier Transform Infrared Spectroscopy (FTIR), was used to identify four types of bacteria (Hassan, 2016). This was not performed in solution but onto a sensor's surface. Also, Principal Component Analysis was used by the authors to successfully separate types of bacteria, to discriminate between Gram⁺ and Gram⁻ bacteria and to analyze mixtures. Another cited work presents the use of several sensors (Turbidity, particle counting, temperature, pH and spectrophotometry) to detect contaminants in drinking water (Ikonen, 2017). However, methods described in this paper seem to be extremely difficult to automate. Also in this paper, analysis was performed using only 2 wavelengths which drastically reduces the amount of available information. Note that 2 reduced wavelength ranges were used by Hassan (Hassan, 2016).

It should be noted that in our work, we exploit the total amount of information contained in the whole 450 nm to 1120 nm wavelength range (except for associated-color investigations) which represents over 900 data per spectrum with the spectrometer we used. Also, because measurements are performed in a cuvette and not on a biosensor's surface, the method can easily be used to provide a real time analysis of what happens in the bioreactor, in a closed-loop environment and without the need for any biosensor cleaning and/or regeneration.

From a practical point of view indeed, white light spectroscopy and conventional cuvettes makes possible an easy adaptation in a closed system configuration as illustrated in figure 18.

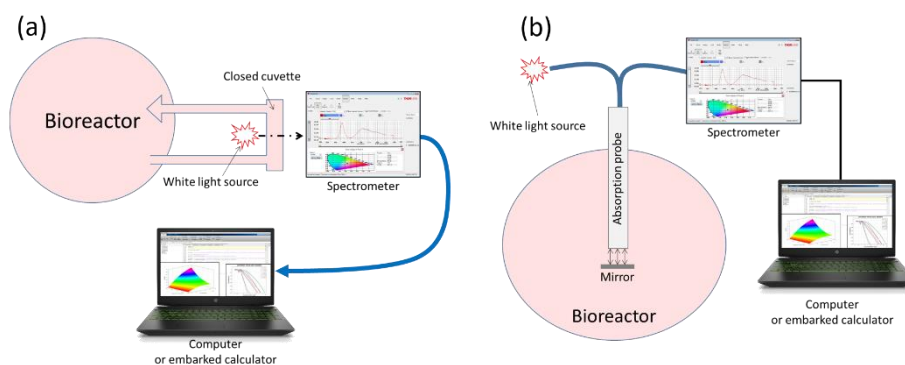


Fig. 18. Possible integrations of the method in a closed system environment. (a) Using a derivation. (b) Using a sterilized reflection probe coupled to a mirror. Adapted from (Wacogne, 2020).

In this work, only a contamination due to *E. coli* is considered. Cases of other bacteria or other types of containments like yeasts and fungi are currently being investigated. But as long as the shapes of the absorption spectra of contaminants are different enough from the ones of lymphocytes, the methods presented here should still be valid.

4.2 Socio-economic Impacts

The expansion phase lasts several days and each day increases the price of ATMPs. The quality control imposes regular samplings, themselves risk of contamination. Being able to monitor cell expansion and to perform quality controls without sampling is particularly interesting. In this way, expansion phase can be stopped rapidly which reduces the global cost of ATMPs. In this work, methods able to issue a warning signal about 3 hours after a contamination with 1000 *E. coli*/mL are described.

ATMPs are likely to create a real therapeutic revolution in the coming years because they are designed to treat patients with pathologies that are currently incurable. It is right now difficult to estimate the number of pathologies that these medicines from the living can address and the number of patients likely to benefit.

Only a few ATMPs are available on the market. The enormous production cost also led to marketing authorization cancellation of some of them (Glybera, Sipuleucel-T, ChondoCelect). At the date of manuscript writing, only the following treatments are available (note the price for a single treatment):

- KYMRIA (\$ 475,000)
- YESCARTA (\$ 373,000)
- Strimvelis (\$ 594,000)

Included in these costs are losses due to fabrications, which are found to be contaminated during the final conformity test and those due to late-stage fabrication stops due to the scheduling of controls at precise times. Having a real-time and closed system monitoring and quality control method is of great interest in terms of research, industrial manufacturing and more importantly in terms of benefit to the patients.

5 Conclusion

In this paper, we have presented spectroscopic methods to continuously perform a quality control during the expansion phase of ATMP production. These methods are based on analyzing the transmission or absorption optical spectra of the content of a bioreactor. Colorimetric based techniques can be used to monitor lymphocyte multiplication during the expansion phase of ATMP production. They can also be used to monitor *E. coli* proliferation in other applications. However, the colorimetric method cannot be used to simultaneously monitor the concentration of lymphocytes and bacteria when a contamination occurs.

The originality of this work is that we do not need to measure the contaminant concentration but only to detect its presence. Consequently, spectral shape analysis can be used not only to monitor the lymphocytes multiplication but also to issue an alert signal about 4h52min post contamination. Warning can even be issued earlier by using Principal Component Analysis. The warning signal can then be issued 3h19min post contamination. However, Principal Component Analysis cannot be used alone as it fails to measure lymphocyte concentrations when no contamination is detected. Note that the

warning delays are calculated for a contamination due to 1000 *E. coli*/mL. Advantages of such methods can be summarized as follows.

- Regular sampling of the content of the bioreactor is no longer required. This further reduces the risk of sampling related contaminations.
- The use of planned sampling at a fixed date and time without even knowing whether contamination is avoided.
- The goal being to stop the production as soon as a contamination is detected and not to identify the pathogen responsible for it, this identification can be performed later.
- As long as the culture is normal, the lymphocyte expansion is monitored in real time. Because contamination is detected extremely early the production can be stopped instantly which greatly contribute to reduce the production cost.

Indeed, in order to guarantee access of ATMPs to the largest number of patients, a new conception of the current mode of production and qualification of these living drugs is necessary. Currently, our studies are focused on the validation of these above described methods considering other types of pathogens.

Acknowledgements. This work was supported by the MiMedi project funded by BPI France (grant No. DOS0060162/00) and the European Union through the European Regional Development Fund of the Region Bourgogne-Franche-Comte (grant No. FC0013440).

References

1. Cacopardo, L., Costa, J., Giusti, S., Buoncompagni, L., Meucci, S., Corti, A., Mattei, G., Ahluwalia, A.: Real-time cellular impedance monitoring and imaging of biological barriers in a dual-flow membrane bioreactor. *Biosensors and Bioelectronics*. 140, 111340 (2019). <https://doi.org/10.1016/j.bios.2019.111340>
2. Cesewski, E., Johnson, B.N.: Electrochemical biosensors for pathogen detection. *Biosens Bioelectron*. 159, 112214 (2020). <https://doi.org/10.1016/j.bios.2020.112214>
3. Chang, K.-S., Jang, H.-D., Lee, C.-F., Lee, Y.-G., Yuan, C.-J., Lee, S.-H.: Series quartz crystal sensor for remote bacteria population monitoring in raw milk via the Internet. *Biosensors and Bioelectronics*. 21, 1581-1590 (2006). <https://doi.org/10.1016/j.bios.2005.07.015>
4. Hassan, M., Gonzalez, E., Hitchins, V., Ilev, I.: Detecting bacteria contamination on medical device surfaces using an integrated fiber-optic mid-infrared spectroscopy sensing method. *Sensors and Actuators B: Chemical*. 231, 646-654 (2016). <https://doi.org/10.1016/j.snb.2016.03.044>
5. Ikonen, J., Pitkänen, T., Kosse, P., Cizek, R., Kolehmainen, M., Miettinen, I.T.: On-line detection of *Escherichia coli* intrusion in a pilot-scale drinking water distribution system. *J. Environ. Manage*. 198, 384-392 (2017). <https://doi.org/10.1016/j.jenvman.2017.04.090>
6. Janné, K., Pettersen, J., Lindberg, N.-O., Lundstedt, T.: Hierarchical principal component analysis (PCA) and projection to latent structure (PLS) technique on spectroscopic data as a data pretreatment for calibration. *Journal of Chemometrics*. 15, 203-213 (2001). <https://doi.org/10.1002/cem.677>

7. Lee, S.-M., Han, N., Lee, R., Choi, I.-H., Park, Y.-B., Shin, J.-S., Yoo, K.-H.: Real-time monitoring of 3D cell culture using a 3D capacitance biosensor. *Biosensors and Bioelectronics*. 77, 56-61 (2016). <https://doi.org/10.1016/j.bios.2015.09.005>
8. Lei, K.F., Wu, M.-H., Hsu, C.-W., Chen, Y.-D.: Real-time and non-invasive impedimetric monitoring of cell proliferation and chemosensitivity in a perfusion 3D cell culture microfluidic chip. *Biosens Bioelectron*. 51, 16-21 (2014). <https://doi.org/10.1016/j.bios.2013.07.031>
9. Leva-Bueno, J., Peyman, S.A., Millner, P.A.: A review on impedimetric immunosensors for pathogen and biomarker detection. *Med Microbiol Immunol*. 209, 343-362 (2020). <https://doi.org/10.1007/s00430-020-00668-0>
10. Liu, Y.-J., André, S., Saint Cristau, L., Lagresle, S., Hannas, Z., Calvosa, É., Devos, O., Duponchel, L.: Multivariate statistical process control (MSPC) using Raman spectroscopy for in-line culture cell monitoring considering time-varying batches synchronized with correlation optimized warping (COW). *Anal. Chim. Acta*. 952, 9-17 (2017). <https://doi.org/10.1016/j.aca.2016.11.064>
11. Mansencal, T., Mauderer, M., Parsons, M., Shaw, N., Wheatley, K., Cooper, S., Vandenberg, J.D., Canavan, L., Crowson, K., Lev, O., Leinweber, K., Sharma, S., Sobotka, T.J., Moritz, D., Pppp, M., Rane, C., Eswaramoorthy, P., Mertic, J., Pearlstine, B., Leonhardt, M., Olli Niemitalo, Marek Szymanski, Maximilian Schambach: Colour 0.3.15. Zenodo (2020)
12. Melchor, J., López-Ruiz, E., Soto, J., Jiménez, G., Antich, C., Perán, M., Baena, J.M., Marchal, J.A., Rus, G.: In-bioreactor ultrasonic monitoring of 3D culture human engineered cartilage. *Sensors and Actuators B: Chemical*. 266, 841-852 (2018). <https://doi.org/10.1016/j.snb.2018.03.152>
13. Safavieh, M., Ahmed, M.U., Ng, A., Zourob, M.: High-throughput real-time electrochemical monitoring of LAMP for pathogenic bacteria detection. *Biosens Bioelectron*. 58, 101-106 (2014). <https://doi.org/10.1016/j.bios.2014.02.002>
14. Teixeira, A.P., Oliveira, R., Alves, P.M., Carrondo, M.J.T.: Advances in on-line monitoring and control of mammalian cell cultures: Supporting the PAT initiative. *Biotechnology Advances*. 27, 726-732 (2009). <https://doi.org/10.1016/j.biotechadv.2009.05.003>
15. Thakur, B., Zhou, G., Chang, J., Pu, H., Jin, B., Sui, X., Yuan, X., Yang, C.-H., Magruder, M., Chen, J.: Rapid detection of single E. coli bacteria using a graphene-based field-effect transistor device. *Biosensors and Bioelectronics*. 110, 16-22 (2018). <https://doi.org/10.1016/j.bios.2018.03.014>
16. Theint, H.T., Walsh, J.E., Wong, S.T., Voon, K., Shitan, M.: Development of an optical biosensor for the detection of *Trypanosoma evansi* and *Plasmodium berghei*. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*. 218, 348-358 (2019). <https://doi.org/10.1016/j.saa.2019.04.008>
17. Wacogne, B., Legrand, D., Pieralli, C., Frelet-Barrand, A.: Optical Spectroscopy for the Quality Control of ATMP Fabrication: A New Method to Monitor Cell Expansion and to Detect Contaminations. In: *Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies - Volume 1: BIODEVICES*. p. 64-72. SciTePress (2020)
18. Wang, X., Rivière, I.: Clinical manufacturing of CAR T cells: foundation of a promising therapy. *Mol Ther Oncolytics*. 3, 16015 (2016). <https://doi.org/10.1038/mto.2016.15>