



## Monitoring the water content in NADES extracts from spirulina biomass by means of ATR-IR spectroscopy

Suha Elderderi, Soukaina Hilali, Laura Wils, Igor Chourpa, Martin Soucé, Barbara Clément-Larosière, Abdalla Elbashir, Hugh Byrne, Emilie Munnier, Leslie Boudesocque-Delaye, et al.

### ► To cite this version:

Suha Elderderi, Soukaina Hilali, Laura Wils, Igor Chourpa, Martin Soucé, et al.. Monitoring the water content in NADES extracts from spirulina biomass by means of ATR-IR spectroscopy. *Analytical Methods*, 2022, 14 (20), pp.1973-1981. 10.1039/D2AY00234E . hal-04302019

**HAL Id: hal-04302019**

**<https://hal.science/hal-04302019>**

Submitted on 23 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## **Monitoring water content in NADES extracts from *Spirulina* biomass by means of ATR-IR spectroscopy**

Suha Elderderi<sup>a,b</sup>, Soukaina Hilali<sup>c</sup>, Laura Wils<sup>c</sup>, Igor Chourpa<sup>a</sup>, Martin Soucé<sup>a</sup>, Barbara Clément-Larosière<sup>d</sup>, Abdalla A. Elbashir<sup>e,f</sup>, Hugh J. Byrne<sup>g</sup>, Emilie Munnier<sup>c</sup>, Leslie Boudesocque-Delaye<sup>c</sup>, Franck Bonnier<sup>a,\*</sup>

<sup>a</sup> Université de Tours, Faculté de pharmacie, EA 6295 Nanomédicaments et Nanosondes, 31 avenue Monge, 37200 Tours, France

<sup>b</sup> University of Gezira, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, P.O. Box 20, 21111 Wad Madani, Sudan.

<sup>c</sup> Université de Tours, EA 7502 Synthèse et Isolement de Molécules BioActives (SIMBA), 31 avenue Monge, 37200 Tours, France

<sup>d</sup> AquaEcoCulture, 7 rue des Blossières Maroue BP 60328, 22400 Lamballe, France

<sup>e</sup> King Faisal University, College of Science, Department of Chemistry, P.O. Box 400, 31982 Al-Ahsa, Saudi Arabia

<sup>f</sup> University of Khartoum, Faculty of Science, Department of Chemistry, P.O. Box 321, 11115 Khartoum, Sudan

<sup>g</sup> FOCAS Research Institute, TU Dublin, City Campus, Camden Row, Dublin 8, Ireland

**\*Corresponding author**

franck.bonnier@univ-tours.fr

**Key words:** Label free water quantification, natural deep eutectic solvent, Extraction, *Spirulina*, Partial least squares regression, Attenuated Total Reflectance–infrared spectroscopy.

## **Abstract**

Attenuated Total Reflectance–infrared spectroscopy (ATR-IR) coupled to Partial Least Squares Regression (PLSR), was evaluated as a rapid, label free and cost-effective tool to quantify water content in extracts obtained from *Spirulina* wet biomass using a Glucose Glycerol Natural Deep Eutectic solvent (NADES). NADES are green, renewable and biodegradable solvents with unique properties outcompeting existing organic solvents, for instance for plant or biomass extraction. The properties of NADES depend critically on their water concentration, and therefore it is essential to develop methods to monitor it, to ensure optimal extraction efficiency and experimental repeatability to achieve a better standardization of extraction protocols. Firstly, Karl Fischer titration was performed on a set of 20 NADES extracts in order to obtain reference water concentrations. Secondly, ATR-IR spectra were collected and subjected to datamining to construct PLSR predictive models. The  $R^2$  values of 0.9996, mean Root Mean Square Error of Cross Validation of 0.136% w/w and Root Mean Square Error of prediction of 0.1304% w/w highlight the feasibility and reliability to perform quantitative analysis using ATR-IR. Moreover, the mean relative error percentage achieved, ~0.5%, confirms the high accuracy of water concentration determination in NADES extracts. This work demonstrates that powerful alternatives are available to provide more environmentally responsible analytical protocols. ATR-IR spectroscopy applied to NADES extracts does not require any sample preparation, no reagents or solvent and has minimal requirements for single use consumables. The technique is consistent with current concerns to develop greener chemistry, especially in the field of extraction of natural compounds from plants which currently represents a major focus of interest in both research and industry.

## 1. Introduction

Natural Deep Eutectic solvents (NADES) are firstly announced by Chio and co-workers in 2011<sup>1</sup>. NADES have been studied as a sustainable green alternative to organic solvents<sup>2</sup>. NADES are prepared by mixing a hydrogen bond donor (HBD), such as polyols or sugars, and a hydrogen bond acceptor (HBA), such as choline chloride betaine, glucose or alanine, that results in a hydrogen bonded donor-acceptor network<sup>1</sup>. NADES reported in the literature are mostly hydrophilic, with varying polarity, depending on the combination of HBA and HBD.

NADES are renewable, biodegradable and biocompatible solvents<sup>4</sup>, whose properties are easily tunable<sup>5</sup> for applications in a broad range of fields, including analytical chemistry<sup>6</sup>, organic synthesis<sup>7</sup>, biotechnology<sup>8,9</sup>, electrochemistry<sup>10</sup>, nanotechnology<sup>11-13</sup>, energy<sup>14</sup>, water remediation<sup>15</sup>, food industry<sup>16-18</sup>, cosmetics and pharmaceuticals industries<sup>19,20</sup>. The high solubilisation power and their high extraction ability for some natural products make them promising media for green processing<sup>6,21-23</sup>. NADES have been studied as solvents for solid/liquid (S/L) extraction of polyphenols and flavonoid compounds from plants<sup>4,6,24</sup>, or extraction of lipids and pigments from microalgae<sup>25,26</sup>. Moreover, the stabilizing ability of NADES suggest them as perfect media for sustainable valorisation of vegetal biomasses. Hence, they are valuable environmental friendly alternatives in extraction processes to organics solvents, which are increasingly challenged by greener methods<sup>24</sup>.

Water can be added in certain ratios to tailor NADES properties, commonly between 5% w/w and 30 % w/w, according to the desired application<sup>27,28</sup>. Water affects physicochemical properties like viscosity, polarity, and density of NADES<sup>29</sup> and it strongly affects the extraction yield<sup>27,29-31</sup>. Most NADES are composed from hygroscopic materials, which absorb atmospheric moisture during storage or from biomass during extraction<sup>30</sup>. Variations in water content of NADES can lead to loss of efficiency, while high content can even result in the disruption of the hydrogen network<sup>29</sup>. Therefore, being able to rapidly determine and continuously monitor water concentration of NADES is key to ensuring repeatable and reproducible use and achieving a better standardisation of extraction protocols.

The gravimetric quantification of water or moisture by weighing and drying is the simplest, solvent-free and cost-effective approach. The main drawback is the lack of repeatability in results due to thermal decomposition of sample or for volatile samples<sup>32</sup>. It has also been demonstrated that, in materials with high viscosity, the rubbery matrix formed during drying makes water diffusion and evaporation difficult, leading to underestimation of moisture levels<sup>33,34</sup>. Karl Fischer titration (KFT) is the gold standard for accurate quantification of water traces<sup>35</sup>. However, the technique consumes large volumes of reagents and solvents during titration (especially for high water content), in contradiction with the concept of developing green solvents like NADES. Moreover, the time requirements for analysis of large sample cohorts<sup>32</sup> further motivates the development of alternative methods.

To this end, a rapid, and label free analysis performed by Attenuated Total Reflectance infrared spectroscopy (ATR-IR) coupled to multivariate data mining protocols is investigated. ATR-IR is a powerful non-destructive characterisation technique providing specific spectral signatures without the requirement for pre-analytical sample preparation<sup>36</sup>. It enables specific probing of samples at the molecular level, for functional groups identification and structural characterization of samples. It is a well-established technique in organic chemistry<sup>37</sup>, biochemistry<sup>38-40</sup>, nanotechnology<sup>41</sup> and pharmacy<sup>42,43</sup>. ATR-IR spectroscopy shows the potentials to deliver reliable quantitative information with applications in the food industry<sup>44</sup>, clinical and biomedical sciences<sup>45</sup>, pharmaceuticals science<sup>46-48</sup>, chemotherapeutic solutions<sup>49,50</sup> and quantification of cosmetic ingredients in complex formulations<sup>51</sup>. As ATR-IR does not require reagents or solvent for analysis, it is a strong candidate as an analytical alternative, consistent with environmental concerns.

Recently, ATR-IR coupled to partial least squares regression (PLSR) analysis has been studied for quantification of water in 3 NADES systems; Betaine: Glycerol (1: 8), Choline Chloride: Glycerol (1: 2) and Glucose: Glycerol (1: 3), each with systematically varying added water concentrations in the range 0% w/w to 40% w/w<sup>36</sup>. The study highlighted the potential of ATR-IR spectroscopy to monitor the water content of *in vitro* models, simulating prolonged storage periods. However, it is particularly relevant to be able to monitor water transfer from a wet biomass to the NADES during extraction. Therefore, in the present study, a set of extractions was performed from *Spirulina* biomass<sup>25</sup> using

Glucose Glycerol (GG) NADES, a hydrophilic NADES commonly used with initially water content of  $\sim 20\%$  w/w<sup>25,26</sup>.

Based on the previous demonstrations of the potential of the technique, the aim of this work is to investigate the ability of ATR-IR to deliver accurate water quantification in a practical example of the NADES extraction process. Firstly, KFT was applied to determine water concentration in extracts to provide reference concentrations. Secondly, infrared spectra were collected and data mining protocols were applied to construct predictive models and evaluate the accuracy on the analysis performed.

## 2. Materials and Methods

### 2.1. Reagents

*Preparation of NADES:* Glycerol 98% was purchased from Fischer Scientific (Illkirch, France). Glucose 99% were purchased from Acros Organics (Geel, Belgium). Water was purified using a Milli-Q system (Millipore Corporation, Bedford, MA, USA).

*Karl Fischer titration:* Hydranal<sup>TM</sup> Composite 5K (Honeywell, Fluka<sup>TM</sup>, Germany), Methanol  $\geq 99,9\%$  (Honeywell, Riedel-de Haën<sup>TM</sup>, France) and Sodium Tartrate (Carlo ERBA reagents, Milano, Italy) were purchased from Sigma Aldrich (Burlington, MA, USA).

Spirulina (*Arthrospira platensis*) biomass (9.5% dry matter) was kindly provided by Aqua Eco Culture SA (Lamballe, France) and stored at  $-20^{\circ}\text{C}$ .

### 2.2. Preparation of NADES samples from extraction

*NADES preparation:* The NADES selected for this study is Glucose-Glycerol (GG) NADES, which is a hydrophilic system commonly used for extraction with  $\approx 20\%$  w/w added water. GG NADES was prepared using the heating and stirring method<sup>28</sup>, by mixing a HBA (i.e. glucose), a HBD (i.e. glycerol) in the ratio (1:8) and the 20% added water. Then mixture was heated at  $70^{\circ}\text{C}$  and stirred until a colourless liquid was obtained. GG NADES were prepared with an absolute initial water concentration (i.e. prior to extraction) =  $16.6\%$  w/w.

*Extracts from Spirulina biomass:* The purpose of the study is to investigate variations of water concentration in NADES during extractions from wet biomass. Therefore, a set of 20 extractions was conducted by adding different masses of *Spirulina* biomass, ranging over 0.2g - 8.33g, to 20g of GG NADES. The biomass/NADES ratio thus ranged from 0.01 to 0.4165. Firstly, the mixtures were prepared by stirring the biomass and the NADES in appropriate ratio for 5 min, at room temperature. Then, samples were centrifugated for 20 minutes at 16.200 g (Rotanta 460R, Hettich) and the supernatant were recovered. NADES extracts are henceforth named SPGG. All extracts were analyzed in terms of phycocyanin, chlorophylls and carotenoids according to Wils et al.<sup>26</sup>

### **2.3. Data collection**

#### **2.3.1. Karl Fischer titration (KFT)**

Karl Fischer titration was performed using a Titroline® 7500KF (Xylem SI analytics, Germany). Firstly, the KF reagent (Hydranal™ Composite 5K) has been standardised daily, before SPGG extracts analysis. While the water equivalent of the reagent is specified by the manufacturer to be between 4.5-5.5 mg/ml, sodium tartrate was used as primary standard to determine its exact titer. Approximately 45 mg were added to the titration vessel. The KFT was conducted in triplicate (n=3) for all samples. Water concentrations calculated from the KFT were used as reference values to construct the PLSR models from the ATR-IR data (See section 2.4).

#### **2.3.2. ATR-IR Spectroscopy**

IR spectra were acquired using a Frontier spectrometer (Perkin Elmer, France) equipped with a Quest single reflection diamond attenuated total reflectance (ATR) accessory (Specac, UK). The spectral range was set between 4000-400  $\text{cm}^{-1}$  and the spectral resolution at 4  $\text{cm}^{-1}$ . Drops of  $\approx 100 \mu\text{L}$  were deposited directly onto the diamond surface and spectroscopic measurements were performed immediately. **Prior to sample measurement, a background spectrum was recorded in the air (4 scans) and automatically ratioed with the sample spectrum (4 averaged scans) by the software.** For each sample, 5 deposits have been measured and 3 spectra per drop have been collected. Ultimately, 15

spectra were recorded from each sample, capturing the inter-and intra-variability during measurements. Spectra from pure compounds have also been collected using similar parameters. The entire operation including cleaning the ATR crystal, collection of background, and collecting the IR spectrum from the sample takes less than 30 seconds.

#### **2.4. Data handling and analysis**

Data pre-processing and analysis were performed using MATLAB® (The Mathworks, USA). It was observed that pre-processing of data does not improve the reliability of the quantitative analysis. This is consistent with previous investigations conducted with ATR-IR for analysis of NADES<sup>36</sup> and cosmetic formulations<sup>52</sup>. Therefore, the raw data were used throughout this study. Partial Least Squares Regression (PLSR) is used to demonstrate the quantitative performances of ATR-IR applied to the determination of water concentrations in NADES extracts from biomass. In this study, the data set is limited to 20 samples, and therefore a Leave One Out Cross Validation (LOOCV) protocol has been applied to construct PLSR models. In this approach, all the spectra of one sample are selected as the test set, i.e. unknown sample to be determined, while all other samples (n=19) are used as the training set. In turn, all 20 samples will be used as test against all others. The training set (n= 19 samples) is further divided into the calibration (2/3 of data randomly selected) and the validation set (remaining 1/3 of the data). A 100-fold iteration was implemented to evaluate the robustness of the analysis with multiple random combinations of calibration/validation sets and to provide an overall estimation of the reliability of the data collected. The PLSR output is evaluated using the linearity of the regression between the measured and predicted concentrations ( $R^2$ ), the mean of the Root Mean Square Error of Cross-Validation (RMSECV) calculated from the training data (i.e. n= 19 samples), the Root Mean Square Error of Prediction (RMSEP) calculated from test samples and the % of relative error between the mean predicted concentrations and the target concentration (observed concentration) for each test sample. The regression co-efficient, which represents the variables (wavenumbers) used to create the regression model is used to highlight the specificity of the analysis performed.

### **3. Results and Discussions**



### 3.1. Karl Fischer titration (KFT)

Results from the KFT are presented in Table 1. 20 extractions were conducted, using GG NADES with varying masses of SP biomass in order to cover a sufficiently broad range of water concentrations to apply PLSR. For each extract, the mass of Spirulina added to GG NADES and the corresponding ratio biomass/NADES are provided. The mean and standard deviation (SD) are calculated from the KFT performed in triplicate on each extract. GG NADES is used with a starting water concentration of  $\approx 16.6\%$  w/w. After extraction of SP, the water concentrations were found to range from 16.79% w/w to 38.12% w/w, illustrating the water transfer from the wet biomass to the SPGG extracts. It is confirmed that water concentration found in extracts increases according to the biomass/NADES ratio. In the present study, KFT has been used as the reference technique to provide known concentrations of water in extracts. At the later stage, these concentrations were used to construct PLSR models from the ATR-IR spectra collected.

*Table 1. List of SPGG NADES extracts analysed with their corresponding biomass weight (g), biomass/NADES ratio and water concentration determined by KFT.*

Sample no	Biomass weight (g)	Biomass/NADES ratio	water content (mean $\pm$ SD) %w/w
1	0.22	0.011	17.31 $\pm$ 0.14
2	0.24	0.012	16.79 $\pm$ 0.14
3	0.27	0.014	17.00 $\pm$ 0.14
4	0.28	0.014	17.20 $\pm$ 0.16
5	0.30	0.015	17.67 $\pm$ 0.07
6	0.33	0.017	17.51 $\pm$ 0.05
7	0.37	0.019	17.80 $\pm$ 0.04
8	0.42	0.021	18.47 $\pm$ 0.09
9	0.48	0.024	18.49 $\pm$ 0.06
10	0.56	0.028	18.50 $\pm$ 0.14
11	0.67	0.034	19.26 $\pm$ 0.04
12	0.83	0.042	20.25 $\pm$ 0.05
13	1.11	0.056	21.02 $\pm$ 0.05
14	1.67	0.084	21.80 $\pm$ 0.04
15	3.33	0.167	25.90 $\pm$ 0.09
16	3.33	0.167	25.98 $\pm$ 0.04
17	3.33	0.167	24.97 $\pm$ 0.02
18	4.99	0.250	30.82 $\pm$ 0.16
19	6.66	0.333	35.01 $\pm$ 0.05

20	8.33	0.417	38.12 ± 0.24
----	------	-------	--------------

## 3.2. ATR-IR spectroscopy

### 3.2.1. Quantification of water in SPGG extract

#### A. Spectral characterisation

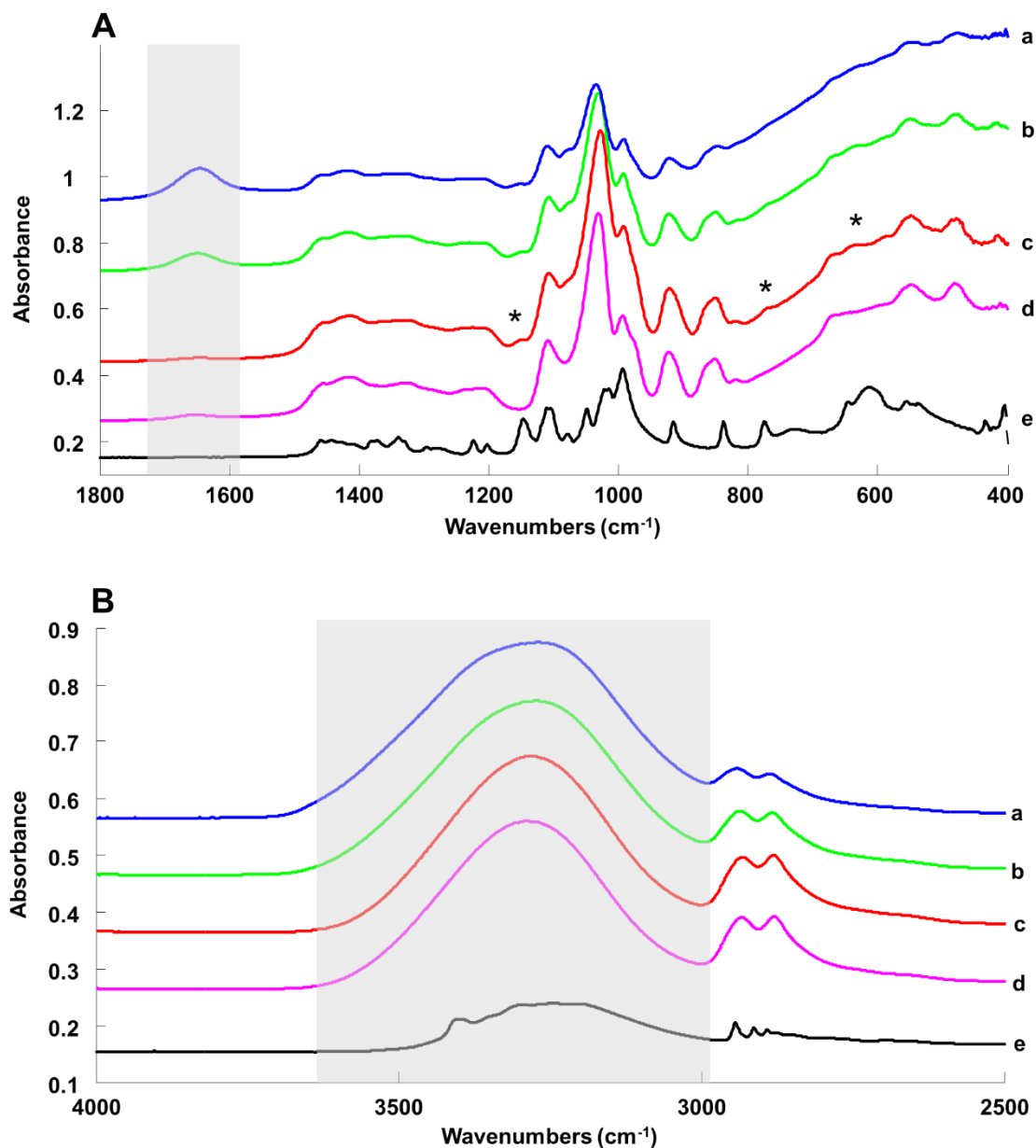
Figure 1 compares ATR-IR mean spectra obtained from SPGG extracts with the highest (38.12% w/w - figure 1-a) and the lowest (16.79 % w/w - figure 1-b) water content, to those of GG NADES with 0 % w/w added water, although not used in this study, but it is to characterise NADES without influence from water (figure 1-c), glycerol (figure 1-d) and glucose (figure 1-e). The water features are typically observed in the spectral range 1500-1750  $\text{cm}^{-1}$  with a maximum at  $\sim 1635 \text{ cm}^{-1}$  (HOH scissoring) and in the spectral range 3000-3700 with a maximum at  $\sim 3330 \text{ cm}^{-1}$  (OH symmetric and asymmetric stretching) (figure 1 A and 1 B, highlighted in grey).

In the fingerprint region (400-1800  $\text{cm}^{-1}$ ), compared to the spectrum of the GG NADES with 0% w/w added water, it is observed that the  $\text{H}_2\text{O}$  feature at  $1635 \text{ cm}^{-1}$  does not overlap with other bands from glycerol (figure 1-d) and glucose (figure 1-e). For glycerol, the main features are observed at 993  $\text{cm}^{-1}$  (C–O stretching vibration in  $\text{CH}_2\text{OH}$ ), 1031  $\text{cm}^{-1}$  (C–O stretching vibration in  $\text{CH}_2\text{OH}$ ), 1108  $\text{cm}^{-1}$  (C–O stretching vibration in  $\text{CHOH}$ ), 1208  $\text{cm}^{-1}$  (C–O vibration), 1327  $\text{cm}^{-1}$  (in-plane rocking of OH vibration) and 1410  $\text{cm}^{-1}$  (stretching of C–O in  $\text{CH}_2\text{OH}$ ) (figure 1A-d) <sup>53,54</sup>. For glucose, the main features are found at 612  $\text{cm}^{-1}$  ( $\text{CH}_2$  bending vibration), 774  $\text{cm}^{-1}$  (CCO + CCH bending vibration), 838  $\text{cm}^{-1}$  (CH bending vibration), 915  $\text{cm}^{-1}$  (CO + CCH stretching vibration), 994  $\text{cm}^{-1}$  (CO + CC stretching vibration), 1020  $\text{cm}^{-1}$  (CO stretching vibration), 1050  $\text{cm}^{-1}$  (CO + CC stretching vibration), 1109  $\text{cm}^{-1}$  (CO stretching vibration vibration), 1147  $\text{cm}^{-1}$  (CO + CC stretching vibration vibration), 1203  $\text{cm}^{-1}$  (CH + OH in plane bending vibration), 1337  $\text{cm}^{-1}$  (CCH + OCH bending vibration) and 1441  $\text{cm}^{-1}$  ( $\text{CH}_2$  + OCH+ CCH bending vibration) (figure 1A-e) <sup>55</sup>. Due to the glucose-glycerol 1:8 molar ratio used to prepare the GG NADES, features from glucose are not clearly observed in

its spectrum, except for the weak bands at  $600\text{ cm}^{-1}$ ,  $771\text{ cm}^{-1}$  and  $1151\text{ cm}^{-1}$  (figure 1A-c labelled with \*).

In the high-wavenumber region ( $2500\text{--}4000\text{ cm}^{-1}$ ), the contribution from glycerol is identifiable as the broad band with a maximum at  $3292\text{ cm}^{-1}$  (OH stretching vibration), overlapping significantly with the  $\text{H}_2\text{O}$  feature. Two weaker features at  $2880\text{ cm}^{-1}$  (symmetric  $\text{CH}_2$  stretching vibration) and  $2934\text{ cm}^{-1}$  (asymmetric  $\text{CH}_2$  stretching vibration)<sup>53,54</sup> can be specifically assigned to glycerol. In this spectral range, the contribution from glucose is not perceptible. In neither spectral region are features of the biomass itself observed.

As the water content increases, changes in shape and/or width of features from water are observed. The Area Under the Curve (AUC) calculated from the water band in the finger region ( $1580\text{--}1740\text{ cm}^{-1}$ ) gives values of 11.04 for the spectrum corresponding to the SPGG extract with the highest water content (figure 1A-a), compared to an AUC of 6.70 for the spectrum of the extract with the lowest water content (figure 1A-b). In the high wavenumber region ( $3000\text{--}3700\text{ cm}^{-1}$ ), AUCs were found to be 123.6 (figure 1B-a) and 113.7 (figure 1B-b). The lesser difference between the samples in this region can be attributed to the overlap of the water band with the main glycerol band. It is observed that the initial absorbance for GG NADES with 0% added water (figure 1B-c) is already strong with  $A = 0.311$ . As a result, in this region the variation in intensities measured is reduced with respectively  $A = 0.313$  and  $A = 0.310$  for the highest water content (figure 1A-a) and the lowest water content (figure 1A-b). However, the increase in water concentration in extracts also induces a modification of the shape of the broad water bands, evident for example as a change in the bandwidth at a half height, from  $391\text{ cm}^{-1}$  (Figure 1B-a) to  $358\text{ cm}^{-1}$  (Figure 1B-b). Moreover, the features from glycerol in the fingerprint and the high wavenumbers regions show inversed behaviour to those of water, decreasing in intensity as the water content increases (figure 1 A, B).



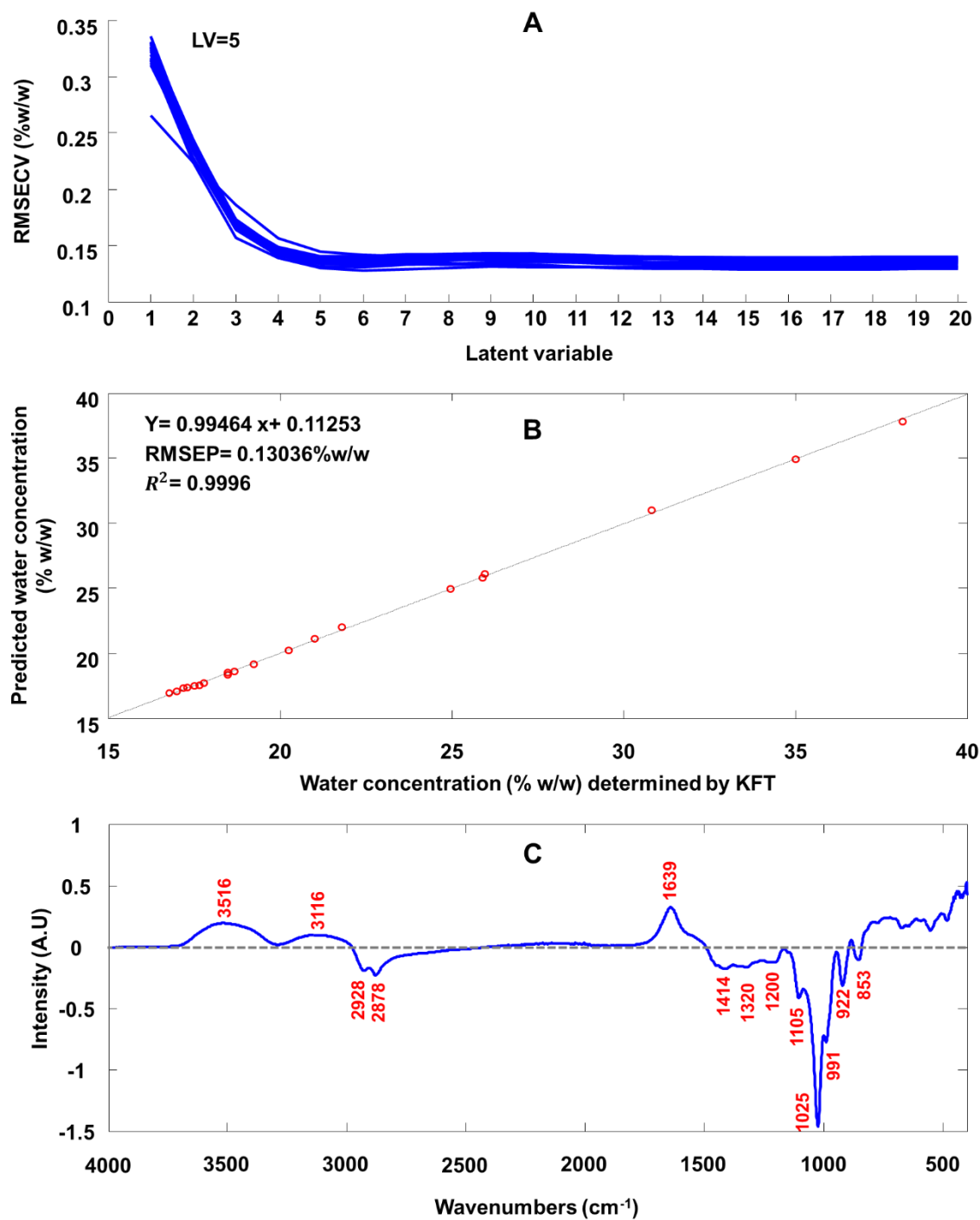
**Figure 1** Mean ATR-IR spectra in range 400-1800 cm<sup>-1</sup> (A) and 2500- 4000 cm<sup>-1</sup> (B) collected from SPGG 38.12% w/w water (a), SPGG 16.79 % w/w (b), GG 0 % w/w added water (c), glycerol (d) and glucose (e). \* Indicate features from glucose observed in NADES. Spectra have been offset for clarity.

## B. PLSR analysis

PLSR results are presented in figure 2. A LOOCV approach was preferred to perform the analysis due to the limited number of samples,  $n = 20$  SPGG extracts, included in this study. In turn each sample was tested as an unknown to be determined,

the other 19 acting as training set, resulting in 20 calibration models. Figure 2A presents the plots of RMSECV as a function of number of latent variables (LVs). It is observed that all curves except for one exhibit a similar profile of an initial RMSECV of  $\approx 0.3$ - $0.35\%$  w/w, which decreases to an approximately constant value of  $\approx 0.12$ - $0.15\%$  w/w within the 5 first LVs. Therefore, 5 LVs were chosen to construct all PLSR models for this study, delivering a mean RMSECV =  $0.136\%$  w/w, considering that beyond this point no improvement in the reliability of the quantification can be reached. Figure 2B shows the PLSR fitting as a regression of the % w/w water concentration obtained from the KFT against the % w/w water concentrations predicted from ATR-IR data. The  $R^2$  of  $0.999$  highlights the linearity of the PLSR fitting. The RMSEP =  $0.130\%$  w/w, representing  $0.47\%$  of the median concentration of the analysis range, is an indicator of the accuracy of the quantification achieved.

Figure 2C displays the regression coefficient, obtained from one example of LOOCV PLSR analysis, confirming that the regression is based on combined contributions from water and GG NADES features. In the fingerprint region, the positive feature at  $1639\text{ cm}^{-1}$  is assigned to the water feature. Negative features at  $853\text{ cm}^{-1}$ ,  $922\text{ cm}^{-1}$ ,  $991\text{ cm}^{-1}$ ,  $1025\text{ cm}^{-1}$ ,  $1105\text{ cm}^{-1}$ ,  $1200\text{ cm}^{-1}$ ,  $1320\text{ cm}^{-1}$  and  $1414\text{ cm}^{-1}$  are assigned to glycerol (figure 1A-d), while weak peaks at  $672\text{ cm}^{-1}$ ,  $772\text{ cm}^{-1}$  are assigned to glucose (figure 1A-e). In the high-wavenumber region, positive features at  $3516\text{ cm}^{-1}$  and  $3116\text{ cm}^{-1}$  are assigned to modification of the water band. The negative features at  $2928\text{ cm}^{-1}$  and  $2878\text{ cm}^{-1}$  are assigned to glycerol. No feature from glucose is detected in this spectral region, due to its weak contribution, as observed in figure 1B.



*Figure 2* PLSR outcomes obtained using ATR-IR Raw Spectra from SPGG extracts. A: Overlaid RMSECV plot from the 20 training sets, B: PLS regression prediction plot, and C: Example of first regression coefficient from PLSR.

Percent relative errors obtained for SPGG extracts range from 0.111% to 1.075% (Table 2). They further highlight the accuracy of the predicted water concentrations. Overall, the mean percent relative error is found to be 0.507%. The water concentrations were determined with less than 0.5% relative error for 10 out of 20 samples analysed. The other ten are distributed between 0.5% and 1.25% percent relative errors. There is no correlation between water concentrations and observed percent relative errors. The 2 highest values are 1.075% for sample 14 (21.79% w/w) and 1.006% for sample 2 (16.79% w/w). The lowest % relative error, 0.111%, is found for sample 17 (24.97 % w/w) and sample 1 (17.31% w/w) delivers a comparable error of 0.170%. It appears the distribution of % relative error is random over the range of concentrations analysed. In a previous study, Elderderi *et al.* have reported water quantification in GG NADES<sup>36,56</sup>. PLSR analysis performed on NADES samples with systematically varying added % water concentration between 0% and 40% gave RMSECV = 0.801% w/w and RMSEP = 1.157% w/w<sup>36,56</sup>. Percent relative errors were determined to be between 0.74% - 11.16%, with a mean equal to 4.081%. Results obtained from SPGG extracts appear more reliable, considering that the RMSCV and RMSEP are respectively 83.01% and 88.7% lower. Water is a strong absorber, making infrared spectroscopy highly sensitive to modification in its content in extracts. Moreover, extraction from wet biomass further increases the % w/w water concentration in samples, positioning the analysis within the range delivering the best accuracy as water concentrations in SPGG extracts range from  $\approx 17\%$  w/w to  $\approx 38\%$  w/w (Table 1) that explains the high performances for ATR-IR witnessed presently.

**Table 2:** Summary of PLSR accuracy for SPGG extracts

Range	% Relative Error						Min- max value
	< 0.25%	< 0.5%	< 0.75%	< 1%	< 1.25%	Mean	
Full range	4	6	6	2	2	0.5066	0.111-1.075

## Conclusion

ATR-IR coupled to multivariate data analysis is demonstrated to be a powerful analytical tool for water quantification in biomass extracts. Spectral variations can be specifically correlated to modifications in water concentrations in extracts. The high

degree of accuracy of the PLSR predictive models confirms the reliability of the analysis and the suitability of the technique to monitor water in NADES based extracts. ATR-IR spectra are specific molecular fingerprints of SPGG extracts, enabling observation of variations in both water and NADES features. Ultimately, the concentrations can be determined with % relative errors below 1.25% w/w. ATR-IR is therefore a suitable tool for monitoring water content in NADES extracts and the technique offers a rapid, cost-effective and reagent free alternatives to current gold standards such as KFT. Ultimately, this study highlights the potential of ATR-IR to support the establishment of green chemistry protocols, further supporting optimisation of processes for extraction of active molecules. It addresses current imperatives to develop environmentally responsible solutions, including all facets, i.e. extractions performed with natural solvents such as NADES, and analytical tools which are less demanding in terms of reagents, solvents and consumables. Considering the context, ATR-IR has the potential to become the next gold standard methods in both research and industry.

**Conflicts of interest** There are no conflicts to declare

**Acknowledgements** This work received financial support from the French National Research Agency (DERMIC Project ANR-19-CE43-0001-01) and the Ambition Recherche et Développement Centre Val de Loire (PIERIC project, ARD CVL 2020-00141275).

Suha Elderderi thanks the Ministry of Higher Education (Sudan), University of Gezira (Sudan), the French embassy (Sudan) and Campus France and Laura Wils thanks the Région Centre-Val de Loire for their respective PhD grants.

## References

- 1 H. Vanda, Y. Dai, E. G. Wilson, R. Verpoorte and Y. H. Choi, *Comptes Rendus Chimie*, 2018, **21**, 628–638.
- 2 D. J. G. P. van Osch, L. F. Zubeir, A. van den Bruinhorst, M. A. A. Rocha and M. C. Kroon, *Green Chem.*, 2015, **17**, 4518–4521.
- 3 A. K. Dwamena, *Separations*, 2019, **6**, 9.
- 4 Y. Liu, J. B. Friesen, J. B. McAlpine, D. C. Lankin, S.-N. Chen and G. F. Pauli, *J. Nat. Prod.*, 2018, **81**, 679–690.



- 5 Y. Dai, G.-J. Witkamp, R. Verpoorte and Y. H. Choi, *Food Chemistry*, 2015, **187**, 14–19.
- 6 M. de L. Á. Fernández, J. Boiteux, M. Espino, F. J. V. Gomez and M. F. Silva, *Anal. Chim. Acta*, 2018, **1038**, 1–10.
- 7 P.-O. Delaye, M. Pénichon, L. Boudesocque-Delaye, C. Enguehard-Gueiffier and A. Gueiffier, *SynOpen*, 2018, **02**, 306–311.
- 8 Y. P. Mbous, M. Hayyan, A. Hayyan, W. F. Wong, M. A. Hashim and C. Y. Looi, *Biotechnology Advances*, 2017, **35**, 105–134.
- 9 Z. Yang, *Adv Biochem Eng Biotechnol*, 2019, **168**, 31–59.
- 10 F. S. Ghareh Bagh, K. Shahbaz, F. S. Mjalli, M. A. Hashim and I. M. AlNashef, *Journal of Molecular Liquids*, 2015, **204**, 76–83.
- 11 A. Abo-Hamad, M. Hayyan, M. A. AlSaadi and M. A. Hashim, *Chemical Engineering Journal*, 2015, **273**, 551–567.
- 12 M. Tohidi, F. A. Mahyari and A. Safavi, *RSC Adv.*, 2015, **5**, 32744–32754.
- 13 H.-G. Liao, Y.-X. Jiang, Z.-Y. Zhou, S.-P. Chen and S.-G. Sun, *Angewandte Chemie International Edition*, 2008, **47**, 9100–9103.
- 14 M. Atilhan and S. Aparicio, *Energy Fuels*, 2021, **35**, 6402–6419.
- 15 S. H. Chang, *Environ Sci Pollut Res Int*, 2020, **27**, 32371–32388.
- 16 A. Mišan, J. Nađpal, A. Stupar, M. Pojić, A. Mandić, R. Verpoorte and Y. H. Choi, *Crit Rev Food Sci Nutr*, 2020, **60**, 2564–2592.
- 17 Y. Dai, R. Verpoorte and Y. H. Choi, *Food Chemistry*, 2014, **159**, 116–121.
- 18 K. Radošević, N. Ćurko, V. Gaurina Srček, M. Cvjetko Bubalo, M. Tomašević, K. Kovačević Ganić and I. Radojčić Redovniković, *Food Science and Technology*, 2016, **73**, 45–51.
- 19 C. Benoit, C. Virginie and V. Boris, in *Advances in Botanical Research*, eds. R. Verpoorte, G.-J. Witkamp and Y. H. Choi, Academic Press, 2021, vol. 97, pp. 309–332.
- 20 B. I. J. N, S. Venkatachalam and D. Datta, *International Journal of Research in Pharmaceutical Sciences*, 2020, **11**, 3017–3023.
- 21 N. Delgado-Mellado, M. Larriba, P. Navarro, V. Rigual, M. Ayuso, J. García and F. Rodríguez, *Journal of Molecular Liquids*, 2018, **260**, 37–43.
- 22 A. K. Kumar, B. S. Parikh, L. Z. Liu and M. A. Cotta, *Materials Today: Proceedings*, 2018, **5**, 23057–23063.
- 23 Y. H. Choi and R. Verpoorte, *Current Opinion in Food Science*, 2019, **26**, 87–93.
- 24 L. Wils, S. Hilali and L. Boudesocque-Delaye, *Molecules*, 2021, **26**, 6556.
- 25 L. Wils, C. Leman-Loubière, N. Bellin, B. Clément-Larosière, M. Pinault, S. Chevalier, C. Enguehard-Gueiffier, C. Bodet and L. Boudesocque-Delaye, *Algal Research*, 2021, **56**, 102317.
- 26 S. Hilali, L. Wils, A. Chevalley, B. Clément-Larosière and L. Boudesocque-Delaye, *Biomass Conv. Bioref.*, , DOI:10.1007/s13399-021-02263-6.
- 27 Y. H. Choi, J. van Spronsen, Y. Dai, M. Verberne, F. Hollmann, I. W. C. E. Arends, G.-J. Witkamp and R. Verpoorte, *Plant Physiology*, 2011, **156**, 1701–1705.
- 28 Y. Dai, J. van Spronsen, G.-J. Witkamp, R. Verpoorte and Y. H. Choi, *Anal. Chim. Acta*, 2013, **766**, 61–68.
- 29 M. Vilková, J. Płotka-Wasyłka and V. Andrich, *Journal of Molecular Liquids*, 2020, **304**, 112747.

- 30 Y. Dai, J. van Spronsen, G.-J. Witkamp, R. Verpoorte and Y. H. Choi, *Anal. Chim. Acta*, 2013, **766**, 61–68.
- 31 G. C. Dugoni, A. Mezzetta, L. Guazzelli, C. Chiappe, M. Ferro and A. Mele, *Green Chem.*, 2020, **22**, 8680–8691.
- 32 R. E. Wrolstad, T. E. Acree, E. A. Decker, M. H. Penner, D. S. Reid, S. J. Schwartz, C. F. Shoemaker, D. M. Smith and P. Sporns, *Handbook of Food Analytical Chemistry, Volume 1: Water, Proteins, Enzymes, Lipids, and Carbohydrates*, John Wiley & Sons, 2005.
- 33 V. Sanchez, R. Baeza, C. Ciappini, M. C. Zamora and J. Chirife, *Food Control*, 2010, **21**, 339–341.
- 34 H.-D. Isengard, D. Schultheiß, B. Radović and E. Anklam, *Food Control*, 2001, **12**, 459–466.
- 35 N. Dantan, W. Frenzel and S. Küppers, *Talanta*, 2000, **52**, 101–109.
- 36 S. Elderderi, C. Leman-Loubière, L. Wils, S. Henry, D. Bertrand, H. J. Byrne, I. Chourpa, C. Enguehard-Gueiffier, E. Munnier, A. A. Elbashir, L. Boudesocque-Delaye and F. Bonnier, *Journal of Molecular Liquids*, 2020, **311**, 113361.
- 37 A. Ricci, K. J. Olejar, G. P. Parpinello, P. A. Kilmartin and A. Versari, *Applied Spectroscopy Reviews*, 2015, **50**, 407–442.
- 38 M. Jackson and H. H. Mantsch, *Critical Reviews in Biochemistry and Molecular Biology*, 1995, **30**, 95–120.
- 39 I. Amenabar, S. Poly, W. Nuansing, E. H. Hubrich, A. A. Govyadinov, F. Huth, R. Krutokhvostov, L. Zhang, M. Knez, J. Heberle, A. M. Bittner and R. Hillenbrand, *Nat Commun*, 2013, **4**, 2890.
- 40 S. A. Tatulian, in *Lipid-Protein Interactions: Methods and Protocols*, ed. J. H. Kleinschmidt, Humana Press, Totowa, NJ, 2013, pp. 177–218.
- 41 B. Sarmiento, D. Ferreira, F. Veiga and A. Ribeiro, *Carbohydrate Polymers*, 2006, **66**, 1–7.
- 42 B. Van Eerdenbrugh and L. S. Taylor, *International Journal of Pharmaceutics*, 2011, **417**, 3–16.
- 43 D. Lischke, in *Solid State Development and Processing of Pharmaceutical Molecules*, John Wiley & Sons, Ltd, 2021, pp. 151–164.
- 44 J. A. L. Pallone, E. T. dos S. Caramês and P. D. Alamar, *Current Opinion in Food Science*, 2018, **22**, 115–121.
- 45 H. J. Byrne, F. Bonnier, J. McIntyre and D. R. Parachalil, *Clinical Spectroscopy*, 2020, **2**, 100004.
- 46 S. Mazurek and R. Szostak, *J AOAC Int*, 2012, **95**, 744–750.
- 47 C. J. Strachan, T. Rades, K. C. Gordon and J. Rantanen, *Journal of Pharmacy and Pharmacology*, 2007, **59**, 179–192.
- 48 M. A. Mallah, S. T. H. Sherazi, M. I. Bhanger, S. A. Mahesar and M. A. Bajeer, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2015, **141**, 64–70.
- 49 A. A. Makki, F. Bonnier, R. Respaud, F. Chtara, A. Tfayli, C. Tauber, D. Bertrand, H. J. Byrne, E. Mohammed and I. Chourpa, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, **218**, 97–108.
- 50 A. A. Makki, V. Massot, H. J. Byrne, R. Respaud, D. Bertrand, E. Mohammed, I. Chourpa and F. Bonnier, *Vibrational Spectroscopy*, 2021, **113**, 103200.

- 51 L. Miloudi, F. Bonnier, D. Bertrand, H. J. Byrne, X. Perse, I. Chourpa and E. Munnier, *Anal Bioanal Chem*, 2017, **409**, 4593–4605.
- 52 F. Bonnier, L. Miloudi, S. Henry, D. Bertrand, C. Tauber, X. Perse, F. Yvergnaux, H. J. Byrne, I. Chourpa and E. Munnier, *Anal Bioanal Chem*, 2020, **412**, 159–169.
- 53 Y. Kataoka, N. Kitadai, O. Hisatomi and S. Nakashima, *Appl Spectrosc*, 2011, **65**, 436–441.
- 54 A. Habuka, T. Yamada and S. Nakashima, *Appl Spectrosc*, 2020, **74**, 767–779.
- 55 M. Ibrahim, M. Allam, H. Elhaes and A. Leon, *Ecl. Quim., Sao Paulo*, , DOI:10.1590/S0100-46702006000300002.
- 56 S. Elderderi, L. Wils, C. Leman-Loubière, S. Henry, H. J. Byrne, I. Chourpa, E. Munnier, A. A. Elbashir, L. Boudesocque-Delaye and F. Bonnier, *Anal Bioanal Chem*, , DOI:10.1007/s00216-021-03432-2.