



Diagnosing sickle cell disease and iron deficiency anemia in human blood by Raman spectroscopy

Wagner Rafael da Silva¹ · Landulfo Silveira Jr² · Adriana Barrinha Fernandes²

Received: 23 May 2019 / Accepted: 23 September 2019 / Published online: 22 October 2019
© Springer-Verlag London Ltd., part of Springer Nature 2019

Abstract

This work proposed the diagnosis of iron deficiency anemia (IDA) and sickle cell disease (SCD) in human blood caused by iron deficiency and hemoglobin S (HbS), which are among the most common anemias, by means of Raman spectroscopy. Whole blood samples from patients diagnosed with IDA and HbS, as well as from normal subjects (HbA), were obtained and submitted to Raman spectroscopy (830 nm, 150 mW, 400–1800 cm⁻¹ spectral range, 4 cm⁻¹ resolution). Difference spectra of IDA–HbA showed spectral features of hemoglobin with less intensity in the IDA, whereas the difference spectra of SCD–HbA showed spectral features of deoxyhemoglobin increased and of oxyhemoglobin decreased in SCD. An exploratory analysis by principal components analysis (PCA) showed that the peaks referred to oxy- and deoxyhemoglobin markedly differentiated SCD and HbA, as well as the increased amount of hemoglobin features in the SCD group, suggesting increased erythropoiesis. The IDA group showed hemoglobin features with lower intensities as well as peaks referred to the iron bonding to the porphyrin ring with reduced intensities when compared to the HbA. Discriminant analysis based on partial least squares (PLS-DA) and PCA (PCA-DA) showed that the IDA and SCD anemias could be discriminated from the HbA spectra with 95.0% and 93.8% of accuracy, for the PLS and PCA respectively, with sensitivity/specificity of 93.8%/95.7% for the PLS-DA model. The iron depletion and the sickling of erythrocytes could be identified by Raman spectroscopy and a spectral model based on PLS accurately discriminated these IDA and SCD samples from the normal HbA.

Keywords Raman spectroscopy · Diagnosis · Sickle cell disease · Iron deficiency anemia

Introduction

Anemia is a syndrome characterized by a decrease in the quantity of red blood cells, normally associated to reduction of hemoglobin levels or changes in the red blood cell morphology, which leads to a clinical scenario of low-oxygen supply for tissues (hypoxia) as a result of the reduced oxygen transport capacity of the bloodstream [1]. An anemic individual has a variety of physiological changes as a consequence of the reduction in the red blood cells; at the center of these

changes are compensatory responses of the organism itself to anemia that, within limits, help preserve the oxygen supply to the tissues. One of the compensatory mechanisms is associated with increased kidney erythropoietin synthesis, which stimulates the bone marrow to produce the red blood cell [2].

There are numerous types of anemia; among the majors stand out the iron deficiency anemia (IDA) and the sickle cell disease (SCD); the first is the most prevalent nutritional disease in the world and is the most severe form of lack of iron [3]. The second is the most common hereditary monogenic disease in the world with the most damaging consequences for the body and is caused by sickling of red blood cells due to changes in an amino acid in the hemoglobin chain [4]. One of the most serious complications of SCD is high blood pressure in the blood vessels supplying the lungs, thus inducing pulmonary hypertension that occurs in about one-third of adults with SCD and can lead to heart failure [5]. The World Health Organization estimates for 2008 indicated that IDA would affect about 1.62 billion people worldwide [3], whereas the number of people affected by sickle cell disorders in 2003 is

✉ Landulfo Silveira, Jr
landulfo.silveira@gmail.com; lsjunior@anhemb.br

¹ Universidade Brasil, Estrada Projetada F1, S/N,
Fernandópolis, SP 15600-000, Brazil

² Center for Innovation Technology and Education–CITE,
Universidade Anhembi Morumbi–UAM, Estr. Dr. Altino
Bondensan, 500, São José dos Campos, SP 12247-016, Brazil

on the order of 2.3 per 1000 conceptions [6]. In Brazil, the prevalence of IDA in 2006 was 20.9% for children below 5 years of age and 24.1% for children below 2 years of age; in women of childbearing age, the prevalence was 29.4% [7]. Regarding the SCD, it affects mainly African descendants, but it is not exclusive to this population due to the evident racial miscegenation that occurs in Brazil [8]. The incidence of SCD is approximately 1–3/1000 live births; in states such as Bahia, where African ancestry predominates, this rate reaches 1/650 newborns [9].

The IDA caused by iron deficiency presents a higher prevalence in women and children. IDA can impair mental and psychomotor development, causes increased maternal and infant morbidity and mortality, as well as a decrease in the individual's performance at work and reduction of infection's resistance [10–12]. The major risk factors for IDA include low intake of iron, poor absorption of iron from diets with high phytate or phenolic compounds, and period of life when iron requirements are especially high, for example, during the period of growth and pregnancy [3]. The SCD is a genetic disorder where the amino acid glutamic acid is substituted by valine in the hemoglobin β chain, causing a "sickling" (falcization) of the erythrocytes as a consequence of the exposure of these cells to a low $3O_2$ tension. The normal HbA is present in erythrocytes after the initial 6 months of life and throughout the adult phase, being composed of two pairs of polypeptide chains ($\alpha_2\beta_2$). Hemoglobin S (HbS) is the result of abnormal mRNA template and the abnormal β^S globin chain complexes with α chains to form the sickle hemoglobin tetramer ($\alpha_2\beta^S_2$) [13]. Other hemoglobin variants also include hemoglobin C and E as well as thalassemia that also leads to hemoglobinopathies [14].

The diagnosis of anemia is based on the patient's history and laboratory evaluation, by analyzing the amount of hemoglobin and the number of erythrocytes in whole blood [15]. As IDA is due to the lack of sufficient iron to form normal red blood cells, it is characterized by microcytosis, hypochromia, and low level of circulating iron [16]. Serum ferritin is currently the accepted laboratory test for diagnosis of IDA, and a ferritin value < 12 mg/L is a highly specific indicator of iron deficiency [17]. Other laboratory tests used are serum iron, total iron binding capacity, mean corpuscular volume, and transferrin saturation [18–21]. The most employed tests for the diagnosis of SCD include hemoglobin electrophoresis, dithionite solubility tube test for sickling hemoglobin, and metabisulfite slide test. The peripheral smear test normally presents fragmented red cells, polychromasia, and sickle cells. Other methods include HPLC, DNA analysis, and isoelectric focusing [22].

Due to the impact of the anemias to the individual's health and also to the demanding public health services, since untreated anemias can lead to complications that result in irreversible brain damage, such as dementia, loss of brain

function, and cardiovascular problems (cerebrovascular accident mainly for the HbS carrier) [23], and due to the relative complex steps to reach the diagnosis of the HbS, there is a need for a method to directly identify such diseases in the whole blood without laborious preprocessing of the samples. Optical techniques such as Raman spectroscopy can help in the development of an in situ, real-time, and low-cost diagnostic tool for direct analysis of biological tissues including blood analysis [24].

Raman spectroscopy technique is based on the inelastic scattering of the incident monochromatic light by a polarizable molecule and has been used in the evaluation of blood and serum, aimed at the quantification of analytes in serum [25, 26] and diagnosis of SCD in hemolyzed blood samples [27]. Atkins et al. [28] have highlighted the potential of using Raman spectral information from the vibrational energy levels of the molecules for the differentiation between healthy and altered blood and blood components aiming diagnosis and forensic applications. In fact, Raman spectroscopy can provide the vibrational "fingerprint" of the most important biochemicals related to the diseases [24, 29–31] and also reflect the changes in the spectrum due to the oxygenation status of the erythrocytes [32, 33].

Due to the need of different techniques to identify these two types of anemia (electrophoresis and spectrophotometer), this study proposed the use of whole blood samples to identify the biochemical differences between whole blood samples from patients with IDA and SCD compared to the blood of HbA subjects by means of Raman spectroscopy and principal components analysis (PCA), to explain the spectral changes related to the oxygenation status of the erythrocytes and to promote the discrimination of the IDA and SCD from the normal HbA using discriminant analysis models via partial least squares (PLS) and PCA (PLS-DA and PCA-DA).

Material and methods

The study has been approved by the Research Ethics Committee of Universidade Brasil (Fernandópolis, SP, Brazil) under Protocol No. 1.578.298 (CAAE No. 53543516.4.0000.5494). A total of 87 samples of whole blood were obtained, being 28 from HbA, 26 from IDA and 33 from SCD subjects. These samples were obtained from the Instituto de Hematologia de Fernandópolis (Fernandópolis, SP, Brazil) by an expert hematologist after the standard diagnosis by electrophoresis (for the SCD) and serum ferritin (for the IDA). Briefly, approximately 4.5 mL of total blood was collected from each patient with anticoagulant (10% EDTA) to preserve the erythrocytes from the action of coagulation factors and to maintain the erythrocyte physicochemical characteristics. The samples were snap-frozen and kept at -20 °C prior to spectral analyses for approximately 2 weeks to maintain the molecular

viability of hemoglobin, being transported to the Raman Laboratory using adequate container to keep the temperature between -20 and -15 °C. At the time of spectral analysis, these samples were kept to passively reach room temperature (23 °C) and then were subjected to Raman spectroscopy.

A dispersive Raman spectrometer was used (model Dimension P1, Lambda Solutions Inc., MA, USA) with 830-nm laser excitation and 350-mW maximum power, connected to a Raman probe (model Vector Probe, Lambda Solutions Inc.) for sample excitation and scattering collection. The spectrometer has a resolution of about 4 cm^{-1} in the Raman shift range of 400 – 1800 cm^{-1} (fingerprint region). A volume of 80 μL of blood was pipetted in an aluminum sample holder with round-shape wells and the sample was scanned with 3-s exposure time and 10 acquisitions. The laser power was reduced to 150 mW at the probe's distal (excitation) end to avoid burning of blood due to the absorption of the excitation light, which could coagulate the blood. Triplicate spectra were obtained from each blood sample, being these spectra treated as different samples in the further discriminant analysis models. A total of 261 spectra were collected. Some spectra ($n = 3$) were removed due to the signal-to-noise ratio lower than 10, estimated in the spectral region of 1400 – 1600 cm^{-1} , therefore, the number of useful spectra was 258.

After collection, the spectra were pre-processed to remove spikes from cosmic rays, to remove the Raman background (mostly due to fluorescence from blood molecules) by fitting and subtracting a 7th order polynomial over the whole spectrum in the range of 400 – 1800 cm^{-1} Raman shift, and to normalize by the area under the curve (1-norm or taxicab norm) [34]. The mean spectra of each group were calculated and the difference spectra between HbA and IDA and between HbA and SCD were calculated. These spectra were then plotted for the purpose of finding the gross differences between the anemias and the HbA.

An exploratory analysis was performed using PCA (princomp.m function, MatLab 7.4, The Mathworks Inc., Natick, MA, USA), where the principal component loading vectors and scores were used to identify which spectral features were different between the three groups and to assign these features to the blood constituents referred to the HbA, IDA, and SCD. The loading resembles Raman spectra and the first loadings give the Raman features with the highest variance in the dataset (in fact they account for much of the variability of the dataset), while the scores provide the intensities of these loadings in each spectrum [35]. Therefore, the Raman features found in a specific loading can be assigned to a particular or a group of biochemicals presented in the samples, or spectral changes due to the changes in the polarizability of the “new chemical environment” caused by a disease, and high scores in a particular group indicate that these features, meaning biochemicals or changes in the “environment,” are high in

this group. Then the differences in the biochemical constitution or biochemical environment between HbA and anemias, as well as the differences between the anemias, can be precisely identified and can also be quantified.

Discriminant analysis (DA) models based on partial least squares (PLS-DA) and PCA (PCA-DA) were also performed in order to discriminate the IDA and SCD samples from the HbA samples. The discriminant models group the samples according to the biochemical/environmental differences found in the PCA described by the loadings and scores (for the PCA-DA), and differences in the latent variables (for the PLS-DA). While in the PCA-DA model, the grouping is based on the optimization of the inter-group differences in the scores; in the PLS-DA model, the grouping is optimized using the inter-group differences while estimating the intra-group differences [36]. The DA models were developed with the aid of the Chemoface software [37] using the “leave-one-out” cross-validation (LOOcv) approach, where a sample is withdrawn from the dataset, the model is built, and the left-out sample is tested, being this procedure repeated for all samples. DA models using LOOcv have advantages when the number of samples per group is not big enough to allow the split of the data into training and validation dataset. The LOOcv has been used to discriminate Raman spectra of bone marrow smear and peripheral blood from subjects with acute myeloid leukemia and myelodysplastic syndrome [38] and to discriminate skin cancer and keratosis from normal skin sites with Raman spectra taken ex vivo and in vitro [39].

Results and discussion

Raman spectra of blood with IDA, SCD, and normal blood (HbA)

Figure 1 shows the normalized Raman spectra of blood samples: iron deficiency anemia (IDA), sickle cell disease (SCD) and normal HbA. The spectra present Raman bands referred to the human blood constituents—cells (mainly erythrocytes and also leucocytes and platelets) and plasma (mainly albumin and other minor constituents)—which present Raman vibrational band assignments as described elsewhere [28, 32, 33, 40–44]. Despite the very similar spectral features, there were observable differences in the 753 and 1004 cm^{-1} and in the regions of 1130 – 1230 cm^{-1} , 1350 – 1410 cm^{-1} , and 1520 – 1650 cm^{-1} . Difference spectra of IDA and HbA (IDA–HbA) and of SCD and HbA (SCD–HbA) are also presented, showing that the mentioned peaks were increased or reduced depending on the type of anemia.

The difference spectra in Fig. 1 showed that both IDA and SCD presented lower intensity for the peaks at 1227 , 1375 , and 1406 cm^{-1} , which can be assigned to oxyhemoglobin [33, 45, 46]; the IDA presented lower intensity for the peaks at

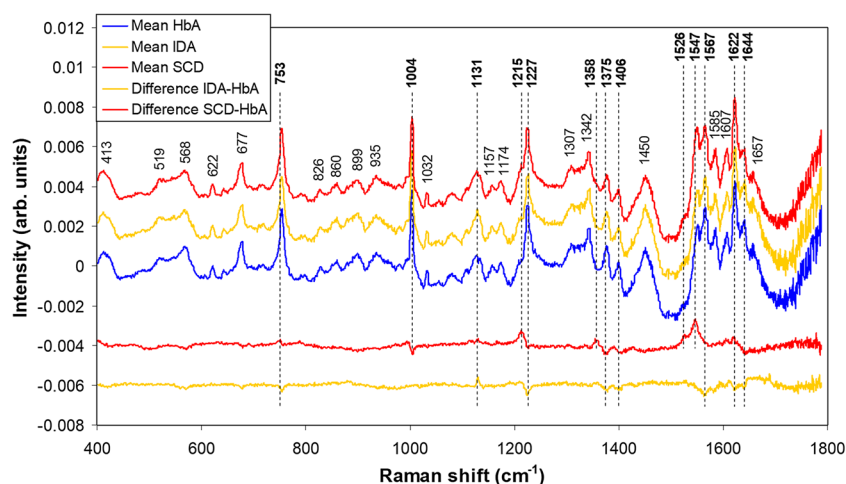


Fig. 1 Mean normalized Raman spectra of whole blood samples from normal (HbA), iron deficiency anemia (IDA) and sickle cell disease (SCD) obtained from 87 subjects. All spectra present the Raman features of blood as reported by the literature (red and white blood cells as well as compounds of the plasma) [27, 31, 32, 39–43] and the assignments of the most relevant bands (bold labels) are presented in the text ([Differences in the spectra of IDA and SCD compared to HbA](#)

[observed by principal component analysis](#) section). The difference spectra IDA–HbA and SCD–HbA are also presented, where the bold-labeled peaks represent the main biochemical differences in the SCD and IDA versus HbA, also discussed in the text ([Differences in the spectra of IDA and SCD compared to HbA](#) observed by principal component analysis section)

753, 1567, 1622, and 1644 cm^{-1} , assigned to the hemoglobin molecule [46–50] and higher intensity for the peak at 1131 cm^{-1} , assigned to the pyrrolic ring of the porphyrin; the SCD presented lower intensity for the peak at 1004 cm^{-1} , assigned to aromatic ring of amino acids (phenylalanine) [51], and higher intensity for the peaks at 1215, 1358, 1526, 1547, and 1622 cm^{-1} , assigned to deoxyhemoglobin [45, 46, 50]. All of these labeled peaks showed statistically significant differences when comparing the three groups (HbA, IDA, and SCD) (one-way ANOVA test, $p < 0.05$). Detailed analysis of these spectral differences is presented in the exploratory analysis of the “[Differences in the spectra of IDA and SCD compared to HbA observed by principal component analysis](#)” section.

Differences in the spectra of IDA and SCD compared to HbA observed by principal component analysis

An exploratory analysis by principal component analysis (PCA) revealed the differences in the spectral features of IDA and SCD compared to HbA furnished by the principal component loadings. Figure 2 presents the first four principal component loadings and the respective scores, which accounted for more than 98% of the spectral variation in the dataset (principal components 1, 2, 3, and 4: 84.3%, 2.5%, 1.4%, and 0.8%, respectively). The number of spectral features revealed by the loadings suggests specific biochemical changes caused by both types of anemia compared to HbA. The loading vector 1 (Loading 1) presents the spectral features referred to whole blood [28, 41, 44, 46], with positive peaks at 753, 1004, 1215, 1227, 1358, 1375, 1406, 1526, 1547, 1567,

1622, and 1644 cm^{-1} assigned to blood cells (mainly erythrocytes) [28, 41] and plasma constituents (mainly albumin) [44]. The Score 1 showed statistically significant differences between the groups IDA versus HbA, SCD versus HbA, and IDA versus SCD (one-way ANOVA, Tukey posttest, $p < 0.05$), with lower intensity for the IDA and higher intensity for the SCD. The lower Score 1 for the IDA can be explained by the fact that Loading 1 presents peaks referred to erythrocytes, as the iron deficiency disease is characterized by the reduction of the produced erythrocytes by lack of iron. With regard the SCD group, the increased Score 1 may indicate that the erythrocytes are in higher amounts, suggesting that a compensative mechanism of erythropoiesis is acting at the marrow level, probably due to the low tissue oxygenation, microvascular occlusion, or lack of function of the erythrocytes in the HbS hemoglobin. Jelkmann [52] showed elevated erythropoietin in SCD due to microvascular occlusion associated with hypoxia.

The loading vector 2 (Loading 2) presented positive spectral features at 1004, 1375, and 1644 cm^{-1} , which can be assigned to hemoglobin and oxyhemoglobin [33, 46, 50, 51], and negative spectral features at 753, 1215, 1358, 1526, and 1547 cm^{-1} , which can be assigned to deoxyhemoglobin [33, 41, 45, 46, 50]. The Score 2 showed statistically significant differences for the SCD versus HbA and SCD versus IDA (one-way ANOVA, Tukey posttest, $p < 0.05$), with positive values for the HbA and IDA and negative values for the SCD. Negative loading and negative score indicate that the spectral features assigned to deoxyhemoglobin are significantly higher in the SCD group. The peaks of deoxyhemoglobin (753, 1215, 1358, 1526, and 1547 cm^{-1}) present a higher

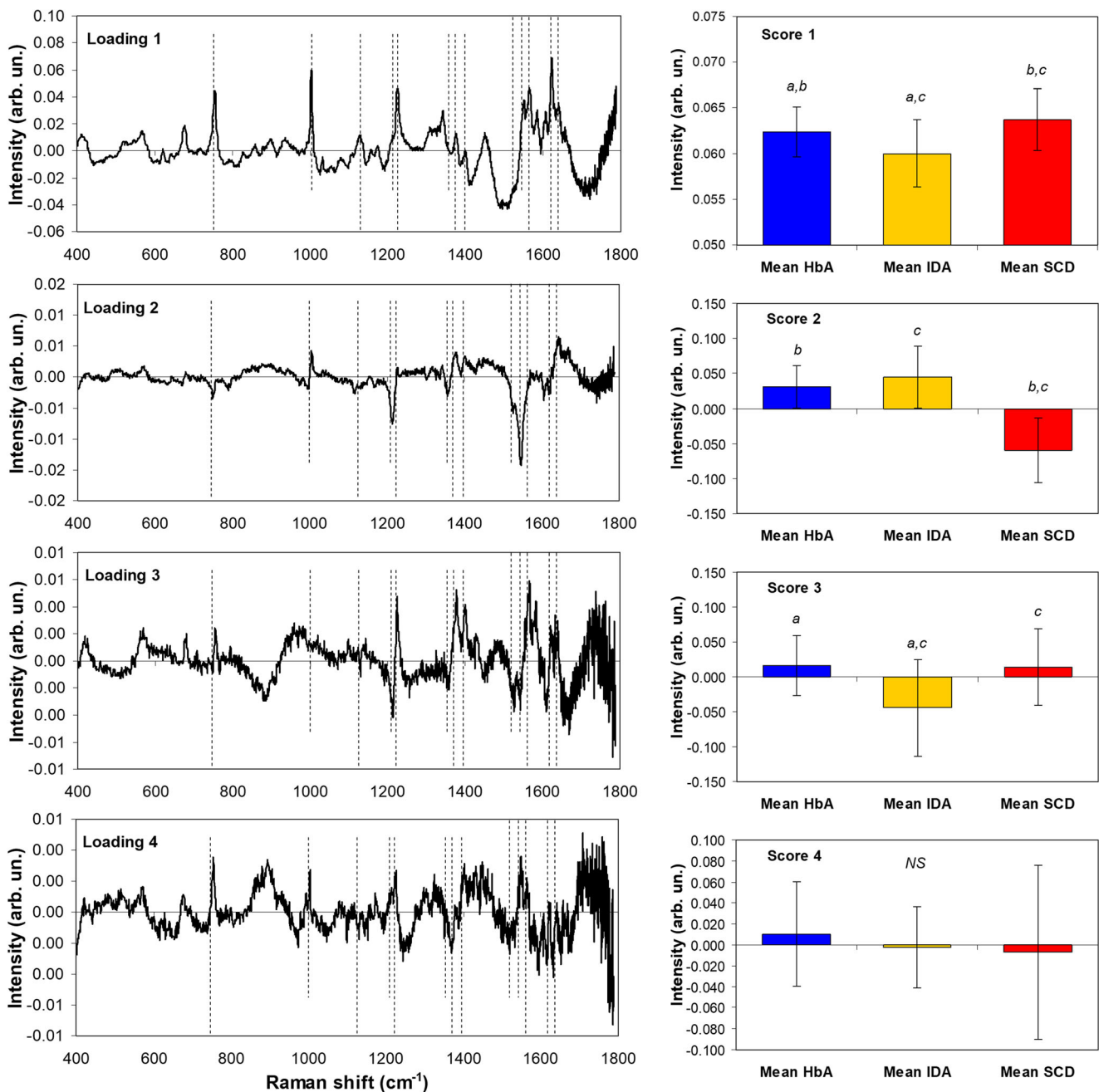


Fig. 2 Plot of the first 4 principal components loading vectors (left) and scores (right) used to identify the spectral differences between the HbA, IDA, and SCD blood samples and to discriminate these samples using the PLS-DA and PCA-DA models. The vertical dashed lines in the Loadings

are in the same positions as the ones presented in Fig. 1. Superscript letters *a*, *b*, and *c* in the Scores indicate statistically significance (one-way ANOVA, Tukey post test, $p < 0.05$) between groups HbA vs IDA, HbA vs SCD, and IDA vs SCD, respectively

intensity for the SCD compared to the hemoglobin and oxy-hemoglobin peaks that presented higher intensity in the HbA and IDA groups. This is interesting because the formation of deoxy-sickle hemoglobin polymers is directly related to the degree of oxygen desaturation, which is a known component of SCD [52]. The decrease in oxygen desaturation results from the marked shift to the right of the oxyhemoglobin dissociation curve that frequently appears in the sickling of the red blood cells in the SCD. A right-shifted curve follows in a

reduction of affinity between oxygen and hemoglobin that is considered a defense mechanism to improve oxygen delivery to the tissues in the presence of anemia [53]. Score 2 shows that the intensity of these peaks is significantly altered in the SCD. Changes in the shape of erythrocytes (discoid red blood cells can be transformed into echinocyte or stomatocyte shape) lead to change in the oxygen affinity, and it was demonstrated that changes in hemoglobin oxygenation lead to changes in the Raman peaks referred to hemoglobin [54].

Atkins et al. [46] showed that these deoxyhemoglobin peaks increase with the time of storage of blood transfusion bags. According to this Atkins's study, the ration between the 1208 and 1222 cm^{-1} peaks correlated with the “morphology scoring” of the erythrocytes and can be used to monitor blood oxygenation status after long-term storage. Ward et al. [32] identified differences in oxy- and deoxyhemoglobin regarded to the peaks at 1376 and 1385 cm^{-1} , respectively, under “normal” conditions, being the peak at 1376 cm^{-1} much more intense than the 1385 cm^{-1} . In contrast, in hemorrhagic conditions, they noticed an increase in the peak at 1385 cm^{-1} compared to the peak at 1376 cm^{-1} due to a reduction in the red blood cells, close to the IDA condition.

Loading vector 3 (Loading 3) presents positive spectral features at 753, 1227, 1375, 1406, 1567, 1585, 1622, and 1644 cm^{-1} , all of them assigned to hemoglobin, and negative features at 1215, 1358, 1526, 1547, and 1607 cm^{-1} assigned to deoxyhemoglobin [46–50]. Score 3 presented statistically significant differences for the IDA versus both HbA and SCD (one-way ANOVA, Tukey posttest, $p < 0.05$), with negative intensity for the IDA compared to both HbA and SCD. The negative Score 3 suggests decrease in the amount of iron in the IDA group since most of these peaks, markedly the peaks at 1375 and 1406 cm^{-1} , assigned to the pyrrolic ring of the porphyrin, and the peak at 1227 cm^{-1} , assigned to the C-H methine deformation of the porphyrin [55], are diminished in the IDA samples due to the absence of iron (by the iron deficiency) to bond, by methine bridges, to the pyrrolic molecules at the porphyrin ring. The peaks at 1585 and 1644 cm^{-1} , assigned to $C_{\alpha}C_m$ vibration of the methine ring [42, 49, 56], are also diminished in these IDA samples, which may be due to the already mentioned decrease in the methine bonds. Loading vector 4 presented positive features at 753, 1004, 1227, 1406, 1547, and 1567 cm^{-1} , assigned to hemoglobin [28, 41], and sparse negative features, whereas Score 4 presented not significant differences between the groups (one-way ANOVA, Tukey posttest, $p = 0.08$), despite the (not

significant) higher intensity of the oxyhemoglobin features in the HbA group.

As observed in Fig. 2 (right), the features of Loadings 1 and 3 can distinguish IDA from HbA and SCD, whereas the features of Loading 2 can distinguish SCD from IDA and HbA. This is clearly seen in the binary plot of Score 1 versus Score 2, Score 1 versus Score 3, and Score 2 versus Score 3 (Fig. 3). These features, related to general hemoglobin features of blood (Loading 1), being reduced in the IDA and increased in the SCD (Score 1), related to deoxyhemoglobin (Loading 2), higher in SCD (Score 2), and related to pyrrolic ring and methine bridge of the porphyrin ring (Loading 3), lower in the IDA, configures the general biochemical changes related to IDA and SCD that can be used to biochemically differentiate these two types of blood anemia by means of Raman spectroscopy.

Discrimination between HbA, IDA, and SCD blood samples

Discriminant analysis (DA) models based on PLS and PCA (PLS-DA and PCA-DA, respectively) that uses the spectral dataset were developed in order to obtain discrimination between the three groups: HbA, IDA, and SCD. The PLS-DA model takes advantage of the PCA-DA model because the PLS uses the annotated label (in fact the group class number) to maximize intra-group variance, i.e., maximizing the separation of classes while reducing the data dimension. The PLS treats the class as the dependent variable and includes the relationship of the independent (spectra) and the dependent variables in its minimization. On the other hand, PCA treats all variables equally while maximizing the between-group variance [36]. In fact, with PLS one can achieve better discrimination than PCA when intra-group variations are relevant. Discrimination models have been implemented using the Chemoface routine described by Nunes et al. [37] (download at <http://www.ufla.br/chemoface/>), where the inputs are the spectra and the real group (or class) each sample

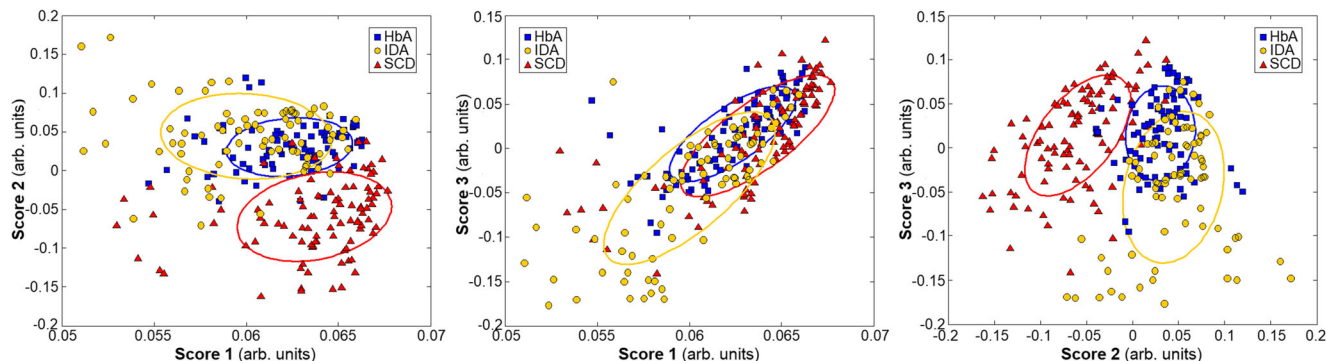


Fig. 3 Binary plot of Score 1 vs Score 2, Score 1 vs Score 3 and Score 2 vs Score 3 and the covariance ellipse error that represents the 90% confidence interval of each group data dispersed over each binary plot.

This plot shows the capability of these Scores to discriminate the SCD group from the IDA and HbA groups

belongs, and the outputs are the number of correct classification per latent variable (for the PLS-DA model) or loading vector (for the PCA-DA model) and the model's predicted group (or class), estimated in a "leave-one-out" cross-validation fashion.

Table 1 presents the contingency table with the discrimination of the spectra into the groups HbA, IDA, and SCD using the "leave-one-out" cross-validation. It has been shown that both PLS-DA and PCA-DA promoted the overall accuracy of 78.3% using the first 6 latent variables and the first 6 loading vectors, respectively. Sensitivity, specificity, and overall accuracy values for the discrimination of anemias (SCD and IDA) from HbA and IDA from SCD and HbA were also presented in Table 1. Due to the high confusion of the HbA samples into the IDA group and vice versa (Table 1), mostly due to the small spectral differences in these two groups (IDA and HbA presented lower spectral difference than the SCD and HbA (Fig. 1)), one can consider grouping the HbA and IDA and use these two groups as "non-HbS" hemoglobin; therefore, the PLS-DA promoted the highest sensitivity, specificity, and overall accuracy of 92.8, 93.8, and 93.4%, respectively, compared to the PCA-DA model (both models with 6 latent variables and loadings) to discriminate subjects with SCD.

The DA model was also developed to discriminate "non-HbS" hemoglobin (HbA + IDA) from SCD, by grouping the spectra of HbA and IDA, the contingency table presented in Table 2 showed that PLS-DA promoted sensitivity, specificity, and overall accuracy of 93.8%, 95.7%, and 95.0%, respectively, using the first 3 latent variables, compared to the sensitivity, specificity, and overall accuracy of 90.7%, 95.7%, and 93.8%, respectively, for the PCA-DA using the first 6 loading vectors. Therefore, the PLS model is suitable to better discriminate the groups by the biochemical changes associated with both anemias

and presenting accuracy as high as 95.0% when discriminating biochemical changes associated with SCD compared to "non-HbS" hemoglobin. The employment of the first and second derivative of the Raman spectra in the discriminant models changed only the number of latent variables or loadings used in the models, without changing the discrimination capability.

Final remarks

The differences in the Raman features of the different types of anemia revealed by PCA were related mostly to differences in the oxygenation status of the hemoglobin (lower oxy- and higher deoxyhemoglobin) for the SCD and reduction of general hemoglobin features and lack of iron in the porphyrin ring of hemoglobin for the IDA. In fact, most of the differences of HbS compared to HbA can be related to the difference in the oxygenation status of the Hb molecule [56]; mainly the increased peaks at 1215, 1358, and 1547 cm^{-1} and the decreased peaks at 1227 and 1375 cm^{-1} with the decrease in the hemoglobin oxygenation as mentioned by Ward et al. [32], Torres Filho et al. [33], and Atkins et al. [28]. The differences in the spectra of IDA compared to HbA, with lower intensity of the peaks at 753, 1227, 1375, 1406, and 1567 cm^{-1} , can be attributed to the lack of the iron ion bonded to the porphyrin ring in the iron deficiency anemia samples, since these peaks are attributed to the methine bridge and pyrrolic ring in the hemoglobin [42, 49, 56].

The results described here are in accordance with the recent literature of Raman spectroscopy applied to the diagnosis of anemias [27, 43, 57]. De Luca et al. [27] performed the spectroscopic and mechanical characterization of normal and thalassaemic red blood cells by Raman tweezers. Bueno Filho et al.

Table 1 Contingency table showing the discrimination of the HbA, IDA, and SCD based on PLS and PCA discriminant analysis (PLS-DA and PCA-DA) considering both anemias as independent groups

Blood type	Discrimination PLS-DA			Discrimination PCA-DA		
	HbA	IDA	SCD	HbA	IDA	SCD
HbA (84 spectra)	62	16	6	57	14	13
IDA (77 spectra)	23	50	4	21	54	2
SCD (97 spectra)	6	1	90	6	0	91
Number of variables	6 latent variables			6 loading vectors		
Overall accuracy	78.3%			78.3%		
Sensitivity/specificity*	83.3%/73.8%			84.5%/67.9%		
Overall accuracy*	80.2%			79.1%		
Sensitivity/specificity**	64.9%/90.6%			70.1%/92.3%		
Sensitivity/specificity***	92.8%/93.8%			93.8%/90.7%		
Overall accuracy***	93.4%			91.9%		

*Sensitivity, specificity, and overall accuracy for diagnosis of anemia by grouping the classification data of the HbS and IDA groups as anemic group versus normal HbA group

**Sensitivity and specificity for diagnosis of IDA by grouping the classification data of the HbA and HbS groups as a non-IDA hemoglobin group

***Sensitivity, specificity, and overall accuracy for diagnosis of SCD by grouping the classification data of the HbA and IDA groups as a "non-HbS" hemoglobin group

Table 2 Contingency table showing the discrimination of the spectra in the HbA and IDA grouped together as “non-HbS” hemoglobin versus SCD group based on PLS and PCA discriminant analysis

Blood type	Discrimination PLS-DA		Discrimination PCA-DA	
	“Non-HbS” hemoglobin	SCD	“Non-HbS” hemoglobin	SCD
“Non-HbS” hemoglobin (161 spectra)	154	7	154	7
SCD (97 spectra)	6	91	9	88
Number of variables	3 latent variables		6 loading vectors	
Sensitivity/specificity	93.8%/95.7%		90.7%/95.7%	
Overall accuracy	95.0%		93.8%	

[43] proposed a method to discriminate hemolyzed blood samples diagnosed as SCD from HbA based on PCA scores and Mahalanobis distance, reaching 100% accuracy. Yan et al. [58] proposed a method based on fiber-enhanced Raman spectroscopic analysis to evaluate the biomolecules hematin, hemoglobin, biliverdin, and bilirubin in whole intact, oxygenated erythrocytes, thus demonstrating the potential for the diagnosis of red blood cell-related diseases, such as different types of anemia and hemolytic disorders. Wood and McNaughton [59] proposed a method based on Raman microspectroscopy to monitor porphyrin perturbation related to the tense-relaxed state transition of hemoglobin within a single erythrocyte, illustrating the dynamics of porphyrin perturbation during erythrocyte respiration with applications in the diagnosis of red blood cell disorders including thalassemia and sickle cell anemia. Therefore, Raman spectroscopy shows promising results for a direct, rapid, and precise analysis of hemoglobin and red blood cells for disease characterization and diagnosis.

The results are encouraging in proposing the Raman as a technique for diagnosis of anemias (IDA and SCD) in whole blood, since high sensitivity, specificity, and accuracy with respect to the discrimination of both anemias from HbA (93.8%, 95.7%, and 95.0%, respectively), where a single drop of peripheral blood could be used to identify the spectral changes associated to the diseases and a discrimination model could promote diagnosis. The limitations of the study is related to the “not so high” discrimination capability of IDA compared to the HbA, since the remarkable difference between these two groups is related to the reduced number of erythrocytes and microcytic and hypochromic red blood cells in the IDA samples, which may be challenging to be measured by the Raman spectroscopy methodology as presented and can be exploited in future studies aiming differential diagnosis of IDA and SCD.

Advantages of Raman spectroscopy in detecting the spectral changes associated with SCD and IDA compared to HbA rests on the no need for preparation of blood samples prior analysis, fast spectrum collection, and processing, and the possibility of obtaining biochemical information for real-time diagnostics with a drop of peripheral blood sample or even directly in capillary vessels in vivo. The study could be extended to other types of anemia (for instance the

thalassemia) in order to identify the particular changes in the Raman features related to each disease.

Conclusion

This study showed that the Raman spectral differences between iron deficiency anemia (IDA) and sickle cell disease (SCD) compared to normal hemoglobin (HbA) in human blood samples were related to the lower oxyhemoglobin and higher deoxyhemoglobin content, as well as the possible high number of erythrocytes (increased erythropoiesis) in SCD samples, and low hemoglobin content, as well as to the lack of Raman peaks of iron bonded to the porphyrin ring in IDA samples. Both discrimination models (PLS-DA and PCA-DA) could discriminate all the three groups individually (HbA, IDA, and SCD) with overall accuracy of 78.3%, and the PLS-DA model could discriminate SCD (HbS hemoglobin) from “non-HbS” hemoglobin (HbA and IDA) better than the PCA-DA, with accuracy of 95.0%, showing that Raman spectroscopy is a promising technique for the discrimination of sickle cell disease anemia, with rapidness, no need for reagents, and ease to obtain the signal, thus opening the possibility for in vivo diagnosis.

Funding information L. Silveira Jr. acknowledges FAPESP (São Paulo Research Foundation) for granting the Raman spectrometer (Grant No. 2009/01788-5). L. Silveira Jr. and A. B. Fernandes acknowledge CNPq (National Council for Scientific and Technological Development) for financial support (Grant No. 460014/2014-5). L. Silveira Jr. acknowledges CNPq for the Productivity fellowship (Process No. 306344/2017-3).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval Authors declare compliance with ethical standards when using human blood samples, being the study approved by the Committee for Ethics in Research of Universidade Brasil, Protocol No. 1.578.298 (CAAE No. 53543516.4.0000.5494) following Brazilian regulations for use of human subjects or materials in research (Brazilian Ministry of Health, National Health Council, Resolution CNS No. 466/2012) and have been performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments.

References

- Kassebaum NJ, Jasrasaria R, Naghavi M, Wulf SK, Johns N, Lozano R, Regan M, Weatherall D, Chou DP, Eisele TP, Flaxman SR, Pullan RL, Brooker SJ, Murray CJ (2014) A systematic analysis of global anemia burden from 1990 to 2010. *Blood* 123:615–624. <https://doi.org/10.1182/blood-2013-06-508325>
- World Health Organization, Blood Transfusion Safety (2002) The clinical use of blood in medicine, obstetrics, paediatrics, surgery and anaesthesia, trauma and burns. WHO. <https://apps.who.int/iris/bitstream/handle/10665/42397/a72894.pdf?sequence=1&isAllowed=y>. Accessed 30 Nov 2018
- Benoist B, McLean E, Egli I, Cogswell M (2008) Worldwide prevalence of anaemia 1993–2005: WHO Global Database on Anaemia. WHO. http://apps.who.int/iris/bitstream/handle/10665/43894/9789241596657_eng.pdf?ua=1. Accessed 30 Nov 2018
- Fernandes AP, Januário JN, Cangussu CB, Macedo DL, Viana MB (2010) Mortality of children with sickle cell disease: a population study. *J Pediatr* 86(4):279–284. <https://doi.org/10.2223/JPED.2005>
- U. S. National Institute of Health, U. S. National Library of Medicine (2019) Sickle cell disease. NIH. <https://ghr.nlm.nih.gov/condition/sickle-cell-disease#genes>. Accessed 6 Feb 2019
- Modell B, Darlison M (2008) Global epidemiology of haemoglobin disorders and derived service indicators. *Bull World Health Organ* 86(6):480–487. <https://doi.org/10.2471/blt.06.036673>
- Brazilian Ministry of Health, Pesquisa Nacional de Demografia e Saúde da Criança e da Mulher (2006) Anemia e hipovitaminose A. Brazilian Ministry of Health. <http://bvsmms.saude.gov.br/bvms/pnds/anemia.php>. Accessed 2 Dec 2018
- Arduini GA, Rodrigues LP, Trovó de Marqui AB (2017) Mortality by sickle cell disease in Brazil. *Rev Bras Hematol Hemoter* 39(1): 52–56. <https://doi.org/10.1016/j.bjhh.2016.09.008>
- Brazilian Ministry of Health, Secretaria de Atenção à Saúde, Departamento de Atenção Especializada (2012) Doença falciforme: condutas básicas para tratamento. Brazilian Ministry of Health. http://bvsmms.saude.gov.br/bvms/publicacoes/doenca_falciforme_condutas_basicas.pdf. Accessed 2 Dec 2018
- Brabin BJ, Hakimi M, Pelletier D (2001) An analysis of anemia and pregnancy-related maternal mortality. *J Nutr* 131(2):604S–614S. <https://doi.org/10.1093/jn/131.2.604S>
- Haas JD, Brownlie T IV (2001) Iron deficiency and reduced work capacity: a critical review of the research to determine a causal relationship. *J Nutr* 131(2):676S–690S. <https://doi.org/10.1093/jn/131.2.676S>
- World Health Organization Regional Office for Europe, UNICEF Regional Office for Central and Eastern Europe the Commonwealth of Independence States and the Baltic States (1999) Prevention and control of iron-deficiency anaemia in women and children. UNICEF/WHO. http://www.who.int/nutrition/publications/micronutrients/anaemia_iron_deficiency/e73102/en/. Accessed 2 Dec 2018
- Dean J, Schechter AN (1978) Sickle-cell anemia: molecular and cellular bases of therapeutic approaches. *N Engl J Med* 299(14): 752–763. <https://doi.org/10.1056/NEJM197810052991405>
- Bryan LJ, Zakai NA (2012) Why is my patient anemic? *Hematol Oncol Clin North Am* 26(2):205–230. <https://doi.org/10.1016/j.hoc.2012.02.008>
- Archer NM, Brugnara C (2015) Diagnosis of iron-deficient states. *Crit Rev Clin Lab Sci* 52(5):256–272. <https://doi.org/10.3109/10408363.2015.1038744>
- Johnson-Wimbley TD, Graham DY (2011) Diagnosis and management of iron deficiency anemia in the 21st century. *Ther Adv Gastroenterol* 4(3):177–184. <https://doi.org/10.1177/1756283X11398736>
- Ali MA, Luxton AW, Walker WH (1978) Serum ferritin concentration and bone marrow iron stores: a prospective study. *Can Med Assoc J* 118(8):945–946
- Burns ER, Goldberg SN, Lawrence C, Wenz B (1990) Clinical utility of serum tests for iron deficiency in hospitalized patients. *Am J Clin Pathol* 93(2):240–245. <https://doi.org/10.1093/ajcp/93.2.240>
- Guyatt GH, Patterson C, Ali M, Singer J, Levine M, Turpie I, Meyer R (1990) Diagnosis of iron-deficiency anemia in the elderly. *Am J Med* 88(3):205–209. [https://doi.org/10.1016/0002-9343\(90\)90143-2](https://doi.org/10.1016/0002-9343(90)90143-2)
- Baumann Kurer S, Seifert B, Michel B, Ruegg R, Fehr J (1995) Prediction of iron deficiency in chronic inflammatory rheumatic disease anaemia. *Brit J Haematol* 91(4):820–826. <https://doi.org/10.1111/j.1365-2141.1995.tb05395.x>
- Mast AE, Blinder MA, Gronowski AM, Chumley C, Scott MG (1998) Clinical utility of the soluble transferrin receptor and comparison with serum ferritin in several populations. *Clin Chem* 44(1): 45–51
- Ogedegbe HO, Csury L, Simmons BH (2004) Anemias: a clinical laboratory perspective. *Lab Med* 35(3):177–185. <https://doi.org/10.1309/21XVW414MQRTKH14>
- Gonçalves MS, Bomfim GC, Maciel E, Cerqueira I, Lyra I, Zanette A, Bomfim G, Adorno EV, Albuquerque AL, Pontes A, Dupuit MF, Fernandes GB, dos Reis MG (2003) β^S -Haplotypes in sickle cell anemia patients from Salvador, Bahia, Northeastern Brazil. *Braz J Med Biol Res* 36(10):1283–1288. <https://doi.org/10.1590/S0100-879X2003001000001>
- Hanlon EB, Manoharan R, Koo TW, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, Feld MS (2000) Prospects for in vivo Raman spectroscopy. *Phys Med Biol* 45(2): R1–R59. <https://doi.org/10.1088/0031-9155/45/2/201>
- de Almeida ML, Saatkamp CJ, Fernandes AB, Pinheiro AL, Silveira L (2016) Estimating the concentration of urea and creatinine in the human serum of normal and dialysis patients through Raman spectroscopy. *Lasers Med Sci* 31(7):1415–1423. <https://doi.org/10.1007/s10103-016-2003-y>
- Silveira L, Borges RCF, Navarro RS, Giana HE, Zângaro RA, Pacheco MTT, Fernandes AB (2017) Quantifying glucose and lipid components in human serum by Raman spectroscopy and multivariate statistics. *Lasers Med Sci* 32(4):787–795. <https://doi.org/10.1007/s10103-017-2173-2>
- Bueno Filho AC, Silveira L, Yanai AL, Fernandes AB (2015) Raman spectroscopy for a rapid diagnosis of sickle cell disease in human blood samples: a preliminary study. *Lasers Med Sci* 30(1): 247–253. <https://doi.org/10.1007/s10103-014-1635-z>
- Atkins CG, Buckley K, Blades MW, Turner RFB (2017) Raman spectroscopy of blood and blood components. *Appl Spectrosc* 71(5):767–793. <https://doi.org/10.1177/0003702816686593>
- Matousek P, Stone N (2013) Recent advances in the development of Raman spectroscopy for deep non-invasive medical diagnosis. *J Biophotonics* 6(1):7–19. <https://doi.org/10.1002/jbio.201200141>
- Diem M, Mazur A, Lenau K, Schubert J, Bird B, Miljković M, Krafft C, Popp J (2013) Molecular pathology via IR and Raman spectral imaging. *J Biophotonics* 6(11–12):855–886. <https://doi.org/10.1002/jbio.201300131>
- Silveira L, Silveira FL, Bodanese B, Zângaro RA, Pacheco MT (2012) Discriminating model for diagnosis of basal cell carcinoma and melanoma in vitro based on the Raman spectra of selected biochemicals. *J Biomed Opt* 17(7):077003. <https://doi.org/10.1117/1.JBO.17.7.077003>
- Ward KR, Torres Filho I, Barbee RW, Torres L, Tiba MH, Reynolds PS, Pittman RN, Ivatury RR, Terner J (2006) Resonance Raman spectroscopy: a new technology for tissue oxygenation monitoring. *Crit Care Med* 34(3):792–799. <https://doi.org/10.1097/01.CCM.0000201898.43135.3F>

33. Torres Filho IP, Terner J, Pittman RN, Somera LG 3rd, Ward KR (2005) Hemoglobin oxygen saturation measurements using resonance Raman intravital microscopy. *Am J Physiol Heart Circ Physiol* 289(1):H488–H495. <https://doi.org/10.1152/ajpheart.01171.2004>
34. Eigenvector Research, Eigenvector Documentation Wiki (2017) Advanced preprocessing: sample normalization. Eigenvector Research. http://wiki.eigenvector.com/index.php?title=Advanced_Preprocessing:_Sample_Normalization. Accessed 10 June 2017
35. Chatfield C, Collins AJ (1980) Introduction to multivariate statistics. Chapman & Hall, London
36. Barker M, Rayens W (2003) Partial least squares for discrimination. *J Chemom* 17(3):166–173. <https://doi.org/10.1002/cem.785>
37. Nunes CA, Freitas MP, Pinheiro ACM, Bastos SC (2012) Chemoface: a novel free user-friendly interface for chemometrics. *J Braz Chem Soc* 23(11):2003–2010. <https://doi.org/10.1590/S0103-50532012005000073>
38. Vanna R, Ronchi P, Lenferink AT, Tresoldi C, Morasso C, Mehn D, Bedoni M, Picciolini S, Terstappen LW, Ciceri F, Otto C, Gramatica F (2015) Label-free imaging and identification of typical cells of acute myeloid leukaemia and myelodysplastic syndrome by Raman microspectroscopy. *Analyst* 140(4):1054–1064. <https://doi.org/10.1039/c4an02127d>
39. Lima AMF, Daniel CR, Navarro RS, Bodanese B, Pasqualucci CA, Pacheco MTT, Zângaro RA, Silveira L (2019) Discrimination of non-melanoma skin cancer and keratosis from normal skin tissue in vivo and ex vivo by Raman spectroscopy. *Vib Spectrosc* 100(1): 131–141. <https://doi.org/10.1016/j.vibspec.2018.11.009>
40. Brunner H, Sussner H (1973) Resonance Raman scattering on haemoglobin. *Biochim Biophys Acta* 310(1):20–31. [https://doi.org/10.1016/0005-2795\(73\)90004-4](https://doi.org/10.1016/0005-2795(73)90004-4)
41. Bankapur A, Zachariah E, Chidangil S, Valiathan M, Mathur D (2010) Raman tweezers spectroscopy of live, single red and white blood cells. *PLoS One* 5(4):e10427. <https://doi.org/10.1371/journal.pone.0010427>
42. Wood BR, McNaughton D (2002) Raman excitation wavelength investigation of single red blood cells in vivo. *J Raman Spectrosc* 33(7):517–523. <https://doi.org/10.1002/jrs.870>
43. De Luca AC, Rusciano G, Ciancia R, Martinelli V, Pesce G, Rotoli B, Selvaggi L, Sasso A (2008) Spectroscopical and mechanical characterization of normal and thalassemic red blood cells by Raman tweezers. *Opt Express* 16(11):7943–7957. <https://doi.org/10.1364/oe.16.007943>
44. Dingari NC, Horowitz GL, Kang JW, Dasari RR, Barman I (2012) Raman spectroscopy provides a powerful diagnostic tool for accurate determination of albumin glycation. *PLoS One* 7(2):e32406. <https://doi.org/10.1371/journal.pone.0032406>
45. Perez-Guaita D, de Veij M, Marzec KM, Almohammed ARD, McNaughton D, Hudson AJ, Wood BR (2017) Resonance Raman and UV-visible microscopy reveals that conditioning red blood cells with repeated doses of sodium dithionite increases haemoglobin oxygen uptake. *ChemistrySelect* 2(11):3342–3346. <https://doi.org/10.1002/slct.201700190>
46. Atkins CG, Schulze HG, Chen D, Devine DV, Blades MW, Turner RFB (2017) Using Raman spectroscopy to assess hemoglobin oxygenation in red blood cell concentrate: an objective proxy for morphological index to gauge the quality of stored blood? *Analyst* 142(12):2199–2210. <https://doi.org/10.1039/c7an00349h>
47. Wood BR, Tait B, McNaughton D (2001) Micro-Raman characterisation of the R to T state transition of haemoglobin within a single living erythrocyte. *Biochim Biophys Acta* 1539(1-2):58–70. [https://doi.org/10.1016/S0167-4889\(01\)00089-1](https://doi.org/10.1016/S0167-4889(01)00089-1)
48. Wood BR, Caspers P, Puppels GJ, Pandiancherri S, McNaughton D (2007) Resonance Raman spectroscopy of red blood cells using near-infrared laser excitation. *Anal Bioanal Chem* 387(5):1691–1703. <https://doi.org/10.1007/s00216-006-0881-8>
49. Rao S, Bálint S, Cossins B, Guallar V, Petrov D (2009) Raman study of mechanically induced oxygenation state transition of red blood cells using optical tweezers. *Biophys J* 96(1):209–216. <https://doi.org/10.1529/biophysj.108.139097>
50. Gautam R, Oh JY, Marques MB, Dluhy RA, Patel RP (2018) Characterization of storage-induced red blood cell hemolysis using Raman spectroscopy. *Lab Med* 49(4):298–310. <https://doi.org/10.1093/labmed/lmy018>
51. Talari ACS, Movasaghi Z, Rehman S, Rehman I (2015) Raman spectroscopy of biological tissues. *Appl Spectrosc Rev* 50(1):46–111. <https://doi.org/10.1080/05704928.2014.923902>
52. Jelkmann W (2007) Erythropoietin after a century of research: younger than ever. *Eur J Haematol* 78(3):183–205. <https://doi.org/10.1111/j.1600-0609.2007.00818.x>
53. Farrell K, Dent L, Buchowski M, Aguinaga MP (2009) Sickle cell anemia as a disease of oxygen transport: possible implications for the prevention of sickle cell crises. *Blood* 114(22):4617
54. Chowdhury A, Dasgupta R, Majumder SK (2017) Changes in hemoglobin-oxygen affinity with shape variations of red blood cells. *J Biomed Opt* 22(10):105006. <https://doi.org/10.1117/1.JBO.22.10.105006>
55. Brunner H, Mayer A, Sussner H (1972) Resonance Raman scattering on the haem group of oxy- and deoxyhaemoglobin. *J Mol Biol* 70(1):153–156. [https://doi.org/10.1016/0022-2836\(72\)90169-6](https://doi.org/10.1016/0022-2836(72)90169-6)
56. Weselucha-Birczyńska A, Kozicki M, Czepiel J, Łabanowska M, Nowak P, Kowalczyk G, Kurdziel M, Birczyńska M, Biesiada G, Mach T, Garlicki A (2014) Human erythrocytes analyzed by generalized 2D Raman correlation spectroscopy. *J Mol Struct* 1069: 305–312. <https://doi.org/10.1016/j.molstruc.2014.03.023>
57. Liu R, Mao Z, Matthews DL, Li CS, Chan JW, Satake N (2013) Novel single-cell functional analysis of red blood cells using laser tweezers Raman spectroscopy: application for sickle cell disease. *Exp Hematol* 41(7):656–661. <https://doi.org/10.1016/j.exphem.2013.02.012>
58. Yan D, Domes C, Domes R, Frosch T, Popp J, Pletz MW, Frosch T (2016) Fiber enhanced Raman spectroscopic analysis as a novel method for diagnosis and monitoring of diseases related to hyperbilirubinemia and hyperbiliverdinemia. *Analyst* 141(21): 6104–6115. <https://doi.org/10.1039/c6an01670g>
59. Wood BR, McNaughton D (2002) Micro-Raman characterization of high- and low-spin heme moieties within single living erythrocytes. *Biopolymers* 67(4-5):259–262. <https://doi.org/10.1002/bip.10120>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.