Talanta 174 (2017) 587–598

Contents lists available at ScienceDirect

Talanta

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

Elimination of interference from water in KBr disk FT-IR spectra of solid biomaterials by chemometrics solved with kinetic modeling☆

Sherald H. Gordon[a,*], Rogers E. Harry-O'kuru[b], Abdellatif A. Mohamed[c]

[a] Plant Polymer Research National Center for Agricultural Utilization Research, Agricultural Research Service, United States Department of Agriculture, 1815 N. University Street, Peoria, IL 61604, USA

[b] Bio-Oils Research, National Center for Agricultural Utilization Research, Agricultural Research Service, United States Department of Agriculture, 1815 N. University Street, Peoria, IL 61604, USA

[c] Department of Food Science & Nutrition, King Saud University, P.O. Box 2460, Riyadh 11451, Saudi Arabia

ARTICLE INFO

Keywords:
Biomaterial
FT-IR
Water interference
KBr disk
R-matrix method
Chemometric method

ABSTRACT

Infrared analysis of proteins and polysaccharides by the well known KBr disk technique is notoriously frustrated and defeated by absorbed water interference in the important amide and hydroxyl regions of spectra. This interference has too often been overlooked or ignored even when the resulting distortion is critical or even fatal, as in quantitative analyses of protein secondary structure, because the water has been impossible to measure or eliminate. Therefore, a new chemometric method was devised that corrects spectra of materials in KBr disks by mathematically eliminating the water interference. A new concept termed the Beer-Lambert law absorbance ratio (R-matrix) model was augmented with water concentration ratios computed via an exponential decay kinetic model of the water absorption process in KBr, which rendered the otherwise indeterminate system of linear equations determinate and thus possible to solve in a formal analytic manner. Consequently, the heretofore baffling KBr water elimination problem is now solved once and for all. Using the new formal solution, efforts to eliminate water interference from KBr disks in research will be defeated no longer. Resulting spectra of protein were much more accurate than attenuated total reflection (ATR) spectra corrected using the well-accepted Advanced ATR Correction Algorithm.

1. Introduction

For nearly seventy years, since 1947–1949, chemical research on solid materials has depended heavily on infrared spectroscopy using the pressed potassium bromide (KBr) disk technique [1–3]. A more modern technique, attenuated total reflection (ATR), which has become popular for routine analyses because of its speed and ease of sample preparation, is being adopted for quantitative analysis in research on solid materials despite its well known deficiencies in accuracy as well as its unavoidably poor reproducibility. Still, for advanced chemical research and quantitative analyses that demand precision Fourier transform infrared (FT-IR) spectrometry, the pressed KBr disk technique remains the method of choice and its use has expanded widely [4–11]. Indeed, KBr is still the only supporting medium that can be pressed into clear disks that closely approach the crucial solid solution condition required by the Bear-Lambert law for accurate chemometric analyses of solids. However, quantitative FT-IR analysis of biomaterials, such as proteins and polysaccharides, by the traditional KBr disk technique is always compromised if not completely frustrated and defeated by interference from absorbed water in the very important amide and hydroxyl regions of their spectra where water bands will distort or even obliterate these amide and hydroxyl bands. This is a major problem that has perplexed chemists and infrared spectroscopists for over 70

☆ Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

* Correspondence to: USDA-ARS, PO Box 10438, Peoria, IL 61612.
E-mail address: sherald_gordon@icloud.com (S.H. Gordon).

http://dx.doi.org/10.1016/j.talanta.2017.06.043
Received 13 May 2017; Received in revised form 15 June 2017; Accepted 16 June 2017
Available online 23 June 2017
0039-9140/ Published by Elsevier B.V.

years [1–3,11]. The obstacle for quantitative FT-IR analysis of biomaterials in real samples is the strong interference bands arising from water absorbed by the hygroscopic KBr in sample preparation. KBr is notorious in infrared spectroscopy because of intractable water interference [12–19].

This interference has all too often been overlooked or ignored even where the resulting spectral distortion is critical or even fatal, as in quantitative analyses of protein secondary structure, because it has been impossible to measure the water and correct for or eliminate the interference. In addition, since the water concentration in KBr is too minute and too impracticable to measure, known chemometric methods like the classical K-matrix, P-matrix and Q-matrix least squares methods are rendered mathematically indeterminate for solution of this problem [11,20]. Therefore, the true and accurate infrared spectra of pure biomaterials in KBr are unknown and unavailable at present. Because the amide and hydroxyl bands in proteins and polysaccharides are increasingly being scrutinized in biochemical, environmental and agricultural research on their conformations, secondary and tertiary structures, intermolecular interactions, hydrogen bonding and other effects, more precise quantitative analysis of these bands in such biomaterials is very much needed.

Therefore, a chemometric method was recently invented that corrects FT-IR spectra of solid biomaterials in KBr disks by mathematically eliminating the absorbed water interference [11]. The method employed the Beer-Lambert law to directly compute individual water and solids concentrations from a system of linear chemometric equations derived to solve the water interference problem in a rigorous and objective way. It introduced a new concept termed the Beer-Lambert law absorbance ratio (R-matrix) model. By combining the model with water concentration ratios obtained from spectral peak resolution (called deconvolution or curve fitting) of the infrared bands affected by water absorption, the otherwise indeterminate system of linear Beer-Lambert law equations was rendered amenable to solution. This chemometric approach has already demonstrated its utility. The R-matrix model plus spectral peak resolution not only made the equations determinate, it also solved the problem from only two KBr disk sample preparations. However, as a practical tool, this approach has certain disadvantages and controversial assumptions which will be discussed.

In the present work, a new and more direct approach was discovered that augmented a further expanded R-matrix model with water concentration ratios computed via an exponential decay kinetic model of the water absorption process in KBr, which not only made the otherwise indeterminate system of linear equations possible to solve rigorously and objectively, but also more efficiently from a single KBr disk sample preparation.

2. Theory and derivation of the R-matrix model

Consider, for example, the FT-IR spectra of a set of biomaterials consisting of a number of different mixtures of four component materials, as depicted in Fig. 1. The Beer-Lambert law [21–24] assumes the infrared absorbances of the individual components in a mixture spectrum are additive and the total absorbance is a linear function of the component concentrations. Thus, the Beer-Lambert law is expressed by a system of multicomponent linear equations that relate the total absorbance (A) of a mixture at a given infrared frequency (or wavenumber) to the concentrations (**c**) of its components as:

$$A_{vm} = \sum_{n=1}^{p} k_{vn}\, c_{nm} \tag{1}$$

for frequencies v = 1, 2, 3, ..., w; components n = 1, 2, 3, ..., p; and mixtures m = 1, 2, 3, ..., q; where $\mathbf{k}_{vn}$ is the absorption coefficient (absorptivity) at frequency v of component n, and $\mathbf{c}_{nm}$ is the concentration of component n in mixture m. The FT-IR spectral data form a w by q matrix of absorbances (A) which contains the spectra at w frequencies for each of the q mixtures. A matrix of w x p absorption coefficients (**K**) of the p components in each mixture at the w frequencies is formed from known mixtures. A matrix of $p \times q$ concentrations (**C**) of the p components in each of the q mixtures is also formed when the concentrations of components in the mixtures are known.

Expressed in matrix-vector notation, the system of equations for the Beer-Lambert law is

$$\mathbf{K}_{vn}\mathbf{C}_{nm} = \mathbf{A}_{vn} \tag{2}$$

If the example illustrated in Fig. 1 is extended to five or more different mixtures of four components, the system has $q \geq 5$, $p = 4$, and $w = 4$, where the $\mathbf{k}_{vn}$ are absorption coefficients for each of the component concentrations $\mathbf{c}_{nm}$ in mixtures having total absorbances $\mathbf{A}_{vm}$, where the subscripts are the integer indices of the w frequencies, p components and q mixtures as denoted above, and where $q \geq p$ and $w \geq p$ are constraints necessary for unique mathematical solution.

Thus, at the first four frequencies in infrared spectra, the absorbances of the first five mixtures, with four component concentrations $\mathbf{c}_{11}$, $\mathbf{c}_{21}$, $\mathbf{c}_{31}$ and $\mathbf{c}_{41,}$ are defined by the Beer-Lambert law as:

$$\begin{aligned}
&k11\ c11 + k12\ c21 + k13c31 + k14\ c41 = A11\\
&k21\ c11 + k22\ c21 + k23\ c31 + k24\ c41 = A21\\
&k31\ c11 + k32\ c21 + k33\ c31 + k34\ c41 = A31\\
&k41\ c11 + k42\ c21 + k43\ c31 + k44\ c41 = A41\\
&\\
&k11\ c12 + k12\ c22 + k13\ c32 + k14\ c42 = A12\\
&k21\ c12 + k22\ c22 + k23\ c32 + k24\ c42 = A22\\
&k31\ c12 + k32\ c22 + k33\ c32 + k34\ c42 = A32\\
&k41\ c12 + k42\ c22 + k43\ c32 + k44\ c42 = A42\\
&\\
&k11\ c13 + k12\ c23 + k13\ c33 + k14\ c43 = A13\\
&k21\ c13 + k22\ c23 + k23\ c33 + k24\ c43 = A23\\
&k31\ c13 + k32\ c23 + k33\ c33 + k34\ c43 = A33\\
&k41\ c13 + k42\ c23 + k43\ c33 + k44\ c43 = A43\\
&\\
&k11\ c14 + k12\ c24 + k13\ c34 + k14\ c44 = A14\\
&k21\ c14 + k22\ c24 + k23\ c34 + k24\ c44 = A24\\
&k31\ c14 + k32\ c24 + k33\ c34 + k34\ c44 = A34\\
&k41\ c14 + k42\ c24 + k43\ c34 + k44\ c44 = A44\\
&\\
&k11\ c15 + k12\ c25 + k13c35 + k14\ c45 = A15\\
&k21\ c15 + k22\ c25 + k23\ c35 + k24\ c45 = A25\\
&k31\ c15 + k32\ c25 + k33\ c35 + k34\ c45 = A35\\
&k41\ c15 + k42\ c25 + k43\ c35 + k44\ c45 = A45
\end{aligned} \tag{3}$$

where **c**11, $\mathbf{c}_{21}$, **c**31 and **c**41 are the four component concentrations in the first mixture, **c**12, **c**22, **c**32 and **c**42 are the concentrations in the second mixture, **c**13, **c**23, **c**33 and **c**43 are concentrations in the third mixture, **c**14, **c**24, **c**34 and **c**44 are concentrations in the fourth mixture, and **c**15, **c**25, **c**35 and **c**45 are concentrations in the fifth mixture. If there are no other components present in each mixture, then these concentrations, expressed as weight fractions, sum to unity:

$$\begin{aligned}
&c11 + c21 + c31 + c41 = 1\\
&c12 + c22 + c32 + c42 = 1\\
&c13 + c23 + c33 + c43 = 1\\
&c14 + c24 + c34 + c44 = 1\\
&c15 + c25 + c35 + c45 = 1
\end{aligned} \tag{4}$$

When all concentrations of the components in all of the mixtures are known, the absorption coefficients $\mathbf{k}_{vn}$ for each frequency can be calibrated by multiple linear regression using classical chemometric

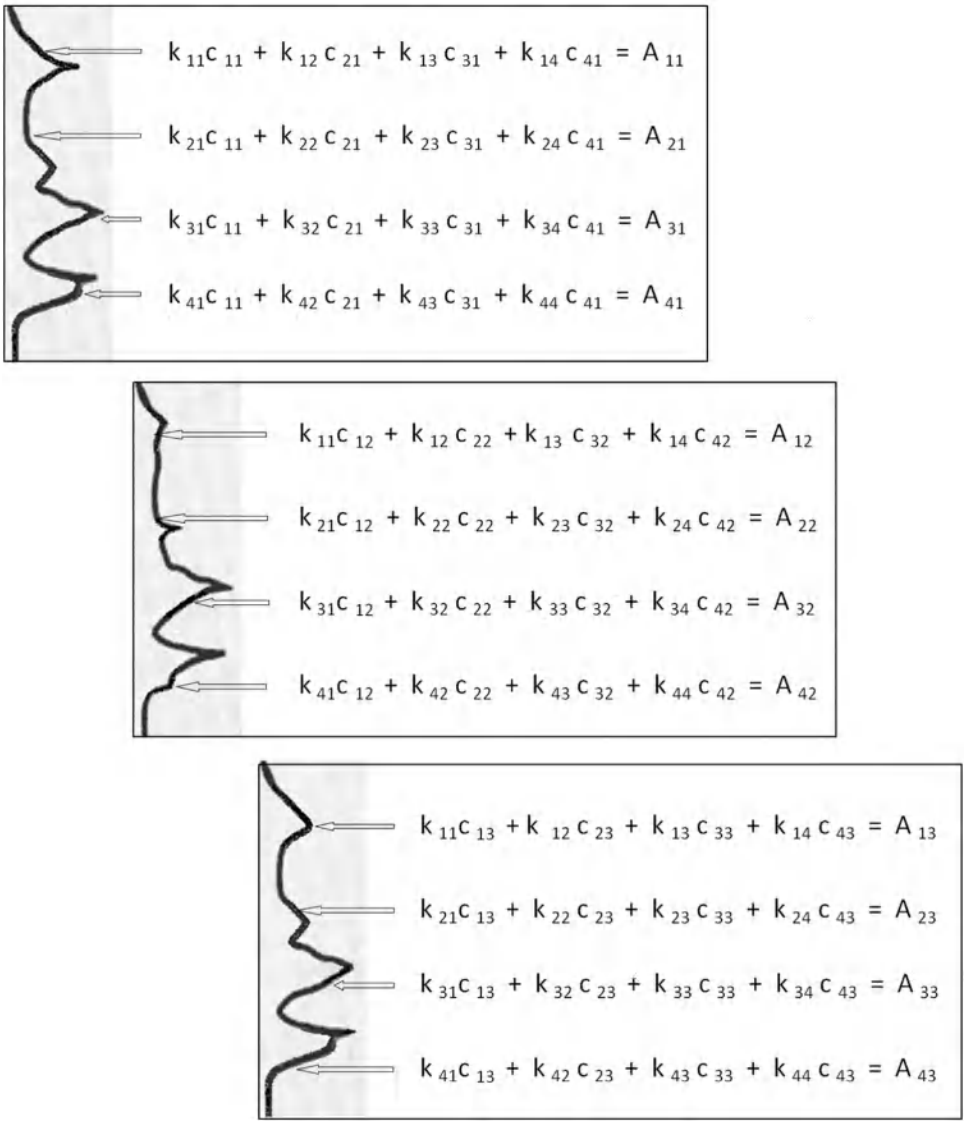

Fig. 1. FT-IR spectra of three mixtures of three components with Beer-Lambert law equations for absorbances at four frequencies.

methods [23,25]. In the analysis of an unknown mixture, the absorbances $\mathbf{A}_{vm}$ at the same frequencies are measured and the desired concentrations $\mathbf{c}_{nm}$ of the components in the mixture can be directly calculated from $\mathbf{k}_{vn}$ values obtained in prior calibration of known mixtures.

However, for the mixtures of a single solid biomaterial and absorbed water in KBr considered in this work, where the concentration of the solid biomaterial is unknown and the concentration of the absorbed water is also unknown and unmeasurable, the absorption coefficients, $\mathbf{k}$, for the solids at each frequency can be calculated by the R-matrix method conceived earlier at this laboratory [10].

If the infrared spectrum of the pure biomaterial in KBr were available, the ratios of the absorbance bands for any two frequencies in a given spectrum would be known (measurable) even though the absorption coefficients $\mathbf{k}_{vn}$ would be unknown. And, because the concentrations of the pure biomaterial at any two frequencies are identical, the absorbance ratios, $\boldsymbol{r}_{vn}$, for the two frequencies would be equal to the ratios of the unknown absorption coefficients $\mathbf{k}_{vn}$ as follows

$$\begin{aligned}
\boldsymbol{r}11 &= \frac{k11}{k11} \quad \boldsymbol{r}22 = \frac{k12}{k12} \quad \boldsymbol{r}23 = \frac{k13}{k13} \quad \boldsymbol{r}24 = \frac{k14}{k14} \\
\boldsymbol{r}21 &= \frac{k21}{k11} \quad \boldsymbol{r}22 = \frac{k22}{k12} \quad \boldsymbol{r}23 = \frac{k23}{k13} \quad \boldsymbol{r}24 = \frac{k24}{k14} \\
\boldsymbol{r}31 &= \frac{k31}{k11} \quad \boldsymbol{r}32 = \frac{k32}{k12} \quad \boldsymbol{r}33 = \frac{k33}{k13} \quad \boldsymbol{r}34 = \frac{k34}{k14} \\
\boldsymbol{r}41 &= \frac{k41}{k11} \quad \boldsymbol{r}42 = \frac{k42}{k12} \quad \boldsymbol{r}43 = \frac{k43}{k13} \quad \boldsymbol{r}44 = \frac{k44}{k14}
\end{aligned} \tag{5}$$

where each absorbance ratio, $\boldsymbol{r}_{vn}$, is the ratio of the absorbance of component n at frequency v to the absorbance of the same component n at the first frequency, $v = 1$. Thus, all absorbance band ratios are unknown except $\boldsymbol{r}_{11} = \boldsymbol{r}_{12} = \boldsymbol{r}_{13} = \boldsymbol{r}_{14} = 1$ given.

Rewriting Eq. (3) in terms of these absorbance ratios, $\boldsymbol{r}_{vn}$, gives the new set of absorbance

equations for the two selected frequencies as

$$
\begin{aligned}
&r11\ \ k11c11 + r12\ \ k12c21 + r13\ \ k13c31 + r14\ \ k14c41 = A11\\
&r21\ \ k11c11 + r22\ \ k12c21 + r23\ \ k13c31 + r24\ \ k14c41 = A21\\
&r31\ \ k11c11 + r32\ \ k12c21 + r33\ \ k13c31 + r34\ \ k14c41 = A31\\
&r41\ \ k11c11 + r42\ \ k12c21 + r43\ \ k13c31 + r44\ \ k14c41 = A41\\
&r11\ \ k11c12 + r12\ \ k12c22 + r13\ \ k13c32 + r14\ \ k14c42 = A12\\
&r21\ \ k11c12 + r22\ \ k12c22 + r23\ \ k13c32 + r24\ \ k14c42 = A22\\
&r31\ \ k11c12 + r32\ \ k12c22 + r33\ \ k13c32 + r34\ \ k14c42 = A32\\
&r41\ \ k11c12 + r42\ \ k12c22 + r43\ \ k13c32 + r44\ \ k14c42 = A42\\
&r11\ \ k11c13 + r12\ \ k12c23 + r13\ \ k13c33 + r14\ \ k14c43 = A13\\
&r21\ \ k11c13 + r22\ \ k12c23 + r23\ \ k13c33 + r24\ \ k14c43 = A23\\
&r31\ \ k11c13 + r32\ \ k12c23 + r33\ \ k13c33 + r34\ \ k14c43 = A33\\
&r41\ \ k11c13 + r42\ \ k12c23 + r43\ \ k13c33 + r44\ \ k14c43 = A43\\
&r11\ \ k11c14 + r12\ \ k12c24 + r13\ \ k13c34 + r14\ \ k14c44 = A14\\
&r21\ \ k11c14 + r22\ \ k12c24 + r23\ \ k13c34 + r24\ \ k14c44 = A24\\
&r31\ \ k11c14 + r32\ \ k12c24 + r33\ \ k13c34 + r34\ \ k14c44 = A34\\
&r41\ \ k11c14 + r42\ \ k12c24 + r43\ \ k13c34 + r44\ \ k14c44 = A44\\
&r11\ \ k11c15 + r12\ \ k12c25 + r13\ \ k13c35 + r14\ \ k14c45 = A15\\
&r21\ \ k11c15 + r22\ \ k12c25 + r23\ \ k13c35 + r24\ \ k14c45 = A25\\
&r31\ \ k11c15 + r32\ \ k12c25 + r33\ \ k13c35 + r34\ \ k14c45 = A35\\
&r41\ \ k11c15 + r42\ \ k12c25 + r43\ \ k13c35 + r44\ \ k14c45 = A45
\end{aligned}
\tag{6}
$$

which is written in matrix form as

$$
\begin{bmatrix} r11 & r12 & r13 & r14\\ r21 & r22 & r23 & r24\\ r31 & r32 & r33 & r34\\ r41 & r42 & r43 & r44 \end{bmatrix}
\begin{bmatrix} k11c11 & k11c12 & k11c13 & \ldots & k11c1m\\ k12c21 & k12c22 & k12c23 & \ldots & k12c2m\\ k13c31 & k13c32 & k13c33 & \ldots & k13c3m\\ k14c41 & k14c42 & k14c43 & \ldots & k14c4m \end{bmatrix}
= \begin{bmatrix} A11 & A12 & A13 & \ldots & A1m\\ A21 & A22 & A23 & \ldots & A2m\\ A31 & A32 & A33 & \ldots & A3m\\ A41 & A42 & A43 & \ldots & A4m \end{bmatrix}
\tag{7}
$$

where the $m = 5$ or more systems of equations, at the four frequencies, contain the same

four by four matrix R of absorbance ratios. This ratio matrix (R-matrix) R and its determinant Det(R) are designated by different brackets. Thus,

$$
R = \begin{bmatrix} r11 & r12 & r13 & r14\\ r21 & r22 & r23 & r24\\ r31 & r32 & r33 & r34\\ r41 & r42 & r43 & r44 \end{bmatrix}
$$

$$
\mathrm{Det}(R) = \begin{vmatrix} r11 & r12 & r13 & r14\\ r21 & r22 & r23 & r24\\ r31 & r32 & r33 & r34\\ r41 & r42 & r43 & r44 \end{vmatrix}
\tag{8}
$$

Provided the frequencies are selected with absorbance ratios such that R is nonsingular, i.e., Det(R) ≠ 0, the five or more systems of four equations in four unknowns (Eq. (7)) will be determinate and therefore capable of being solved using Cramer's rule [26] as follows:

$$
\begin{aligned}
k11c11 &= \frac{\begin{vmatrix} A11 & r12 & r13 & r14\\ A21 & r22 & r23 & r24\\ A31 & r32 & r33 & r34\\ A41 & r42 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k12c21 = \frac{\begin{vmatrix} r11 & A11 & r13 & r14\\ r21 & A21 & r23 & r24\\ r31 & A31 & r33 & r34\\ r41 & A41 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)}\\
k11c12 &= \frac{\begin{vmatrix} A12 & r12 & r13 & r14\\ A22 & r22 & r23 & r24\\ A32 & r32 & r33 & r34\\ A42 & r42 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k12c22 = \frac{\begin{vmatrix} r11 & A12 & r13 & r14\\ r21 & A22 & r23 & r24\\ r31 & A32 & r33 & r34\\ r41 & A42 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)}\\
k11c13 &= \frac{\begin{vmatrix} A13 & r12 & r13 & r14\\ A23 & r22 & r23 & r24\\ A33 & r32 & r33 & r34\\ A43 & r42 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k12c23 = \frac{\begin{vmatrix} r11 & A13 & r13 & r14\\ r21 & A23 & r23 & r24\\ r31 & A33 & r33 & r34\\ r41 & A43 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)}\\
k11c14 &= \frac{\begin{vmatrix} A14 & r12 & r13 & r14\\ A24 & r22 & r23 & r24\\ A34 & r32 & r33 & r34\\ A44 & r42 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k12c24 = \frac{\begin{vmatrix} r11 & A14 & r13 & r14\\ r21 & A24 & r23 & r24\\ r31 & A34 & r33 & r34\\ r41 & A44 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)}\\
k11c15 &= \frac{\begin{vmatrix} A15 & r12 & r13 & r14\\ A25 & r22 & r23 & r24\\ A35 & r32 & r33 & r34\\ A45 & r42 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k12c25 = \frac{\begin{vmatrix} r11 & A15 & r13 & r14\\ r21 & A25 & r23 & r24\\ r31 & A35 & r33 & r34\\ r41 & A45 & r43 & r44 \end{vmatrix}}{\mathrm{Det}(R)}
\end{aligned}
\tag{9}
$$

and

$$
\begin{aligned}
k13c31 &= \frac{\begin{vmatrix} r11 & r12 & A11 & r14\\ r21 & r22 & A21 & r24\\ r31 & r32 & A31 & r34\\ r41 & r42 & A41 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k14c41 = \frac{\begin{vmatrix} r11 & r12 & r13 & A11\\ r21 & r22 & r23 & A21\\ r31 & r32 & r33 & A31\\ r41 & r42 & r43 & A41 \end{vmatrix}}{\mathrm{Det}(R)}\\
k13c32 &= \frac{\begin{vmatrix} r11 & r12 & A12 & r14\\ r21 & r22 & A22 & r24\\ r31 & r32 & A32 & r34\\ r41 & r42 & A42 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k14c42 = \frac{\begin{vmatrix} r11 & r12 & r13 & A12\\ r21 & r22 & r23 & A22\\ r31 & r32 & r33 & A32\\ r41 & r42 & r43 & A42 \end{vmatrix}}{\mathrm{Det}(R)}\\
k13c33 &= \frac{\begin{vmatrix} r11 & r12 & A13 & r14\\ r21 & r22 & A23 & r24\\ r31 & r32 & A33 & r34\\ r41 & r42 & A43 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k14c43 = \frac{\begin{vmatrix} r11 & r12 & r13 & A13\\ r21 & r22 & r23 & A23\\ r31 & r32 & r33 & A33\\ r41 & r42 & r43 & A43 \end{vmatrix}}{\mathrm{Det}(R)}\\
k13c34 &= \frac{\begin{vmatrix} r11 & r12 & A14 & r14\\ r21 & r22 & A24 & r24\\ r31 & r32 & A34 & r34\\ r41 & r42 & A44 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k14c44 = \frac{\begin{vmatrix} r11 & r12 & r13 & A14\\ r21 & r22 & r23 & A24\\ r31 & r32 & r33 & A34\\ r41 & r42 & r43 & A44 \end{vmatrix}}{\mathrm{Det}(R)}\\
k13c35 &= \frac{\begin{vmatrix} r11 & r12 & A15 & r14\\ r21 & r22 & A25 & r24\\ r31 & r32 & A35 & r34\\ r41 & r42 & A45 & r44 \end{vmatrix}}{\mathrm{Det}(R)} \qquad
k14c45 = \frac{\begin{vmatrix} r11 & r12 & r13 & A15\\ r21 & r22 & r23 & A25\\ r31 & r32 & r33 & A35\\ r41 & r42 & r43 & A45 \end{vmatrix}}{\mathrm{Det}(R)}
\end{aligned}
\tag{10}
$$

If the concentrations are measured as relative weights and it is known that the weights of one or more of the four components in the mixtures are constant while the other components vary in their weights, this fact can be modeled into the above system of equations. As given in this present work, the weight of only one component, $\mathbf{c}_{11,}$ is known to be constant in all of five mixtures, in which case

$$
k11c11 = k11c12 = k11c13 = k11c14 = k11c15
\tag{11}
$$

Then, all of the unknown absorption coefficients, $\boldsymbol{k}$, can be eliminated (cancelled out) by dividing by the corresponding absorbance $\mathbf{k}_{vn}\mathbf{c}_{nm}$ computed from Eqs. (9) and (10). Dividing the absorbance of one mixture by the corresponding absorbance of another mixture will give the ratios of the concentrations of the two components in terms of the absorbance ratios, $\boldsymbol{r}$, for the components at the two given frequencies. Dividing $\mathbf{k}_{11}\mathbf{c}_{11}$ by $\mathbf{k}_{11}\mathbf{c}_{12,}$ which are assumed to be equal in this example (Eq. (11)), will give

$$
\frac{c11}{c12} = \frac{\begin{vmatrix} A11 & r12 & r13 & r14\\ A21 & r22 & r23 & r24\\ A31 & r32 & r33 & r34\\ A41 & r42 & r43 & r44 \end{vmatrix}}{\begin{vmatrix} A12 & r12 & r13 & r14\\ A22 & r22 & r23 & r24\\ A32 & r32 & r33 & r34\\ A42 & r42 & r43 & r44 \end{vmatrix}} = 1
\tag{12}
$$

Similarly, dividing other absorbances, such as $\mathbf{k}_{11}\mathbf{c}_{11}$ by $\mathbf{k}_{11}\mathbf{c}_{13}$ and $\mathbf{k}_{11}\mathbf{c}_{11}$ by $\mathbf{k}_{11}\mathbf{c}_{13}$ will give

$$\frac{c11}{c13} = \frac{\begin{vmatrix} A11 & r12 & r13 & r14 \\ A21 & r22 & r23 & r24 \\ A31 & r32 & r33 & r34 \\ A41 & r42 & r43 & r44 \end{vmatrix}}{\begin{vmatrix} A13 & r12 & r13 & r14 \\ A23 & r22 & r23 & r24 \\ A33 & r32 & r33 & r34 \\ A43 & r42 & r43 & r44 \end{vmatrix}} = 1$$

$$\frac{c12}{c15} = \frac{\begin{vmatrix} A12 & r12 & r13 & r14 \\ A22 & r22 & r23 & r24 \\ A32 & r32 & r33 & r34 \\ A42 & r42 & r43 & r44 \end{vmatrix}}{\begin{vmatrix} A15 & r12 & r13 & r14 \\ A25 & r22 & r23 & r24 \\ A35 & r32 & r33 & r34 \\ A45 & r42 & r43 & r44 \end{vmatrix}} = 1 \qquad (13)$$

Since these are ratios of two component concentrations in the five mixtures of this model, the following set of 10 (combinations of five taken two at a time) concentration ratios can be computed, of which Eqs. (12) and (13) are three members.

$$\frac{c11}{c12} = \frac{c11}{c13} = \frac{c11}{c14} = \frac{c11}{c15} = \frac{c12}{c13} = \frac{c12}{c14} = \frac{c12}{c15} = \frac{c13}{c14} = \frac{c13}{c15} = \frac{c14}{c15} = 1 \qquad (14)$$

These contain the necessary and sufficient information to calculate the nine absorbance ratios, $\boldsymbol{r}$, in Eqs. (12) and (13) from 10 different simultaneous equations derived from Eq. (14). Thus, if the nine absorbance band ratios, $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{34}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{43}$, and $\boldsymbol{r}_{44}$ are unknown or unavailable (not measurable), they can be obtained explicitly from this model by solving the system of 10 simultaneous nonlinear equations in nine unknowns. For example, the first of the 10 nonlinear equations, derived by cofactor expansion of the determinants in Eq. (12), is

$$\frac{c11}{c12} = \frac{A11\begin{vmatrix} r22 & r23 & r24 \\ r32 & r33 & r34 \\ r42 & r43 & r44 \end{vmatrix} - A21\begin{vmatrix} r12 & r13 & r14 \\ r32 & r33 & r34 \\ r42 & r43 & r44 \end{vmatrix} + A31\begin{vmatrix} r12 & r13 & r14 \\ r22 & r23 & r24 \\ r42 & r43 & r44 \end{vmatrix} - A41\begin{vmatrix} r12 & r13 & r14 \\ r22 & r23 & r24 \\ r32 & r33 & r34 \end{vmatrix}}{A12\begin{vmatrix} r22 & r23 & r24 \\ r32 & r33 & r34 \\ r42 & r43 & r44 \end{vmatrix} - A22\begin{vmatrix} r12 & r13 & r14 \\ r32 & r33 & r34 \\ r42 & r43 & r44 \end{vmatrix} + A32\begin{vmatrix} r12 & r13 & r14 \\ r22 & r23 & r24 \\ r42 & r43 & r44 \end{vmatrix} - A42\begin{vmatrix} r12 & r13 & r14 \\ r22 & r23 & r24 \\ r32 & r33 & r34 \end{vmatrix}} = 1 \qquad (15)$$

On the other hand, dividing $\mathbf{k}_{12}\mathbf{c}_{21}$ by $\mathbf{k}_{12}\mathbf{c}_{22}$, which in this example are not assumed to be equal, will give

$$\frac{c21}{c22} = \frac{\begin{vmatrix} r11 & A11 & r13 & r14 \\ r21 & A21 & r23 & r24 \\ r31 & A31 & r33 & r34 \\ r41 & A41 & r43 & r44 \end{vmatrix}}{\begin{vmatrix} r11 & A12 & r13 & r14 \\ r21 & A22 & r23 & r24 \\ r31 & A32 & r33 & r34 \\ r41 & A42 & r43 & r44 \end{vmatrix}} \qquad (16)$$

and similarly,

$$\frac{c31}{c33} = \frac{\begin{vmatrix} r11 & r12 & A11 & r14 \\ r21 & r22 & A21 & r24 \\ r31 & r32 & A31 & r34 \\ r41 & r42 & A41 & r44 \end{vmatrix}}{\begin{vmatrix} r11 & r12 & A13 & r14 \\ r21 & r22 & A23 & r24 \\ r31 & r32 & A33 & r34 \\ r41 & r42 & A43 & r44 \end{vmatrix}} \qquad \frac{c42}{c45} = \frac{\begin{vmatrix} r11 & r12 & r13 & A12 \\ r21 & r22 & r23 & A22 \\ r31 & r32 & r33 & A32 \\ r41 & r42 & r43 & A42 \end{vmatrix}}{\begin{vmatrix} r11 & r12 & r13 & A15 \\ r21 & r22 & r23 & A25 \\ r31 & r32 & r33 & A35 \\ r41 & r42 & r43 & A45 \end{vmatrix}} \qquad (17)$$

However, unlike Eqs. (12) and (13), these Eqs. (16) and (17) contain insufficient information to calculate the absorbance ratios $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$, even when $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{34}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{43}$, and $\boldsymbol{r}_{44}$ are known or obtained directly from Eq. (14), because the concentration ratios are unknown. Therefore, the desired ratios of the absorbance bands $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ cannot be obtained explicitly by calculation from this R-matrix model alone.

Any equivalent chemometric model that uses physical measurements of $\mathbf{k}_{12}\mathbf{c}_{21}$, $\mathbf{k}_{22}\mathbf{c}_{21}$, $\mathbf{k}_{12}\mathbf{c}_{22}$, $\mathbf{k}_{22}\mathbf{c}_{22}$, or the other component absorbances as external inputs will give $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ explicitly. Hence, the ratios, $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{34}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{43}$, and $\boldsymbol{r}_{44}$, of absorbances at the four frequencies would be obtained from the infrared absorbance measurements of five or more different mixtures. However, without external measurements of the variable component concentration ratios like Eqs. (16) and (17) ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{33}$ and $\mathbf{c}_{42}/\mathbf{c}_{45}$, which are just three of 30 ($(\boldsymbol{n}-1)\times 100$) possible combinations in the five mixtures), or without their absorbance equivalents, the absorbance ratios $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ for the constant component cannot be obtained. In fact, it can be proved mathematically that without external measurement of the variable component concentrations or their absorbances independent of the Beer-Lambert law model, the systems of R-matrix equations (Eq. 6) are indeterminate, i.e., each system has an infinite number of solutions because any matrix expression of Eq. (6) is singular (Determinant of Determinant of R = 0, whether computed analytically or numerically = 0, whether computed analytically or numerically).

Therefore, in order to solve this system of R-matrix equations, the unknown concentration ratios ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, $\mathbf{c}_{41}/\mathbf{c}_{42....}$) or their absorbance equivalents must be measured by methods that do not derive solely from the Beer-Lambert law as already defined. In other words, the unknown ratios ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, $\mathbf{c}_{41}/\mathbf{c}_{42}$, ...) must be measured from data external to the model (Eq. (6)), i.e., not from the infrared absorbance measurements ($\mathbf{A}_{11}$, $\mathbf{A}_{12}$, $\mathbf{A}_{13}$, $\mathbf{A}_{21}$, $\mathbf{A}_{22}$, $\mathbf{A}_{23...}$) alone.

Once the unknown ratios ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, $\mathbf{c}_{41}/\mathbf{c}_{42}$ ....) are determined, their values can be substituted into Eqs. (16) and (17) to compute the three absorbance ratios $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ directly from three simultaneous nonlinear equations (three nonlinear equations in three unknowns) of the nine absorbance ratios, $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{34}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{43}$, and $\boldsymbol{r}_{44}$ calculated as in Eq. (15).

For example, one of the three simultaneous nonlinear equations, derived by cofactor expansion of the determinants in Eq. (16), is

$$\left(\frac{c21}{c22}\right)^{*} = \frac{-A11\begin{vmatrix} r21 & r23 & r24 \\ r31 & r33 & r34 \\ r41 & r43 & r44 \end{vmatrix} + A21\begin{vmatrix} r11 & r13 & r14 \\ r31 & r33 & r34 \\ r41 & r43 & r44 \end{vmatrix} - A31\begin{vmatrix} r11 & r13 & r14 \\ r21 & r23 & r24 \\ r41 & r43 & r44 \end{vmatrix} + A41\begin{vmatrix} r11 & r13 & r14 \\ r21 & r23 & r24 \\ r31 & r33 & r34 \end{vmatrix}}{-A12\begin{vmatrix} r21 & r23 & r24 \\ r31 & r33 & r34 \\ r41 & r43 & r44 \end{vmatrix} + A22\begin{vmatrix} r11 & r13 & r14 \\ r31 & r33 & r34 \\ r41 & r43 & r44 \end{vmatrix} - A32\begin{vmatrix} r11 & r13 & r14 \\ r21 & r23 & r24 \\ r41 & r43 & r44 \end{vmatrix} + A42\begin{vmatrix} r11 & r13 & r14 \\ r21 & r23 & r24 \\ r31 & r33 & r34 \end{vmatrix}} \qquad (18)$$

where $(\mathbf{c}_{21}/\mathbf{c}_{22})^*$ is measured from data external to the model (Eq. (6)).

The solution described above for the four component model (Eqs. (16) and (18)) can be extended to higher dimension models or reduced to lower dimensions as needed. The simplest alternative model with only two components (one solid and one water), which was derived previously [11], has no interactions or cross-products of any of the variables, so the solutions are the two linear equations

$$\frac{c11}{c12} = \frac{A21 - r22A11}{A22 - r22A12} = 1 \quad \text{and} \quad \left(\frac{c21}{c22}\right)^* = \frac{r21A11 - A21}{r21A12 - A22} \tag{19}$$

2.1. Theory and derivation of the exponential decay model

According to early studies, water is retained in KBr disks in three types [15]. The three modes of these water vibrations appear in FT-IR spectra as three overlapping absorbance bands in the OH stretching region as confirmed from second derivative spectra [11]. It was observed, in two independent studies, that the magnitudes of these water bands in KBr decrease over time [11,15]. In the early work, isothermal desorption of water from KBr disks was measured by monitoring the infrared absorbance of the water over time. Various kinetic equations were used to examine the kinetics of the water desorption process in KBr. Ten different kinetic equations were studied to find a theoretical model that best described the mechanism of the desorption process, which was determined to be a three-dimensional diffusion mechanism in the porous structure of KBr disks [15].

In the present study, it was noticed that the decrease in the magnitude of the water band in FT-IR spectra, observed in both studies, was essentially exponential over time. This means the kinetics of this process, whatever the absorption/desorption mechanism, can be best described by an exponential decay equation. Therefore, the changes in the relative concentrations of water in KBr disks over time can be directly measured from an exponential decay empirical model. Therefore, in this work, isothermal measurements on time-resolved FT-IR spectra of biomaterial samples in KBr disks were undertaken to determine the water concentration ratios ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, $\mathbf{c}_{41}/\mathbf{c}_{42}$ ....) needed to solve the system of R-matrix equations (Eq. 6) which, as explained above, would otherwise be indeterminate.

If the three types of water known to exist [11,15] are present together with a solid biomaterial in KBr disks, the FT-IR spectra will show four overlapping absorbance bands in the OH region, three from the water OH and one from the biomaterial OH. While the band from the solid biomaterial remains constant over time, the three water bands will decrease exponentially. This produces a complex triple exponential decay pattern that is difficult to visualize graphically. Therefore, for simplicity, in Fig. 2, the exponential decay is plotted for a theoretical solid biomaterial with only one of the three water bands. The solution for this simplified two component model would be Eq. (19).

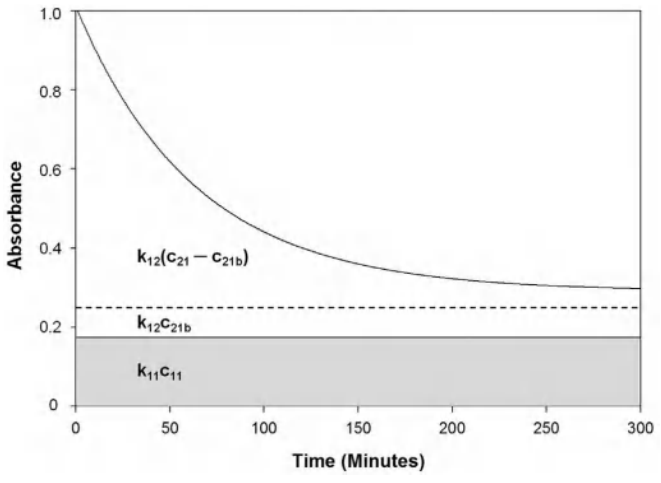

Fig. 2. Exponential decay of the absorbance of a theoretical solid biomaterial with one water band.

As depicted in Fig. 2, the absorbance of the solid biomaterial is $\mathbf{k}_{11}\mathbf{c}_{11}$ which remains constant. Also, in experiments with blank KBr disks, it was observed that some level of the water absorbance remains virtually constant for more than a year. This component is labeled $\mathbf{k}_{12}\mathbf{c}_{21b}$, the absorbance of water with concentration $\mathbf{c}_{21b}$ bound in KBr permanently, for all practical purposes. The decreasing component in Fig. 2 is the difference between the total water and the bound water, $\mathbf{k}_{12}(\mathbf{c}_{21} - \mathbf{c}_{21b})$, the absorbance of water that is not permanently bound to the KBr and is free to be completely desorbed in infinite time.

Expressed as an exponential decay equation, the sum, A_1t, of these three absorbance components at time t is

$$\mathrm{K11\ c11 + K12\ c21}b + \mathrm{K12\ (c21 - c21}b\ )e^{(-\mathrm{a1}\ t)} = \mathrm{A1}t \tag{20}$$

where the new parameter, a_1, is the slope constant of the exponential decay curve.

After this equation is fitted by nonlinear regression to FT-IR spectral data collected from the same KBr disk over time, the relative water concentration ratios for any two points in the time period can be calculated as

$$\frac{c2t1}{c2t2} = \frac{\mathrm{k12\ c21}\ e^{(-\mathrm{a1}\ t1)}}{\mathrm{k12\ c21}\ e^{(-\mathrm{a1}\ t2)}} = \frac{e^{(-\mathrm{a1}\ t1)}}{e^{(-\mathrm{a1}\ t2)}} \tag{21}$$

Thus the unknown water concentration ratio ($\mathbf{c}_{21}/\mathbf{c}_{22}$) needed to solve a simplified version of the R-matrix system (Eq. (6)) with two components (one biomaterial band and one water band) can be simply calculated from the slopes of the exponential decay curve at the times the two spectra were taken, $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$.

The four component model (one biomaterial and three water absorbance bands) that best describes the actual desorption mechanism in KBr disks is a more complex triple exponential decay expansion of Eq. (20) that sums seven absorbance components as

$$\begin{aligned}\mathrm{k11\ c11 + \ k12\ c21}b\ + \mathrm{k13\ c31}b + \mathrm{k14\ c41}b + \mathrm{k12(\ c21 - c21}b)e^{(-\mathrm{a1}t)} \\ + \mathrm{k13(\ c31 - c31}b)e^{(-\mathrm{a2}t)} + \mathrm{k14(\ c41 - c41}b)e^{(-\mathrm{a3}t)} = \mathrm{A_1}t\end{aligned} \tag{22}$$

where a_1, a_2 and a_3 are the slope constants of the exponential decay curves for each of the three types of water components with initial concentrations $\mathbf{c}_{21}$, $\mathbf{c}_{31}$ and $\mathbf{c}_{41}$.

Although not as simply as in Eq. (19), it can be easily shown that the ratio of the three water absorbance bands combined at times $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$,where $((\mathbf{c}_{21+}\mathbf{c}_{31+}\mathbf{c}_{41})/(\mathbf{c}_{22+}\mathbf{c}_{32+}\mathbf{c}_{42}))^*$ is measured from exponential decay data external to the model of Eq. (6), is just the ratio of the sum of the absorbances of the three water desorption curves calculated as

$$\begin{aligned}&\left(\frac{(\mathrm{c21} + \ \mathrm{c31} + \ \mathrm{c41})}{(\mathrm{c22} + \ \mathrm{c32} + \ \mathrm{c42})}\right)^* \\ &= \frac{\mathrm{k12\ c21}\ e^{(-\mathrm{a1}\ t1)} + \mathrm{k13\ c31}\ e^{(-\mathrm{a2}\ t1)} + \mathrm{k14\ c41}\ e^{(-\mathrm{a3}\ t1)}}{\mathrm{k12\ c21}\ e^{(-\mathrm{a1}\ t2)} + \mathrm{k13\ c31}\ e^{(-\mathrm{a2}\ t2)} + \mathrm{k14\ c41}\ e^{(-\mathrm{a3}\ t2)}}\end{aligned} \tag{23}$$

where the absorbance terms $\mathbf{k}_{12}\mathbf{c}_{21}$, $\mathbf{k}_{13}\mathbf{c}_{31}$ and $\mathbf{k}_{14}\mathbf{c}_{41}$ as well as the slope constants a_1, a_2 and a_3 are obtained as fitted parameters in the nonlinear regression of Eq. (22), which is a ten-parameter expansion of Eq. (20). This nonlinear regression is tightly constrained to the conditions and imperatives of the R-matrix model defined above so that of the ten fitted parameters only the three slope constants a_1, a_2 and a_3 are actually unconstrained (free to be any positive values > 0). Eq. (23) can be used to compute $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ directly from Eq. (19) by taking the sum of the water absorbances combined as a single band composed of overlapping contributions from $\mathbf{c}_{21}$, $\mathbf{c}_{31}$ and $\mathbf{c}_{41}$, and employing the simplest reduced model of just two components [11] as shown in Fig. 2.

Eq. (23) is useful to measure the ratio of the three different water absorbances individually, which will be needed in research, for example, to advance the early findings [15] on the nature of the water absorption/desorption mechanism in KBr. This model also provides the ratios of the individual water concentrations ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, $\mathbf{c}_{41}/\mathbf{c}_{42}$ ....) needed to solve the system of R-matrix equations (Eq. (6)), as in Eqs. (16) and (17).

$$\begin{aligned}
\frac{c21}{c22} &= \frac{k12\ c21\ e^{(-a1\ t1)}}{k12\ c21\ e^{(-a1\ t2)}} = \frac{e^{(-a1\ t1)}}{e^{(-a1\ t2)}} \\
\frac{c31}{c32} &= \frac{k13\ c31\ e^{(-a2\ t1)}}{k13\ c31\ e^{(-a2\ t2)}} = \frac{e^{(-a2\ t1)}}{e^{(-a2\ t2)}} \\
\frac{c41}{c42} &= \frac{k14\ c41\ e^{(-a3\ t1)}}{k14\ c41\ e^{(-a3\ t2)}} = \frac{e^{(-a3\ t1)}}{e^{(-a3\ t2)}}
\end{aligned} \tag{24}$$

For this current study and other research made possible by the R-matrix method, the following more complete model is presented.

Combining terms in Eq. (22) (the ten-parameter expansion of Eq. (20)) gives the system of R-matrix - Exponential decay equations

$$\begin{bmatrix} r11 & r12 & r13 & r14 \\ r21 & r22 & r23 & r24 \\ r31 & r32 & r33 & r34 \\ r41 & r42 & r43 & r44 \end{bmatrix} \begin{bmatrix} k11c11 & k11c12 & k11c13 & \ldots & k11c1m \\ k12c21 & k12c22 & k12c23 & \ldots & k12c2m \\ k13c31 & k13c32 & k13c33 & \ldots & k13c3m \\ k14c41 & k14c42 & k14c43 & \ldots & k14c4m \end{bmatrix} = \begin{bmatrix} A11 & A12 & A13 & \ldots & A1m \\ A21 & A22 & A23 & \ldots & A2m \\ A31 & A32 & A33 & \ldots & A3m \\ A41 & A42 & A43 & \ldots & A4m \end{bmatrix}$$

where: $k11c1\,l = k11c11$ for $l = 1, 2, 3, \ldots, m$

$k12c2tl = k12(c21b + c21\ e^{(-A1\ tl)} - c21b\ e^{(-A1\ tl)})$ for $l = 1, 2, 3\ldots, m$

$k13c3tl = k13(c31b + c31\ e^{(-A2\ tl)} - c31b\ e^{(-A2\ tl)})$ for $l = 1, 2, 3\ldots, m$

and $k14c4tl = k14(c41b + c41\ e^{(-A3\ tl)} - c41b\ e^{(-A3\ tl)})$ for $l = 1, 2, 3\ldots, m$

(25)

and where $\mathbf{k}_{12}\mathbf{c}_{21}$, $\mathbf{k}_{13}\mathbf{c}_{31}$, $\mathbf{k}_{14}\mathbf{c}_{41}$ and $\mathbf{k}_{12}\mathbf{c}_{21b}$, $\mathbf{k}_{13}\mathbf{c}_{31b}$, $\mathbf{k}_{14}\mathbf{c}_{41b}$ as well as the slope constants a_1, a_2, a_3 are obtained as fitted parameters in the nonlinear regression of the ten-parameter Eq. (22).

Thus, the total (free + bound) absorbance ratios for the three types of water at times t_1 and t_2 are

$$\begin{aligned}
\frac{c21}{c22} &= \frac{k12(c21b + c21\ e^{(-a1\ t1)} - c21b\ e^{(-a1\ t1)})}{k12(c21b + c21\ e^{(-a1\ t2)} - c21b\ e^{(-a1\ t2)})} \\
\frac{c31}{c32} &= \frac{k13(c31b + c31\ e^{(-a2\ t1)} - c31b\ e^{(-a2\ t1)})}{k13(c31b + c31\ e^{(-a2\ t2)} - c31b\ e^{(-a2\ t2)})} \\
\frac{c41}{c42} &= \frac{k14(c41b + c41\ e^{(-a3\ tl)} - c41b\ e^{(-a3\ tl)})}{k14(c41b + c41\ e^{(-a3\ tl)} - c41b\ e^{(-a3\ tl)})}
\end{aligned} \tag{26}$$

The expanded Eq. (22) with its ten parameters is fitted to the exponential decay FT-IR data measured at 10 or more times (t = 1,2,3,..., q ≥ 10), when there are that many different mixtures of protein and water available in the time-resolved spectra, by least squares minimization as follows:

$$\underset{\text{all a}}{Min} \sum_{t=1}^{q} (a0 + a4\ e^{(-a1\ t)} - a5\ e^{(-a1\ t)} + a6e^{(-a2\ t)} - a7\ e^{(-a2\ t)} + a8e^{(-a3\ t)} - a9e^{(-a3\ t)} - a1t)^2$$

subject to these linear constraints as computed from the R−matrix model:

$a0 = k11\ c11 + k12\ c21b + k13\ c31b + k14\ c41b$
$a4 = k12\ c21$
$a5 = k12\ c21b$
$a6 = k13\ c31$
$a7 = k13\ c31b$
$a8 = k14\ c41$
$a9 = k14\ c41b$
and all $a \geq 0$

(27)

Since it is reasonable to assume that the absorbed water binds almost exclusively with KBr at the extremely minute levels of biomaterial and water in the crystalline KBr matrix, no measurable biomaterial-water interaction is likely in such sparse mixtures. Therefore, this exponential decay model includes no provision for significant biomaterial-water interaction in KBr disks. Also, it should be noted that for purposes of the nonlinear regression other functions of time such as power law functions (t^{-a}) will work equally as well as the exponential function (e^{-at}) used here. However, such functions do not have the kind of useful physical meaning [15] that exponential functions provide to the research.

3. Materials and methods

3.1. Materials

Protein, vital wheat gluten, which consists of two major fractions, glutenin and gliadin, was obtained from Sigma Chemical Co. (St. Louis, MO). KBr was spectral grade potassium bromide, Spectrosol®, purchased from Thermo Fischer Scientific, Inc. (Madison, WI).

3.2. Preparation of biomaterial samples

Protein was vacuum dried for 24 h. at 30 °C before weighing and mixing with KBr. Samples of protein were ground to fine powders by ball-milling (Brinkmann Instruments, Inc., Westbury, NY) in sealed stainless steel vials under liquid nitrogen for several minutes to reach the minimum particle size. Liquid nitrogen (− 196 °C) ball-milling shatters the frozen glass-like biomaterial into micron-sized fragments while preserving the original crystalline nature of the protein and preventing chemical changes or significant disruption of the secondary structure that would alter the infrared spectrum.

Test samples of protein in pressed KBr disks were prepared to obtain infrared spectra of the protein with absorbed water interference.

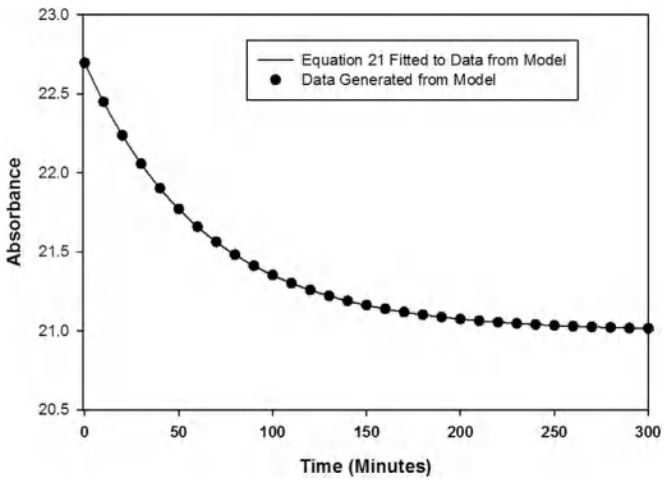

Fig. 3. Model of exponential decay of absorbed water in KBr disk containing solid biomaterial.

Protein was pulverized with KBr to produce water interference with the amide I band of a typical protein. A sample of protein was pulverized in KBr at liquid nitrogen temperature for 2 mins, a period of time that is much longer than the 10–30 s typically used in the KBr disk preparation technique. The longer time of pulverization was necessary to produce a strong water desorption kinetic effect in pressed KBr disks, where infrared bands due to water absorbance decrease exponentially over time. Also, the long pulverization time of protein and KBr at liquid nitrogen temperature insured the pressed KBr disk would be homogeneous and would approach the perfectly transparent solid solution condition required by the Beer-Lambert law. This crucial requirement which was previously semi-empirical and intuitive has now been confirmed on theoretical grounds in a recent and important proof [24]. Blank KBr disks were prepared in the same way but without protein to test the desorption kinetic effect of the absorbed water alone.

3.3. Fourier transform infrared (FT-IR) spectrometry

For analysis by FT-IR spectrometry, a sample of dried protein (2.00 mg) was mixed with KBr (798 mg) and pulverized at liquid nitrogen temperature in a sealed stainless steel vial containing two stainless steel ball pestles for 2 min on a Wig-L-Bug amalgamator (Crescent Dental Manufacturing, Lyons, IL). The vial was allowed to warm to room temperature before 300 mg of the pulverized protein-KBr mixture was transferred to a KBr die (Perkin-Elmer Corp., Norwalk, CT), evacuated 120 s and pressed for 60 s under vacuum at 110 MPa on a laboratory press (Fred S. Carver, Menominee Falls, WI).

FT-IR spectra were obtained on a Bomem Arid Zone FTIR spectrometer (ABB MB-Series, Houston, TX) equipped with a DTGS detector. Time-resolved spectra were acquired in 5–10 min intervals at 4 cm^{-1} resolution and signal-averaged over 16 scans. Interferograms were Fourier transformed using triangular apodization. All spectra were baseline corrected and truncated to display only those infrared absorbance bands that need correction for water interference by the method proposed in this paper. All such spectral manipulations were performed with routines provided in GRAMS AI software (Thermo Galactic, Inc., Salem, NH). FT-IR spectral absorbance of water measured in KBr over time is modeled in Fig. 3 to graphically illustrate the strong water interference that forms in pressed KBr disks and how it decreases exponentially with time.

3.4. Attenuated total reflection (ATR) spectrometry

An ATR spectrum of a sample of the same dried protein was obtained on a Thermo Scientific (Thermo Electron Corp., Madison, WI) NEXUS™ 470 FT-IR spectrometer using the Smart Orbit™ diamond ATR accessory. Interferograms collected at 4 cm^{-1} resolution were signal-averaged over 256 scans. The spectrum was smoothed via Savitzky-Golay function and then baseline corrected. The smoothed spectrum was further processed using the Advanced ATR Correction Algorithm provided in the spectrometer software (OMNIC™ Version 6.2) to correct for the wavelength dependence of penetration depth and refractive index dispersion and to compare with the spectrum corrected for water interference.

3.5. Chemometric calculations

Multivariate quantitative analyses via an algorithm written by the author (S.H. Gordon) containing Eqs. (7) through (18) were conducted using the matrix algebra functions (add-ins) of Microsoft Excel®. However, all of these essential matrix manipulations can be computed much more efficiently using MATLAB® (The MathWorks, Inc., Natick, MA). The software contains functions ideally suited for R-matrix computations from the Beer-Lambert law models and for analyses of protein + absorbed water spectra as four-component mixtures at four frequencies.

3.6. Exponential decay calculations

Using a nonlinear regression algorithm implemented as a user defined function in SigmaPlot® 10.0 (Systat Software, Inc., San Jose, CA), Eq. (22) (the ten-parameter expansion of Eq. (20)) was fitted to data points collected in the time-resolved FT-IR spectra of the protein + absorbed water mixture absorbance over 2 h time. All absorbance measurements (A_1t) in the time-resolved series of spectra were taken at the same frequency, which in this work was the position of the hydroxyl peak at 3340 cm^{-1}. The algorithm was written as a weighted least squares problem of minimizing the sum of squares of the residuals to find the best-fit parameters, a, under the statistical assumption that the weight factors are equal to the reciprocals of the variances of Gaussian distributions of the measurements. Although it was not demonstrated experimentally in this study it was proved mathematically that the ten-parameter exponential decay model (Eq. (22)) can be reduced to only four parameters and solved by nonlinear regression fitted over as few as four time-resolved data points. In this work, however, the algorithm returned the water absorbance values, $\mathbf{k}_{12}\mathbf{c}_{21}$, $\mathbf{k}_{13}\mathbf{c}_{31}$, $\mathbf{k}_{14}\mathbf{c}_{41}$ and $\mathbf{k}_{12}\mathbf{c}_{21b}$, $\mathbf{k}_{13}\mathbf{c}_{31b}$, $\mathbf{k}_{14}\mathbf{c}_{41b}$ as well as a_0 and the slope constants a_1, a_2, a_3 as ten fitted parameters in Eq. (27).

Thus, the resulting calculations from Eq. (22) via Eq. (27) provide estimates of the water concentration ratios, $\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, and $\mathbf{c}_{41}/\mathbf{c}_{42}$, needed to compute the desired band absorbance ratios, $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$, in the protein.

3.7. Correction of biomaterial spectrum for water interference

To correct the spectrum, the software program uses input of absorbances, $\mathbf{A}_{11}$, $\mathbf{A}_{12}$, $\mathbf{A}_{21}$ and $\mathbf{A}_{22}$, at two selected frequencies (in this case for protein, the hydroxyl peak at 3340 cm^{-1} and the amide I peak at 1640 cm^{-1}) from FT-IR spectra of the protein-KBr mixtures measured at two different times (say, 10 min and 60 min, for example). The program also uses input of the known protein weights, $\mathbf{c}_{11} = \mathbf{c}_{12}$, in the KBr mixtures. With this input and the total water concentration ratios $(\mathbf{c}_{21}+\mathbf{c}_{31}+\mathbf{c}_{41})/(\mathbf{c}_{22}+\mathbf{c}_{32}+\mathbf{c}_{42})$ and the absorbance ratios, $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{43}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{34}$ and $\boldsymbol{r}_{44}$ from Eq. (20) and the individual water concentration ratios ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, and $\mathbf{c}_{41}/\mathbf{c}_{42}$) calculated from exponential decay regression, Eq. (22), the program computes the band absorbance ratios, ($\boldsymbol{r}_{21}$, $\boldsymbol{r}_{22}$), ($\boldsymbol{r}_{31}$, $\boldsymbol{r}_{32}$) and ($\boldsymbol{r}_{41}$, $\boldsymbol{r}_{42}$) needed to construct the R-matrix and hence solve the system (Eq. (6)) constraining the solution to comply with the Beer-Lambert law. Thus, in the two mixtures at the two frequencies, the absorbances due to protein, $\mathbf{k}_{11}\mathbf{c}_{11}$, $\mathbf{k}_{21}\mathbf{c}_{11}$, $\mathbf{k}_{31}\mathbf{c}_{11}$, $\mathbf{k}_{41}\mathbf{c}_{11}$, $\mathbf{k}_{11}\mathbf{c}_{12}$, $\mathbf{k}_{21}\mathbf{c}_{12}$, $\mathbf{k}_{31}\mathbf{c}_{12}$, $\mathbf{k}_{41}\mathbf{c}_{12}$ and the absorbances due to water, $\mathbf{k}_{12}\mathbf{c}_{21}$, $\mathbf{k}_{22}\mathbf{c}_{21}$,, $\mathbf{k}_{32}\mathbf{c}_{21}$,, $\mathbf{k}_{42}\mathbf{c}_{21}$, $\mathbf{k}_{12}\mathbf{c}_{22}$, $\mathbf{k}_{22}\mathbf{c}_{22}$, $\mathbf{k}_{32}\mathbf{c}_{22}$, $\mathbf{k}_{42}\mathbf{c}_{22}$, as well as the individual absorption coefficients for protein, $\mathbf{k}_{11}$, $\mathbf{k}_{21}$, $\mathbf{k}_{31}$ and $\mathbf{k}_{41}$, are all precisely determined. Only the absorption coefficients for water, $\mathbf{k}_{12}$ and $\mathbf{k}_{22}$, $\mathbf{k}_{32}$ and $\mathbf{k}_{42}$ at the two frequencies remain undetermined since the absolute weights of water, ($\mathbf{c}_{21}$,$\mathbf{c}_{22}$), ($\mathbf{c}_{31}$, $\mathbf{c}_{32}$), and ($\mathbf{c}_{41}$, $\mathbf{c}_{42}$), are unknown.

Because there is always a certain amount of error in experimental data, the two frequency solution above will not fit the Beer-Lambert law perfectly at the other frequencies that are also part of the model. Therefore, any correction algorithm must include all of the other frequencies. This is accomplished for this model by finding the best solution to Eq. (22) which satisfies the Beer-Lambert law at all frequencies in the model simultaneously. A search for the best solution to Eq. (22) requires the results of Eq. (18) (the 3 protein absorbance ratios $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ and the nine water absorbance ratios, $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{34}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{43}$, and $\boldsymbol{r}_{44}$) to propagate as approximations to the iterative solution of the equations as a nonlinear optimization problem constrained in such a way that Beer-Lambert equations (Eq. (25)) are always satisfied. For the exponential decay work, the models of Eqs. (22) through (26) were computed using algorithms written by the author (S.H. Gordon) in SigmaPlot® 10.0. A Levenberg-Marquardt [27,28] algorithm, also implemented in SigmaPlot®, was applied to

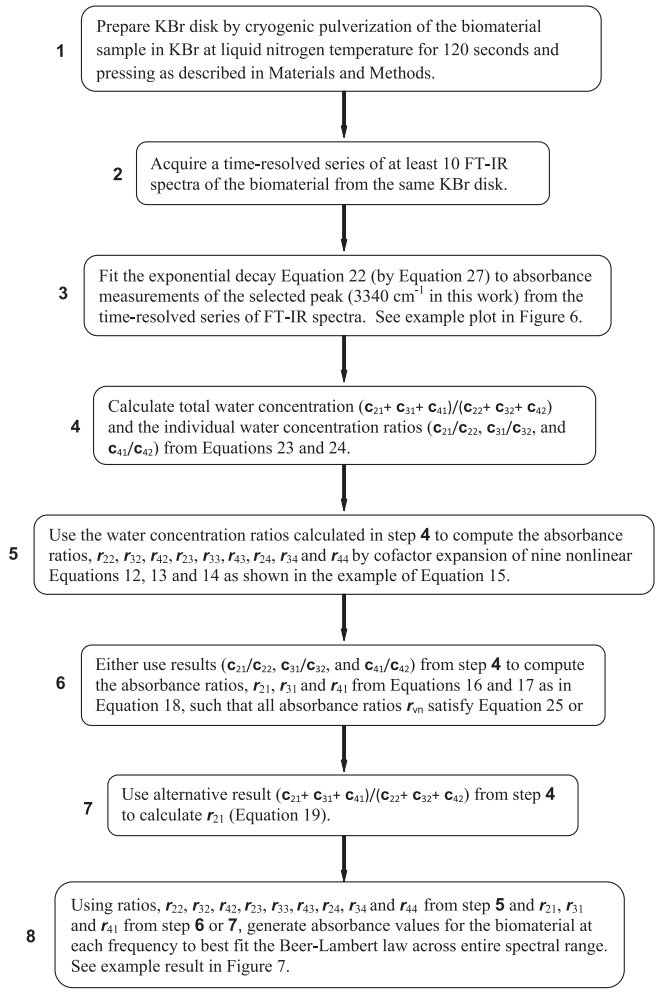

Fig. 4. Flowchart summarizing the eight main steps of the KBr water correction method.

solve the constrained nonlinear optimization problem of Eq. (27). However, other optimization methods, ranging from brute force, random, stochastic algorithms to gradient search, quasi-Newton algorithms [29–31], that are commercially available such as the function (fmincon) in MATLAB® (The MathWorks, Inc., Natick, MA), will work as well. Also, the formal analytic solution to systems like Eqs. (16)–(18), (22) and (27) can sometimes be derived algebraically. For example, it can be proved algebraically that equations exist to solve the constrained nonlinear optimization problem of Eq. (27) efficiently and robustly in as few as four fitted parameters. Potentially, this problem can be solved by measuring only four time-resolved FT-IR spectra on a single KBr disk. By the same exponential decay method, the simplest alternative model in the problem of Eq. (19), can be solved from just two time-resolved FT-IR spectra of a single KBr disk.

The exponential decay-R-matrix chemometric algorithm described above is summarized in a flowchart (Fig. 4) which outlines the eight main steps in the KBr water correction method.

This exponential decay-R-matrix chemometric algorithm can be applied to the infrared absorbance bands from which only elimination of water interference is needed, and it also can be applied across the entire spectrum to generate absorbances that best fit the Beer-Lambert law at all points. Using the hydroxyl peak as a pivot in the algorithm for automatic generation of the corrected spectrum by iteration over all frequencies, the same correction is made in all regions of the spectrum that contain water interference.

The program returns as output the computed absorbance values for protein and water at each frequency in each spectrum. From these absorbance values, each selected spectrum in the time-resolved series is corrected to eliminate the water interference in both the water OH stretching region and the water OH bending region, where the hydroxyl and amide I bands of biomaterials appear.

The procedure described above generates a solution that fits the Beer-Lambert law perfectly at all selected frequencies. Once the water ratios ($\mathbf{c}_{21}$, $\mathbf{c}_{22}$), ($\mathbf{c}_{31}$, $\mathbf{c}_{32}$), and ($\mathbf{c}_{41}$, $\mathbf{c}_{42}$) have been determined, all of the band absorbance ratios ($\boldsymbol{r}_{21}$ and $\boldsymbol{r}_{22}$), ($\boldsymbol{r}_{31}$ and $\boldsymbol{r}_{32}$) or ($\boldsymbol{r}_{41}$ and $\boldsymbol{r}_{42}$) for any two given frequencies can then be computed explicitly, as separated variables, from the Beer-Lambert law (Eq. (6)). Thus, by measuring the exponential decay of a single band (in this case the OH stretching band) and using this band as a pivot for applying the chemometric equations derived in this work, it is possible to obtain the infrared spectrum of the protein in the mixture, corrected for water interference. The resulting band profiles (band absorbance ratios) in this corrected spectrum will be quantitative and exact at all frequencies, which is crucial for subsequent analytical purposes, such as computation of secondary structures.

4. Results and discussion

A new chemometric model based on the Beer-Lambert law for correcting infrared spectra of biomaterials for water interference in pressed KBr disks was derived. The equations that were derived (Eq. (6) or its R-matrix equivalent, Eq. (9)) were proved to be mathematically indeterminate for the needed water absorbance ratios $\boldsymbol{r}_{21}$ and $\boldsymbol{r}_{22}$ unless an independent measure of the water ratio $\mathbf{c}_{21}/\mathbf{c}_{22}$ or its absorbance equivalent is obtained by some means external to and independent of the Beer-Lambert law model. Therefore, a search for an independent measure of $\mathbf{c}_{21}/\mathbf{c}_{22}$ was conducted. In the previous work [11], the independent measure was a constrained differential spectral curve-fitting technique that gave an empirical estimate of the needed water concentration ratio from infrared spectra of the biomaterial in two pressed KBr disks containing different levels of absorbed water.

In that work [11], although the R-matrix model was not expanded and generalized, the method used was similar to the method presented here, insofar as the independent measure used in both works was constrained by the Beer-Lambert conditions of Eq. (6). In addition, as in earlier studies [15,32–34], both works confirm and employ the three OH stretching bands first resolved by Malhotra et al. [15]. However, in none of the previous or earlier work was exponential decay kinetic modeling attempted to measure the water concentrations ratios ($\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, and $\mathbf{c}_{41}/\mathbf{c}_{42}$) as was done here. The challenge of measuring $\mathbf{c}_{21}/\mathbf{c}_{22}$ was accomplished using constrained differential spectral curve-fitting which is a powerful and straightforward, albeit somewhat arbitrary, analytical technique.

When that technique was published [11], it was believed to be the only way to determine the water concentration ratio $\mathbf{c}_{21}/\mathbf{c}_{22}$ needed to compute the absorbance ratios ($\boldsymbol{r}_{21}$ and $\boldsymbol{r}_{22}$) and correct the protein spectrum. Other experimental techniques being investigated for determining the $\mathbf{c}_{21}/\mathbf{c}_{22}$ ratio were thought to be unsuccessful because the water absorbed by pulverized KBr is not constant over time. However, as it turns out, this variation with time is exponential and canonical in form which is the very thing needed to obtain simultaneous data for determination of $\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, and $\mathbf{c}_{41}/\mathbf{c}_{42}$ while collecting time-resolved FT-IR spectra. It directly measures the infrared absorbance of the infrared active water as it decreases exponentially in KBr with time. This is particularly fortuitous since only that infrared-active portion of the water that produces the measured infrared absorbance can properly be analytical [11] for calculation of unambiguous $\mathbf{c}_{21}/\mathbf{c}_{22}$, $\mathbf{c}_{31}/\mathbf{c}_{32}$, and $\mathbf{c}_{41}/\mathbf{c}_{42}$ ratios.

4.1. Computer simulation of spectrum with water interference

Simulated FT-IR spectra of a solid biomaterial and its mixtures with water based on the Beer-Lambert law were generated by a computer to test and prove the chemometric R-matrix – exponential decay model

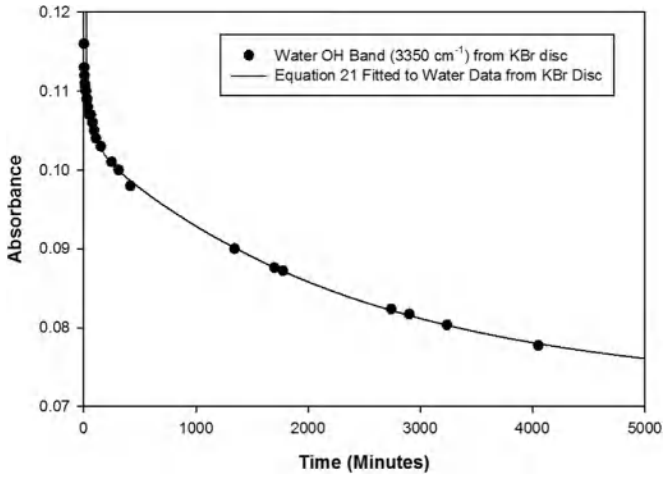

Fig. 5. Experimental exponential decay of water in blank KBr disk.

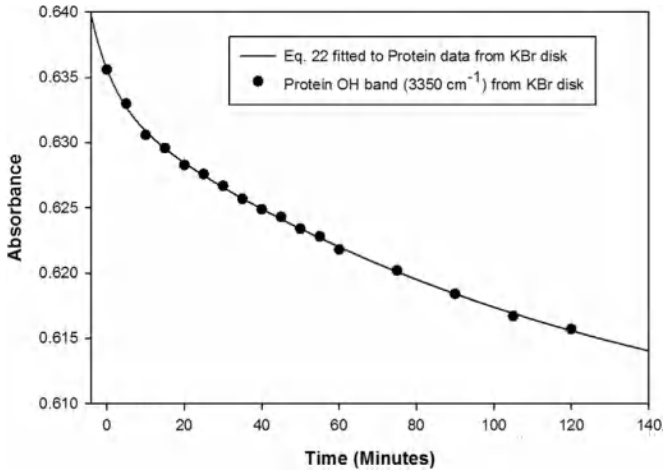

Fig. 6. Typical exponential decay of absorbed water in KBr disk containing solid protein.

and theory presented in this paper. Infrared absorbances were computed for simulated mixtures of a solid protein and three types of water assuming known absorption coefficients and concentrations of the four components.

To test the validity and accuracy of the model, exact model data was first generated by computer so that no error was present, and then the solution algorithms were applied to determine whether they correctly gave the solution known to be exact for the given model data. This was done to test both the R-matrix and the exponential decay algorithms in this work. The results showed that the exponential decay model was a perfect fit to the data (Fig. 3), and solution algorithms reproduced the generated model data precisely. It was shown that the solution algorithms exactly reproduced all of the computer-generated data (absorption coefficients and concentrations) exactly, within round-off error of the computer arithmetic. Thus, this computer simulation, together with the formal analytical solution already derived in the Theory and Derivation of the Model section, proved the model and algorithms are theoretically valid and completely accurate for spectra that conform to the Beer-Lambert law.

Therefore, it is clear that the R-matrix and the exponential decay algorithms conceived in this work represent a new and unobvious solution to a long-standing chemometric problem in the spectrometry of solid biomaterials. Consequently, because of the R-matrix - exponential decay algorithms, it is now possible to extract unknown absorption coefficients and concentrations of the multiple components from a single sample without calibration against any external samples. This method, which is the result that was the objective of the present research, is summarized in the exponential decay R-matrix chemometric algorithm as outlined in the flowchart (Fig. 4).

Unlike exact computer generated data, all laboratory research data contain experimental error. Because of this error and because infrared spectra of solid biomaterials are notorious for deviation from Beer's law [35], the practical application of the R-matrix – exponential decay algorithms for research on real-world data was also tested in this work. The performance of the R-matrix - exponential decay model and solution algorithms with data from a solid protein in KBr disks prepared in the laboratory was tested using protein. Fig. 5 shows the typical exponential decay of absorbed water in blank KBr disks. Fig. 6 shows the typical exponential decay of absorbed water in KBr disks containing protein. In both Figs. 5 and 6, the exponential decay model (Eq. (22)) was fitted to the time-resolved FT-IR data. In both cases, the correlation coefficients were consistently high ($R^2 > 0.99$).

One reason for this high accuracy is the validity of Eq. (22) due to the flexibility it brings with three different exponential terms. Another more important reason for the high accuracy is the excellent quality of the KBr disks prepared due to the cryogenic sample preparation technique. Protein was reduced to a fine powder with KBr in the liquid nitrogen pulverization step. Hence, the pressed KBr disk approached a solid solution of the protein in KBr, having particle sizes in the order of the shortest infrared wavelengths, which gave essentially linear infrared absorbance that met the requirements [24] of the Beer-Lambert law.

Spectroscopists may wonder why such a complex method is required when simple digital subtraction of a known (reference) water spectrum is possible. The infrared spectrometry rationale for using this method instead of digital subtraction is fully explained in the previous paper [11]. Suffice it to say here, simple spectral subtraction is not applicable because one cannot know how much water absorbance to subtract at any frequency, i.e., since neither the intensity nor the profile of the water band in KBr is known, a method considerably more complex than digital subtraction is needed to eliminate the water absorbance. This definitely necessitates and validates the exponential decay kinetic method used in this work. Part of the beauty and elegance of this chemometric solution is the fact that it enables and uses direct measurement of the actual unknown water absorbance in KBr in situ and subtracts precisely that correct water absorbance measurement at each and every frequency. Thus, this solution achieves the absolute certainty of eliminating not just a computed estimate of the interference but the actual interfering water absorbance itself, for the first time.

Chemometricians may also wonder why the exponential decay kinetic method was not simply used by itself, without the Beer-Lambert law R-matrix equations, to determine absorbances needed to correct the spectra. The reason for this is that although exponential decay kinetic model is pivotal, only the R-matrix chemometric equations derived in this work can yield the three critical band absorbance ratios, $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ and the nine water absorbance ratios, $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{34}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{43}$, and $\boldsymbol{r}_{44}$, that are crucial to the success of this method. Exponential decay kinetics only yields the water absorbances from nonlinear regression of a single infrared spectral band. This provides no information whatsoever about the relative absorbances of the many different bands across the entire spectrum. Without the R-matrix chemometric solution derived in this work, simple exponential decay kinetics of any single absorption band cannot possibly yield the entire infrared absorbance spectrum of the biomaterial in total compliance with the Beer-Lambert law.

Fortunately, the R-matrix-exponential decay kinetic model works because the OH stretching band in water is very strong, reproducible and well known to consist of only a few underlying bands. Unlike the constrained differential spectral curve-fitting technique, which is not an exact science and requires some arbitrary selections and decisions about underlying bands, the exponential decay kinetics regression

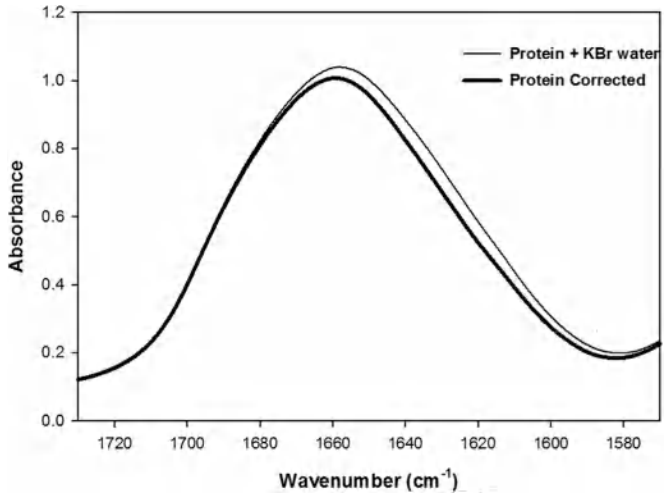

Fig. 7. FTIR spectra of protein in KBr before (gray) and after (black) water interference correction.

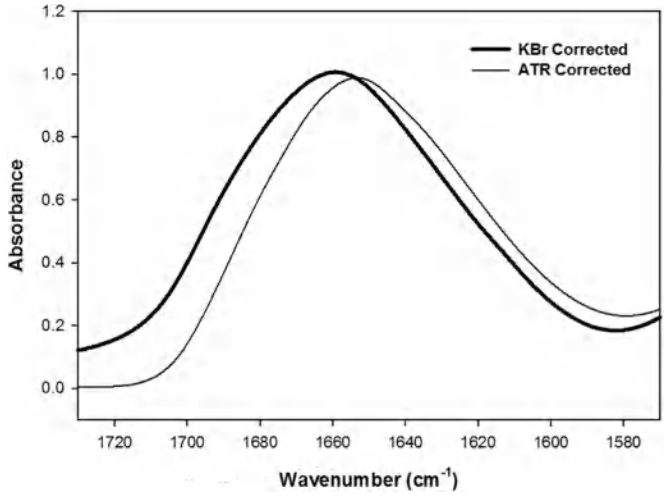

Fig. 8. Amide region of protein in KBr (black) after correction for water interference compared with amide region of protein by ATR (gray) after refractive index correction.

technique is exact, reproducible and reliable as a measure of changes in the single OH stretching band over time. Spectral curve fitting requires theoretical assumptions of symmetric Gaussian/Lorentzian absorbance bands which, although they might closely approximate real absorbance bands, certainly contain some error. Spectral deconvolution curve fitting also assumes the inevitable differences in absorptivities between adjacent bands underlying the fitted curve are small and can therefore be ignored. This assumption is controversial. Unlike spectral curve fitting, exponential decay kinetics regression requires no such theoretical assumptions about band shapes or absorptivities and involves only direct experimental measurements of true profiles of actual absorbance band changes occurring in real time. Also, valid and reliable constraints on the nonlinear regression process can be applied to minimize experimental error and eliminate much uncertainty.

By normalizing the solution to the strongest and most reliable band in the water spectrum (the OH stretching band), as was done here, the R-matrix solution for the corrected spectrum is guaranteed to have the highest probability of accuracy. Determination of the unknown band absorbance ratios, $\boldsymbol{r}_{21}$, $\boldsymbol{r}_{31}$, and $\boldsymbol{r}_{41}$ and the water absorbance ratios, $\boldsymbol{r}_{22}$, $\boldsymbol{r}_{23}$, $\boldsymbol{r}_{24}$, $\boldsymbol{r}_{32}$, $\boldsymbol{r}_{33}$, $\boldsymbol{r}_{34}$, $\boldsymbol{r}_{42}$, $\boldsymbol{r}_{43}$, and $\boldsymbol{r}_{44}$, is central to this method and the most important result. These ratios enable generation of multiple corrected band profiles across the entire spectrum including the amide regions in proteins. These crucial band absorbance ratios could not have been determined without the R-matrix-exponential decay kinetic model presented in this paper.

To demonstrate the utility of the R-matrix-exponential decay kinetic model for biomaterials in the laboratory, the FT-IR spectrum of a solid protein in KBr was corrected by this method. With the $\mathbf{c}_{21}/\mathbf{c}_{22}$ ratio determined by R-matrix-exponential decay of the OH stretching band, the corrected spectrum was computed using the resulting $\boldsymbol{r}_{21}$ and $\boldsymbol{r}_{22}$ band absorbance ratios as constraints. Fig. 7 shows the FT-IR spectrum of dried protein before and after the correction. It shows that the OH stretching (centered at 3340 cm^{-1}) in the corrected spectrum of protein is significantly reduced and peak-shifted due to elimination of the interference from overlapping water OH bands. Also, the corrected spectrum of protein reveals smaller but very significant changes in the profile of the critical amide I band (1660 cm^{-1}). This is the true amide I band profile of protein required for subsequent quantitation of the protein secondary structures [36–39] by spectral curve-fitting.

To demonstrate the full advantage and need for this R-matrix-exponential decay correction method, the corrected KBr spectrum of protein was compared with the corrected ATR spectrum in the Amide I and II regions used in quantitating secondary structures of the protein in the solid state by spectral curve-fitting. The ATR spectrum (Fig. 8) was first corrected using the conventional and well-accepted Advanced ATR Correction Algorithm introduced by Nunn and Nishikida [40] to correct for the notorious band intensity distortion caused by inherent frequency dependence of the depth of penetration and to correct for the concomitant shifts to lower frequencies caused by refractive index dispersion in ATR spectrometry.

In Fig. 8, the KBr spectrum corrected by the R-matrix-exponential decay algorithm is compared with the ATR spectrum corrected by the Advanced ATR Correction Algorithm. The corrected spectra show large differences in the profiles of the important amide bands that will produce large differences in the spectral curve-fitting results in subsequent analyses of the protein secondary structure. It has been demonstrated that virtually all of this difference consists of errors contained in the corrected ATR spectrum, not the corrected KBr spectrum, and that the corrected KBr spectrum of protein does not contain significant error [11,39,41–44] or reflect the serious drawbacks inherent to ATR spectrometry [45]. Therefore, compared to the corrected ATR spectra, the corrected KBr spectra of the present work are expected to be much closer to the true infrared spectra of proteins and other solid biomaterials. Furthermore, when applied to correlations between x-ray analyses and spectral curve-fitting, the R-matrix-exponential decay method will provide a much more accurate determination of secondary and tertiary structures in solid proteins, polysaccharides and other biomaterials than was previously possible.

5. Conclusions

A new chemometric method was devised to correct FT-IR spectra of solid biomaterials for interference from water absorbed in the KBr disk preparation. The correction method uses a novel R-matrix algorithm and exponential decay desorption kinetics to free spectra of biomaterials from water bands in the important amide and hydroxyl regions. The method provides more accurate quantitative infrared spectra of components in solid biomaterials in KBr disks than was previously possible. Its treatment of band absorbances as ratios in a matrix (R-matrix) not only renders a multivariate system of Beer-Lambert law equations amenable to solution, but also affords the unknown absorption coefficients of the solid component from a single KBr disk sample that exhibits exponentially decreasing levels of absorbed water over time. The water absorbed by or contained in KBr, which has frustrated chemists and defeated infrared spectroscopists for more than 70 years, no longer presents a problem. Now this problem is finally solved once and for all. Using this current formal and uncontroversial solution, efforts to eliminate water interference from KBr disks in research will be defeated no longer. Infrared spectral profiles corrected by this method are accurate and reproducible. Moreover, the new chemometric method gives the pressed KBr disk technique a decided

advantage over ATR for reliable quantitative secondary and tertiary structure analyses.

The R-matrix chemometric algorithm solved with kinetic modeling as presented here marks an important step toward valid spectrometric analysis of many solid biomaterials in nature. Consequently, the long-sought advancements in research methods for precise determination of secondary and tertiary structures of solid proteins and polysaccharides are now feasible. This breakthrough removes a major barrier to reliable quantitative spectrometric analyses of solid biomaterials and will enable research to obtain knowledge crucial to progress in many areas of science.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

The authors gratefully acknowledge the smart technical assistance and helpful editorial advice of Jason Adkins, Amelia Ruth Griffin, Mark Klokkenga and Kathleen Payne-Wahl.

Helpful advice from the preeminent FT-IR spectroscopist and world-renowned emeritus professor, Peter R. Griffiths, who described this work as "important" is greatly appreciated.

References

[1] E.S. Cook, C.W. Kreke, E.B. Barnes, W. Motzel, Infra-red and ultra-violet absorption spectra of proteins in the solid state, Nature 174 (1954) 1144–1145.
[2] M.M. Stimson, M.J. O'Donnell, The infrared and ultraviolet absorption spectra of cytosine and isocytosine in the solid state, J. Am. Chem. Soc. 74 (1952) 1805–1808.
[3] D.N. Ingebrigtson, A.L. Smith, Infrared analysis of solids by potassium bromide pellet technique, Anal. Chem. 26 (1954) 1765–1768.
[4] R.G. Milkey, Potassium bromide method of infrared sampling, Anal. Chem. 30 (1958) 1931–1933.
[5] R.S. Browning, S.E. Wiberley, F.C. Nachod, Application of infrared spectrophotometry to quantitative analysis in solid phase, Anal. Chem. 2 (1955) 7–11.
[6] H. Rosenkrantz, The infrared analysis of vitamins, hormones, and coenzymes, Methods Biochem. Anal. 5 (1957) 407–416.
[7] O. Yavuz Ataman, H.B. Mark Jr., Alkali halide pelleting technique for solid sampling in infrared spectroscopy, Appl. Spectrosc. Rev. 13 (1977) 1–13.
[8] V. Viswanath, B. Narasimharao, M. Purushothaman, G. Sowjanya, S. Sreevani, Formulation and evaluation of fluconazole ocular inserts, World J. Pharm. Res. 4 (2015) 828–841.
[9] F.J. Stevenson, K.M. Goh, Infrared spectra of humic acids: elimination of interference due to hygroscopic moisture and structural changes accompanying heating with KBr, Soil Sci. 117 (1974) 34.
[10] S.H. Gordon, S.H. Imam, C. James, A method for measuring degradation of individual components in multicomponent biodegradable plastics by Fourier transform infrared spectrometry, J. Polym. Environ. 8 (2000) 125–134.
[11] S.H. Gordon, A.A. Mohamed, R.E. Harry-Okuru, S.H. Imam, A chemometric method for correcting FTIR spectra of biomaterials for interference from water in KBr discs, Appl. Spectrosc. 64 (2010) 448–457.
[12] F.J. Stevenson, K.M. Goh, Infrared spectra of humic acids: elimination of interference due to hygroscopic moisture and structural changes accompanying heating with KBr, Soil Sci. 117 (1974) 34.
[13] K. Schwochau, H.J. Eichhoff, Untersuchungen zur trübungserscheinung von kaliumbromid-preßlingen für die ir-spektroskopie, Z. Nat. 17b (1962) 582.
[14] L. Borka, Hygroscopic properties of potassium bromide in infrared spectrophotometry, Anal. Chem. 47 (1975) 1212.
[15] V.M. Malhotra, S. Jasty, R. Mu, FT-IR spectra of water in microporous KBr pellets and water's desorption kinetics, Appl. Spectrosc. 43 (1989) 638.
[16] W.C. Price, W.F. Sherman, G.R. Wilkinson, Infra-red studies of water adsorbed on alkali halides, Proc. R. Soc. A 247 (1958) 467.
[17] W.K. Thompson, An infrared absorption study of the state of adsorbed water molecules and the effect of sodium ions on potassium bromide pressed discs, and on dipotassium hydrogen phosphate dispersed in potassium bromide pressed discs, J. Chem. Soc. (1964) 3658.
[18] H.V. Walter, Adsorption von wasser an pulvern von alkalihalogenidkristallen vom *NaCl*-typ, Z. Phys. Chem. 75 (1971) 287.
[19] W.P. Norris, A.L. Olsen, R.G. Brophy, Clouding of pressed potassium bromide powder, Appl. Spectrosc. 26 (1972) 247.
[20] S.H. Gordon, Chemometric analysis of multicomponent biodegradable plastics by Fourier transform infrared spectrometry: the Rmatrix method, in: J. Ahmed, B.K. Tiwari, S.H. Imam, M.A. Rao, (Eds.), Starch-Based Polymeric Materials and Nanocomposites, 2012, pp. 271–285.
[21] D.F. Swinehart, The Beer-Lambert law, J. Chem. Educ. 39 (1962) 333.
[22] R.L. Pecsok, D.L. Shields, Modern Methods of Chemical Analysis, John Wiley & Sons, New York, 1968.
[23] B. C. Smith, Quantitative Spectroscopy: Theory and Practice. Chapters 1–3, 2002.
[24] G.M. Thomas, H. Mutschke, J. Popp, Employing theories far beyond their limits – the case of the (Boguer-) Beer–Lambert law, Chem. Phys. Chem. 17 (2016) 1–9.
[25] G.L. McClure, Computerized Quantitative Infrared analysis, ASTM, Philadelphia, 1987.
[26] H. Anton, Elementary Linear Algebra, John Wiley & Sons, New York, 1987.
[27] D.W. Marquardt, An algorithm for least squares estimation of parameters, J. Soc. Ind. Appl. Math. 11 (1963) 431–441.
[28] M.K. Transtrum, J.P. Sethna, Improvements to the Levenberg-Marquardt algorithm for nonlinear least-squares minimization, J. Comput. Phys. (2012) (Preprint submitted).
[29] J.C. Spall, Introduction to Stochastic Search and Optimization: Estimation, Simulation and Control, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley, 2003.
[30] J.F. Bonnans, J. Ch. Gilbert, C. Lemarèchal, C.A. Sagastizábal, Numerical Optimization, Theoretical and Numerical Aspects, Second ed., Springer, 2006.
[31] D.M. Himmelblau, Applied Nonlinear Programming, McGraw-Hill Companies, 1972.
[32] R.D.B. Fraser, E. Suzuki, Secondary structure of a channel-forming protein: porin from E. coli outer membranes, in: S.J. Leach (Ed.)Physical Principles and Techniques of Protein Chemistry, Academic Press, New York, 1973, pp. 301–355 (Part C).
[33] M. Falk, T.A. Ford, Infrared spectrum and structure of liquid water, Can. J. Chem. 44 (1966) 1699.
[34] R. Mu, M.S. Thesis, Department of Physics, SIU-Carbondale, Illinois, 1987, pp. 1–129.
[35] J.L. Koenig, Vibrational spectroscopy of polymers, Spectroscopy of Polymers (Chapter 2), American Chemical Society, Washington, 1992.
[36] J.D. Meyer, M.C. Manning, J.F. Carpenter, Effects of KBr disk formation on the infrared spectra of model proteins, J. Pharm. Sci. 93 (2004) 496–506.
[37] D.M. Byler, H. Susi, Examination of the secondary structure of proteins by deconvoluted FTIR spectra, Biopolymers 25 (1986) 469–487.
[38] M. Jackson, H.H. Mantsch, The use and misuse of FTIR spectroscopy in the determination of protein structure, Crit. Rev. Biochem. Mol. Biol. 30 (1995) 95–120.
[39] M.C. Manning, Use of infrared spectroscopy to monitor protein structure and stability, Expert Rev. Proteom. 2 (2005) 731–743.
[40] S. Nunn, K. Nishikida, Advanced ATR correction algorithm; Thermo Fisher Scientific Application AN50581_E 05/08M, Thermo Fisher Scientific Instruments LLC, Madison, WI, 2008. ⟨https://tools.thermofisher.com/content/sfs/brochures/D10241~.pdf⟩, (Accessed on June 12, 2017).
[41] M. van de Weert, P.I. Haris, W.E. Hennink, D.J. Crommelin, Fourier transform infrared spectrometric analysis of protein conformation: effect of sampling method and stress factorsFourier transform infrared spectrometric analysis of protein conformation effect of sampling method and stress factors, Anal. Biochem. 297 (2001) 160.
[42] B.S. Kendrick, A. Dong, S.D. Allison, M.C. Manning, J.F. Carpenter, Quantitation of the area of overlap between second-derivative amide I infrared spectra to determine the structural similarity of a protein in different states, J. Pharm. Sci. 85 (1996) 155.
[43] H.R. Costantino, J.D. Andya, S.J. Shire, C.C. Hsu, Fourier-transform infrared spectroscopic analysis of the secondary structure of recombinant humanized immunoglobulin G, Pharm. Sci. 3 (1997) 121.
[44] S. Sinha, Ph.D. Dissertation, Dept. of Pharmaceutical Chemistry, The University of Kansas, 2008, pp. 12–14.
[45] Shimadzu Corporation, ATR Precautions, FTIR: Fundamentals of FTIR Analysis, Part2, 2017. ⟨http://www.shimadzu.com/an/ftir/support/ftirtalk/letter2/atr2.html⟩. (Accessed on May 7, 2017).