Detection of Albumin Unfolding Preceding Proteolysis Using Fourier Transform Infrared Spectroscopy and Chemometric Data Analysis

ARTICLE *in* ANALYTICAL CHEMISTRY · JUNE 2006

Impact Factor: 5.64 · DOI: 10.1021/ac0520137 · Source: PubMed

4 AUTHORS:

Ana Dominguez-Vidal
Universidad de Jaén
43 PUBLICATIONS **597** CITATIONS
SEE PROFILE

María-Pilar Sáenz-Navajas
University of Zaragoza
29 PUBLICATIONS **349** CITATIONS
SEE PROFILE

Maria Jose Ayora Cañada
Universidad de Jaén
59 PUBLICATIONS **1,226** CITATIONS
SEE PROFILE

Bernhard Lendl
TU Wien
270 PUBLICATIONS **4,317** CITATIONS
SEE PROFILE

Anal. Chem. **2006,** *78,* 3257−3264

Detection of Albumin Unfolding Preceding Proteolysis Using Fourier Transform Infrared Spectroscopy and Chemometric Data Analysis

Ana Domínguez-Vidal,† María P. Saenz-Navajas,† María José Ayora-Cañada,‡ and Bernhard Lendl*,†

Institute of Chemical Technologies and Analytics, Vienna University of Technology, Vienna, Austria, and Department of Physical and Analytical Chemistry, University of Jaén, Jaén, Spain

The hydrolysis of bovine serum albumin with protease K at 60 °C has been studied by means of infrared spectroscopy. Two-dimensional correlation spectroscopy (2DCoS) has been used to study spectral changes in the reaction. The use of the multivariate curve resolution-alternating least-squares method applied to infrared measurements allowed the recovery of pure infrared spectra and concentration profiles of the different species involved in the reaction. Special attention was paid to the careful inspection of residuals again using 2DCoS. In this way, a heat-induced unfolding step previous to protein hydrolysis was identified. The infrared spectra of the intermediate species showed a more disordered structure than native albumin, the decrease in α-helix conformation being especially noticeable. The formation of *β*-sheet aggregates due to heating was detected too.

Serum albumin is one of the most important and abundant blood proteins in mammalians. It is a versatile carrier protein, active against a variety of substances with widely differing properties. Bovine serum albumin (BSA), which molecular mass is 66 500 Da, is a single polypeptide chain built from 583 amino acid residues. The secondary structure of BSA is composed of 67% α-helix, 10% turn, and 23% extended chain, and no *β*-sheet is present.[1−3] Conformational changes of BSA are reversible in the temperature range of 42−50 °C, but unfolding of α-helices of BSA is irreversible in the temperature range of 52−60 °C. From this temperature, the unfolding of BSA progresses and *β*-aggregation of the molecule begins. Above 70 °C, the gel formation by unfolding of BSA advances further; additionally, BSA forms heat-induced gels in an aqueous environment.[4]

Proteases have been used widely through the field of biochemistry in many areas of study. Proteinase K is a typical member of the subtilisin family of proteases.[5] Like other subtilisins, it does not exhibit any pronounced cleavage specificity, but rather a slight preference in cleaving peptide bonds at the carboxylic sides of aliphatic, aromatic, or hydrophobic amino acids. Proteinase K maintains full enzymatic activity in a broad range of pH (6.5−9.5),[6] and its peak catalytic activity is seen in the range from 50 to 60 °C. The smallest peptide to be hydrolyzed is a tetrapeptide. Hydrolysis by proteases often occurs under denaturing conditions, so every peptide bond is cut where the local amino acid sequence satisfies the specific requirements of the enzyme.[7]

The combination of Fourier transform infrared (FT-IR) spectroscopy with appropriate data evaluation methods is an effective method for analyzing the secondary structure of proteins. The specific stretching and bending vibrations of the peptide backbone in amide I, II, and III bands provide information about the different types of secondary structure, since these vibrational modes are sensitive to hydrogen bonding and coupling between the transition dipole of adjacent peptide bonds. It is well known that the infrared amide I band is particularly more sensitive to protein secondary structure than other amide bands.[8] Infrared spectroscopy provides a global insight into the overall secondary structure, but the analysis of the spectra is not straightforward. Data processing is necessary to extract the descriptive chemical information from the spectra. The most common methods to separate the broad amide I band into different contributions are second derivative, Fourier self-deconvolution, and curve fitting. More recently other methods, such as multivariate curve resolution-alternating least squares (MCR-ALS)[9] and two-dimensional correlation spectroscopy (2DCoS)[10] have been introduced to infrared studies of proteins. The major advantage of 2DCoS is that, besides the resolution enhancement, it provides information about the relationship between bands, which aids band assignments.[8,11] On the other hand, multivariate curve resolution allows for the mathematical resolution of concentration and spectra profiles of pure chemical species from raw data sets recorded in multicomponent

* To whom correspondence should be addressed. Tel. +43 15880115140. Fax: +43 15880115199. E-mail: blendl@ mail.zserv.tuwien.ac.at.

† Vienna University of Technology.

‡ University of Jaén.

(1) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000**, *28*, 235−242.
(2) Murayama, K.; Tomida, M. *Biochemistry* **2004**, *43*, 11526−11532.
(3) Peters, T. Jr. *All about albumin. Biochemistry, genetics and medical applications*; Academics Press Inc. Ltd: London, 1995.
(4) Yamasaki, M.; Yano, H. *Int. J. Biol. Macromol.* **1990**, *12*, 263−268.
(5) Kolattukudy, P. E.; Sirakova, T. D. In *Handbook of proteolytic enzymes*; Barret, A. J., Rawlings, N. D., Woessner, J. F., Eds.; Academic Press: San Diego, 1998.
(6) Kats, M.; St Germain, M. E. *Anal. Biochem.* **2002**, *307*, 212−218.
(7) Hubbard, S. J. *Biochim. Biophys. Acta* **1998**, *1382*, 191−206.
(8) Barth, A.; Zscherp, C. *Q. Rev. Biophys.* **2002**, *35*, 369−430.
(9) Navea, S.; de Juan, A.; Tauler, R. *Anal. Chem.* **2003**, *75*, 5592−5601.
(10) Noda, I. *Appl. Spectrosc.* **2005**, *54*, 994−999.
(11) Murayama, K.; Wu, Y. Q.; Czarnik-Matusewicz, B.; Ozaki, Y. *J. Phys. Chem. B* **2001**, *105*, 4763−4769.

10.1021/ac0520137 CCC: $33.50
Published on Web 04/14/2006

mixtures or evolving chemical processes. It is a flexible method, which can be applied simultaneously to several matrices collected in different experiments to improve the resolution of the system. It has been proven to be very useful for the modeling of protein processes with different spectroscopic techniques.[12]

In this work, the enzymatic hydrolysis of BSA by proteinase K is monitored using FT-IR spectroscopy. MCR-ALS and 2DCoS are used to analyze the data in order to determine the number of species involved in the process and detect the existence of different protein conformations, as well as possible conformational changes. Multivariate curve resolution-alternating least squares is a powerful method to study chemical reactions. Since using MCR, pure spectra of the different components are recovered; it is especially useful when structural information about intermediates and products needs to be obtained. The number of components has to be chosen to minimize the residual data variance, leaving in it, if possible, only the experimental error or noise.[13] Estimation of the number of components in reaction-based systems is difficult because it is usually lower than the number of chemical contributions present in the system.[14] This difference is related to the fact that in this kind of system the rank depends on the number of independent reactions present in the system. Here, we propose a careful study of the residuals obtained from MCR analysis using 2DCoS to provide more information about species or processes not modeled by MCR.

EXPERIMENTAL SECTION

Materials. All chemicals were of analytical reagent grade. BSA was purchased from Sigma (Steinheim, Germany) and the enzyme proteinase K (299 units/mg) from Fluka (Steinheim, Germany). Sodium dihydrogen phosphate dehydrate and disodium hydrogen phosphate dihydrate, also from Fluka, were used to prepare 20 mM phosphate buffer at pH 7.0 in deuterium oxide.

Experimental Procedure. Preliminary studies were performed to select the conditions for the BSA proteolysis reaction. Phosphate buffer at pD 7.4 (pD = pH + 0.4[15]) and 60 °C as the temperature were selected as optimum. To avoid water contamination, sample handling was carried out inside a glovebox purged with dry nitrogen. Samples were prepared by diluting BSA in the appropriate amount of buffer to obtain the desired concentrations. The enzyme, proteinase K, was prepared also in buffer. The enzyme was added to the BSA solutions, to obtain a final concentration of 0.5 mg mL^{-1} enzyme and 30 or 50 mg mL^{-1} protein. After the addition of the enzyme, the obtained solutions were injected into the thermostatized flow cell for measurement (60 °C). A blank solution was prepared in the same way with 30 mg mL^{-1} BSA but without enzyme.

Infrared Spectroscopy. All measurements were performed on a Bruker Equinox 55 FT-IR spectrometer (Bruker Optik GmbH, Ettlingen, Germany) equipped with a narrow-band mercury cadmium telluride detector. Measurements were performed in a thermostated flow-through cell equipped with CaF_2 windows (each 4 mm thick) and a poly(tetrafluoroethylene) spacer providing an optical path length of 50 μm. An oil bath was used to control the temperature in the flow cell.

All spectra were recorded with a resolution of 2 cm^{-1} averaging 128 scans. The background spectrum was recorded with the flow cell filled with buffer. OPUS 5.0/IR software (Bruker) was used to collect the spectra in the range of 4000−900 cm^{-1}. Blackman−Harris three-term apodization and a scanner velocity of 100 kHz HeNe modulation frequency were used throughout. Infrared spectra were recorded every 2 min during 320 min.

Multivariate curve resolution-alternating least squares is based on the mathematical decomposition of a data matrix into a concentration profile and a pure spectrum for each modeled component. As spectroscopic measurements follow the Beer−Lambert law and as the MCR-ALS algorithm assumes that the experimental data follow a bilinear model, the matrix **D** can be decomposed into two matrices, **C** and $\mathbf{S}_T$, following the model

$$\mathbf{D} = \mathbf{C}\mathbf{S}_T + \mathbf{E} \quad (1)$$

where **C** represents the concentration profiles of the components during the reaction and $\mathbf{S}_T$ the matrix of the pure spectra of these species. **E** is the matrix of residuals not explained by the modeled components. The general steps in MCR-ALS analysis of evolving data sets are summarized below.

(a) Determination of the number of components. Singular value decomposition or principal component analysis (PCA) can be used for this purpose.

(b) Initial estimation of either **C** or $\mathbf{S}_T$ matrices derived from chemometric methods, such as evolving factor analysis (EFA)[16] or methods based on the selection of the purest variables such as SIMPLISMA (simple-to-use interactive self-modeling mixture analysis).[17]

(c) Optimization of the initial estimates by an iterative alternating least-squares process. Although MCR-ALS does not require any previous knowledge about the system, either chemical or mathematical, the addition of information helps to decrease the ambiguity in the results. This information is given by constraining the concentration profiles (**C**) and spectra ($\mathbf{S}_T$) during the iterative optimization. Constraints used in this study have been nonnegativity of spectra and concentration profiles, unimodality of concentration profiles, and closure in the concentration direction. Local rank, which imposes absence of the product in the beginning of the experiment, has been used too. The resolution process ends when the reproduced data matrix (**D***) obtained from the product of the resolved concentrations profiles (**C**) and spectra ($\mathbf{S}_T$) is similar enough to the original matrix **D**. The optimal percent of lack of fit in relative standard deviation units is defined by eq 2 and compares the matrix obtained from the resolution results, **D**, with the experimental data.

$$\%\ \text{lack of fit} = 100\ [(\sum r_{ij}^2)/(\sum d_{ij}^2)]^{1/2} \quad (2)$$

where d_{ij} designates an element of the data set **D** and r_{ij} is its

(12) Yan, Y. B.; Wang, Q.; He, H. W.; Hu, X. Y.; Zhang, R. Q.; Zhou, H. M. *Biophys. J.* **2003**, *85*, 1959−1967.
(13) Mendieta, J.; Diaz-Cruz, M. S.; Esteban, M.; Tauler, R. *Biophys. J.* **1998**, *74*, 2876−2888.
(14) Amrhein, M.; Srinivasan, B.; Bonvin, D.; Schumacher, M. M. *Chemom. Intell. Lab. Syst.* **1996**, *33*, 17−33.
(15) Militello, V.; Casarino, C.; Emanuele, A.; Giostra, A.; Pullara, F.; Leone, M. *Biophys. Chem.* **2004**, *107*, 175−187.
(16) Maeder, M.; Zuberbühler, A. D. *Anal. Chim. Acta* **1986**, *181*, 287−291.
(17) Windig, W.; Antalek, B.; Lippert, J. L.; Batonneau, Y.; Bremard, C. *Anal. Chem.* **2002**, *74*, 1371−1379.

related residual. Convergence is assumed to be fulfilled when the difference in fit between two consecutive iterations is less than 0.1%.

Sometimes, a single data matrix does not provide enough information to properly model the chemical system under study. In this situation, MCR-ALS can be applied to the simultaneous analysis of several matrices recorded under the same or different conditions. For doing so, the individual matrices are arranged either in an augmented columnwise matrix (wavenumbers-wise, keeping wavenumbers in common) or in an augmented rowwise matrix (timewise, keeping time intervals in common). In our case, we use augmented columnwise matrices, which have a number of rows equal to the total number of acquired spectra in the different experiments and a number of columns equal to the number of wavenumbers present in each single matrix. In the analysis of augmented matrices, different constraints can be applied to the individual $\mathbf{C}$ or $\mathbf{S}_T$ submatrices when needed.

MCR-ALS analysis was performed in Matlab 7.0 (The Math Works Inc., Natick, MA) using the freely available program (Matlab code) by de Juan and Tauler.[18]

Two-Dimensional Correlation Spectroscopy. 2DCoS performs cross-correlation analysis of spectra series of a system that changes with some modulation variable (reaction time in our case). It spreads the one-dimensional spectra into a second spectral dimension, thus yielding two-dimensional correlation maps with the same wavenumber or wavelength axis in both dimensions.[19] The 2D wavenumber−wavenumber correlation analysis provides two different correlation maps. The synchronous map shows correlations between all spectral bands changing in the experiment and whether they increase or decrease relative to each other. The asynchronous correlation map relates wavenumbers that change at different rates, and it also contains information on the sequence of the events taking place.

2DCoS analysis was also performed in Matlab 7.0 on the basis of the freely available 2DCoS toolbox by Berry and Ozaki,[20] which was properly modified by the authors.

Data Sets. The proteolysis of BSA by proteinase K at 60 °C was studied for two different concentrations of substrate (30 and 50 mg mL^{-1}). A blank run (30 mg mL^{-1}) in the absence of enzyme was measured too. Data matrices ($\mathbf{r} \times \mathbf{c}$) were built with wavenumbers (cm^{-1}) as columns and absorbance intensity at each time as rows.

RESULTS AND DISCUSSION

Most significant spectral changes during the hydrolysis of BSA by proteinase K, took place in the spectral region between 1700 and 1510 cm^{-1}. In Figure 1, spectral changes occurring during proteolysis at 60 °C are shown. The most significant changes are the decrease of the amide I band[8] located around 1651 cm^{-1} and the increase of a broad band centered at 1594 cm^{-1}, which corresponds to the antisymmetric stretching of free carboxylates[21]

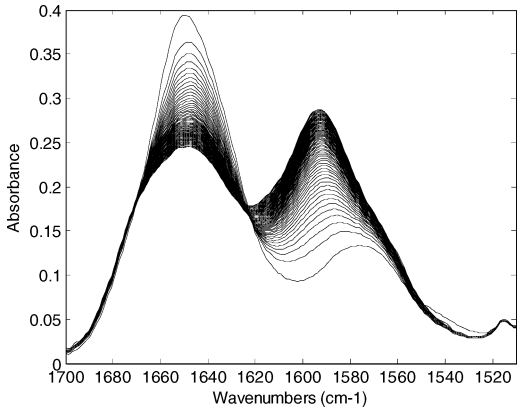

Figure 1. Original IR spectra taken during reaction time (final time 320 min, interval showed 4 min), for an experiment at 60 °C. Sample: BSA, 30 mg mL^{-1}; proteinase K, 0.5 mg mL^{-1}.

appearing due to the cleavage of peptidic bonds at the working pH (7.0).

Synchronous and asynchronous correlation maps obtained by 2D correlation analysis of these data are given in Figure 2. The synchronous map is dominated by the mentioned major peaks at 1651 and 1594 cm^{-1}, reflecting the overall changes occurring as proteolysis proceeds. In the asynchronous map, these two bands develop correlation peaks with 1616 and 1675 cm^{-1}. The first band at 1616 cm^{-1} is usually assigned to the intermolecular β-sheet structures in aggregates appearing in heated protein solutions.[22] Absorption at 1675 cm^{-1} is assigned to β-turn structures.[2] These two bands are not directly observable in the original spectra and could be attributed to conformational changes taking place with kinetics different from the proteolysis process. Additionally, an asynchronous peak develops between 1641 and 1654 cm^{-1}. These bands could be attributed to disordered structures and α-helix, respectively.[8] According to Noda rules about the sign of the correlation peaks,[19,23] the order of spectral changes occurring during proteolysis can be established as follows: 1654/1641, 1594/1675, and 1616 cm^{-1}. From this sequence it can be said that the first event occurs for the band due α-helix structures, followed by the intensity changes in the bands assigned to disordered structures (1641 cm^{-1}) and carboxylate anions (1594 cm^{-1}). Finally, bands of β-turn and β-sheet structures change.

Multivariate curve resolution was used to obtain the pure spectra and the concentration profiles of the components involved in the reaction. PCA was employed to select the appropriate number of components to describe the system. Two major components were found, explaining 99.99% of accumulative variance. EFA was used to generate initial estimates. The results obtained using MCR-ALS are shown in Figure 3. The most significant spectral contribution in the spectrum of the component that decreases corresponded to the amide I. Another significant band can be found at 1575 cm^{-1}, which corresponds to amino acids such as glutamic acid.[2] The proteolysis product, represented by the increasing component, showed the characteristic band at 1594 cm^{-1}, which corresponds to the antisymmetric stretching of carboxylates.

This model explains one process being proteolysis by mainly considering the bands at 1651 and 1594 cm^{-1}. Nevertheless, it

(18) Tauler, R.; de Juan, A. Multivariate Curve Resolution−Alternating Least Squares (MCR-ALS), MATLAB code. University of Barcelona, Barcelona, Spain, 1999.
(19) Noda, I. *Appl. Spectrosc.* **1993**, *47*, 1329−1336.
(20) Berry, J. R.; Ozaki, Y. 2DCoS Toolbox, Matlab code. Kwansei-Gakuin University, Uegahra, Japan, 2001.
(21) Sieler, G.; Schweitzer-Stenner, R.; W.Holtz, J. S.; Pajcini, V.; Asher, S. A. *J. Phys. Chem. B* **1999**, *103*, 372−384.
(22) Zhang, J.; Yan, Y. B. *Anal. Biochem.* **2005**, *340*, 89−98.

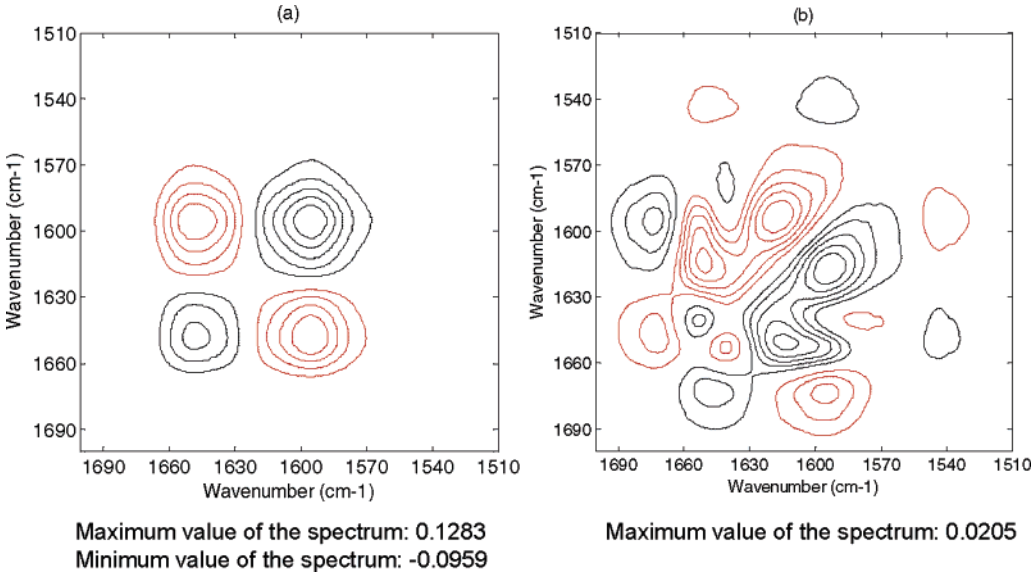

Figure 2. Two-dimensional FT-IR correlation maps for an experiment at 60 °C with enzyme. (a) Synchronous map; (b) asynchronous map. Black lines show positive peaks whereas red ones represent negative peaks. [BSA]= 30 mg mL^{-1}.

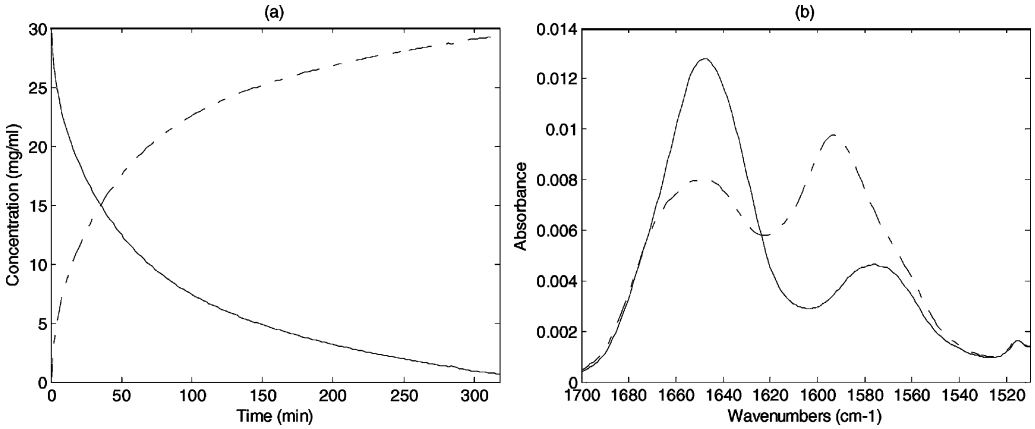

Figure 3. Results of the performed multivariate curve resolution analysis for the experiment run at 60 °C. (a) Concentration profiles for [BSA] = 30 mg mL^{-1} and (b) spectra. First component is represented by solid lines and second component by chain dotted lines.

does not agree with the results obtained from 2D correlation analysis that revealed additional processes occurring at different rates. Therefore, the residuals obtained from MCR analysis have been further investigated. The matrix of residuals was obtained according to eq 3 using the experimental reproduced data set (**D**) and the pure resolved concentration and spectra profiles.

$$\mathbf{E} = \mathbf{D} - (\mathbf{C}\cdot\mathbf{S}_T) \qquad (3)$$

The inspection of the residual matrix revealed the presence of bands in the beginning of the experiment that are not explained by the proposed model. In Figure 4A, a three-dimensional representation of the residuals shows two bands, one centered at 1654 cm^{-1}, typical for α-helix structures, and another at 1616 cm^{-1}, due to intermolecular β-sheets. By analyzing the residual matrix with 2DCoS (Figure 4B) an intense autopeak centered at 1616 cm^{-1} is found in the synchronous plot showing the increase of this band. The most intense peak in the asynchronous map relates the wavenumbers 1654 and 1616 cm^{-1}. This means that two different processes have been excluded from the model, one implies the formation of β-sheet aggregates, and another is related to changes in the amide I band involving α-helix conformation.

From this, it can be concluded that two components were insufficient to model the system and at least a third minor component should be considered. By adding this component, 99.99% of accumulative variance is explained.

MCR-ALS analysis was performed by using EFA results with three factors as initial estimates, and the results are shown in Figure 5. The concentration profiles show a component that decreased as hydrolysis proceeded. The main spectral contribution of this component is the amide I band centered at 1648 cm^{-1}. An intermediate forms initially very fast and then its concentration decreases. The intermediate is a proteolysis product since the characteristic band at 1594 cm^{-1}, which corresponds to the antisymmetric stretching of carboxylates, is present in the spectrum. Furthermore, the amide I band is also present in the spectrum of the intermediate and centered at 1645 cm^{-1}. It is slightly shifted with respect to the native albumin, which indicates that the intermediate is richer in disordered structures. This seems to be in agreement with the observations made by 2D correlation that revealed amide bands at 1654 and 1641 cm^{-1}. A second proteolysis product increases continuously in concentration. In the spectrum of this product, the relative intensity of the carboxy-

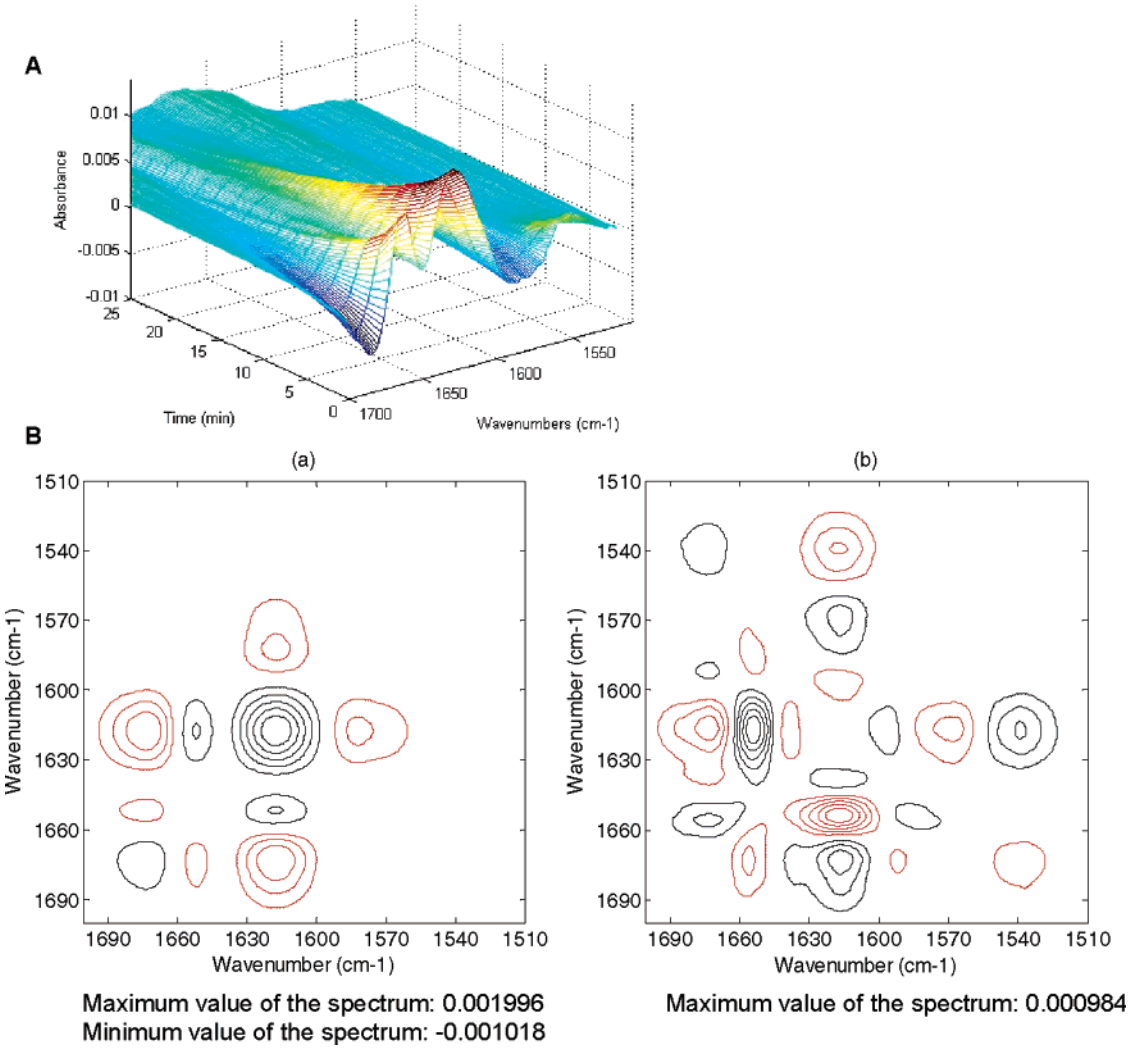

Figure 4. Residuals obtained from MCR analysis of the proteolysis of BSA by proteinase K at 60 °C. (A) 3-D plot of **E** (residual matrix) and (B) 2DCoS analysis of the residuals matrix, synchronous (a) and asynchronous plot (b).

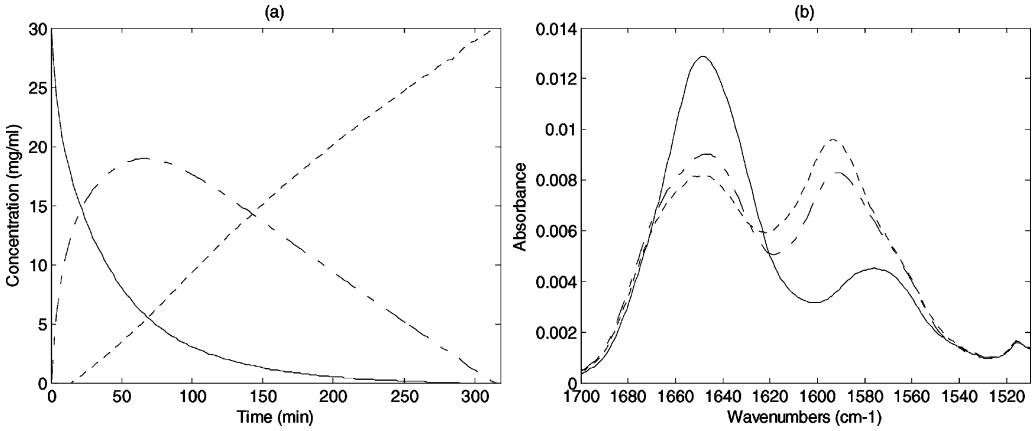

Figure 5. Multivariate curve resolution results for the experiment run at 60 °C with three components. (a) Concentration profiles for [BSA] = 30 mg mL^{-1} and (b) spectra. Components: native albumin (solid line), intermediate (chain dotted line), and product (dotted line).

late band with respect to the amide I band is higher than in the intermediate. Both hydrolysis products are richer in β-turns than the native albumin, as can be seen from the broadening of the amide I band at ~1670 cm^{-1}.

According to these results, the proteolysis of the albumin can be modeled by considering the substrate and two hydrolysis products. In principle, this could be interpreted in terms of peptides of different size: the intermediate species could correspond to peptides larger than the final hydrolysis product. But a different interpretation is more plausible. It is well known from literature that preceding denaturation is necessary for proteolysis to start. An increase in denaturated protein contents thus enhances the rate of proteolysis.[24,25] The fast reaction rate at the beginning of the experiment, coinciding with the heating to 60 °C, must be

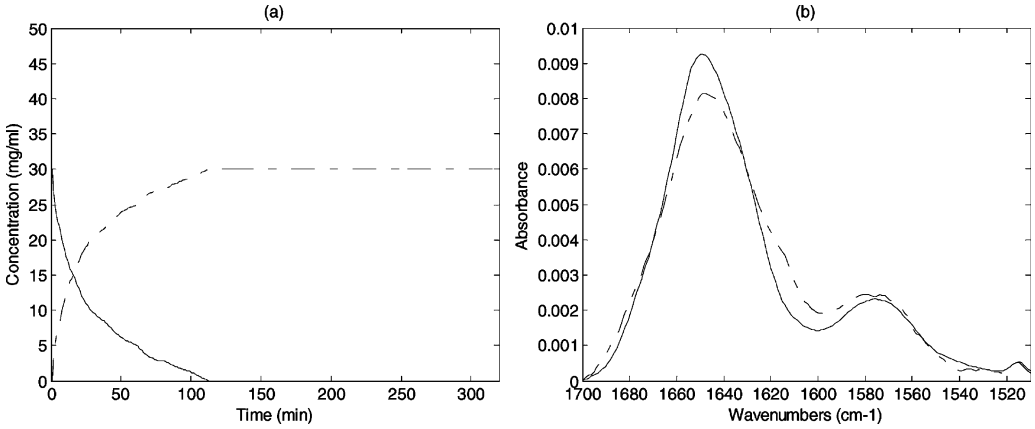

Figure 6. Concentration profiles (a) and spectra (b) obtained for MCR analysis of the blank run without enzyme. [BSA] = 30 mg mL^{-1}. Components: native albumin (solid line) and denaturated albumin (chain dotted line).

a consequence of protein denaturation. But as soon as denaturation takes place, proteolysis starts and the recovered spectrum of the intermediate reflects conformational changes and proteolysis. In reaction networks, it is often encountered that the rank of the matrix is lower than the number of chemical species involved due to linear dependence of the concentration profiles or high spectral correlation. Matrix augmentation[26] can be very helpful to circumvent this problem, and it will be used here to separate the spectral contributions from the thermal unfolding step and the proteolysis reaction.

To do this, a blank experiment, BSA heated at 60 °C without the enzyme, was carried out to get more information about the conformational changes occurring in the absence of proteolysis. Visual inspection of the spectra revealed mainly a decrease of amide I centered at 1651 cm^{-1} that also presented a slight shift to lower wavenumbers and an increasing shoulder at 1614 cm^{-1} (also slightly seen at 1685 cm^{-1}) assigned to β-aggregation. Multivariate curve resolution analysis was performed for this single experiment, and two components were found to be sufficient to fully describe the system as they explained 99.99% of the accumulative variance. Figure 6 shows the obtained results. The decreasing component contains the spectral features from the native albumin, whereas the increasing component reflects the conformational changes occurring due to heating during 5 h.[15] The main characteristics of this spectrum are the decrease of the α-helix band and the shift to lower frequencies concordant with the increase of disordered structures. The appearance of the bands at 1615 and 1685 cm^{-1} (assigned based on second-derivative spectra not shown) is also noticeable because these two bands are diagnostic for inter- or intrapeptide chain aggregation leading to antiparallel pleated β-sheet structures.[2] The analysis of the residuals from MCR in this case shows that they are basically noise, obtaining maximum values for the 2DCoS maps of 2×10^{-4}, which means that every important feature has been explained by the MCR model.

Matrix augmentation was used to introduce this information in the analysis of the proteolysis reaction. A data matrix built by columnwise augmentation with two experiments of BSA (30 and 50 mg/mL) and proteinase K (0.5 mg/mL) at 60 °C and the blank run (30 mg/mL) without enzyme was analyzed by MCR. Again, EFA was used to calculate initial estimates, and in this case, at least three components, explaining 99.85% of accumulative variance, were necessary to model the data. A fourth minor component could also have been considered. The results obtained using MCR-ALS are shown in Figure 7.

The three species identified by the model, taking into account all the assignments mentioned in the previous analysis, can be assigned to native albumin, "unfolded" albumin, and proteolysis product. Looking at the concentration profile, native albumin suffers the heat denaturation in the few first minutes of the experiment, showing a fast decrease in its concentration. The amide I band of the native albumin is centered at 1651 cm^{-1}. Simultaneously, the component attributed to the unfolded albumin, being richer in disordered structures as reflected in the amide I band at 1648 cm^{-1}, appears and then the proteolysis starts, consuming it and developing the proteolysis product with its major spectral contribution at ~1594 cm^{-1}. The rate of the unfolding process necessary for proteolysis is very fast in the beginning, which allows the observation of the unfolded species. Afterwards, the concentration of the component attributed to unfolded albumin is close to zero, probably because it is hydrolyzed at the same rate as it is being produced. The spectrum of the unfolded albumin also presents two shoulders at 1616 and 1685 cm^{-1} corresponding, as mentioned before, to β-sheets. In the experiment without enzyme, a certain degree of autoproteolysis is observed probably due to the effect of heating. These results also agree with the sequence of events revealed by 2DCoS. The first changes are due to the unfolding of α-helix structures followed by the consumption of the unfolded albumin to give free carboxylate groups.

The residual matrix obtained when considering three components in augmented matrix was also investigated, directly and by 2DCoS. Again, the obtained results showed interesting spectral features. The synchronous plots of the residuals still showed an autopeak centered at 1651 cm^{-1}. This residual intensity due to the native albumin can be justified because the denaturation is so fast that it is very difficult to model. From the asynchronous plots (Figure 8), the most interesting point is the presence of the band at 1616 cm^{-1}. This band, assigned to aggregated β-sheet structures, was more significant in the experiment without

(23) Noda, I. *Appl. Spectrosc.* **1990**, *44*, 550−561.
(24) Rupley, J. A. *Methods in Enzymology*; Hirs, C. H. W., Ed.; Academic Press: New York, 1967; Chapter XI.
(25) Vandermeulen, D. L.; Ressler, N. *Arch. Biochem. Biophys.* **1980**, *205*, 180−190.
(26) Saurina, J.; Hernandez-Cassou, S.; Tauler, R.; Izquierdo-Ridorsa, A. *J. Chemom.* **1998**, *12*, 183−203.

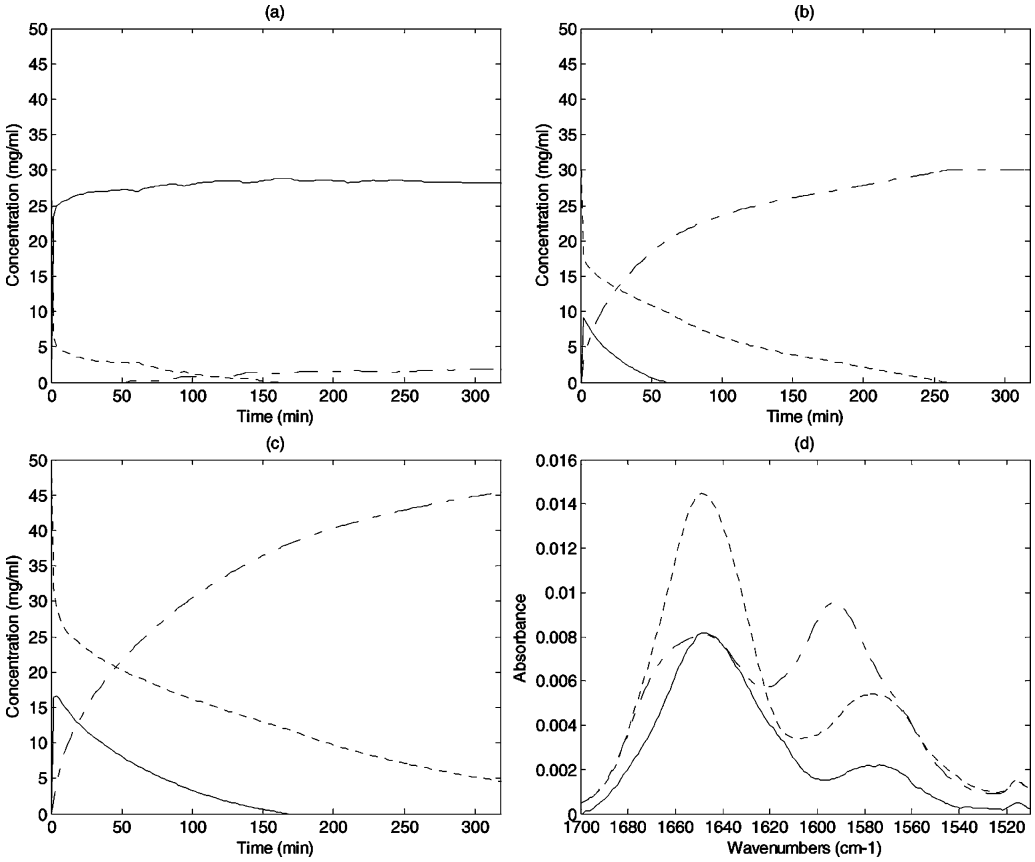

Figure 7. MCR results obtained from applying MCR to the columnwise augmented matrix composed from the experiments run at 60 °C. Concentration profiles: (a) [BSA] = 30 mg mL^{-1}, no enzyme; (b) [BSA] = 30 mg mL^{-1}, [Prot K] = 0.5 mg mL^{-1}; (c) [BSA]= 50 mg mL^{-1}, [Prot K] = 0.5 mg mL^{-1}, and spectra (d). Components: unfolded albumin (solid line), proteolysis product (chain dotted line), and native albumin (dotted line).

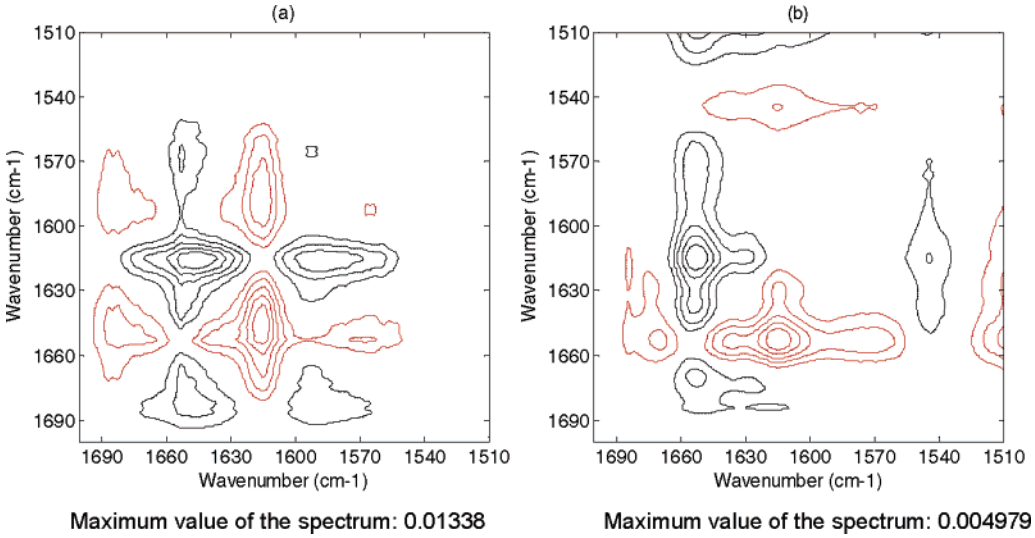

Figure 8. Asynchronous plots from 2DCoS analysis of the residual matrix, (a) [BSA] = 30 mg mL^{-1} and (b) [BSA] = 30 mg mL^{-1}, [Prot K] = 0.5 mg mL^{-1}. Black lines show positive peaks whereas red ones represent negative peaks.

enzyme, as can be seen from the maximum values of the asynchronous spectra, 1.3×10^{-2} for the experiment without proteinase K and 4.9×10^{-3} for the same experiment but with 0.5 mg/mL enzyme. This means that the heat-induced conformational changes producing β-sheet aggregated structures have not been completely modeled in the blank experiment. This also explains the difference in the denaturation rate observed for the blank experiment in Figures 6a and 7a. When the experiment was treated alone (Figure 6a), the observed kinetic profiles are linear combinations of the fast unfolding step and the slow formation of β-sheet aggregates. When analyzed using matrix augmentation, this information is split. The fast process is modeled by the "unfolded albumin" kinetic profile, but the information of the slow aggregate formation cannot be completely modeled and remains

mainly in the residuals. This is probably because the spectral contributions of β-sheet structures are of minor importance in the experiments involving the enzyme. Considering that β-sheet structures have been found to be very stable and resistant to proteinase K digestion,[27,28] it must be concluded that the formation of these structures is disabled by the proteolysis process. The possibility of modeling the contribution of β-sheet aggregation by the inclusion of a fourth component in the MCR analysis was tested. However, this was not successful probably due to the extremely low contribution of this component to the spectral variance and because of the insufficient rank for the resolution. This made it impossible to separate the spectral features of β-sheet aggregation from those of the proteolysis product.

CONCLUSIONS

From the results obtained in this study it may be concluded that the combined use of MCR and 2DCoS is a very powerful approach in the analysis of FT-IR spectra obtained from protein reactions. The complementary nature of these two chemometric techniques provides opportunities to develop and test models for complex biochemical processes. Considering that these two techniques are essentially different and rely on distinct mathematical approaches, the coherence between the results demonstrates the reliability of the obtained conclusions. Furthermore, 2DCoS was applied to the residuals from MCR to get more information about the modeling process. In this particular example, the preceding unfolding of bovine serum albumin to proteolysis could be detected by applying both techniques to the analysis of the experimental data. In addition, the appearance of β-sheet aggregates was observed.

(27) Zhang, H.; Kaneko, K.; Nguyen, J. T.; Livshits, T. L.; Baldwin, M. A.; Cohen, F. E.; James, T. L.; Prusiner, S. B. *J. Mol. Biol.* **1995**, *250*, 514−526.
(28) Manning, M.; Colon, W. *Biochemistry* **2004**, *43*, 11248−11254.

ACKNOWLEDGMENT

This work was financed by Project 15531 from the Austrian Science Fund. A.D.V. and M.J.A.C. were respectively supported by a postdoctoral fellowship and a mobility grant from the Ministerio de Educación y Ciencia, Secretaría de Estado de Universidades e Investigación (Spain).

Received for review November 12, 2005. Accepted March 20, 2006.

AC0520137