# Techniques of two-dimensional (2D) correlation spectroscopy useful in life science research [1]

Isao Noda [*]

*Department of Materials Science & Engineering, University of Delaware, Newark, DE, USA*

**Abstract.** Several illustrative examples are presented in this tutorial review to demonstrate the utility of 2D correlation spectroscopy, especially in life science applications. A set of IR spectra for a model protein system, which is undergoing complex thermally induced changes in the secondary structures, is analyzed by 2D correlation spectroscopy. The method of constructing 2D correlation maps from temperature-dependent IR spectra and their interpretation procedure are described. The resolution of overlapped protein IR bands by 2D correlation is demonstrated, and sequential order of spectral intensity changes is determined. Newly emerging techniques in the field, such as Pareto scaling, positive null-space projection, and 2D codistribution analysis, are discussed in some detail, in addition to the traditional application of 2D correlation spectroscopy.

Keywords: Two-dimensional correlation spectroscopy, protein secondary structures, Pareto scaling, vector projection, codistribution analysis

## 1. Introduction

Two-dimensional (2D) correlation spectroscopy [15–17,29,30] has become a very popular analytical technique in the last several decades, especially in the field of life science. A series of comprehensive survey reviews have been tracking the progress of 2D correlation spectroscopy in various applications [18,20,21,23,25,26,28]. In essence, 2D correlation spectroscopy is a tool to extract useful information from a set of spectral data, which are obtained from a sample under the influence of some form of external perturbation. The nature of the perturbation can be a time-dependent phenomenon, like chemical reactions or physical relaxation processes. Static effects, such as the temperature variation or change in the concentration of constituents, are also routinely employed as an effective form of perturbation. Systematic variations in the spectral intensities induced by such perturbations are used to construct 2D correlation spectra by employing a form of cross correlation analysis [17,29,30]. The technique can be applied to a number of analytical probes, such as IR, Raman, NIR, fluorescence, NMR, X-ray, and the like [26]. The basic concept of 2D correlation can also be utilized in the field of applications other than spectroscopy, such as chromatography or microscopy [30].

Notable features of 2D correlation spectroscopy that have been realized are: simplification of complex spectra consisting of many overlapped peaks, enhancement of apparent spectral resolution by spreading peaks over the second dimension, establishment of unambiguous assignment through correlation

---

[1] This was presented at the Asian Regional Conference on 2DCOS (2DCOS-2014).

[*] Address for correspondence: Isao Noda, Department of Materials Science & Engineering, University of Delaware, Newark, DE 19716, USA. E-mail: noda@udel.edu.

of bands selectively coupled by various interaction mechanisms, and determination of the sequence of events represented by the variations of spectral intensities [16,17,29,30]. Such unique advantages of 2D correlation spectroscopy are especially attractive in the study of highly complicated biological samples, like proteins [26,31]. In fact, the very first report on 2D IR correlation spectroscopy on the record was indeed based on the study of a protein sample under the influence of mechanical perturbation [14]. Further future growth in the usage of 2D correlation spectroscopy is expected in the biological and biomedical applications.

In this tutorial review, several illustrative examples are presented to demonstrate the utility of 2D correlation spectroscopy, especially in the life science applications. A set of IR spectra for a model protein system, which is undergoing complex thermally induced changes in the secondary structures, are analyzed by 2D correlation spectroscopy. The method of constructing 2D correlation maps from temperature-dependent IR spectra and their interpretation are described. The resolution of overlapped protein IR bands by 2D correlation is demonstrated, and sequential order of spectral intensity changes is determined. Newly emerging techniques in the field, such as Pareto scaling, null-space projection and codistribution analysis, will be discussed in some detail, in addition to the traditional application of 2D correlation spectroscopy.

## 2. Background

### 2.1. Perturbation-induced dynamic spectra

The basic theory of generalized 2D correlation spectroscopy has been described many times in the past [17,19,29,30], so only a brief summary which is pertinent to practical applications is provided here. In the so-called generalized 2D correlation scheme, a series of perturbation-induced dynamic spectra are collected first in a systematic manner. Suppose a set of $m$ spectra $A(\nu_j, t_i)$ is obtained as a function of the spectral variable $\nu_j$ with $j = 1, 2, \ldots, n$ and some perturbation variable $t_i$ with $i = 1, 2, \ldots, m$ during a well-defined observation interval between $t_1$ and $t_m$. The spectral variable $\nu$ can be, for example, IR wavenumber, Raman shift, X-ray diffraction angle, UV wavelength, etc., depending on the specific analytical probe employed. Perturbation variable $t$ can be one of many reasonable measures of the effect of perturbation, such as time, temperature, composition, pressure or even electric potential. The spectral dataset obtained under the influence of a perturbation can be readily transformed into 2D correlation spectra by applying a form of cross correlation analysis.

The *dynamic spectrum* $\tilde{A}(\nu_j, t_i)$ of a system, induced by the application of an external perturbation within the observation interval between $t_1$ and $t_m$, is formally defined as

$$\tilde{A}(\nu_j, t_i) = A(\nu_j, t_i) - \overline{A}(\nu_j), \tag{1}$$

where $\overline{A}(\nu_j)$ is the *reference spectrum* of the system. For most applications, the reference spectrum is selected as the *stationary* or *average spectrum* given by

$$\overline{A}(\nu_j) = \frac{1}{m} \sum_{i=1}^{m} A(\nu_j, t_i). \tag{2}$$

## 2.2. Synchronous 2D correlation spectrum

Synchronous correlation spectrum $\Phi(\nu_1, \nu_2)$, which represents the correlation of simultaneous variations of spectral intensities measured at two different wavenumbers, $\nu_1$ and $\nu_2$, is given by

$$\Phi(\nu_1, \nu_2) = \frac{1}{m-1} \sum_{i=1}^{m} \tilde{A}(\nu_1, t_i) \cdot \tilde{A}(\nu_2, t_i). \tag{3}$$

This quantity is a measure of the similarity among spectral intensity variations along the perturbation variable $t_i$. A large value of $\Phi(\nu_1, \nu_2)$ implies that spectral intensities at $\nu_1$ and $\nu_2$ vary in a coordinated manner, suggesting the common or coupled origin of the spectral signals. It can be easily seen from Eq. (3) that synchronous spectrum is a form of covariance of spectral intensity fluctuations observed at two different wavenumbers, $\nu_1$ and $\nu_2$.

2D correlation spectrum is often displayed as a contour map of correlation intensity on a spectral plane defined by two independent wavenumber axes. A synchronous spectrum plotted in this manner is a symmetric map, i.e., $\Phi(\nu_a, \nu_b) = \Phi(\nu_b, \nu_a)$, with respect to the main diagonal line of the map located at the coordinate $\nu_1 = \nu_2$. Correlation intensity at the diagonal position corresponds to the autocorrelation function of spectral intensity variations. Peaks on the main diagonal line are therefore referred to as auto-peaks. The intensities of auto-peaks represent the overall extent of dynamic fluctuations of spectral signals induced by the perturbation.

Cross peaks located at the off-diagonal positions of a synchronous 2D correlation spectrum, in turn, represent the simultaneous change in signals at two different wavenumbers. The sign of a cross peak can be either negative or positive. It is customary to indicate negative cross peaks with either shading or broken contour lines. If the sign of a synchronous cross peak is positive, the two intensities observed at corresponding wavenumbers increase or decrease together in the same direction. On the other hand, the sign of a cross peak becomes negative if one of the spectral intensities is increasing while the other is decreasing.

## 2.3. Asynchronous 2D correlation spectrum

Asynchronous correlation spectrum $\Psi(\nu_1, \nu_2)$ represents the correlation of sequential variations of spectral intensities measured at $\nu_1$ and $\nu_2$. This quantity is given by

$$\Psi(\nu_1, \nu_2) = \frac{1}{m-1} \sum_{i=1}^{m} \tilde{A}(\nu_1, t_i) \cdot \sum_{k=1}^{m} N_{ik} \tilde{A}(\nu_2, t_k) \tag{4}$$

with

$$N_{ik} = \begin{cases} 0 & \text{if } i = k, \\ \dfrac{1}{\pi(k-i)} & \text{otherwise.} \end{cases} \tag{5}$$

The intensity of asynchronous correlation spectrum can be used as a measure of dissimilarity of spectral intensity variations. A large value of $\Psi(\nu_1, \nu_2)$ implies the spectral intensities at $\nu_1$ and $\nu_2$ vary independently of each other. This phenomenon, in turn, suggests that corresponding spectral signals are originating from separate moieties, capable of responding differently to a common perturbation.

An asynchronous 2D correlation spectrum map consists exclusively of off-diagonal cross peaks. It is an anti-symmetric map, i.e., $\Psi(\nu_a, \nu_b) = -\Psi(\nu_b, \nu_a)$, such that for every positive peak there is a negative peak counterpart at the opposite side of the main diagonal line located at $\nu_1 = \nu_2$. Asynchronous correlation represents sequential changes, instead of simultaneous ones, in spectral intensity variations measured at $\nu_1$ and $\nu_2$, which are observed at different points along the perturbation axis. In other words, asynchronous cross peaks develop only if two dynamic spectral intensities vary out of phase with each other. This feature becomes especially useful in differentiating overlapped bands arising from different moieties which are capable of responding independently of each other to a given perturbation.

By comparing the signs of cross peaks, it becomes possible to determine the sequential order of spectral intensities variations by following a simple set of rules. If the signs of both synchronous and asynchronous peaks located at the same spectral coordinate $(\nu_1, \nu_2)$ are the same, the intensity change at $\nu_1$ is occurring predominantly before that at $\nu_2$ along the perturbation axis. On the other hand, if the signs of cross peaks are different, the opposite becomes true, i.e., intensity variation at $\nu_1$ takes place predominantly after $\nu_2$. If an asynchronous cross peak is not observed but a sizable synchronous cross peak appears at a spectral coordinate $(\nu_1, \nu_2)$, then the spectral intensity variations observed at $\nu_1$ and $\nu_2$ occur simultaneously. And finally, in the rare case of observing a finite asynchronous peak but no synchronous peak at the same coordinate, the sequential order cannot be determined.

### *2.4. Matrix representation of* 2D *correlation spectra*

It is often useful to represent a spectral dataset in a form of $m$ by $n$ matrix $\tilde{\mathbf{A}}$ with it element being $\tilde{A}_{ij} = \tilde{A}(\nu_j, t_i)$. With this notation, the synchronous and asynchronous spectra are conveniently given in the form of $n$ by $n$ matrices

$$\boldsymbol{\Phi} = \frac{1}{m-1}\tilde{\mathbf{A}}^{\mathrm{T}}\tilde{\mathbf{A}}, \tag{6}$$

$$\boldsymbol{\Psi} = \frac{1}{m-1}\tilde{\mathbf{A}}^{\mathrm{T}}\mathbf{N}\tilde{\mathbf{A}}. \tag{7}$$

The superscript $^{\mathrm{T}}$ stands for the matrix transpose operation. Elements of the so-called Hilbert–Noda transformation matrix $\mathbf{N}$ are defined in Eq. (5) [19]. The matrix notation of 2D correlation spectra becomes especially useful in describing the vector projection operation to be discussed later.

## 3. 2D IR correlation spectroscopy studies of proteins

### *3.1. Protein secondary structures*

IR and Raman probes are very often used in the study of proteins and peptides [1,3,9,10,13,40]. Vibrational spectra of proteins, especially in the Amide I and II regions, contain very rich information about the conformational secondary structure of various proteins. Due to specific transition dipole coupling phenomena, different local conformational structures, such as $\alpha$-helix and $\beta$-sheet moieties, give bands with characteristic frequencies which can in principle be distinguished from each other. In reality, however, such conformation sensitive IR bands are substantially overlapped with each other to make the unambiguous assignment sometimes difficult.

2D correlation spectroscopy has been extensively used in the IR and Raman study of various proteins [2,5–8,11,32,34,37,39]. The apparent enhancement of spectral resolution by spreading the overlapped bands along the second dimension clearly is an advantage in interpreting complex protein spectra. The assignment of the sequential order of intensity variations of overlapped conformation-sensitive bands also becomes helpful. It is often used in gaining the mechanistic understanding of the dynamic evolution of secondary structures of protein, especially during the denaturation and aggregation processes induced by thermal or chemical perturbation.

In this section, a set of simulated model protein IR spectra, which are meant to somehow mimic the behavior of naturally occurring protein undergoing thermal denaturation, is constructed for 2D correlation study. The temperature-dependent dynamics of individual secondary structures are modeled to follow the patterns similar to those observed in the actual experimental measurement [40] to make this study as realistic as possible. The real advantage of using the simulated data instead of the actual experimental protein IR spectra is that the assumed component spectra and their dynamics are a priori known without any ambiguity. Thus, the performance of various analysis techniques associated with 2D correlation spectroscopy can be validated without making any further assumptions.

### 3.2. Simulated protein IR spectra

Figure 1(a) shows a simulated protein IR spectrum in the Amide I region, which reasonably resembles many known experimental protein spectra with various secondary structures. The spectrum is a composite of combined contributions from four separate secondary structure component spectra, corresponding to the $\alpha$-helix, $\beta$-sheet, $\beta$-turn and random (disordered) conformation. Such component spectra are also individually shown in Fig. 1(a) under the envelope of the simulated protein spectrum.

For the simulated component spectrum, each absorption peak is constructed by assuming a Lorentzian profile

$$L_i(\nu) = a_i \left\{ 1 + \left( \frac{\nu - \nu_i}{w_i} \right)^2 \right\}^{-1} \tag{8}$$

with appropriate parameters representing the peak $i$, given by the peak maximum wavenumber position $\nu_i$, maximum intensity $a_i$, and half width at half height $w_i$ as listed in Table 1. Peaks of component spectra are substantially overlapped when combined into the final protein spectrum. In other words, unambiguous identification of different peaks of component spectra from the protein spectrum is not always straightforward.

It is now assumed that the conformational structures of this model protein undergo temperature-induced transformations due to denaturation. Assumed population changes of the secondary structures are graphically depicted in Fig. 1(b). They are based on the following temperature response functions $F_j(T)$ of the $j$th conformation given by

$$F_{\alpha\text{-}helix}(T) = \left\{ 1 - 0.1G_1(T) \right\} \left\{ 1 - 0.1G_2(T) \right\} \left\{ 1 - 0.9G_3(T) \right\}, \tag{9}$$

$$F_{\beta\text{-}turn}(T) = G_1 \left\{ 1 - 0.9G_3(T) \right\}, \tag{10}$$

$$F_{\beta\text{-}sheet}(T) = G_2 \left\{ 1 - 0.9G_3(T) \right\}, \tag{11}$$

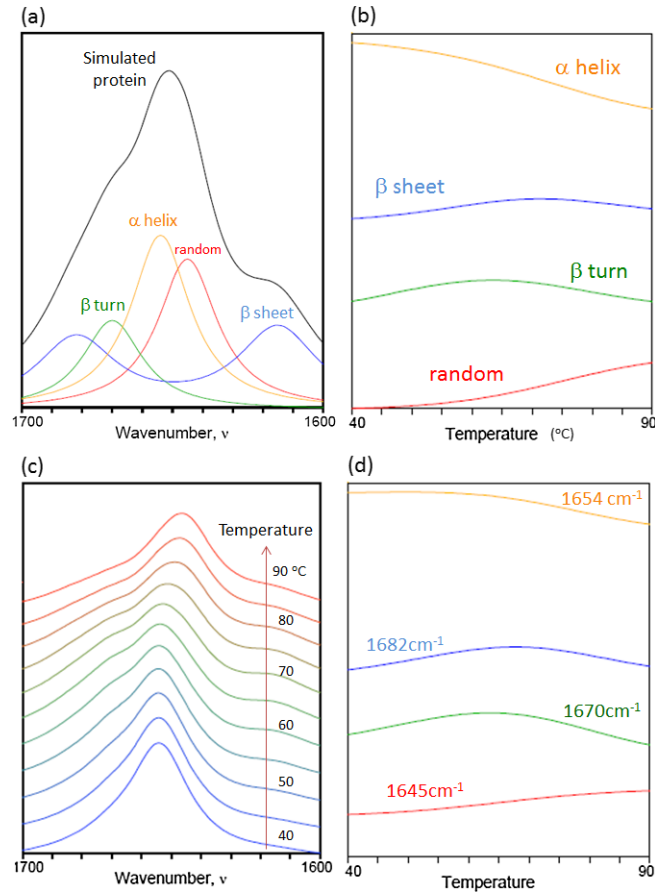$$F_{random}(T) = G_3(T). \tag{12}$$

Fig. 1. Simulated typical protein IR spectra, somewhat mimicking the thermally induced changes of naturally occurring protein: (a) simulated IR spectrum and assumed constituent spectra of individual conformational secondary structures; (b) assumed temperature-dependent population profiles of the secondary structures; (c) temperature-dependent IR spectra of the model; and (d) temperature profiles of the spectral intensities at the wavenumbers corresponding to the band positions of conformational secondary structures.

Table 1

Parameters defining the simulated IR absorption peak profiles of the conformational secondary structures of model protein

| Conformation | $\nu_i$ (cm$^{-1}$) | $a_i$ | $w_i$ (cm$^{-1}$) | Temperature dependence |
|---|---|---|---|---|
| $\alpha$-helix | 1654 | 0.15 | 12 | $F_{\alpha\text{-helix}}$ |
| High $\beta$-sheet | 1682 | 0.07 | 16 | $F_{\beta\text{-sheet}}$ |
| Low $\beta$-sheet | 1615 | 0.08 | 16 | $F_{\beta\text{-sheet}}$ |
| $\beta$-turn | 1670 | 0.07 | 12 | $F_{\beta\text{-turn}}$ |
| Random | 1645 | 0.12 | 12 | $F_{random}$ |

The response functions are made of several Gaussian sigmoid transition functions $G_k(T)$ with the parameters of the transition temperature $T_j$ and transition interval $d_j$ provided in Table 2,

$$G_j(T) = \left\{ 1 + \exp\left( \frac{T_j - T}{d_j} \right) \right\}^{-1}. \tag{13}$$

Table 2

Parameters defining the assumed transition functions for the temperature-induced conformational transformations of model protein

| Transformation | Transition function | $T_j$ (°C) | $d_j$ (°C) |
|---|---|---|---|
| $\alpha$-helix $\rightarrow$ $\beta$-turn | $G_1(T)$ | 50 | 10 |
| $\alpha$-helix $\rightarrow$ $\beta$-sheet | $G_2(T)$ | 65 | 10 |
| All $\rightarrow$ random | $G_3(T)$ | 75 | 10 |

Here, $G_1(T)$ represents the assumed transformation of $\alpha$-helix to $\beta$-turn predominantly occurring around 50°C, $G_2(T)$ is for $\alpha$-helix to $\beta$-sheet near 65°C and $G_3(T)$ is for the formation of random or disordered conformation around 75°C.

Given the spectra and population dynamics of the secondary structures of the model, the evolution of the temperature-dependent IR spectra of this model protein can be simulated as shown in Fig. 1(c). The simulated IR spectra are intentionally designed to closely mimic that of the real protein denaturation process, for example [13]. The protein spectrum shown in Fig. 1(a) actually corresponds to the simulated spectrum at 75°C in Fig. 1(c). We will use this set of simulated temperature dependent protein IR spectra to demonstrate the utility of 2D correlation analysis.

Figure 1(d) shows the temperature dependent profiles of IR absorption intensities at 1654, 1682, 1670 and 1645 cm$^{-1}$. These band positions correspond, respectively, to the peak maximum positions of $\alpha$-helix, $\beta$-sheet, $\beta$-turn, and random conformations of this model protein. Indeed, the curves shown here are reasonably similar to those found in Fig. 1(b). However, there are clearly some differences in the detailed shapes of the curves between Fig. 1(b) and (d). For example, the decrease in the intensity at 1654 cm$^{-1}$ seems to occur at a higher temperature range compared to that for the population of $\alpha$-helix conformation. Such discrepancy is due to the influence of contributions from overlapped neighboring bands. In this case, the contribution from the increasing intensity of the random conformation band centered around 1645 cm$^{-1}$ is making the intensity decrease at 1654 cm$^{-1}$ to occur at a higher temperature.

### 3.3. 2D correlation analysis

Figure 2 shows the contour map representations of 2D IR correlation spectra constructed from the temperature-dependent IR spectra of Fig. 1(c). Negative correlation peaks are indicated by the blue shade. The reference spectrum obtained as the average of temperature dependent spectra is provided at the top and side of correlation maps. Band locations of the component spectra for individual secondary structures, i.e., 1654 cm$^{-1}$ for $\alpha$-helix, 1682 and 1615 cm$^{-1}$ for $\beta$-sheet, 1670 cm$^{-1}$ for $\beta$-turn, and 1645 cm$^{-1}$ for random conformation, are also indicated by the color-coded vertical and horizontal bars placed along with the reference spectrum.

Auto-peaks found along with the main diagonal line and corresponding cross peaks of the synchronous spectrum (Fig. 2(a)) are all located very closely to the wavenumber positions of the secondary structure absorption bands of component spectra. Thus, the dynamics of individual components can be resolved and analyzed, even though the spectral features within the original protein IR spectra are heavily overlapped. High intensities of auto-peaks at 1654 and 1645 cm$^{-1}$ indicate strong susceptibility of these bands, corresponding to the large population changes of $\alpha$-helix and random conformations with increasing temperature.

Negative synchronous cross peaks found near the band position of $\alpha$-helix at 1654 cm$^{-1}$ indicate that the direction of intensity change of $\alpha$-helix conformation is opposite of the behavior of random conformation at 1645 cm$^{-1}$. As the intensity at 1654 cm$^{-1}$ is consistently decreasing with the rising
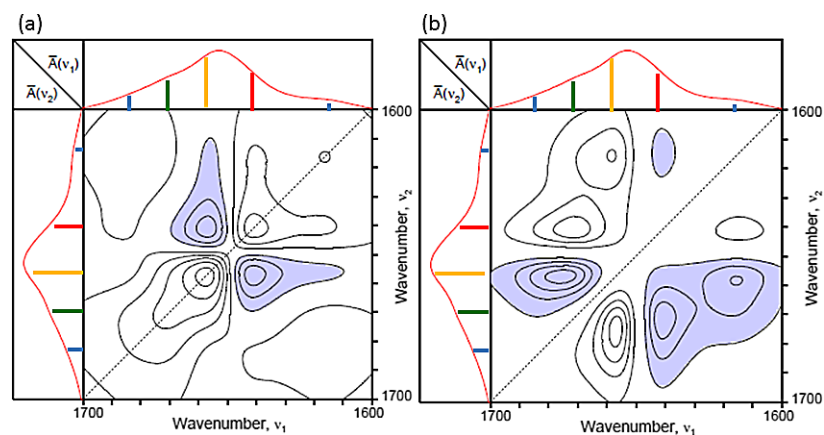
Fig. 2. Synchronous (a) and asynchronous (b) 2D correlation spectrum of the model protein constructed from the simulated temperature-dependent IR spectra shown in Fig. 1(c) ranging from 40 to 90°C. Average spectrum is placed at the top and side of the 2D contour map as a reference. Short color-coded bars are provided with the reference spectrum to indicate the band position of conformational secondary structures. 2D cross peaks with negative correlation intensities are indicated by the blue shading.

temperature, these cross peaks correctly reflect the increase in the amount of random conformation with temperature. The population dynamics of $\beta$-sheet at 1670 and 1615 cm$^{-1}$ may be somewhat more complicated, as it increases first but then decreases later at a higher temperature. Almost near zero synchronous correlation intensity between $\alpha$-helix at 1654 cm$^{-1}$ and $\beta$-sheet at 1670 cm$^{-1}$ indeed suggests the cancellation of the two competing trends. As discussed later, selecting different observation intervals focused on either the lower or higher temperature range provides much less ambiguous synchronous correlation results.

Spectral features observed in the asynchronous spectrum (Fig. 2(b)) are even better resolved than those in the synchronous spectrum. Cross peaks are observed at every positions of individual absorption bands of the protein secondary structure conformations. The result immediately reveals that the secondary structures are evolving asynchronously, i.e., semi-independently with each other, at different temperatures. This finding is consistent with the fact that the temperature response functions of individual conformations are all designed to behave differently as described in Eqs (9)–(12).

It is interesting to note that the position of the maximum point of an asynchronous cross peak does not always match with the position of individual conformation band. This discrepancy is due primarily to spectral overlap. For example, the maximum point of the asynchronous cross peak at (1676, 1654 cm$^{-1}$) is a result of the overlapped contributions from $\beta$-turn (1670 cm$^{-1}$) and $\beta$-sheet (1682 cm$^{-1}$), both behaving asynchronously with $\alpha$-helix (1654 cm$^{-1}$). Yet, the asymmetric shape of the cross peak and the pointed lobe around (1682, 1670 cm$^{-1}$) clearly indicate the presence of two overlapped contributions, and furthermore the possible location of the $\beta$-sheet (1682 cm$^{-1}$) band. Similarly, the location of the $\beta$-turn (1670 cm$^{-1}$) can be more accurately deduced from the cross peak with random conformation (1645 cm$^{-1}$).

Sequential order of the spectral intensity changes can be determined from the signs of 2D correlation cross peaks. For example, both synchronous and asynchronous correlation intensities observed at the coordinate (1654, 1670 cm$^{-1}$) are positive. The result indicates the continuous decrease in $\alpha$-helix at 1654 cm$^{-1}$ starts earlier at a lower temperature than the decrease in $\beta$-turn at 1670 cm$^{-1}$. However, the assignment of the precise order of intensity changes may become a complicated matter, if the temperature

interval of observation is selected too wide to include multiple competing population dynamics, such as the increase and decrease in the absorption intensities of $\beta$-sheet and $\beta$-turn conformations during the temperature scan.

The initial increase in the population of $\beta$-turn conformation observed at 1670 cm$^{-1}$ obviously happens before the eventual decrease of the same secondary structure with rising temperature. According to the assumed model, the creation of $\beta$-turn actually occurs at a temperature lower than the overall consumption of $\alpha$-helix, which is continuously converted to other structures, like $\beta$-sheet and random conformations, in a higher temperature range. In short, the actual sequential order of intensity changes can be summarized as: $\beta$-turn$\uparrow$ $\rightarrow$ $\alpha$-helix$\downarrow$ $\rightarrow$ $\beta$-turn$\downarrow$. The horizontal arrow $\rightarrow$ denotes the sequence of occurrence along with increasing temperature, while up or down arrow indicates, respectively, the increase or decrease in their population.

It should be noted that a simple 2D correlation analysis of the temperature-dependent spectral data, encompassing the full range from 40 to 90°C, has so far highlighted only the higher temperature process ($\alpha$-helix$\downarrow$ $\rightarrow$ $\beta$-turn$\downarrow$) which was the most dominant features. In order to capture individual processes without any ambiguity, it seems more helpful to split the dataset into smaller segments of temperature ranges.

### 3.4. 2D correlation analysis of segmented data

Segmentation of dataset into small blocks has been practiced for many years to unambiguously capture the fine features of population dynamics and assignment of sequential orders of intensity changes of complex processes [30]. By limiting the observation interval small enough, confusing results arising from multiple processes, like the one observed above, can be substantially simplified. Here we demonstrate the utility of such dataset segmentation by splitting the temperature range of the model protein IR data into two separate blocks: one from 40 to 65°C and the other from 65 to 90°C.

Figure 3 shows the 2D correlation spectra constructed from the simulated temperature-dependent protein IR spectra for the range from 40 to 65°C. The entire synchronous spectrum (Fig. 3(a)) is dominated by the strong auto-peaks and positive cross peaks, except for the elongated streaks of negative cross peaks running at the spectral coordinate of 1654 cm$^{-1}$ which corresponds to the decrease in the $\alpha$-helix
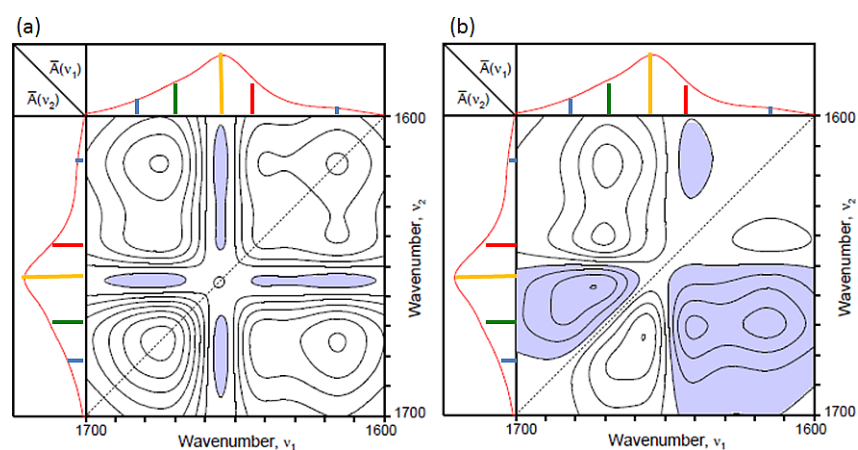


Fig. 3. Synchronous (a) and asynchronous (b) 2D correlation spectrum of the model protein constructed from the simulated temperature-dependent IR spectra ranging from 40 to 65°C.

conformation by heating. The result suggests that contents of all other secondary structures, like $\beta$-sheet, $\beta$-turn and random conformation, are increasing. In other words, they are created by the conversion from $\alpha$-helix structure, which is consistent with the assumption made during the building of this model protein system. The merging of features for $\beta$-sheet and $\beta$-turn at 1682 and 1670 cm$^{-1}$, arising from the relative similarity of the pattern of intensity changes, makes the distinction of these bands difficult.

Features of the asynchronous spectrum for the narrower temperature range from 40 to 65°C (Fig. 3(b)) actually is very similar to the result obtained from the dataset (Fig. 2(b)) encompassing the entire temperature range from 40 to 90°C. Development of cross peaks is observed at the spectral coordinates corresponding to individual conformation bands. Signs of cross peaks also are identical to the ones found in Fig. 2(b). That means the apparent ambiguity in assigning the sequential order actually arises from the peak signs of synchronous cross peaks, which are influenced by the direction of signal intensity variations. The sequential order of spectral intensity changes determined by the signs of cross peaks can be summarized as: 1670 cm$^{-1}$ ($\beta$-turn$\uparrow$) $\rightarrow$ 1682, 1615 cm$^{-1}$ ($\beta$-sheet$\uparrow$) $\rightarrow$ 1654 cm$^{-1}$ ($\alpha$-helix$\downarrow$) $\sim$ 1645 cm$^{-1}$ (random$\uparrow$). The result is consistent with the lower temperature portion of the model depicted in Fig. 1(b).

Figure 4 shows the 2D correlation spectra constructed by using the IR spectra obtained at a higher temperature range from 65 to 90°C. The reference spectrum obtained by averaging the spectra along the higher temperature range is modified due to the presence of a substantial amount of disordered structure contribution. Features of synchronous spectrum (Fig. 4(a)) are also noticeably different from the previous one for lower temperature range (Fig. 3(a)). Signs of cross peaks between $\alpha$-helix (1654 cm$^{-1}$) and $\beta$-turn (1670 cm$^{-1}$) or $\beta$-sheet (1682 and 1615 cm$^{-1}$) are all positive, indicating that the populations of these secondary structures are changing in the same direction, i.e., decreasing together. Negative synchronous cross peaks are observed at the spectral coordinate near 1645 cm$^{-1}$ which corresponds to the random or disordered conformation. The result is very much consistent with the assumption used in building the model spectra based on the disordering of the protein secondary structures, which is taking place at an elevated temperature.

Features of the new asynchronous spectrum (Fig. 4(b)) are again very similar to the previous ones (Figs 2(b) and 3(b)) obtained from different temperature ranges. However, the distinction between 1682 and 1670 cm$^{-1}$ bands, corresponding to $\beta$-sheet and $\beta$-turn conformations, has become clearer in this
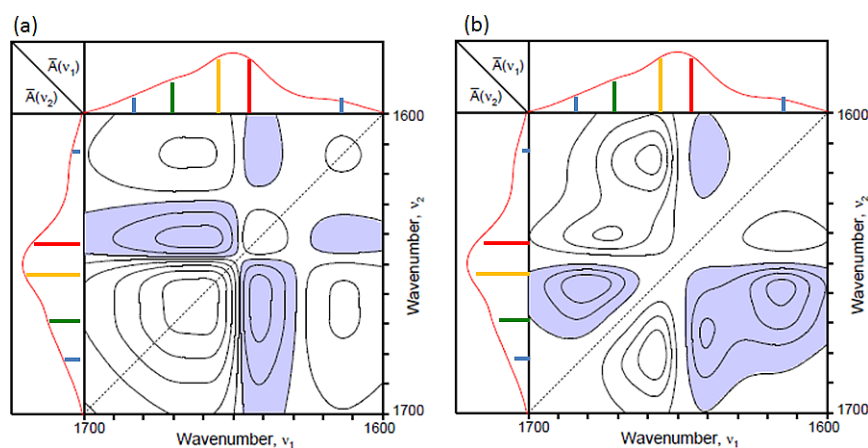


Fig. 4. Synchronous (a) and asynchronous (b) 2D correlation spectrum of the model protein constructed from the simulated temperature-dependent IR spectra ranging from 65 to 90°C.

asynchronous correlation spectrum, indicating much more pronounced difference in the temperature dependence of these bands in the higher temperature range as observed in Fig. 1(b). The sequential order of spectral intensity changes from peak signs now become: 1670 cm$^{-1}$ ($\beta$-turn$\downarrow$) $\rightarrow$ 1682 and 1615 cm$^{-1}$ ($\beta$-sheet$\downarrow$) $\rightarrow$ 1654 cm$^{-1}$ ($\alpha$-helix$\downarrow$) $\rightarrow$ 1645 cm$^{-1}$ (random$\uparrow$). The result is consistent with the higher temperature portion of the model depicted in Fig. 1(b).

2D correlation analysis of simulated protein IR spectra in this section has clearly illustrated that the technique is a powerful tool to be able to sort out the evolution of highly overlapped protein secondary structures undergoing thermally induced transformations. Overlapped peaks of conformation-sensitive IR absorption bands associated with the secondary structures are readily resolved by spreading the peaks along the second spectral dimension. The sequence of spectral intensity variations can be determined by examining the sings of cross peaks appearing at various spectral coordinates of 2D correlation spectra. The practical utility of this technique has been reflected in the increasing popularity in the use of 2D correlation spectroscopy for protein research [2,5–8,11,32,34,37,39].

## 4. New techniques in 2D correlation spectroscopy

In recent years, a number of additional techniques in 2D correlation spectroscopy have been introduced to further enhance the performance of this versatile tool. Some of the new techniques involve the pretreatment of spectral dataset prior to 2D correlation analysis. Other more recent development is based on a new methodology which is substantially different from the traditional 2D correlation analysis approach. Several noteworthy new types of 2D correlation spectroscopy techniques, which should be useful in the study of biological systems, are now highlighted here. More specifically, scaling treatment of spectral dataset commonly used in 2D correlation spectroscopy [22], a useful form of projection analysis to simplify congested 2D spectra [24], and newly introduced codistribution analysis to determine the sequential order of the presence of constituent species [27] are discussed in some detail.

### 4.1. Scaling of dataset used in 2D correlation spectroscopy

The remarkable sensitivity of 2D correlation spectra in detecting very subtle perturbation-induced spectral intensity changes can often reveal features not readily observable in the original set of spectral data. Sometimes it becomes desirable to further enhance subtle features of 2D correlation spectra by applying additional data pretreatment to selectively enhanced finer spectral details. Scaling operation, i.e., transformation of spectral data to attenuate or amplify the signals by multiplying the original spectra with an appropriate wavenumber-dependent weight function, is an effective technique. The treatment often brings out subtle features of spectral intensity variations, which are sometimes obscured by the dominating strong signals [22]. The merit of 2D correlation analysis of scaled spectral dataset is discussed here with an illustrative example.

Scaling of spectral data can potentially be a powerful and versatile technique to enhance the quality of spectra, but only if carried out judiciously [36]. Blind application of some scaling operation to spectral data may sometimes generate unwanted artifacts. There are many different ways to scale raw spectral data. The so-called unit-variance scaling or auto-scaling of data prior to 2D correlation analysis is one of the oldest and well-known techniques in the field. The idea of obtaining 2D correlation spectra by using the auto-scaled data was first proposed by Barton et al. [4] and later popularized by Šašić et al. [33].

Suppose a set of mean-centered dynamic spectra $\tilde{A}(\nu_j, t_i)$ are obtained by subtracting the average spectrum $\overline{A}(\nu_j)$ from the original perturbation-dependent spectral data $A(\nu_j, t_i)$, according to Eqs (1)

and (2). In the unit-variance scaling operation, dynamic spectra are scaled by the standard deviation $\sigma(\nu_j)$ at each wavenumber point. As it has already been pointed out, 2D synchronous spectrum $\Phi(\nu_1, \nu_2)$ is a form of covariance of spectral intensity fluctuations observed at two different wavenumbers, $\nu_1$ and $\nu_2$, along the perturbation axis. Thus, the standard deviation $\sigma(\nu_j)$ is readily obtained from the synchronous correlation intensity observed at the main diagonal position $\Phi(\nu_j, \nu_j)$, which directly corresponds to the variance $\sigma(\nu_j)^2$ of the spectral intensity fluctuations,

$$\sigma(\nu_j) = \sqrt{\frac{1}{m-1} \sum_{i=1}^{m} \tilde{A}(\nu_j, t_i)^2} = \sqrt{\Phi(\nu_j, \nu_j)}. \tag{14}$$

A new set of 2D correlation spectra for the unit-variance scaled dynamic spectra $\tilde{A}(\nu_j, t)/\sigma(\nu_j)$ are then given by

$$\Phi(\nu_1, \nu_2)^{Unit\ variance} = \Phi(\nu_1, \nu_2)/\big[\sigma(\nu_1) \cdot \sigma(\nu_2)\big], \tag{15}$$

$$\Psi(\nu_1, \nu_2)^{Unit\ variance} = \Psi(\nu_1, \nu_2)/\big[\sigma(\nu_1) \cdot \sigma(\nu_2)\big]. \tag{16}$$

It should be noted that the unit-variance scaled synchronous spectrum $\Phi(\nu_1, \nu_2)^{Unit\ variance}$ is equivalent to the Pearson's correlation coefficient for the pair of spectral intensity fluctuations observed at $\nu_1$ and $\nu_2$.

Figure 5 shows the unit-variance scaled 2D correlation spectra obtained from the simulated temperature-dependent IR spectra of the model protein (Fig. 1(c)). The appearance of the unit-variance scaled 2D spectra is substantially different from that of conventional 2D correlation spectra constructed without applying any scaling operation (Fig. 2). The auto-peaks prominent in the synchronous spectrum are no longer observable, as the correlation coefficient intensity at the main diagonal position is now normalized to unity. Correlation peaks tend to be much broader and have less distinct features. Even for asynchronous spectrum, it is difficult to identify the existence of two separate bands at 1682 and 1670 cm$^{-1}$, corresponding to $\beta$-sheet and $\beta$-turn conformations.

The auto-scaling or unit-variance scaling operation should provide purely correlational or coincidental aspect of the analysis without being influenced by the magnitude of signals. The expected advantage of
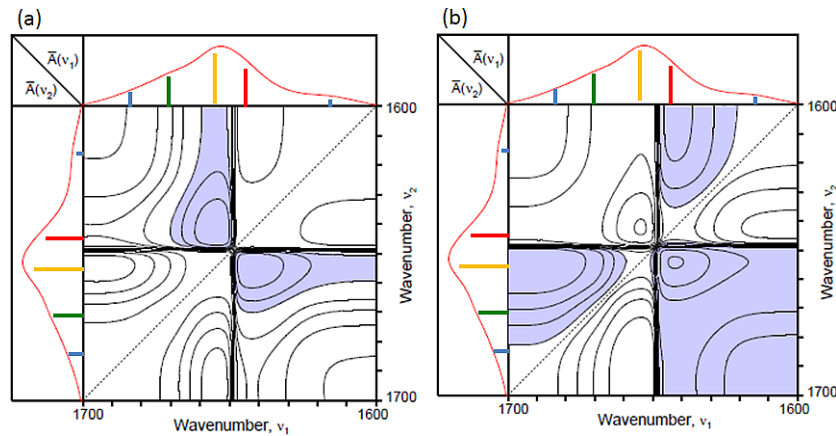


Fig. 5. Unit-variance scaled synchronous (a) and asynchronous (b) 2D correlation spectrum of the model protein constructed from the entire simulated temperature-dependent IR spectra.

normalizing the data prior to the correlation analysis is to mathematically suppress the overwhelming effect of strongly varying signals, which tends to dominate and obscure the contributions from equally important but more subtle spectral intensity variations. Indeed, the feature associated with the relatively weak spectral intensity variations of $\beta$-sheet at 1615 cm$^{-1}$ has become more visible in Fig. 5, while those for $\alpha$-helix and random conformation at 1654 and 1645 cm$^{-1}$ are less dominating in the auto-scaled 2D correlation spectra.

One of the major limitations of using unit-variance scaled data for 2D correlation is the unintended enlargement of the effect of small noise in the spectra, which is not so obvious in a clean simulated dataset, like the one used here. However, even a small amplitude noise typically present in real-world experimental spectra can become a very serious issue. The noise effect is especially severe in the spectral region with little absorption bands, as usually negligible small variations in the measurement can be dramatically amplified to create wild fluctuations in the resulting correlation intensities. To make the matter worse, such enlarged noise contributions may partially correlate with the real signals to produce numerous artifacts. Fortunately, there is an alternative scaling technique to circumvent the apparent limitation of auto-scaling operation.

Pareto scaling is a pragmatic compromise between somewhat problematical unit-variance scaling operation mentioned above vs. avoiding any scaling pretreatment at all for raw data. The concept of Pareto scaling was first introduced [38] by Svante Vold in 1993, who also coined this term, named after an influential Italian economist of the 19th century Vilfredo Pareto. In Pareto scaling, the spectral data are scaled not by the standard deviation but by the square root of the standard deviation. Thus, Pareto-scaled 2D correlation spectra become

$$\Phi(\nu_1, \nu_2)^{Pareto} = \Phi(\nu_1, \nu_2) / \sqrt{\sigma(\nu_1) \cdot \sigma(\nu_2)}, \tag{17}$$

$$\Psi(\nu_1, \nu_2)^{Pareto} = \Psi(\nu_1, \nu_2) / \sqrt{\sigma(\nu_1) \cdot \sigma(\nu_2)}. \tag{18}$$

Pareto scaling avoids some of the annoying limitations of unit-variance scaling operation, such as exaggeration of the effect of small noise or the complete loss of useful auto-peak information, while achieving the intended goal of suppressing the overwhelming effect of a few dominant signals to bring out many other subtle details of weaker signals without being obscured.

Figure 6 shows the 2D correlation spectra constructed from the Pareto-scaled spectral data. Auto-peaks disappeared by the unit-variance scaling in Fig. 5(a) are now brought back in the Pareto-scaled synchronous spectrum (Fig. 6(a)). So are the finer and more distinct features of cross peaks. Relative strength of correlation intensities has been modified. Weaker signal contributions from $\beta$-sheet and $\beta$-turn, especially from the band for $\beta$-sheet at 1615 cm$^{-1}$, have become more visible by applying Pareto scaling compared to the conventional 2D correlation spectra (Fig. 2) constructed without any scaling pretreatment operation. Although most of the pertinent information can be seen in the conventional 2D correlation spectra, Pareto-scaled 2D spectra tend to bring out finer spectral features more clearly to assist easier interpretation of data.

### 4.2. Projection analysis

Projection 2D correlation analysis [24] is a useful tool to dramatically simplify highly congested 2D correlation spectra often encountered in practice. As already indicated in Eqs (6) and (7), 2D correlation spectra can be expressed in terms of matrices $\Phi$ and $\Psi$ obtained by the simple matrix multiplications
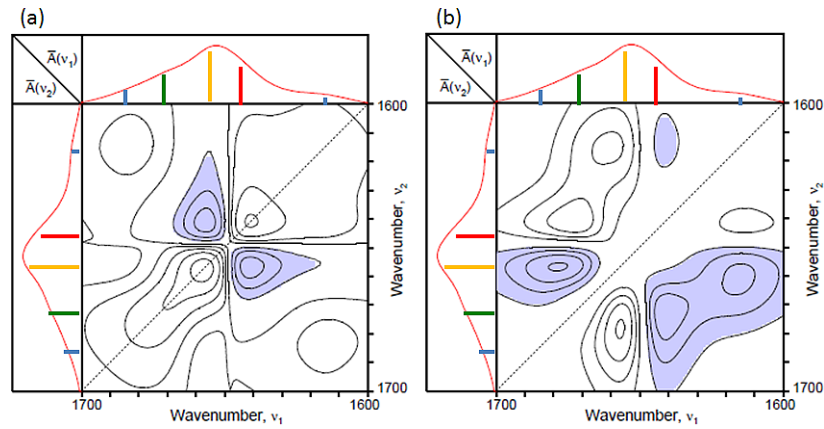
Fig. 6. Pareto-scaled synchronous (a) and asynchronous (b) 2D correlation spectrum of the model protein constructed from the entire simulated temperature-dependent IR spectra.

of $m$ by $n$ spectral data matrix $\tilde{\mathbf{A}}$ with $m$ dynamic spectra along the perturbation axis and $n$ wavenumber points for each spectrum. One can apply matrix-based mathematical projection operations as an effective filtering method to transform spectral data to those containing much simplified features. This technique is gaining popularity in streamlining 2D correlation spectra for easier interpretation [12,35]. Detailed discussion on the subject is found in the original article [24]. Only a very brief description of the technique which is pertinent to the specific example is provided here.

Projecting vector used in the projection operation can be obtained from various sources. One of the most useful sources of a projecting vector actually is a column of the original spectral data matrix. Such a projecting vector can be used to carry out the so-called *positive null-space projection* to filter out specific contributions which are overlapping with other features of 2D correlation spectra. In this operation, a column $\mathbf{y}$ of the data matrix $\tilde{\mathbf{A}}$ at a target wavenumber is selected that contains the specific feature to be filtered out. For example, it can be the temperature-dependent spectral intensity variations observed at 1670 cm$^{-1}$ which should be containing predominantly the contribution from $\beta$-turn structure (Fig. 1(d)).

The normalized vector of $\mathbf{y}$ is given by

$$\mathbf{u}_y = \mathbf{y}/\left(\mathbf{y}^{\mathrm{T}}\mathbf{y}\right)^{-1/2} \tag{19}$$

which can be used to calculate the loading vector $\mathbf{v}_A$ of the portion of the data of matrix $\tilde{\mathbf{A}}$ projected onto the space spanned by the column $\mathbf{y}$,

$$\mathbf{v}_A = \tilde{\mathbf{A}}\mathbf{u}_y. \tag{20}$$

Now the positive loading vector $\mathbf{v}_{+A}$ is obtained by simply replacing all negative elements of $\mathbf{v}_A$ with zero. The positive null-space projected dynamic spectra $\tilde{\mathbf{A}}_{+N}$ and the corresponding 2D correlation spectra are then given by

$$\tilde{\mathbf{A}}_{+N} = \tilde{\mathbf{A}} - \mathbf{u}_y\mathbf{v}_{+A}^{\mathrm{T}}, \tag{21}$$

$$\boldsymbol{\Phi}_{+N} = \frac{1}{m-1}\tilde{\mathbf{A}}_{+N}^{\mathrm{T}}\tilde{\mathbf{A}}_{+N}, \tag{22}$$
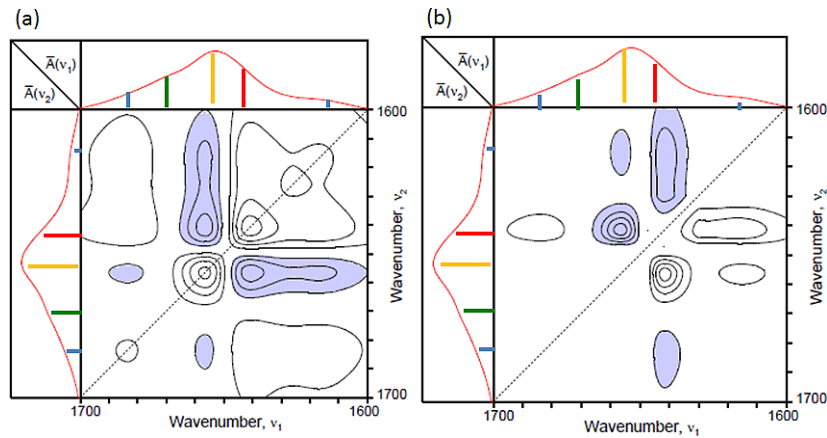
Fig. 7. Synchronous (a) and asynchronous (b) 2D correlation spectrum of the model protein constructed from the entire simulated temperature-dependent IR spectra after the positive null-space projection operation with the spectral intensity variation of $\beta$-turn at 1670 cm$^{-1}$.

$$\mathbf{\Psi}_{+N} = \frac{1}{m-1}\tilde{\mathbf{A}}_{+N}^{\mathrm{T}}\mathbf{N}\tilde{\mathbf{A}}_{+N}. \tag{23}$$

It should be noted that the positive null-space projected data matrix $\tilde{\mathbf{A}}_{+N}$ no longer contains any feature of spectral intensity variation pattern which is similar (not only in the shape but also in the direction) to the one represented by the projecting column vector $\mathbf{y}$.

Figure 7 shows 2D correlation spectra constructed from the temperature-dependent IR spectra of model protein, which has been subjected to the positive null-space projection with the intensity variation of $\beta$-turn at 1670 cm$^{-1}$ as the projecting vector $\mathbf{y}$. It is clear that features related to $\beta$-turn conformation are completely removed from the correlation spectra. By removing the overlapping $\beta$-turn contribution, neighboring bands at 1682 and 1654 cm$^{-1}$ for $\beta$-sheet and $\alpha$-helix in the synchronous spectrum (Fig. 7(a)) are now much more clearly resolved, as indicated by the appearance of independent auto-peaks and cross peaks. Similar effect of clear peak separation is also observed in the asynchronous spectrum (Fig. 7(b)).

Similarly, by applying the positive null-space projection operation with the intensity variation at 1615 cm$^{-1}$ as the projector vector, the removal of the contribution from $\beta$-sheet conformation to the dataset is achieved. Using this band is preferred to the other option of using 1682 cm$^{-1}$, which has substantial overlap with the contribution from $\beta$-turn conformation. Figure 8 shows the resulting 2D correlation spectra, which no longer contains features associated with $\beta$-sheet. They clearly depict the different population dynamics of three remaining components: $\alpha$-helix, $\beta$-turn and random conformations. Cross peaks at higher wavenumber region extend themselves beyond 1682 cm$^{-1}$, indicating that the projection operation removed only the $\beta$-sheet contribution but retained the overlapped $\beta$-turn contribution.

### 4.3. Codistribution analysis

Two-dimensional codistribution spectroscopy is a technique designed specifically for the analysis of population dynamics, such as temporal distribution of conformational secondary structures within a protein during the denaturation and aggregation process [27]. The technique focuses on the features of
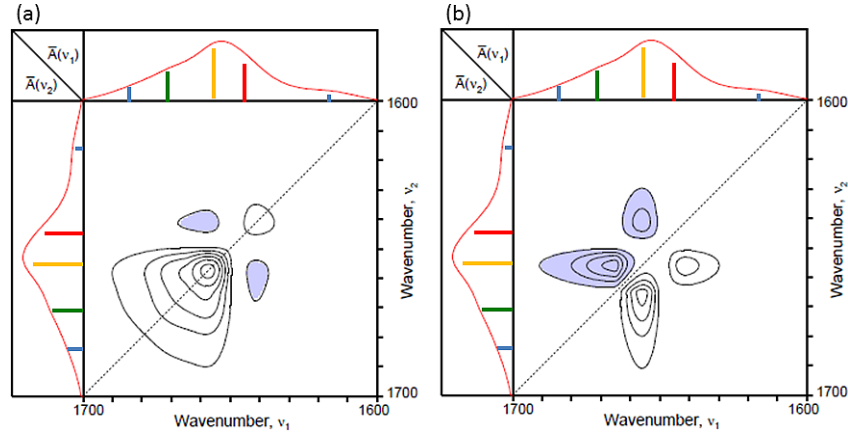
Fig. 8. Synchronous (a) and asynchronous (b) 2D correlation spectrum of the model protein constructed from the entire simulated temperature-dependent IR spectra after the positive null-space projection operation with the spectral intensity variation of $\beta$-sheet at 1615 cm$^{-1}$.

spectral signals reflecting the distributed presence of contributing species, instead of the variation patterns of perturbation-induced spectral intensities traditionally analyzed by 2D correlation spectroscopy. Codistribution analysis utilizes the first moment or the center of gravity of spectral intensity distributions for the comparison of distributions of intensities observed at two different wavenumbers along the perturbation axis. The technique is especially useful in identifying intermediate species arising during the course of a multi-staged reaction process.

Traditional 2D correlation analysis provides the information on the sequential order of spectral intensity variations. For a system involving species which can both increase and decrease in population during the observation, it is sometimes necessary to segment the dataset into smaller blocks, as already demonstrated in Section 3.4 of this paper, to obtain the unambiguous assignment of the population dynamics. In contrast, 2D codistribution spectroscopy, based on the moment analysis of distribution densities of spectral intensities along the perturbation axis, directly provides the information about the sequential order of distributed presence of constituent species, such as different secondary structures of a protein.

For a set of $m$ dynamic spectra $\tilde{A}(\nu_j, t_i)$ with the average $\overline{A}(\nu_j)$ as the reference spectrum obtained according to Eqs (1) and (2), the asynchronous 2D codistribution spectrum $\Delta(\nu_1, \nu_2)$ is given by

$$\Delta(\nu_1, \nu_2) = \frac{T(\nu_1, \nu_2)}{m(m-1)} \sum_{i=1}^{m} i \left\{ \frac{\tilde{A}(\nu_2, t_i)}{\overline{A}(\nu_2)} - \frac{\tilde{A}(\nu_1, t_i)}{\overline{A}(\nu_1)} \right\}. \tag{24}$$

The total joint variance $T(\nu_1, \nu_2)$ is given in terms of the standard deviation $\sigma(\nu)$ of the spectral intensity variation, which in turn is related to the synchronous correlation intensity at the main diagonal position (Eq. (14))

$$T(\nu_1, \nu_2) = \sigma(\nu_1) \cdot \sigma(\nu_2) = \sqrt{\Phi(\nu_1, \nu_1) \cdot \Phi(\nu_2, \nu_2)}. \tag{25}$$

Asynchronous codistribution spectrum has the following properties useful for identifying the sequence of presence of species [27]. For a cross peak with positive sign, i.e., $\Delta(\nu_1, \nu_2) > 0$, the presence of the species reflected by the spectral intensity at $\nu_1$ is distributed predominantly at the earlier stage along
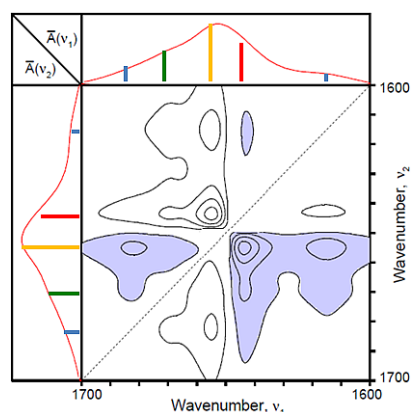
Fig. 9. Asynchronous 2D codistribution spectrum of the model protein constructed from the entire simulated temperature-dependent IR spectra.

the perturbation axis compared to that for $\nu_2$. In contrast, if we have the observation of a negative cross peak, i.e., $\Delta(\nu_1, \nu_2) < 0$, the order of presence is reversed. In the case of $\Delta(\nu_1, \nu_2) \approx 0$, the average distributions of the spectral intensities observed at two wavenumbers over the course of observation are similar, that is they coexist together.

Figure 9 shows the 2D codistribution spectrum of the simulated temperature-dependent model protein IR spectra shown in Fig. 1. The interpretation of this 2D codistribution spectrum is straightforward. The population of $\alpha$-helix secondary structure represented by the band centered around 1654 cm$^{-1}$ is most predominantly present at the earlier stage on the perturbation axis, i.e., at a lower temperature region. As the temperature is raised, the population of $\alpha$-helix is diminished. In turn, the population of $\beta$-turn conformation (1670 cm$^{-1}$) arises followed by that of $\beta$-sheet (1682 and 1615 cm$^{-1}$) structure. Finally, the steady increase in the population of random structure (1645 cm$^{-1}$) at a higher temperature range is observed. The picture depicted on the population dynamics here is very much consistent with the assumed behavior of constituents in the simulation model (Fig. 1(b)).

## 5. Concluding remarks

2D correlation spectroscopy has become a very popular tool in the field of life science due to its versatility and relative ease of use. The technique can be utilized with a number of spectroscopic and other analytical probes, such as IR and Raman spectroscopy, for a very broad range of sample systems, including proteins and other biomolecules, by employing different types of external perturbations, like temperature change or chemical stimuli, to induce spectral variations. 2D correlation analysis of simulated temperature-dependent IR spectra of protein clearly demonstrated that this technique can yield valuable information. Segmentation of dataset becomes helpful in assigning the sequence of spectral intensity variations when the dynamics of some species involve both increase and decrease in the population during the observation.

A number of new types of techniques have lately become available to enhance the performance of 2D correlation spectroscopy. Scaling of spectral data, especially the use of Pareto scaling, brings out the finer details of 2D correlation spectra sometimes obscured by the dominant contribution of strongly varying signals. Projection approach can be a powerful tool in simplifying congested 2D correlation

spectra constructed from highly overlapped spectral data by selectively filtering out the contribution of particular species. Newly introduced 2D codistribution analysis provides the simple determination of the sequence of the presence of constituent species during a complex series of transformation processes, like denaturation and aggregation of a protein. Many other techniques potentially useful in the applications for life science research are also mentioned in the recent review article [25].

## References

[1] J.L.R. Arrondo and F.M. Goñi, Structure and dynamics of membrane proteins as studied by infrared spectroscopy, *Prog. Biophys. Mol. Biol.* **72** (1999), 367–405.

[2] L. Ashton, C. Johannessen and R. Goodacre, The importance of protonation in the investigation of protein phosphorylation using Raman spectroscopy and Raman optical activity, *Anal. Chem.* **83** (2011), 7978–7983.

[3] A. Barth, Infrared spectroscopy of proteins, *Biochim. Biophys. Acta* **1767** (2007), 1073–1101.

[4] F.E. Barton II., D.S. Himelsbach, J.H. Duckworth and M.J. Smith, Two-dimensional vibration spectroscopy: Correlation of mid- and near infrared regions, *Appl. Spectrosc.* **46** (1992), 420–429.

[5] I. de la Arada, J.-P. Julien, B.G. de la Torre, N. Huarte, D. Andreu, E.F. Pai, J.L. Arrondo and J.L. Nieva, Structural constraints imposed by the conserved fusion peptide on the HIV-1 gp-41 epitope recognized by the broadly neutralizing antibody 2F5, *J. Phys. Chem. B* **113** (2009), 13626–13637.

[6] I. de la Arada, C. Seiler and W. Mäntele, Amyloid fibril formation from human and bovine serum albumin followed by quasi-simultaneous Fourier-transform infrared (FT-IR) spectroscopy and static light scattering (SLS), *Eur. Biophys. J.* **41** (2012), 931–938.

[7] L.V. Sosa, E. Alifaro, J. Santiago, D. Naváez, M.C. Rosado, A. Rodríguez, A.M. Gómez, E.R. Schreiter and B. Pastrana-Rios, The structure, molecular dynamics, and energetics of centrin-melittin complex, *Proteins: Struct. Func. Bioinformatics* **72** (2011), 3132.

[8] H. Fabian, H.H. Mantsch and C.P. Schultz, Two-dimensional IR correlation spectroscopy: Sequential events in the unfolding process of the lambda Cro-V55C repressor protein, *Proc. Natl. Acad. Sci. USA* **96** (1999), 13153–13158.

[9] E. Goomaghtigh, V. Cabiaux and J.-M. Ruysschaert, Determination of soluble and membrane protein structure by Fourier transform infrared spectroscopy. I. Assignments and model compounds, in: *Physicochemical Methods in the Study of Biomembranes*, Subcellular Biochemistry, Vol. 23, 1994, pp. 329–362.

[10] M. Jackson and 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.

[11] Y.M. Jung, B. Czarnik-Matusewicz and S.B. Kim, Characterization of concentration-dependent infrared spectral variations of urea aqueous solutions by principal component analysis and two-dimensional correlation spectroscopy, *J. Phys. Chem. B* **108** (2004), 13008–13014.

[12] M.K. Kim, S.R. Ryu, I. Noda and Y.M. Jung, Projection 2D correlation analysis of biodegradable P(HB-*co*-HHx)/PEG blend, *Vib. Spectrosc.* **60** (2012), 163–167.

[13] S. Krimm and J. Bandekar, Vibrational spectroscopy and conformation of peptides, polypeptides, and proteins, *Adv. Prot. Chem.* **38** (1986), 181–367.

[14] I. Noda, Two-dimensional infrared (2D IR) spectroscopy of synthetic and biopolymers, *Bull. Am. Phys. Soc.* **31** (1986), 520.

[15] I. Noda, Two-dimensional infrared (2D IR) spectroscopy, *J. Am. Chem. Soc.* **111** (1989), 8116–8118.

[16] I. Noda, Two-dimensional infrared (2D IR) spectroscopy. Theory and applications, *Appl. Spectrosc.* **44** (1990), 550–561.

[17] I. Noda, A generalized two-dimensional correlation method applicable to infrared, Raman, and other types of spectroscopy, *Appl. Spectrosc.* **47** (1993), 1329–1336.

[18] I. Noda, Progress in 2D correlation spectroscopy, in: *Two-Dimensional Correlation Spectroscopy*, Y. Ozaki and I. Noda, eds, AIP Press, Melville, 2000, pp. 3–17.

[19] I. Noda, Determination of two-dimensional correlation spectra using the Hilbert transform, *Appl. Spectrosc.* **54** (2000), 994–999.

[20] I. Noda, Advances in two-dimensional correlation spectroscopy, *Vib. Spectrosc.* **36** (2004), 143–165.

[21] I. Noda, Progress in two-dimensional (2D) correlation spectroscopy, *J. Molec. Struct.* **799** (2006), 2–15.

[22] I. Noda, Scaling techniques to enhance two-dimensional correlation spectra, *J. Molec. Struct.* **883–884** (2008), 216–227.

[23] I. Noda, Two-dimensional correlation spectroscopy – Biannual survey 2007–2009, *J. Molec. Struct.* **974** (2010), 3–24.

[24] I. Noda, Projection two-dimensional correlation analysis, *J. Molec. Struct.* **974** (2010), 116–126.

[25] I. Noda, Frontiers of two-dimensional correlation spectroscopy. Part 1. New concepts and noteworthy developments, *J. Molec. Struct.* **1069** (2014), 3–22.

[26] I. Noda, Frontiers of two-dimensional correlation spectroscopy. Part 2. Perturbation methods, fields of applications, and types of analytical probes, *J. Molec. Struct.* **1069** (2014), 23–49.

[27] I. Noda, Two-dimensional codistribution analysis to determine the sequential order of distributed presence of species, *J. Molec. Struct.* **1069** (2014), 50–59.

[28] I. Noda, A.E. Dowrey and C. Marcott, Recent developments in two-dimensional infrared (2D IR) correlation spectroscopy, *Appl. Spectrosc.* **47** (1993), 1317–1323.

[29] I. Noda, A.E. Dowrey, C. Marcott, G.M. Story and Y. Ozaki, Generalized two-dimensional correlation spectroscopy, *Appl. Spectrosc.* **54** (2000), 236A–248A.

[30] I. Noda and Y. Ozaki, *Two-Dimensional Correlation Spectroscopy – Applications in Vibrational and Optical Spectroscopy*, Wiley, Chichester, 2004.

[31] Y. Ozaki and I. Noda, Two-dimensional vibrational correlation spectroscopy in biomedical sciences, in: *Encyclopedia of Analytical Chemistry: Instrumentation and Applications*, R.A. Meyers, ed., Wiley, Chichester, 2000, pp. 322–340.

[32] E. Saguer, P.A. Alvarez, J. Sedman and A.A. Ismail, Study of the denaturation/aggregation behavior of whole porcine plasma and its protein fractions during heating under acidic pH by variable-temperature FTIR spectroscopy, *Food Hydro-colloids* **33** (2013), 402–414.

[33] S. Šašić and Y. Ozaki, Statistical two-dimensional correlation spectroscopy: Its theory and applications to sets of vibrational spectra, *Anal. Chem.* **73** (2001), 2294–2301.

[34] V. Shashilov, M. Xu, V.V. Ermolenkov, L. Fredriksen and I.K. Lednev, Probing a fibrillation nucleus directed by deep ultraviolet Raman spectroscopy, *J. Am. Chem. Soc.* **129** (2007), 6972–6973.

[35] H. Shinzawa, K. Awa, I. Noda and Y. Ozaki, Pressure-induced variation of cellulose tablet studied by two-dimensional (2D) near-infrared (NIR) correlation spectroscopy in conjunction with projection pretreatment, *Vib. Spectrosc.* **65** (2013), 28–35.

[36] R.A. van den Berg, H.C.J. Hoefsloot, J.A. Weterhuis, A.K. Smilde and M.J. van den Werf, Centering, scaling and transformations: Improving the biological conformation content of metabolomics data, *BMC Genomics* **7** (2006), 142–157.

[37] Q. Wang, W. Xu, P.W. Wu, H. Zhang, C. Cai and B. Zhao, New insights into the effects of thermal treatment on the catalytic activity and conformational structure of glucose studied by electrochemistry, IR spectroscopy, and theoretical calculation, *J. Phys. Chem. B* **114** (2010), 12754–12764.

[38] S. Wold, E. Johansson and M. Cocchi, PLS-partial least squares projections to latent structures, in: *3D QSAR in Drug Design: Theory, Methods and Applications*, H. Kobinyi, ed., ESCOM Science Publishers, Leiden, 1993, pp. 523–550.

[39] Y. Wu, K. Murayama, B. Czarnik-Matusewicz and Y. Ozaki, Two-dimensional attenuated total reflection/infrared correlation spectroscopy studies on concentration and heat-induced structural changes on human serum albumin in aqueous solutions, *Appl. Spectrosc.* **56** (2002), 1186–1193.

[40] B. Yuan, K. Murayama and H. Yan, Study of thermal dynamics of defatted bovine serum albumin in $D_2O$ solution by Fourier transform infrared spectra and evolving factor analysis, *Appl. Spectrosc.* **61** (2007), 921–927.