Spectroscopy 18 (2004) 49–58
IOS Press

A two-dimensional IR spectroscopic (2D-IR) simulation of protein conformational changes

José Luis R. Arrondo [a,*], Ibon Iloro [a], Julián Aguirre [b] and Félix M. Goñi [a]

[a] *Unidad de Biofísica (Centro Mixto CSIC-UPV/EHU) and Departamento de Bioquímica, Universidad del País Vasco, PO Box 644, E-48080 Bilbao, Spain*

[b] *Departamento de Matemáticas, Facultad de Ciencias, Universidad del País Vasco, PO Box 644, E-48080 Bilbao, Spain*

Abstract. Two-dimensional IR correlation spectroscopy (2D-IR) is a novel method that provides the analysis of infrared spectra with the capacity to differentiate overlapping peaks and to distinguish between in-phase and out-of-phase spectral responses. Artificial spectra originated from protein amide I band component parameters have been used to study their variation in the correlation maps. Using spectra composed of one, two or three Gaussian peaks, behaviour patterns of the bands in the synchronous and asynchronous maps have been originated, with changes in intensity, band position and bandwidth. Intensity changes produce high-intensity spots in the synchronous spectra, whereas only noise is observed in the asynchronous spectra. Band shifting originates more complex patterns. In synchronous spectra, several spots are generated at the beginning and at the end of the shifting band. Also, characteristic asynchronous spectra with butterfly-like shapes are formed showing the trajectory of the shift. Finally, synchronous maps corresponding to band broadening reveal several spots at peak inflection points, related to the zones with higher intensity variation. The asynchronous spectra are very complex but they follow a characteristic symmetric pattern. Furthermore, examples of maps obtained from polypeptides and proteins using temperature as the perturbing factor are interpreted in terms of the patterns obtained from artificial bands.

1. Introduction

Infrared spectroscopy is a choice method, either as an alternative option or as a complement to the high-resolution techniques in protein studies, because of the wealth of information provided and because of the sensitivity of the technique to changes in protein environment. Even if infrared spectroscopy was first applied to proteins as early as in 1952 [14] before any detailed X-ray results were available, its use with proteins in physiological environments has only been possible after the revolution in instrumentation produced by the development of microcomputers that allowed the design of instruments based on the Michelson interferometer and the fast Fourier transform [3]. The increase in signal-to-noise ratio obtained with these instruments allows the subtraction of the aqueous buffer, thus obtaining spectra whose bands contain information on the protein in its native environment, and are free from spectral interference by the solvent. In principle, a structure as large as a protein would give rise to an enormous number of overlapping vibrational modes obscuring the retrievable information, but because of the repeating patterns of the biological molecules, e.g., the secondary structure of the protein backbone, the spectra are much simpler, and useful structural information can be obtained.

* Corresponding author: José Luis R. Arrondo, Tel.: +34 946 012 485; Fax: +34 944 648 500; E-mail: gbproarj@lg.ehu.es.

Structural analysis usually implies a mathematical approach in order to extract the information contained in the conformationally sensitive composite bands, designated in IR spectroscopy as "amide bands", obtained from peptide bond vibrations. Commonly used methods of analysis imply narrowing the intrinsic bandwidths to visualize the overlapping band components and then decomposing the original band contour into its components by means of an iterative process. The various components are finally assigned to protein or subunit structural features [2]. More recently, a new approach has been introduced by Noda (for a review see [10]) in the study of vibrational spectroscopy, the so-called generalized 2D-IR correlational spectroscopy. This approach, essentially different from 2D-NMR spectroscopy, uses correlation analysis of the dynamic fluctuations caused by an external perturbation to enhance spectral resolution without assuming any lineshape models for the bands. The technique was intended for the study of polymers and liquid crystals, but it has been recently applied to proteins [5,7,11]. In the latter case, the perturbation can be achieved through changes in temperature, pH, ligand concentration, lipid-to-protein ratio, etc. The power of the 2D correlation approach results primarily in an increase of the spectral resolution by dispersal of the peaks along a second dimension that also reveals the time-course of the events induced by the perturbation [12]. Correlation between bands are found through the so-called synchronous and asynchronous spectra that correspond to the real and imaginary parts of the cross-correlation of spectral intensity at two wavenumbers. In a synchronous 2D map, the peaks located in the diagonal (autopeaks) correspond to changes in intensity induced (in our) case by temperature, and are always positive. The cross-correlation peaks indicate an in-phase relationship between two bands involved. Asynchronous maps show not-in-phase cross-correlation between the bands and this gives an idea of the time-course of the events produced by the perturbation.

In the present work we have used mathematical simulations of spectra composed of one, two or three Gaussian peaks with parameters obtained from protein amide I bands to originate behaviour patterns in the synchronous and asynchronous maps, examining at changes in intensity, band position and bandwidth, and comparing them with maps obtained from real proteins using temperature as the perturbing agent.

2. Materials and methods

Mathematica 4.1 (Wolfram Research, Inc) was used to generate and manipulate Gaussian peaks from the general Gaussian equation (Eq. (1)) to create the corresponding bidimensional spectrum.

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}}. \tag{1}$$

In all simulations each dynamic spectrum consisted of 13 spectra with 130 points equivalent to a nominal resolution of 2 cm^{-1}. The changes were assumed to be linear.

To obtain two-dimensional infrared correlation spectra the approach proposed by Noda [10] has been followed. From the response of the system to the perturbation a dynamic spectrum is obtained. The type of physical information contained in a dynamic spectrum is determined by the perturbation method. Once the dynamic spectrum has been obtained as a matrix formed by the spectra ordered according to the change produced by the perturbation, the Fourier transform gives two components, the real one corresponding to the synchronous spectrum and the imaginary one corresponding to the asynchronous one. In terms of calculation, instead of using Fourier transform that would require large computation

times, an adequate numerical evaluation of 2D correlation intensity is used. Thus, the synchronous 2D correlation intensity can be expressed as [5]

$$\Phi(v_1, v_2) = \frac{1}{m-1} \sum_{j=1}^{m} \overline{y}_j(v_1) \cdot \overline{y}_j(v_2), \tag{2}$$

where $\overline{y}_j(v_i)$ is the dynamic spectra calculated from the spectral intensities as a deviation from a reference spectrum at a point of physical variable t_j.

The computation of asynchronous 2D correlation intensity is somewhat more complicated. We have used two approaches: (i) using the Hilbert transform and (ii) a direct procedure, obtaining similar results (for a detailed discussion on the asynchronous calculation, see [10]).

Insulin and poly-L-proline were obtained from Sigma (Poole, UK). They were dissolved in 10 mM phosphate buffer, pD 2.0 and measured in a Mattson Galaxy 400 IR spectrometer equipped with a DTGS detector, at a nominal resolution of 4 cm^{-1} with a level of zero filling. The samples were placed onto a Specac cell equipped with CaF_2 windows and a 50 μm spacer. Samples were heated from 30 to 70°C and the buffer contribution subtracted. For the correlation studies, the spectra were deconvolved using a FWHH of 18 cm^{-1} and a $k = 2$ [3].

3. Results and Discussion

3.1. Intensity changes

Changes in intensity were observed in spectra composed of one, two or three peaks. The corresponding correlation maps are represented in Fig. 1. The bottom, middle and top maps correspond respectively to spectra generated from one, two and three peaks. Left-hand traces correspond to synchronous spectra while right-hand traces are from asynchronous spectra. Negative peaks are represented as shaded areas. Asynchronous spectra consist only of noise when only band intensity has been changed. Similar results have been obtained in simulations of 2D electronic correlation spectra [8] or using Lorentzian band-shapes [6]. The lack of bands in the asynchronous spectra is observed in protein thermal denaturation after aggregation has taken place [11], and this is not a noise-artifact due to the decrease in spectral intensity after aggregation, because artificial spectra are noise-free and yet the asynchronous spectra consists only of low-amplitude noise. Thus the lack of bands is a characteristic of the denatured proteins once bandshift due to aggregation has been completed. On the other hand, considering that synchronous spectra reflect in-phase changes, differences in peak intensity are expected to produce changes. It can be seen that the complexity of the synchronous spectra changes as a function of the number of peaks involved. If only one peak is changing, one autopeak is seen. If two or more peaks are involved, the respective cross-peaks show in which way intensities are changing. If both increase or decrease simultaneously, the cross-peaks are positive, and negative if one peak is increasing and the other one is decreasing.

3.2. Band shifting

Changes in band position were analysed in a similar way. Simulated bands were constructed with one, two or three components changing simultaneously, and the synchronous and asynchronous maps are represented in Fig. 2. In all the synchronous maps, autopeaks are located at the initial and final

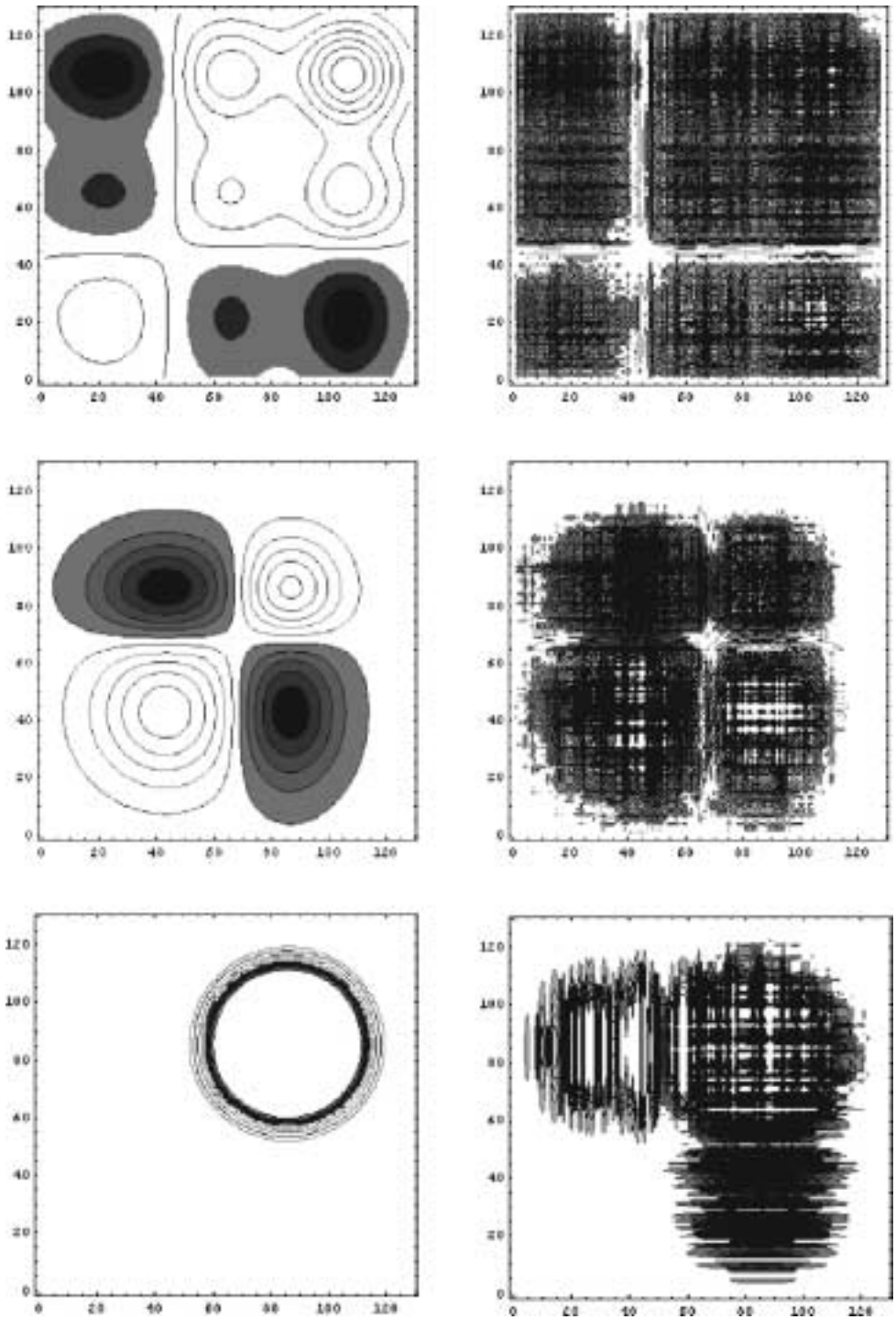

Fig. 1. Bidimensional correlation maps corresponding to changes in intensity of a band composed of one peak (bottom), two peaks (middle) or three peaks (top). Synchronous maps are located at the left and asynchronous to the right. The X and Y axis correspond to the number of the point in the artificial curve and the Z axis is the correlational intensity. Roughly they represent a protein amide I band in a D_2O buffer. Negative peaks are shaded.

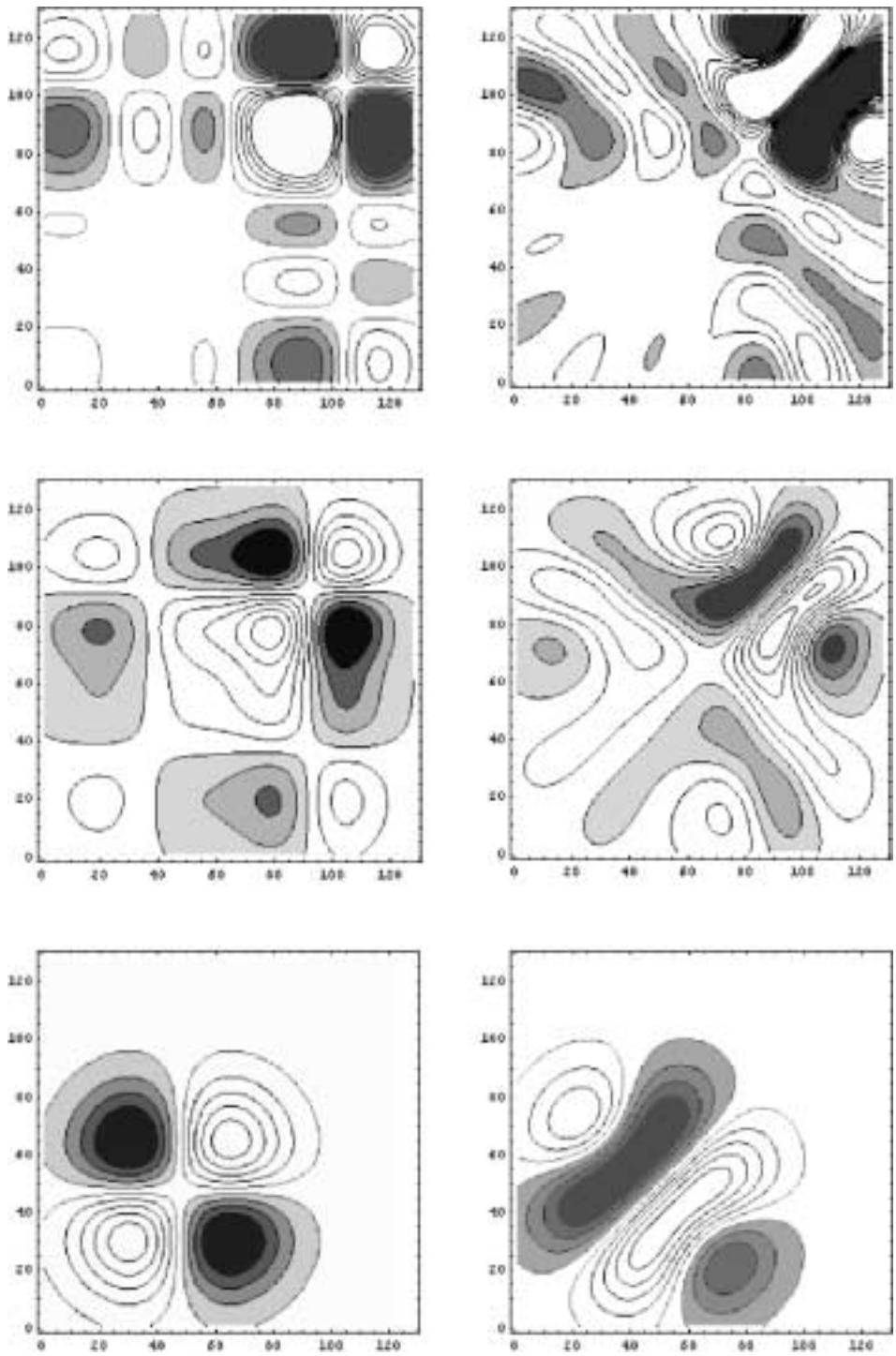

Fig. 2. Bidimensional correlation maps corresponding to changes in position (band shifting) of a band composed of one peak (bottom), two peaks (middle) or three peaks (top). Synchronous maps are located at the left and asynchronous to the right. Axis are as in Fig. 1.

positions of the maxima during band shifting. This is the reason why two peaks are observed in the synchronous map corresponding to a change of band position in one peak. Cross-peaks are positioned next to the equivalent autopeaks with a negative intensity since the intensity at a given frequency at the beginning of the shift decreases, whereas intensity at a different frequency will increase after the shift. The characteristic pattern for a band shift is observed in the asynchronous spectra as a butterfly [9] or banana shape, depending on wether we consider the full symmetric map, or only one side of the diagonal. These peaks include the frequencies through which the maximum has moved throughout the band shifting. As in the synchronous spectra, there are also cross-peaks of the opposite sign between the beginning and the end of the "bananas".

3.3. Bandwidth

The same approach as in the two previous models was used for band broadening. In this case the pattern obtained is more complex. It has to be taken into account that changes in bandwidth in proteins are not usually a result of a phase transition as in the case of lipids [1], but they are rather the consequence of changes in conformation and subsequent variation in the area of the band associated to that conformation. Hence, changes in bandwidth can be clearly discerned because a more complicated correlation map is produced, but they are not easily assigned (see below). In the case of the models studied, the resulting synchronous maps (Fig. 3) show patterns with several autopeaks located at the peak inflection points, since these are the frequencies at which greater intensity variations occur throughout the band broadening. These autopeaks are associated with positive cross-peaks because both points are changing their intensity in the same direction. The asynchronous spectrum is much more complicated, showing an increased number of peaks with different sizes and intensities. A similar result was obtained previously in an analysis of the electronic spectrum with the 2D correlation method [8].

3.4. Studies with experimentally-obtained spectra

The maps obtained from spectral simulations can be used to study changes in the spectra of polypeptides and proteins. Even if these structures are more complex and highly variable, they can be related to 2D maps similar to those obtained from simulations and consequently more information can be extracted from the infrared spectra. Homopolypeptides have been used extensively to assign infrared bands to secondary structures [13] because of their ability to adopt regular-canonical secondary structure conformations without tertiary structure. Polyproline is a particular example since its inability to form hydrogen bonds between peptide bonds produces the so-called poly-proline I and II helices. Poly-proline II helix (PPII), is left-handed with dihedral angles closer to β-sheet than to α-helix. Since no tertiary structure is involved, temperature produces a single transition from the PPII helix to unordered conformation without residual structure. Such transition can be easily studied by 2D-IR spectroscopy. The transition in a D_2O buffer is shown in Fig. 4. The perturbation produces a decrease in the intensity of the peak at 1623 cm^{-1}, associated to the PolyProII helix, correlated with an increase in a band at 1642 cm^{-1} corresponding to unordered structure. Also, a difference in bandwidth between both bands is apparent. In Fig. 5 the 2D-IR maps are shown. The synchronous spectrum displays two autopeaks at 1624 and at 1641 cm^{-1} and negative cross-peaks corresponding to the interaction between the two bands. The presence of two autopeaks and their corresponding negative cross-peaks is a feature observed when two peaks are changing intensities, one increasing and one decreasing (see Fig. 1(middle)). Moreover, the asymmetry of the spots is a feature observed whenever two bands change bandwidth (see Fig. 3(middle)). The asynchronous map

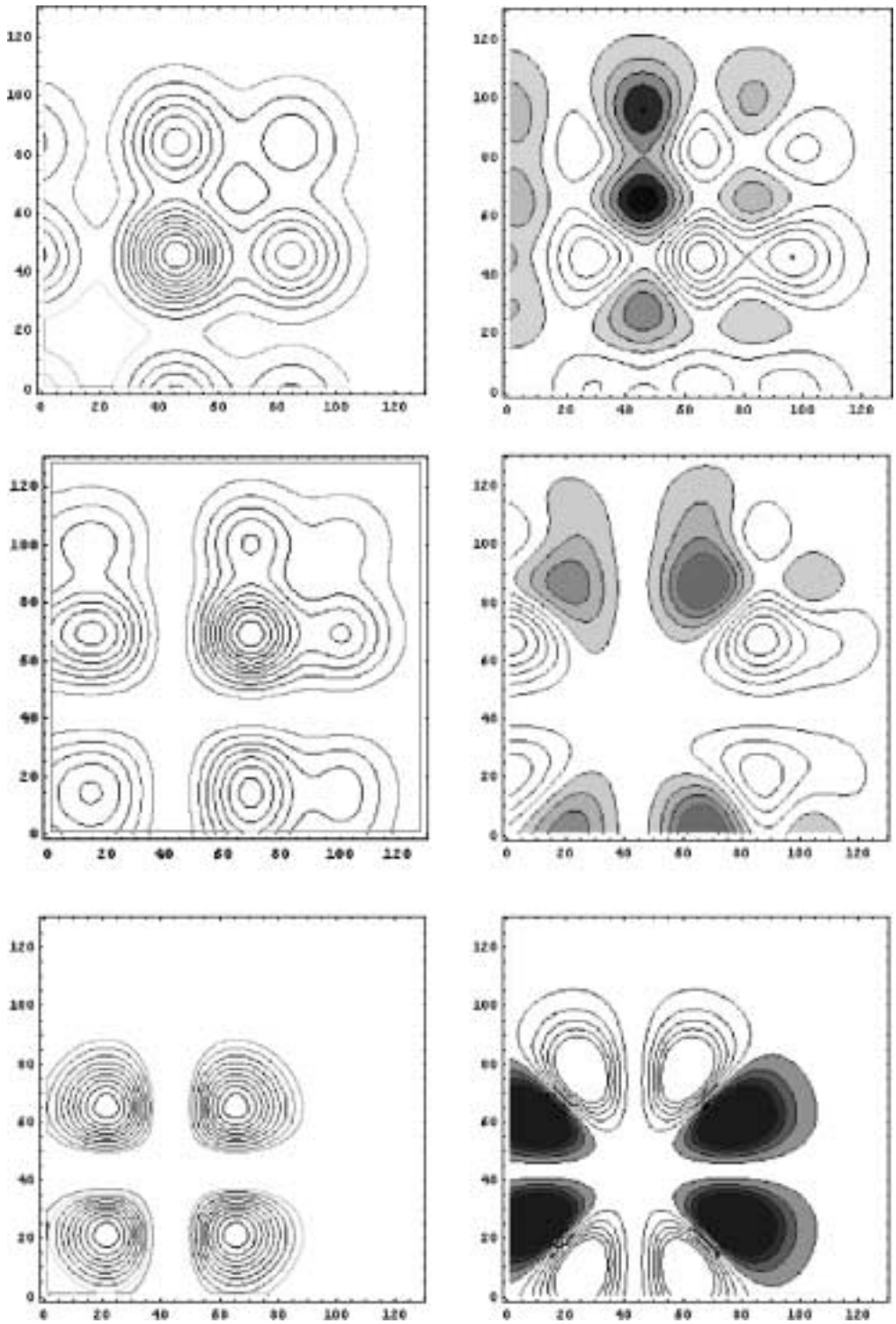

Fig. 3. Bidimensional correlation maps corresponding to changes in width (bandwidth) of a band composed of one peak (bottom), two peaks (middle) or three peaks (top). Synchronous maps are located at the left and asynchronous to the right. Axis are as in Fig. 1.

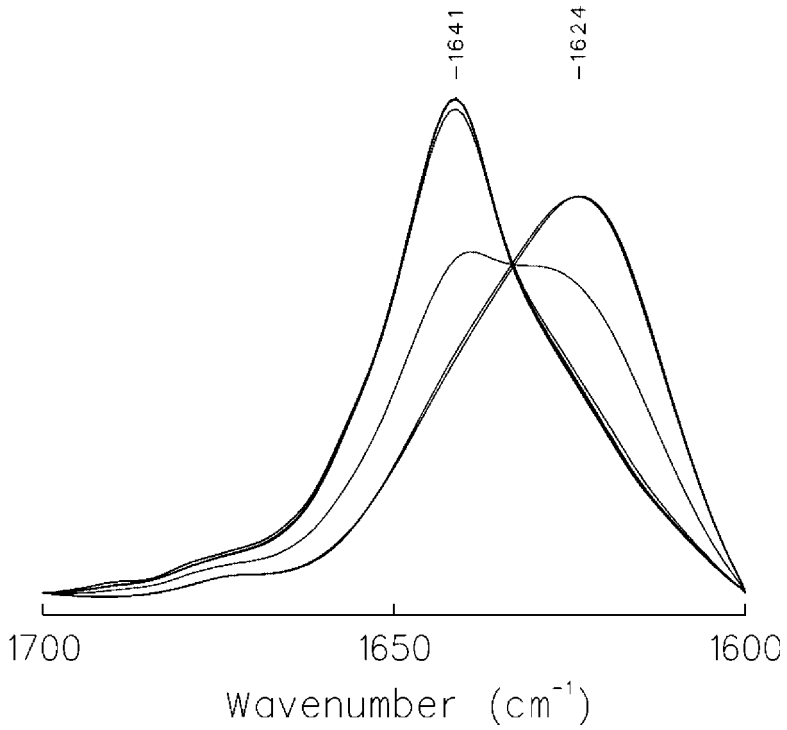

Fig. 4. Amide I band of polyproline in a D_2O solution. Temperature goes from 30°C to 70°C. PolyPro II helix is located at 1624 cm^{-1} and the band corresponding to unordered structure at 1641 cm^{-1}.

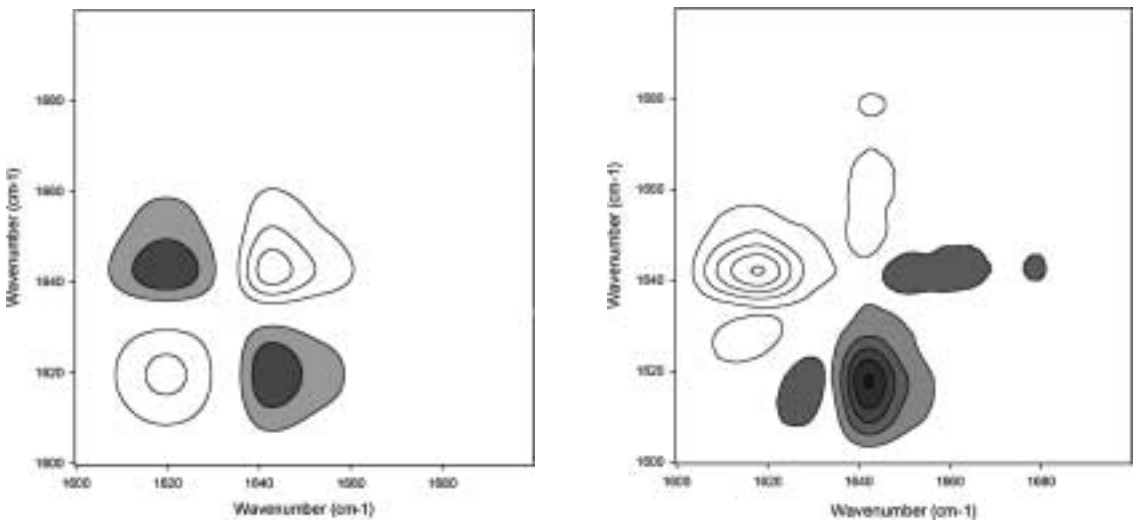

Fig. 5. Polyproline 2D-IR correlation maps corresponding to the spectra shown in Fig. 4. Synchronous (left) and asynchronous (right) maps are represented. Negative peaks are shaded.

shows an intense cross-peak between 1642 and 1623 cm^{-1}. If the process under study consisted only of a change in intensity, this map would show only low-amplitude noise. The existence of a spot in this map means that the assignment of the synchronous map to changes in intensity *and* bandwidth is correct. This is just a simple example of the possibilities of 2D-IR to indicate complex variations in protein spectra.

Insulin is a protein composed of two polypeptide chains A and B, that are linked by disulfide bonds. Upon heating, these bonds are not broken, and extensive heating gives rise to amyloid fibrils [4]. Before fibril formation, an extensive temperature-induced conformation rearrangement takes place in the protein that changes the amide I bandshape. Figure 6 shows the decomposed amide I band of insulin at 30°C and at 70°C after conformational rearrangements have taken place. The 2D-IR correlation maps (Fig. 7) are more complex in this protein as compared to the polyaminoacid, due to the existence of tertiary structure and to the presence of disulfide bonds that do not allow the complete unfolding of the protein. If the synchronous spectrum is considered, two autopeaks at 1627 and 1654 are seen with negative cross-peaks indicating changes in intensity of different sign. The asynchronous map is more complex, the most intense cross-peak appearing at 1654/1627 cm^{-1} confirming the information obtained from

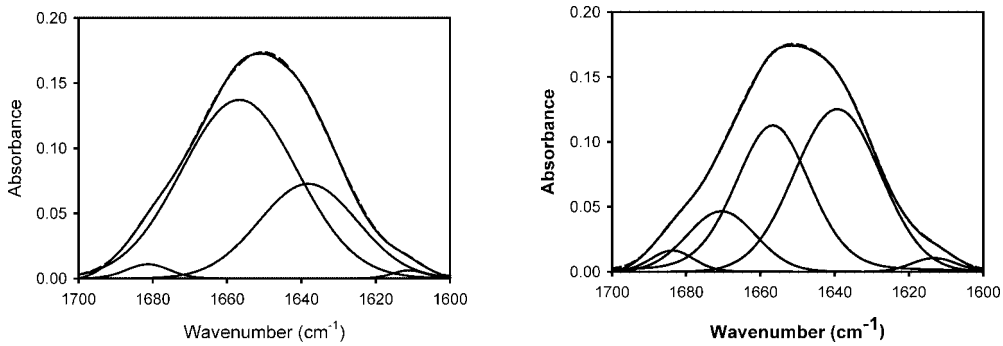

Fig. 6. Insulin amide I band fitting at 30°C (right) and 70°C (left). The spectra were measured in 20 mM phosphate buffer, pD 2.7.

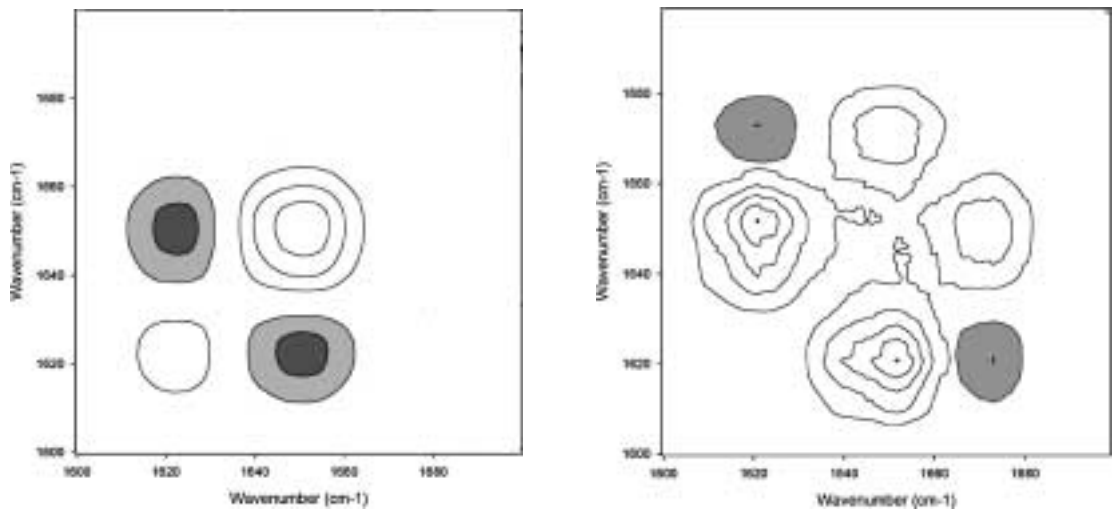

Fig. 7. 2D-IR maps corresponding to insulin thermal transition from 30°C to 70°C. 10 spectra were used to construct the maps. Synchronous (left) and asynchronous (right) maps are represented. Negative peaks are shaded.

the synchronous map. But in the asynchronous plot, cross-peaks at 1670/1654 and 1670/1627 cm^{-1} are also present, the former being more intense. Thus, the asynchronous plot is indicating that after the 1654 $\rightarrow$ 1627 cm^{-1} conversion, a 1670 $\rightarrow$ 1654 cm^{-1} correlation and finally the 1670 $\rightarrow$ 1627 cm^{-1} process are taking place. Bandwidth changes in the asynchronous map are represented by spots at the beginning and the end of the broadening (Fig. 3(middle and bottom)). Consequently, the maps obtained indicate more than one change for the band parameters of insulin.

4. Conclusions

Two-dimensional infrared correlational maps obtained from artificial bands with known changes in band parameters give rise to characteristic patterns. Biological systems of varied complexity such as polypeptides or proteins give rise to 2D-IR correlational maps that reflect the changes induced by the perturbation and can be interpreted as a composite of the band simulation maps. Thus, not only correlation between bands can be characterized but also variations in band parameters are evident in the 2D-IR correlation maps.

Acknowledgements

This work has been supported in part by grant 13552/2001 from the University of Basque Country, PI 1998-33 from the Basque Government and PB96/0171 from DGICYT.

References

[1] J.L.R. Arrondo and F.M. Goñi, Infrared spectroscopic studies of membrane lipids, in: *Biomolecular Structure and Dynamics*, G. Vergoten and T. Theophanldes, eds, Kluwer Academic, Dordrecht/Boston/London, 1997, pp. 229–242.
[2] 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.
[3] J.L.R. Arrondo, A. Muga, J. Castresana and F.M. Goñi, Quantitative studies of the structure of proteins in solution by Fourier-transform infrared spectroscopy, *Prog. Biophys. Mol. Biol.* **59** (1993), 23–56.
[4] M. Bouchard, J. Zurdo, E.J. Nettleton, C.M. Dobson and C.V. Robinson, Formation of insulin amyloid fibrils followed by FTIR simultaneously with CD and electron microscopy, *Protein Sci.* **9** (2000), 1960–1967.
[5] L.M. Contreras, F.J. Aranda, F. Gavilanes, J.M. Gonzalez-Ros and J. Villalain, Structure and interaction with membrane model systems of a peptide derived from the major epitope region of HIV protein gp41: implications on viral fusion mechanism, *Biochemistry* **40** (2001), 3196–3207.
[6] M.A. Czarnecki, Two-dimensional correlation spectroscopy: effect of band position, width and intensity changes on correlation intensities, *Appl. Spectrosc.* **54** (2000), 986–993.
[7] 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.
[8] H. Kim and S.J. Jeon, Simulations of two-dimensional electronic correlation spectra, *Bull. Korean Chem. Soc.* **22** (2001), 807–815.
[9] S. Morita, Y. Ozaki and I. Noda, Global phase angle description of generalized two-dimensional correlation spectroscopy: 1. Theory and its simulation for practical use, *Appl. Spectrosc.* **55** (2001), 1618–1621.
[10] I. Noda, A.E. Dowrey, C. Marcott, G.M. Story and Y. Ozaki, Generalized two-dimensional correlation spectroscopy, *Appl. Spectrosc.* **54** (2000), 236A–248A.
[11] M.J. Paquet, M. Laviolette, M. Pezolet and M. Auger, Two-dimensional infrared correlation spectroscopy study of the aggregation of cytochrome c in the presence of dimyristoylphosphatidylglycerol, *Biophys. J.* **81** (2001), 305–312.
[12] S. Sasic, A. Muszynski and Y. Ozaki, New insight into the mathematical background of generalized two-dimensional correlation spectroscopy and the influence of mean normalization pretreatment on two-dimensional correlation spectra, *Appl. Spectrosc.* **55** (2001), 343–349.
[13] H. Susi, Infrared spectroscopy-conformation, *Methods Enzymol.* **26** (1972), 455–472.
[14] G.B.B.M. Sutherland, Infrared analysis of the structure of amino acids, *Polypeptides and Proteins* (1952), 291–318.