Hz, $PhCH_2$), 1.65 (X part of ABX_3, 3H, J_{BX} = 1 Hz, CH_3).

References

1. The present address: Department of Chemistry, Dongguk University, Seoul.
2. (a) H. Yamanaka, T. Chiba, K. Kawabata, H. Takasugi, T. Masugi, and T. Takaya, *J. Antibiotics*, 38(12), 1738 (1985); (b) Fujisawa Pharmaceutical Company, USP 4,423,213.
3. (a) T. Naito, H. Hoshi, S. Aburaki, Y. Abe, J. Okumura, K. Tomatsu, and H. Kawaguchi, *J. Antibiotics*, 40(7), 991 (1987); (b) Bristol-Myers Company, USP 4,619,925.
4. A. H. Shingler and N. G. Weir, in: Recent Advances in the Chemistry of β-Lactam Antibiotics, ed. J. Elks, The Chemical Society, London, 153 (1977).
5. H. Peter, B. Mueller, and H. Bickel, *Helv. Chim. Acta.*, 58, 2450 (1975).
6. W. J. Kim, K.-Y. Ko, S.-U. Paik, and H. Kim, *Bull. Korean. Chem. Soc.*, 9, 111 (1988).
7. Use of *p*-TsOH (CH_2Cl_2, RT, 1 day) gave also the vinylic cephem.
8. E. W. Collington and A. I. Meyers, *J. Org. Chem.*, 36, 3044 (1971).
9. Attempted dehydration by refluxing in CH_2Cl_2 in the presence of catalytic *p*-TsOH gave the decomposed products. At room temperature the reaction was sluggish. Spry reported that the treatment of **8b** with PBr_3 in pyridine gave a mixture of (E)- and (Z)-alkenes.[10]
10. D. O. Spry and A. R. Bhala, *Heterocycles*, 24, 1799 (1986).
11. W. C. Still, M. Kahn, and A. Mitra, *J. Org. Chem.*, 43, 2923 (1978).

Applications of the Fourier Deconvolution Procedure For Quantitative Analysis of Raman Spectra of Biomolecules

Hong In Lee, Myung-Un Choi, Myung Soo Kim, and Se Won Suh*

Department of Chemistry, Seoul National University, Seoul 151-742. *Received June 5, 1989*

The constrained, iterative Fourier deconvolution procedure was applied to quantitatively analyze the overlapped bands in the Raman spectra of biomolecules. When applied to Raman spectra of lysozyme and α-amylase, this procedure resolved the amide I band into five component peaks. The relative intensities of the resolved peaks can possibly provide the composition of secondary structure elements in proteins. The deconvolution procedure was also useful in monitoring the small changes in relative intensities of C-S stretching modes due to different conformers of L-methionine in aqueous solutions at different pH values. The implemented procedure is generally applicable to the problem of resolution enhancement of spectroscopic, chromatographic, and electrophoretic data.

Introduction

Raman spectra of aqueous solutions of biomolecules often show broad, overlapped bands. Interpretation of these low resolution Raman spectra are greatly facilitated by the ability to resolve the broad bands into their component peaks. Three approaches have been generally followed to achieve the necessary resolution enhancement of the spectral data. The curve-fitting method[1] is conceptually simple and straightforward to implement on a personal computer. However, it requires some a priori imformation about the number of component bands, their shape and width, and the forms of the baselines. The maximum entropy method[2] is capable of providing an exellent resolution enhancement but the computational requirement makes its implementation on a personal computer impractical at present.

Recently, the constrained, iterative Fourier deconvolution method for the analyses of overlapped bands in Raman spectra has been developed[3,4]. The Fourier deconvolution method is computationally efficient and neither a knowledge about the number of components in the complex band nor the starting guesses for peak ampitudes and locations are required.

In this paper, we demonstrate the potential application of the constrained, iterative Fourier deconvolution procedure implemented on a personal computer for quantitative analysis of the overlapped bands in the Raman spectra of some biomolecules.

Experimental

Sample Preparation and Raman Spectroscopy. Lysozyme from hen egg white, α-amylase from *Bacillus amyloliquefaciens*, and L-methionine were purchased from Sigma Chemical Co. An aqueous solution containing lysozyme at 100 mg/ml concentration was maintained at pH 7.0 with 50 mM potassium phosphate buffer. An aqueous solution of α-amylase at 0.6 mM concentration was maintained at pH 7.0 with 5 mM tris buffer. The pH values of aqueous solutions of L-methionine were adjusted with HCl and NaOH. The samples were inserted into glass capillary cells and the cell was sealed at both ends.

The Raman spectra were obtained with the Japan Spectroscopic Company model R-300 laser Raman spectrometer

*To whom correspondence should be addressed.

using the 514.5 nm line of Ar^+ laser (Spectra Physics model 164-06) as an exciting source. The spectrometer was interfaced with an IBM-PC/AT computer and the digital data were stored in the computer. The data were generally taken at 0.7 cm^{-1} interval with an integration time of 0.3 sec. Other spectral conditions were as follows: laser power, 200 mW; scan speed, 144 $cm^{-1}min^{-1}$; and spectral slit width, 9-11 cm^{-1}. In order to improve the quality of the Raman spectrum, it was generally required to average multiple scans. For the background subtraction, Raman data of the aqueous solution containing the buffer alone were recorded under the same condition as many times as the sample spectrum itself.

Data Handling and Deconvolution. In order to minimize the spectral noise, which can be an obstacle in generating the accurately deconvoluted profile of the observed Raman spectrum, each spectrum was obtained by (i) averaging the multiple scans, (ii) subtracting the solvent contributions, and then (iii) smoothing the spectrum according to the Savitzky and Golay method[5]. For the spectra in this paper, one cycle of 13-point smoothing was found to be adequate. The fast Fourier transformation in the deconvolution procedure was done by Cooley and Tukey algorithm[6]. The baseline for the smoothed spectrum was adjusted so that the spectral values at both ends of spectral region to be deconvoluted are zero and the number of data points was made to be 2^n by extending the zero values beyond both ends of spectral region of interest. For the deconvolution, we have employed the modified Janssen-VanCittert[7] constrained, iterative deconvolution scheme developped by Agard *et al.*[8] and applied previously to image enhancement problems[9,10] and Raman spectra[3,4,11].

Briefly, the main features of this deconvolution procedure[8] are as follows. It is assumed that the observed spectrum is produced by convolution of a true spectrum with a single line shape function (band broadening or smearing function) which is slightly narrower than the observed data. The Gaussian-Lorentzian product function (1:1) was found to be the most satisfactory for the tested Raman spectra. Therefore, the goal of the deconvolution procedure is that the true spectrum is extracted from the observed spectrum by the knowledge of the single line shape function. The outline of the iterative deconvolution process is as follows. The starting point is that the observed data ($o(x)$) are considered as the true spectrum ($f(x)$). Then, the calculated observed data ($o'(x)$) are produced by the convolution of $f(x)$ with the band smearing function ($s(x)$). The $f(x)$ is corrected by the proportion of the difference between $o(x)$ and $o'(x)$. The above procedure is iterated until there is no significant difference between the observed data and the calculated observed data ($o'(x)$). Iteration was stopped when E, defined as $\Sigma|o(x)-o'(x)|/\Sigma o(x)$, falls below 0.05. In our computation, the spectral noise resulting from the deconvolution procedure was reduced by a 5-point smoothing in every cycle of iteration. The deconvolution procedure may give more than one solution. This difficulty is further compounded by the presence of experimental noise in the observed profile. Following the suggestion by Agard *et al.*[3,8] that the non-negativity of the true spectrum can effectively solve the problems of uniqueness and the increased noise, the negative values were truncated to be zero in our computation before the next cycle of deconvolution. Throughout the deconvolution computation, the area under each peak is maintained, thus allowing a quantitative analysis of the results[3]. When the deconvolution procedure does not converge rapidly, a different bandwidth of the band smearing function is tried. Usually 50 cycles of iteration were enough with an adequate band smearing function.

All computations were carried out on an IBM-PC/AT computer. 100 cycles of the constrained iteration for 512 data points took less than 10 minutes of computer CPU time. The deconvolution program was written in PASCAL and the program is available from the authors.

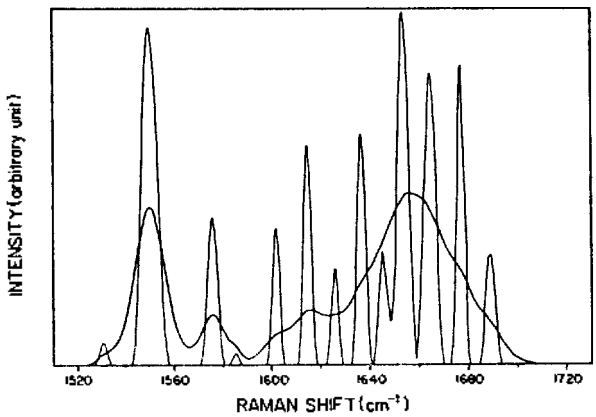

Figure 1. A comparison of the observed Raman spectrum (heavy line) in the amide I band region with the deconvoluted spectrum (light line) of an aqueous solution of lysozyme. The total bandwidth of the smearing Gaussian-Lorentzian product (1:1) line shape function employed was 23 cm^{-1}. The observed spectrum was obtained by averaging of 8 scans, solvent subtraction, and 13-point smoothing.

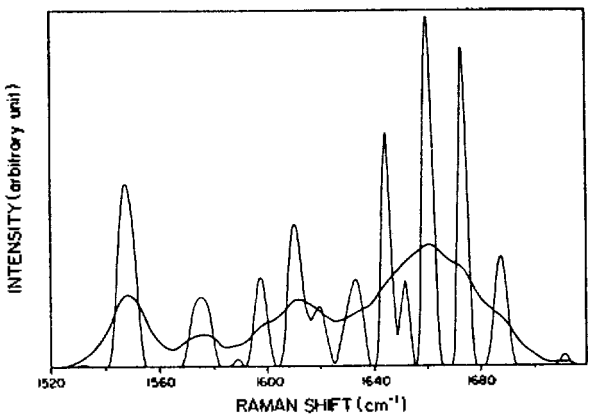

Figure 2. A comparison of the observed Raman spectrum (heavy line) in the amide I band region with the deconvoluted spectrum (light line) of an aqueous solution of α-amylase with essential Ca^{2+} ions bound. The total bandwidth of the smearing Gaussian-Lorentzian product (1:1) line shape function employed was 23 cm^{-1}. The observed spectrum was obtained by averaging of 10 scans, solvent subtraction, and 13-point smoothing.

Results and Discussion

Amide I Bands of Lysozyme and α-Amylase. The observed and deconvoluted Raman spectra of lysozyme and α-amylase in the 1510-1730 cm^{-1} region are compared in Fig-

Table 1. A Tentative Assignment of the Deconvoluted Raman Spectrum of Lysozyme in the Amide I Band Region and the Relative Intensities of Five Component Peaks of the Amide I Band

Peak Position (cm^{-1})	Relative Intensity (%)	Assignment[a] [12,14]
1531		
1550		Trp
1576		Trp
1585		Phe
1602		Phe, Tyr
1614		Trp, Tyr
1625		
1636		
1645	7.9	Amide I (H_2,U)
1654	33.9	Amide I (T,U,H_1)
1664	30.0	Amide I (U,T,H_1)
1677	19.5	Amide I (B,H_1)
1689	8.7	Amide I (H_1,B)
1705		

[a]The notations for amide I band assignment are as follows; H_1 (mono-hydrogen bonded α-helix), H_2 (bi-hydrogen bonded α-helix), T (turn), U (undefined), and B (parallel and antiparallel β-sheet). Phe, Tyr, and Trp are the three letter abbreviations for phenylalanine, tyrosine, and tryptophan residues, respectively.

Table 2. A Tentative Assignment of the Deconvoluted Raman Spectrum of α-Amylase in the Amide I Band Region and the Relative Intensities of Five Component Peaks of the Amide I Band

Peak Position (cm^{-1})	Relative Intensity (%)	Assignment[a] [12,14]
1532		
1547		Trp
1576		Trp
1589		Phe
1598		Phe, Tyr
1611		Trp, Tyr
1620		
1633		
1644	20.4	Amide I (H_2,U)
1652	7.7	Amide I (T,U,H_1)
1660	30.4	Amide I (U,T,H_1)
1673	28.6	Amide I (B,H_1)
1687	12.9	Amide I (H_1,B)
1711		

[a]The notations are the same as in Table 1.

ures 1 and 2, respectively. The tentative assignments of the component peaks are given in Tables 1 and 2. The deconvolution procedure resolved the amide I band into five component peaks, even though the overlap seemed hoplessly severe. Each component peak of the amide I band can be assigned as due to a combination of two or three secondary structure elements in proteins. Assignments of some other peaks to the aromatic amino acids are also given in Tables 1 and 2. The component peak at about 1710 cm^{-1} with a negligible intensity may be disregarded as resulting from the noise in the spectrum, although a similar peak is present in the deconvoluted Raman spectrum of the filamentous bacteriophage Xf[11]. The two extremely weak peaks at about 1532 and 1586-1588 cm^{-1} in Figures 1 and 2 are likely to be artifacts due to the spectral noise. However, the two moderately strong peaks at 1625 and 1636 cm^{-1} in Figure 1 and the corresponding peaks at 1620 and 1630 cm^{-1} in Figure 2 are not likely to be artifacts, although their assignments cannot be made with confidence.

Our assignment of the component peaks in the amide I bands is well correlated with reference Raman intensity profiles of Williams *et al.*[12,13] obtained for the following secondary structures in proteins; mono-hydrogen bonded α-helix, bi-hydrogen bonded α-helix, parallel β-sheet, antiparallel β-sheet, reverse turn, and undefined structure. The reference Raman spectra for parallel and antiparallel β-sheet structures seem to be similar enough to make it difficult to justify a distinction of these two secondary structures by the amide I band analysis. The reference spectra for the turn and undefined structures are also very similar to each other but a side band, present near 1662 cm^{-1} in the reference spectrum of the undefined structure, apparently gives a resolved peak at 1664 and 1660 cm^{-1} in the deconvolued Raman spectra of lysozyme and α-amylase in Figures 1 and 2, respectively.

Since the observed Raman amide I band region can be assigned to be a linear combination of the contributions from each secondary structure elements, one can estimate the composition of secondary structures in a given protein by least squares method. The result of this method is very sensitive to the reference Raman spectra for secondary structure elements. As mentioned above, the similarity of the reference spectra of some secondary structures makes this method prone to error when one attempts to classify the protein secondary structures into more than three or four. That is, Williams *et al.*[12,13] recomended to classify the secondary structures into four groups; α-helix, β-sheet, β-turn, and random coil. Verjot *et al.*[14] used the four reference Raman spectra (α_1- and α_2-helix, β-sheet, and undefined structure) to improve the accuracy of the estimation. Clearly, there are limitations in applying the method of using reference Raman intensity profiles for the secondary structure estimation.

The deconvolution technique has not been examined thoroughly as an alternative to the above method for the estimation of protein secondary structures. However, our limited experiences with lysozyme, α-amylase, and other proteins[15] suggest that the application of the Fourier deconvolution procedure to the amide I band has the potential to be developed into a simple and reliable alternative in estimating the content of five protein secondary structures (mono-hydrogen bonded α-helix, bi-hydrogen bonded α-helix, parallel and antiparallel β-sheet, reverse turn, and undefined structure), if the relative contributions of each secondary structures to the five component peaks of the amide I band can be established by least squares method from a large database of Raman spectra of proteins ($\geq$15) with known structures. This suggestion is based on the comparison of the data in Tables 1 and 2. The discrepancies in Raman shifts of the five component peaks in the amide I bands between lysozyme and α-amylase are 1 to 4 cm^{-1}, whereas the differences between the adjacent component peaks for a given protein are 8 to 14 cm^{-1}. And the relative intensities of the five component peaks in lysozyme and α-amylase are significantly different, reflecting the different secondary structures in two proteins. Therefore, we conclude that the Fourier deconvolution of

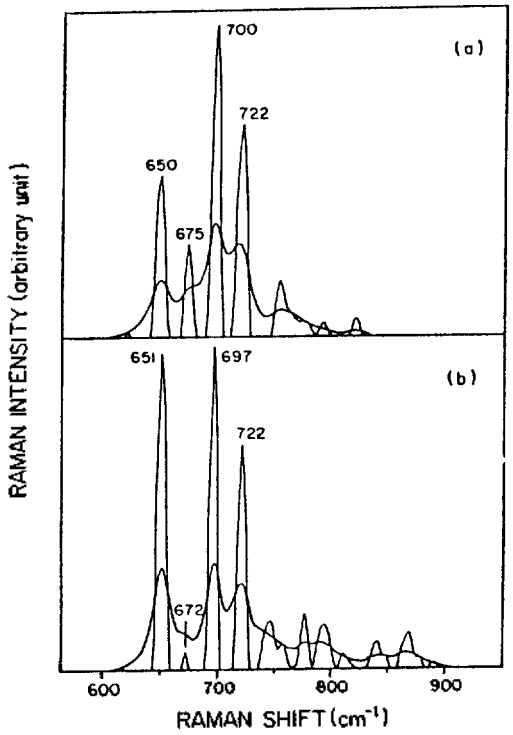

Figure 3. Raman spectra of aqueous solutions of L-methionine at (a) pH 12 and (b) pH 2 in the C-S stretching region. Heavy line indicates the observed spectra and light line indicates the deconvoluted spectra. The total bandwidth of the smearing Gaussian-Lorentzian product (1:1) line shape function was 23 cm^{-1}. The observed spectra were obtained by averaging of 5 scans, solvent subtraction, and 13-point smoothing.

amide I band in the Raman spectra of proteins may become a viable alternative to the use of reference intensity profiles in the estimation of protein secondary structures.

L-Methionine. The side chain C-S stretching Raman band of L-methionine appears generally in 600-750 cm^{-1} region[16]. For the side chain of L-methionine, four distinct conformations are possible, that is, TT, TG, GT, and GG conformations. Here, the first and second letters describe either a trans or a gauche conformation around the S-CH_2 and CH_2-CH_2 bonds, respectively. The position of the C-S stretching mode is known to be sensitive to the conformation around S-CH_2 and CH_2-CH_2 bonds[17,18]. In Figure 3, the observed and deconvoluted Raman bands in the C-S stretching region are compared at (a) pH 12 and (b) PH 2. The C-S stretching modes for each conformation are overlapped and not well resolved in the observed spectra. In the deconvoluted spectra, the bands are well resolved. At pH 12 (Figure 3. (a)), four C-S stretching modes are resolved. The modes appear at 650, 675, 700, and 722 cm^{-1}. Each band is assigned to GG, TG, GT, and TT conformers, respectively[17,18]. And at pH 2 (Figure 3. (b)), the above four bands are appearing at similar positions with somewhat different relative intensities.

The quantitative results of Fourier deconvolution analysis of the four C-S stretching bands at pH 12, 5, and 2 are listed in Table 3. It can be seen from Table 3 that as the pH is lowered, the Raman intensity due to the GG conformer increases while those due to TG and TT conformers decrease.

Table 3. pH Dependence of Relative Intensities of C-S Stretching Modes of Different Conformers in L-Methionine

pH	Relative Intensities (%)			
	GG	TG	GT	TT[a]
12	21.2	11.1	35.3	32.3
5	30.2	4.0	38.6	26.7
2	37.6	1.0	34.2	27.2

[a]Conformations around S-CH_2 and CH_2-CH_2 bonds. See text for details.

The observed change may be due to the change of the relative concentrations of each conformer. L-methionine exists in the anionic form (NH_2-CH(R)-COO^-) at pH 12, in the zwitterionic form (NH_3^+-CH(R)-COO^-) at pH 5, and in the cationic form (NH_3^+-CH(R)-COOH) at pH 2. In each ionic form, the relative thermodynamic stabilities of different conformers may be slightly different, resulting in different relative intensities of each conformer at different pH values.

Acknowledgement. This work was supported by grants from the Korea Science and Engineering Foundation (No. 860614) and S. N. U. Daewoo Research Fund to S. W. S.

References

1. P. M. A. Sherwood, in Practical Surface Analysis by Auger and X-ray Photoelectron Spectroscopy, edited by D. Briggs and M. P. Seah, John Wiley & Sons, New York, pp. 445-475 (1983).
2. F. Ni and H. A. Scheraga, *J. Raman Spectrosc.*, **16**, 337 (1985).
3. G. J. Thomas, Jr. and D. A. Agard, *Biophys. J.*, **46**, 763 (1984).
4. G. J. Thomas, Jr., *J. Mol. Struc.*, **141**, 261 (1986).
5. A. Savitzky and M. J. E. Golay, *Anal. Chem.*, **36**, 1627 (1964).
6. J. W. Cooley, P. A. W. Lewis, and P. D. Welch, *IEEE transact. Ed.*, **12**, 27 (1969).
7. P. A. Jannson, R. H. Hunt, and E. K. Plyer, *J. Opt. Soc. Am.*, **60**, 596 (1970).
8. D. A. Agard, R. A. Steinberg, and R. M. Stroud, *Anal. Biochem.*, **111**, 257 (1981).
9. D. A. Agard and R. M. Stroud, *Acta Crystallogr.*, **B38**, 186 (1982).
10. D. A. Agard and J. W. Sedal, *Nature*, **302**, 676 (1983).
11. G. J. Thomas, Jr., *Spectrochim. Acta*, **41A**, 217 (1985).
12. R. W. Williams and A. K. Dunker, *J. Mol. Biol.*, **152**, 783 (1981).
13. R. W. Williams, *J. Mol. Biol.*, **166**, 581 (1983).
14. M. Berjot, J. Marx, and A. J. P. Alix, *J. Raman Spectrosc.*, **18**, 289 (1987).
15. H. I. Lee and S. W. Suh, unpublished work.
16. F. R. Dollish, W. G. Feteley, and F. F. Bentley, Characteristic Raman Frequencies of Organic Compounds, John Wiley & Sons, New York (1974).
17. T. Shimanouchi, H. Matsuura, Y. Ogawa, and I. Harada, *J. Phys. Chem. Ref. Data*, **7**, 1323 (1978).
18. N. Nogami, H. Sugeta, and T. Miyazawa, *Chem. Lett.*, 147 (1975).