

Raman Spectroscopic Study of Casein Structure

D. MICHAEL BYLER, HAROLD M. FARRELL, JR., and HEINO SUSI
 US Department of Agriculture, Agricultural Research Service
 Eastern Regional Research Center
 Philadelphia, PA 19118

ABSTRACT

The secondary structure of caseins was investigated with resolution-enhanced laser Raman spectroscopy. Raman spectra in the 1580 to 1720 cm^{-1} region were obtained from the following lyophilized proteins: 1) α_{S1} -casein, 2) β -casein, 3) a natural mixture of bovine whole casein, and 4) micelles of the natural mixture in the presence of Ca^{2+} ions. In addition, β -casein was also investigated in D_2O solution. The spectra obtained were Fourier deconvolved and curve fitted with Gaussian components. The results suggest that both α_{S1} and β -casein have around 10% helical structure, around 20% β -structure, and from 20 to 35% turns. The turns are clearly distinguishable from the moiety usually called undefined, random, or structureless. Freeze-dried micelles in the presence of Ca^{2+} ions and submicelles in the presence of K^+ ions appear to contain an increased amount of turns and of β -structure as compared with the α_{S1} - and β -caseins. The increase in turns is at the expense of the amount of undefined structure. All conformational designations here are based on spectroscopic assignments derived from crystallized proteins with well characterized structures. These designations thus have a more qualitative, descriptive meaning for caseins than for other milk proteins, such as α -lactalbumin or β -lactoglobulin.

INTRODUCTION

Caseins constitute about 80% of the proteins of bovine milk. The two major components are α_{S1} - and β -casein, which have different primary structures (5) and, therefore, are expected to have different secondary structures (conformations). In the absence of Ca^{2+} ions, casein

monomers tend to self-associate and form aggregates with Stokes radii of up to 9.4 nm; these particles are referred to as submicelles (11). It has been hypothesized that the submicelles in the presence of Ca^{2+} ions aggregate to form the spherical colloidal complexes of skim milk termed the casein micelles (6, 14). These casein micelles are unique transport colloids that carry the majority of the Ca^{2+} ions and the phosphate of milk.

The conformations of casein molecules under any condition are poorly understood. Creamer et al. (4) have compared data obtained by circular dichroism and optical rotatory dispersion with estimates based on sequence studies. The results for β -casein at pH 6.8 range from 0 to 13% β -structure and 1 to 20% helix. Corresponding numbers for α_{S1} -casein are 8 to 17% β -structure and 13 to 22% helix. Infrared spectroscopy (2, 15) indicates a large amount of unspecified structure and probably some "turns" in the case of α_{S1} -casein in D_2O solution, but the results are too ill-defined to specify any amount of β -structure or helix.

Virtually no reliable information is available on whole caseins with or without Ca^{2+} ions. Raman spectroscopy is a relatively new tool for protein conformation studies (1, 3, 7, 9, 16, 17, 19). A new Raman technique, resolution enhancement of the amide I (the C=O stretching) band via Fourier deconvolution, has already produced a considerable amount of information concerning the conformation of various milk proteins (16). We present here results obtained on dissolved and lyophilized β -casein, lyophilized α_{S1} -casein, and on whole casein lyophilized from solution under micellar (15 mM CaCl_2) and submicellar conditions.

MATERIALS AND METHODS

Preparation of Samples

Two liters of warm milk were obtained from the whole milk of an individual Jersey cow. The animal was in midlactation, in good health, and

RAMAN SPECTROSCOPY OF CASEINS

was part of a commercial herd. Phenylmethylsulfonyl fluoride (.1 g/L), a serine protease inhibitor, was added immediately to retard proteolysis. The milk was transported to the laboratory and skimmed twice by centrifugation at $4000 \times g$ for 10 min at room temperature. Five-hundred milliliters of skim milk were diluted with an equal volume of distilled water and warmed to 37°C . Casein was precipitated by careful addition of 1 N HCl to pH 4.6. The precipitate was homogenized with a polytron ST-10 at low speed and dissolved by addition of NaOH to yield a solution of pH 7.0. Casein was reprecipitated, washed, and then resuspended. The sodium caseinate was cooled to 4°C and centrifuged at $10,000 \times g$ for 30 min to remove residual fat and then freeze-dried. Alkaline urea gel electrophoresis with standard caseins of known structure showed the following genotype of the animal: α_{S1} -BB, β -AA, and κ -BB (18).

For Raman experiments, the whole sodium caseinate prepared as described was dissolved in H_2O at 20 g/L and pH adjusted to 6.80 with NaOH. To one-half the sample a stock CaCl_2 solution dissolved in H_2O was added to give a final concentration of 15 mM; to the second half a stock KCl solution was added to yield a final concentration of 45 mM (comparable ionic strength to CaCl_2). Both solutions were lyophilized; the freeze-dried samples were resuspended gently in 10 ml H_2O and re-lyophilized.

Purified caseins were prepared from whole milk by urea fractionation and chromatography on DEAE-Cellulose as previously described (18). The α_{S1} -casein is the B genetic variant while the β -casein is the A² genetic variant. These samples were lyophilized and used directly or dissolved in D_2O .

Spectroscopic Measurements

The general theory of resolution enhancement of amide I bands via Fourier deconvolution has been described in some detail in connection with infrared studies of protein structure (2, 15). Application of these procedures to Raman spectroscopy is quite similar,

but several specific aspects have to be kept in mind. 1) In contrast to infrared spectra, Raman spectra are most conveniently obtained in the solid state. 2) The signal-to-noise ratio usually obtained is not as good for Raman spectra as it is for Fourier transform infrared spectra. Very careful experimental procedures and mathematical smoothing of the empirical curves are therefore necessary. 3) The overall empirical bandshape of amide I Raman bands is almost 100% Gaussian (16). Computational procedures and algorithms that take the proper band shape into account must therefore be used. 4) Finally, the assignments of Raman band components for solid proteins in the amide I region are not the same as for infrared spectra in D_2O solution because of effects caused by hydration and deuteration.

With these points in mind, the spectra were obtained by standard Raman techniques and mathematically treated essentially as previously described for infrared spectra in D_2O solution (2, 15). Raman spectra of samples sealed in melting point capillaries were recorded from 1450 to 1750 cm^{-1} on a Spex 1401 spectrometer (Spex Industries, Edison, NJ) with the 514.4 nm line of a Spectra-Physics¹ Model 165-3 argon-ion laser (Spectra Physics, Mountain View, CA) used for excitation. The laser power at the sample was about 250 to 350 mW. The spectroscopic slit width was 4 cm^{-1} . The instrument is equipped with a stepping motor and a Spex Datamate microcomputer. One data point was recorded every 1.0 cm^{-1} . Depending on the noise level, 10 to 16 scans were signal averaged and then smoothed by a 9-point Savitzky-Golay smoothing procedure (13).

The spectra were deconvolved with a slightly modified version of program number LI of the National Research Council of Canada (10). A value of 0% Lorentzian (i.e., 100% Gaussian) character was used in the deconvolution equation (10). The band widths (full width at half height, FWHH) were taken as 13 cm^{-1} , in accordance with previous infrared work on proteins in this laboratory and elsewhere. The resolution factor, K, was varied between 2 and 2.5. For most samples a value of $K = 2.3$ gave the best results, with minimum distortion and no overdeconvolution. The deconvolved spectra were fitted with Gaussian components with the help of the program ABACUS [written at Eastern Regional Research Center by William C.

¹Mention of commercial names does not imply any endorsement by the US Department of Agriculture of one product over any others.

Damert (personal communication)]. The agreement between the sum of the calculated components and the deconvolved experimental spectra in the amide I region was exceedingly good in most instances (root mean square errors smaller than 1% of maximum value). The spectra of solid samples and solutions were handled in the same manner, although it took a much longer to obtain acceptable solution spectra, because the Raman scattering intensity increases almost linearly with sample concentration, i.e., solid materials are very much stronger scatterers than dilute solutions.

RESULTS

Figure 1A shows the Raman spectrum of lyophilized β -casein in the amide I spectral region. It is easily shown that the overall contour can be fitted with a few Gaussian

components, as for most amide I Raman bands of proteins (16). The two weak bands close to 1604 and 1616 cm^{-1} are caused by aromatic side chains. The strong band centering close to 1665 cm^{-1} is the unresolved composite amide I band. It consists of a number of different components associated with backbone C=O stretching vibrations of different substructures such as β -segments, helices, and turns (16). Figure 1B depicts the same spectrum after deconvolution using the parameters FWHH = 13 cm^{-1} , $K = 2.4$, percent Lorentzian character = 0, and smoothing function number 8 (10). The deconvolved spectrum shown in Figure 1B is fitted with Gaussian components, as described (see Spectroscopic Measurements). Figure 1C shows the deconvolved and band-fitted amide I band of lyophilized α_{S1} -casein; Figure 1D shows the corresponding band of

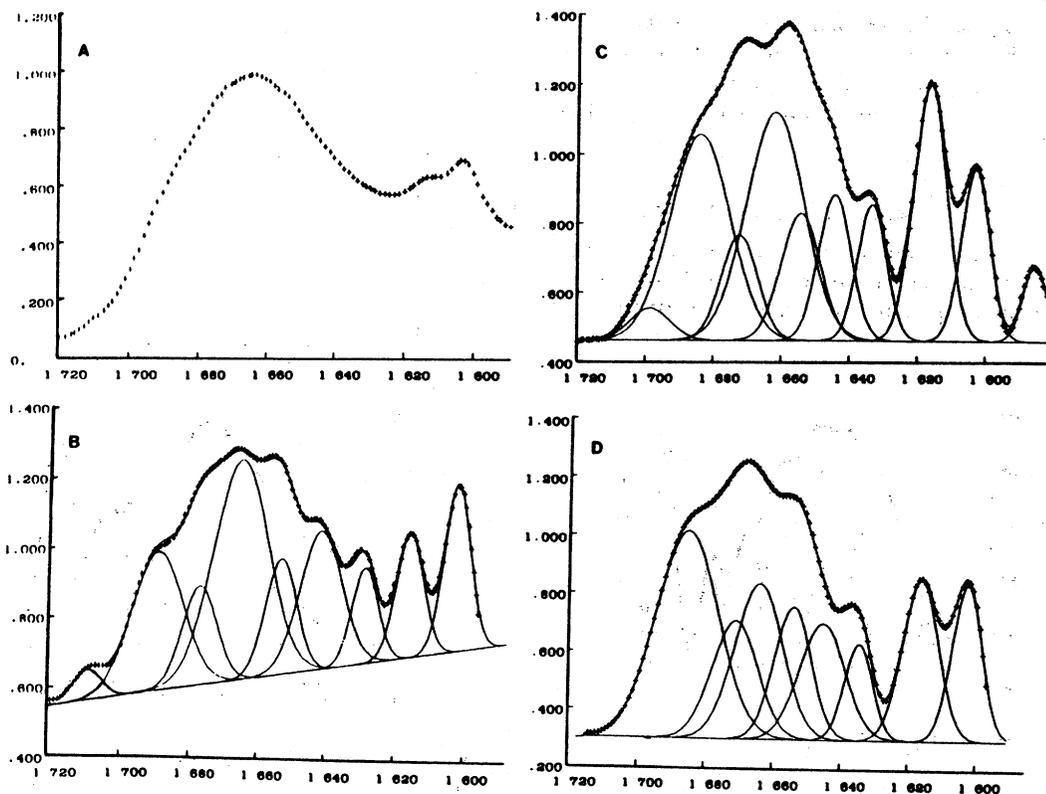


Figure 1. Amide I Raman bands of caseins. Relative intensity (INT) versus Raman shift (cm^{-1}). A) Original spectrum of lyophilized β -casein. B) Deconvolved spectrum of β -casein with added Gaussian curve fitting. C) Deconvolved spectrum of lyophilized α_{S1} -casein with Gaussian curve fitting. D) Deconvolved spectrum of whole bovine casein mixture, in the form of lyophilized submicelles in the presence of K^+ , with Gaussian curve fitting.

lyophilized submicellar casein mixture in the presence of K^+ ions.

Figure 2 shows the deconvolved spectrum of the amide I band of freeze-dried casein micelles in the presence of Ca^{2+} ions. It is easily seen that all band maxima are clearly identified and no artifacts appear to be present. Figure 3 shows the original (undeconvolved) and the deconvolved amide I Raman band of β -casein in D_2O solution. Deuterium oxide is used as a solvent because H_2O displays an H-O-H bending mode, which overlaps with the amide I band. Although Raman spectra of proteins are sometimes reported in H_2O solution, in our judgment, detailed analysis of the amide I band is much more difficult under these conditions because the strong H-O-H bending mode is also observed in this spectral region. In general, deuteration of a protein shifts the amide I frequency by less than 5 cm^{-1} (2, 3, 15), but it is obvious from a comparison of Figures 1B and 3B that in this study much greater changes have taken place. These changes are associated not so much with the deuteration of the protein as with spectral changes caused by protein-solvent interaction through hydrogen bonding (2, 15, 16).

DISCUSSION AND CONCLUSIONS

Table 1 summarizes the observed characteristic frequencies of the amide I component bands of globular proteins. The assignments

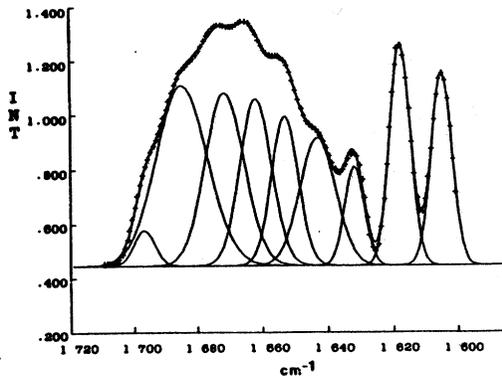


Figure 2. Deconvolved spectrum of lyophilized whole casein micelles in the presence of Ca^{2+} ions. Relative intensity (INT) versus. Raman shift, (cm^{-1}). Gaussian curve fitting as in Figure 1.

indicated in the table are based on a thorough study carried out in this center of Raman (16) and infrared (2, 15) spectra of a considerable number of proteins. The listed frequencies represent the centers of the deconvolved components. These agree in most cases with the observed maxima of the composite deconvolved amide I bands, but this is not necessarily so. A summation of Gaussian components can, in principle, result in a curve with different maxima than the ones of the original subbands (see Figure 3). Table 2 gives the estimated conformational values calculated from the curves given in Figures 1 to 3. The computations are entirely analogous to the ones reported previously for infrared studies (2, 15). The basic assumption is retained that the fractional area intensities of specific components (relative to the total amide I band area) reflect the fraction of respective secondary structure values. Because of the high noise level

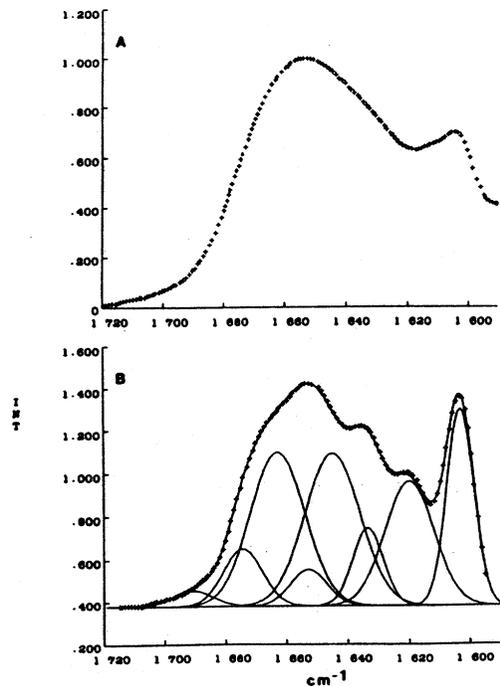


Figure 3. Original spectrum (A) and deconvolved spectrum (B) of β -casein in D_2O solution. Relative intensity (INT) versus. Raman shift (cm^{-1}). Curve fitting was carried out on the basis of characteristic frequencies transferred from infrared spectroscopy (2, 15).

TABLE 1. Assignment of characteristic frequencies (cm^{-1}) for the amide I band components of globular proteins.¹

Wavenumber range	Assignment
Solids	
1602 – 1606	Aromatic sidechains
1615 – 1618	Aromatic sidechains
1631 – 1633	β -Structure, low frequency component (weak)
1641 – 1646	H-O-H bending of bound water
1653 – 1658	Helical segments
1660 – 1663	Unspecified
1668 – 1675	β -Structure, high frequency component (strong)
1680 – 1699	Turns
Deuterium oxide solution	
1621 – 1639	β -Structure, low frequency component (strong)
1640 – 1646	Unspecified
1651 – 1657	Helical segments
1660 – 1665	Turns
1670 – 1680	β -Structure, high frequency component (weak)
1685 – 1695	Turns

¹ Solid state Raman values from H. Susi and D. M. Byler (16). The D_2O solution infrared values from H. Susi and D. M. Byler (2, 15).

of the original spectra the curve fitting procedure does not necessarily lead to unique solutions. Therefore, ranges of percentages are given for the various conformations. These

percentages represent the maximum and minimum values found when the data were curve-fit using a range of different initial parameters for each trial. Therefore, the ranges

TABLE 2. Estimated secondary structure characteristics of casein samples.

Nature of sample	% Turns	% Unspecified	% Helix	% β -Structure
β -Casein, lyophilized				
Range	20 – 26	40 – 60	3 – 11	18 – 24
Best value ¹	23	51	7	19
β -Casein, D_2O solution				
Range	31 – 36	30 – 36	6 – 14	20 – 23
Best value	35	30	13	22
α_{S1} -Casein, lyophilized				
Range	29 – 35	33 – 40	8 – 13	18 – 20
Best value	34	33	13	20
Submicelles, native composition lyophilized, presence of K^+				
Range	36 – 39	16 – 23	8 – 18	24 – 30
Best value	39	16	18	27
Micelles, native composition lyophilized, presence of Ca^{2+}				
Range	36 – 41	18 – 24	10 – 14	26 – 31
Best value	41	18	14	27

¹ The curve-fit with the smallest root mean square error. See text.

RAMAN SPECTROSCOPY OF CASEINS

given roughly indicate the reliability of the measurements. The single "best" values (which add up to 100% in each case) give the results of the "best-fit" experiment (smallest root-mean-square error, usually well below 1% of the maximum absorbance value); they do not represent an average of the observed data. It must be kept firmly in mind that the given results represent estimates.

Casein, to the best of our knowledge, cannot be crystallized. Concepts such as "percentage helix" or "percent β -structure," therefore, cannot be verified to have the same significance for this protein as for proteins where spectroscopic data are directly compared with x-ray analyses (2, 15). The band assignments given here are corroborated by other studies (16) and probably yield the best possible estimates of casein structure. Indeed, the values obtained could simply represent some local order in a generally ill-defined overall conformation. Nevertheless, the observed changes in the spectra do represent real structural differences between the various caseins and casein mixtures.

Figure 3 merits special attention for several reasons. The frequencies and the assignments obtained in D_2O solution are different from the ones reported for lyophilized β -casein (Table 1), although the calculated conformation is quite similar, as shown in Table 2. The assignments are different because in D_2O solution the bulk of amide N-H (actually N-D) groups are hydrogen bonded to the solvent, not to other peptide groups, as in solid samples. The component frequencies in D_2O solution are transferred from corresponding infrared studies (2, 15). Because casein is not expected to have any symmetry, the component frequencies in D_2O solution should be the same in infrared spectra and in the Raman effect, although the intensities could, and indeed do, differ. In fact, agreement is good between conformational data obtained for β -casein in the solid state and in D_2O solution (Table 2), although the number of turns seems to increase in solution. Figure 3B also furnishes a nice example where component frequencies do not agree with the maxima of the overall deconvolved spectrum.

One of the most interesting facts that emerges from these studies is the observation that there are two clearly distinguishable conformations in caseins, in addition to the

short helical sections and β -strands. Segments previously designated as "undefined" are now seen to be composed of two different kinds of substructure. For lyophilized samples, one exhibits bands in the 1680 to 1699 cm^{-1} range, the other in the 1660-1663 cm^{-1} range. Both bands could well represent turns, if infrared spectra obtained in D_2O solution (2, 15) are of any use as a guide. In solid state Raman spectra of undeuterated proteins, the 1660 cm^{-1} region band has usually been assigned to an "unspecified" or "random" conformation (1). We retain this nomenclature although it is quite conceivable that this band, in part, also results from turns, but of a different kind from those associated with the higher frequency component. Such an assumption would help account for the discrepancy seen in the percentage of turns and percentage undefined calculated for the lyophilized protein versus that found for β -casein in D_2O solution (Table 2). Be that as it may, the sections heretofore called "random" (in casein, anyway) are obviously actually composed of two quite different substructures. We choose to call them "turns" (bands ca. 1685 cm^{-1}) and "unspecified" (bands ca. 1660 cm^{-1}) in order to make a clear distinction.

It is also noteworthy that the β -content of all samples appears to be somewhat higher than usually assumed (4). There is a plausible explanation. Extended " β -sheets" rarely exist in actual proteins. Instead, we have more or less regular "strands", sometimes parallel, sometimes antiparallel, sometimes "mixed", and sometimes with no neighboring extended chains at all. All these conformations tend to exhibit similar infrared and Raman bands (2, 15, 16). The same would apply to some disordered extended chains, which are frequently classified as "random". The β -values obtained by infrared or Raman spectroscopy are therefore sometimes a little higher than values derived from, say, x-ray data. Similar qualitative reasoning applies to "turns". "Turns" in the ill-defined casein structure appear to exhibit bands similar to well-defined turns in crystallizable proteins, such as α -lactalbumin or β -lactoglobulin, but they may not be structurally as well-defined.

In contrast to most other proteins, caseins appear to give better resolved spectra in the solid state than in solution, as seen by comparison of Figures 1B and 3B. There is, how-

ever, no indication that the change of state results in a substantial change in the overall conformation, as seen in Table 2.

Finally, it is interesting to note that caseins in the presence of K^+ or Ca^{2+} appear to have different spectra than either pure α_{s1} - or β -casein (Figures 1 to 3, Table 2). There are three possible explanations. 1) The whole caseins, both submicelles and micelles include κ -casein (*ca.* 10% by weight). The latter protein has a different primary structure and therefore possibly a different secondary structure than the species studied here, although computer-generated structural estimates (12) do not predict this to be so. 2) Micelle and submicelle formation influences the secondary structure of all constituents through protein-protein interactions. 3) Another possibility not to be overlooked is that the traditional procedures required to purify the individual caseins have caused alterations in their secondary structures. Because of the extensive self-association exhibited by casein components, successful purifications have been achieved only in the presence of 4 M urea. By contrast, the native mixtures (in K^+ or Ca^{2+}) were never exposed to these conditions. Thus, the differences could be due to either cooperative interactions among the caseins or to structural components sensitive to denaturants and lost during purification. In similar ORD studies, Herskovits did not see pronounced differences between the native mixtures and isolated proteins (8), but the Raman method is probably a more sensitive probe of these structures, particularly turns.

The true explanation probably involves all three of these factors. The difference observed between the structures of the individual proteins and the micelles and submicelles of the native mixture appears to manifest itself in a somewhat higher number of turns and a slightly higher amount of β -structure. As is evident from Table 2, this increase is at the expense of a decrease in the "undefined" substructure. This study thus strongly suggests that the structure of the whole caseins is different than the sum of the structures of the individual components.

ACKNOWLEDGMENTS

The authors thank Thomas M. Muller, Stephen McGady, and Kevin N. Newman for

their technical assistance in obtaining the Raman spectra.

REFERENCES

- 1 Berjot, M., J. Marx, and A.J.P. Alix. 1987. Determination of the secondary structure of proteins from the Raman amide I band: the reference intensity profiles method. *J. Raman Spectrosc.* 18:289.
- 2 Byler, D. M., and H. Susi. 1986. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* 25:469.
- 3 Byler, D. M., and H. Susi. 1986. Amide I region of resolution-enhanced Raman spectra of proteins. *Proc. 10th Int. Conf. Raman Spectrosc.* I-58.
- 4 Creamer, L. K., T. Richardson, and D.A.D. Parry. 1981. Secondary structure of bovine α_{s1} - and β -casein in solution. *Arch. Biophys. Biochem.* 211:689.
- 5 Eigel, W. N., J. E. Butler, C. A. Ernstrom, H. M. Farrell, Jr., V. R. Harwalkar, R. Jenness, and R. McL. Whitney. 1984. Nomenclature of proteins of cow's milk: fifth revision. *J. Dairy Sci.* 67:1599.
- 6 Farrell, H. M., Jr., and M. P. Thompson. 1988. The caseins of milk as calcium binding proteins. Page 117 in *Calcium Binding Proteins*. M. P. Thompson, ed. CRC Press, Boca Raton, FL.
- 7 Frushour, B. G., and J. L. Koenig. 1974. Raman spectroscopic study of tropomyosin denaturation. *Biopolymers* 13:1809.
- 8 Herskovits, T. T. 1966. On the conformation of caseins. Optical rotatory dispersion properties. *Biochemistry* 5:1018.
- 9 Lippert, J. L., D. Tyminski, and P. J. Desmeules. 1976. Determination of the secondary structure of proteins by laser Raman spectroscopy. *J. Am. Chem. Soc.* 98:7075.
- 10 Moffatt, D. J., J. K. Kauppinen, C. G. Cameron, H. H. Mansch, and R. N. Jones. 1986. Computer Programs for Infrared Spectrophotometry. *Bull. 18, Natl. Res. Council. Can., Ottawa, Ont., Can.*
- 11 Pepper, L., and H. M. Farrell, Jr. 1982. Interactions leading to formation of casein submicelles. *J. Dairy Sci.* 65:2259.
- 12 Raap, J., K.E.T. Kerling, H. J. Vreeman, and S. Visser. 1983. Peptide substrates from chymosin (rennin): conformational studies of κ -casein and some κ -casein-related oligopeptides by circular dichroism and secondary structure prediction. *Arch. Biochem. Biophys.* 221:117.
- 13 Savitsky, A., and M.J.E. Golay. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* 36:1627.
- 14 Schmidt, D. G. 1980. Colloidal aspects of casein. *Neth. Milk. Dairy J.* 34:42.
- 15 Susi, H., and D. M. Byler. 1986. Resolution-enhanced Fourier transform infrared spectroscopy of enzymes. *Methods Enzymol.* 130:290.
- 16 Susi, H., and D. M. Byler. 1988. Fourier deconvolution of the amide I Raman band of proteins as related to conformation. *Appl. Spectrosc.*

RAMAN SPECTROSCOPY OF CASEINS

- 42:819.
- 17 Thomas, G. J., Jr., and D. A. Agard. 1984. Quantitative analysis of nucleic acids, proteins and viruses by Raman band deconvolution. *Biophys. J.* 46: 763.
- 18 Thompson, M. P. 1966. DEAE-Cellulose-urea-chromatography of casein. *J. Dairy Sci.* 49:792.
- 19 Williams, R. W. 1986. Protein secondary structure analysis using Raman amide I and amide III spectra. *Methods Enzymol.* 130:311.