Laser-Excited Raman Spectroscopy of Biomolecules. VI. Some Polypeptides as Conformational Models

M. C. Chen[1] and R. C. Lord*

Contribution from the Spectroscopy Laboratory and Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.
Received March 5, 1974

Abstract: Laser Raman spectra of some polypeptides have been obtained in order to study the correlations between various conformations and the frequencies of the peptide linkage. The α-helical conformation in poly-L-alanine and poly-γ-benzyl-L-glutamate shows amide III frequencies in the range 1265–1300 cm^{-1}. The amide III vibration of poly-L-alanine is split into three lines at 1265, 1275, and 1283 cm^{-1}, corresponding to A_1, E_1, and E_2 modes, respectively. The β form in poly-L-valine gives a strong line at 1229 cm^{-1} and a weaker one at 1289 cm^{-1}, while in the random-coil conformation of poly-L-lysine the amide III is centered at about 1245 cm^{-1}. The amide I frequency also correlates with these structures but is found to be less sensitive to the backbone conformation.

In establishing relationships between the Raman spectra and structures of proteins,[2-5] poly-α-amino acids are often useful as model compounds, because homopolypeptides can be obtained in single ordered conformations. The principal lines in the Raman spectrum that contain useful information for this purpose are those of the amide I and III frequencies in the 1650- and 1250-cm^{-1} regions, which are due respectively to C=O stretching and to a mixture of C–N stretching with N–H in-plane bending in the polypeptide backbone. This paper reports a study of the polypeptides with the aim of relating the Raman lines to the polypeptide structure. Spectra of some of these polypeptides have already been reported[6-11] but many have now been obtained at better resolution and much improved background. The additional information thus provided enables us to make a more secure assessment of the structure–spectra relationships.

Experimental Section

Methods: Samples of highly purified poly-γ-benzyl-L-glutamate (PBLG) of molecular weights 30,000 and 275,000 were gifts from Dr. M. Panar and E. I. du Pont de Nemours and Co. Both samples gave identical spectra. Poly-L-alanine (PLA) of molecular weight 94,000 and poly-L-lysine hydrochloride (PLL) of molecular weight 188,000 were purchased from Schwarz/Mann. Poly-L-valine (PLV) and an additional sample of PLA were obtained from the Pilot Chemical Division of New England Nuclear. All samples were used without further purification. The spectroscopic techniques for recording spectra of solid samples have been described previously.[5] All spectra were recorded with a laser power at the sample of about 100 mW and a spectral slit width of 7 cm^{-1} in about 1 hr unless otherwise stated.

Results and Discussion

With the aid of samples of higher purity and improved instrumental techniques, it has been possible to obtain spectra of considerably better signal-to-noise ratio and higher resolution. Assignments of the amide I and III lines in the spectra of the polypeptides can now be based on comparison of these improved spectra with those of the corresponding amino acids in the solid state[10,12] and on data for deuterated polypeptides when available. Original spectra of the polymers are shown in Figures 1–5 in which the lines occurring in the 1250- and 1650-cm^{-1} regions are appropriately labeled.

(a) Polypeptides with α-Helical Conformation. In the solid state PLA is an α helix as determined by X-ray diffraction.[13] The amide I line is strong and very sharp at 1655 cm^{-1} (Figure 1), which indicates uniform hydrogen bonding, as expected in an α helix. The amide III line is a triplet at 1265, 1275, and 1283 cm^{-1}, which shifts to around 945 cm^{-1} upon deuteration.[14] The three amide III components are assigned as the A, E_1, and E_2 modes of this kind of vibration in the $C_{18/5}$ group to which the α helix belongs.[7,14] A theoretical calculation of the frequencies of these modes has been reported by Peticolas, and coworkers[14] for both PLA and the N-deuterated derivative. While the potential-energy distributions given by Peticolas, *et al.*, indicate, as expected, that the amide III vibration is a complicated mixture rather than a simple group motion, it appears that the best choice for the calculated frequencies of the A, E_1, and E_2 amide III modes is 1328, 1333, and 1340 cm^{-1}. While the center of gravity of this triplet is somewhat above the observed value of 1274 cm^{-1}, the spacings are in reasonably good agreement. Moreover, polarization studies[7] by these authors (on a sample with some β-pleated-sheet and random-coil contamination) indicate clearly that the observed component at 1265 cm^{-1} is to be assigned as A and that the 1275 component contains a substantial E_1 contribution. On this basis we feel that the triplet assignment as 1265 (A), 1275 (E_1), and 1283 (E_2) is well established. The observed frequencies just above 1300 cm^{-1} are not significantly

(1) Based in part on the Ph.D. Thesis of M. C. Chen, Department of Chemistry, Massachusetts Institute of Technology, Sept 1973.
(2) R. C. Lord, *Proc. Int. Congr. Pure Appl. Chem., Suppl., 23rd*, 7, 179 (1971).
(3) (a) N. T. Yu, C. S. Liu, and D. C. O'Shea, *J. Mol. Biol.*, **70**, 117 (1972); (b) N. T. Yu and B. H. Jo, *Arch. Biochem. Biophys.*, **156**, 469 (1973).
(4) M. C. Chen, R. C. Lord, and R. Mendelsohn, *Biochim. Biophys. Acta*, 328, 252 (1973).
(5) M. C. Chen, R. C. Lord, and R. Mendelsohn, *J. Amer. Chem. Soc.*, 96, 3038 (1974).
(6) J. L. Koenig and P. Sutton, *Biopolymers*, **8**, 167 (1969).
(7) B. Fanconi, B. Tomlinson, L. A. Nafie, W. Small, and W. L. Peticolas, *J. Chem. Phys.*, **51**, 3993 (1969).
(8) J. L. Koenig and P. Sutton, *Biopolymers*, 9, 1229 (1970).
(9) J. L. Koenig and P. Sutton, *Biopolymers*, **10**, 89 (1971).
(10) L. Simons, G. Bergstrom, G. Blomfelt, S. Forss, H. Stenback, and G. Wansen, *Comment. Phys.-Math.*, 42, 125 (1972).
(11) T. J. Yu, J. L. Lippert, and W. L. Peticolas, *Biopolymers*, 12, 2161 (1973).
(12) M. C. Chen, Ph.D. Thesis, Massachusetts Institute of Technology, 1973.
(13) A. Elliott and B. R. Malcolm, *Proc. Roy. Soc., Ser. A*, 249, 30 (1959).
(14) B. Fanconi, E. W. Small, and W. L. Peticolas, *Biopolymers*, **10**, 1277 (1971).

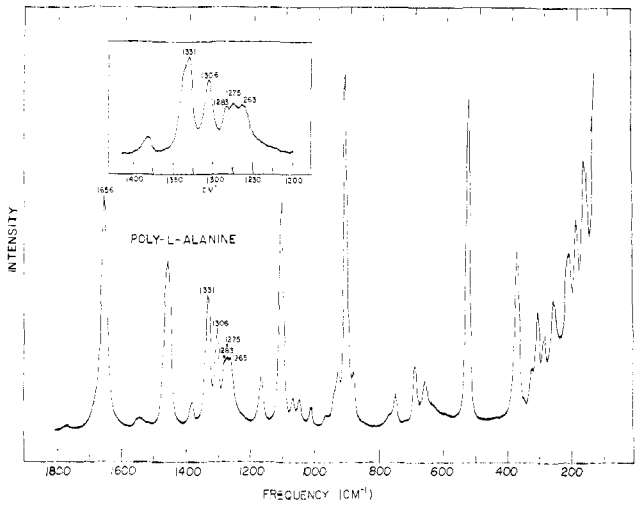

Figure 1. Raman spectrum of poly-L-alanine. The insert is recorded with a spectral slit width of 3 cm^{-1}, time constant 10 sec, and scan speed 0.05 cm^{-1}/sec.

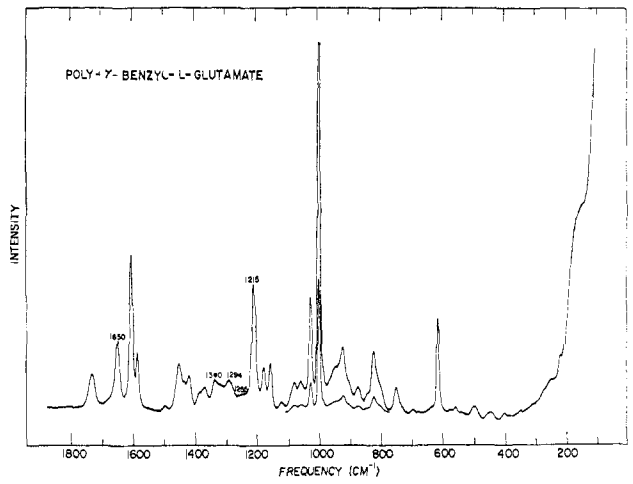

Figure 2. Raman spectrum of poly-γ-benzyl-L-glutamate.

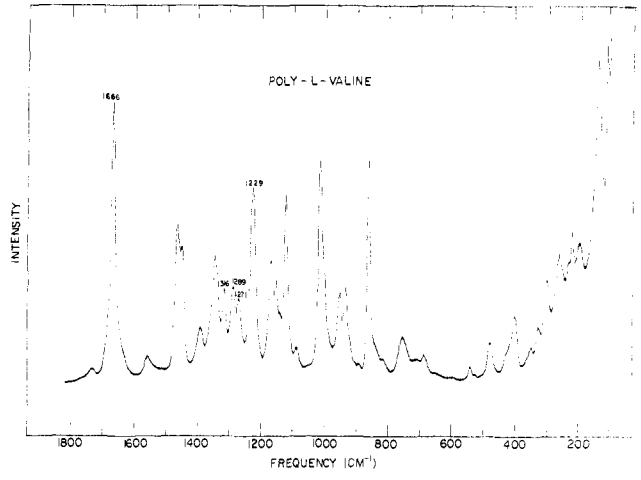

Figure 3. Raman spectrum of poly-L-valine.

affected by N-deuterated substitution and one of them (1306 cm^{-1}) appears in the monomer; thus they cannot be identified with the amide III, as was suggested earlier.[6] The foregoing assignment of the amide III triplet in PLA has been made previously by Simons, *et al.*,[10] although their sample also shows contamination by β-pleated-sheet and random-coil conformations, as evidenced by amide III Raman lines reported at 1227 and 1243 cm^{-1} and by the component of the amide I at 1665 cm^{-1}.

The spectrum of another α-helical polypeptide, PBLG, for which X-ray diffraction data are also available,[15] is shown in Figure 2. Comparison with the spectrum of the monomer established 1650 and 1294 cm^{-1} as the amide I and amide III frequencies, respectively. This is confirmed by the infrared spectra,[16] which show a weak but definite line at 1280 cm^{-1} that shifts to 961 cm^{-1} upon deuteration. Earlier Raman work on the amide III region in PBLG has ascribed the broad line at 1340 cm^{-1} to this mode,[9,10] which seems erroneous in view of the above evidence.

The amide III frequency in PBLG is significantly higher than that in PLA. Although the crystal structures of the α-helical forms of these two polypeptides[13,15] do not suggest a shorter N—H···O=C hydrogen bond for PBLG, the increase in the frequency from 1274 cm^{-1} in PLA to 1294 cm^{-1} in PBLG is accompanied by the expected smaller decrease in the amide I frequency (1655 → 1650 cm^{-1}). Thus we conclude that the side chain in PBLG must contribute to a strengthening of the α helix, leading thereby to stronger N—H···O=C bonding. A similar result is found in the spectrum of the α-helical form of poly-L-lysine,[11] where the amide III frequency is still higher (1311 cm^{-1}).

(b) Polypeptides with β Conformation. The Raman spectrum of a representative solid polypeptide of β-pleated-sheet conformation,[17-19] PLV, is given in Figure 3. Comparison with the spectrum of valine[10,12] reveals that the amide I line occurs at 1666 cm^{-1} while a strong line at 1229 cm^{-1} and a much weaker one at 1289 cm^{-1} are identified as amide III frequencies. In the Raman spectrum[20] of polyglycine I, which has a β conformation of crystal symmetry D_2, the corresponding lines occur at 1674 (A_1 transition of the C_{2v} factor group), 1234 (A_1), and 1295 cm^{-1} (B_1).[21] Koenig and Sutton[9] and Simons, *et al.*,[10] in their studies of PLV assign the weak lines at 1272 and 1291 cm^{-1} to amide III but give no assignment for the strong line at 1230 cm^{-1}. However, the 1272-cm^{-1} line occurs in the spectrum of the monomer and is almost surely due to the side chain. Other β conformations having strong amide III lines near 1230 cm^{-1} include poly-L-serine, where the frequency is given[9] as 1235 cm^{-1} and assigned as amide III, and PLL, in aqueous solution (pH 12, 52°) in the spectrum of which a broad line at 1240 cm^{-1} has also been assigned by Peticolas, *et al.*,[11] to amide III.

(c) Polypeptides with Random-Coil Conformation.

(15) A. Elliott, R. D. B. Fraser, and T. P. MacRae, *J. Mol. Biol.*, **11**, 821 (1965), and the references cited therein.
(16) Y. Masuda and T. Miyazawa, *Makromol. Chem.*, **103**, 261 (1967).
(17) R. D. B. Fraser, H. S. Harrap, T. P. MacRae, F. H. C. Stewart, and E. Suzuki, *J. Mol. Biol.*, **12**, 482 (1965).
(18) E. R. Blout, C. deLoze, S. M. Bloom, and G. D. Fasman, *J. Amer. Chem. Soc.*, 82, 3787 (1960).
(19) S. M. Bloom, G. D. Fasman, C. deLoze, and E. R. Blout, *J. Amer. Chem. Soc.*, **84**, 458 (1962).
(20) (a) E. W. Small, B. Fanconi, and W. L. Peticolas, *J. Chem. Phys.*, **52**, 4369 (1970); (b) S. Krimm and Y. Abe, *Proc. Nat. Acad. Sci. U. S.*, **69**, 2788 (1972).
(21) Small, Fanconi, and Peticolas[20] actually assign a weak line at 1220 cm^{-1} to the B_1 mode but it seems more likely that the weak component at 1295 cm^{-1} is the correct assignment for this mode.

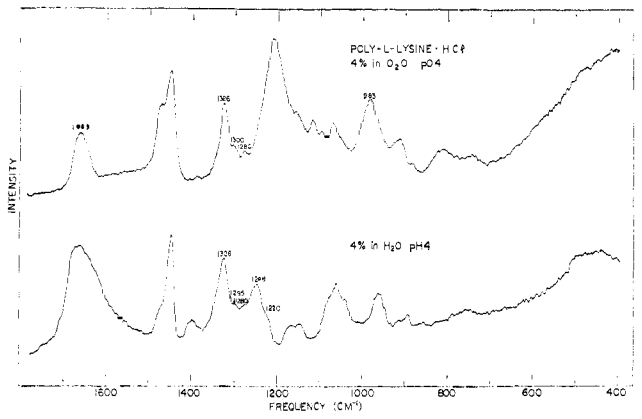

Figure 4. Raman spectra of 4% poly-L-lysine·HCl in H_2O and D_2O at pH (pD) 4.

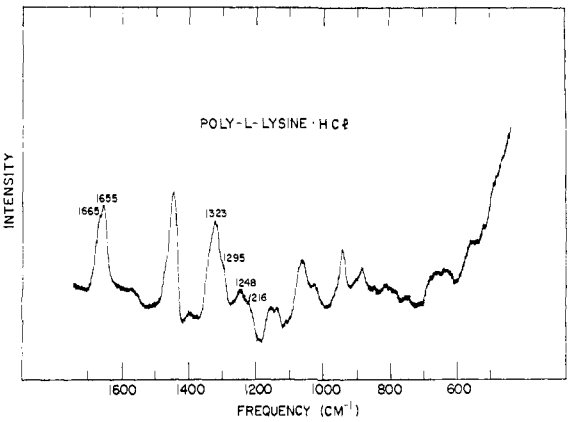

Figure 5. Raman spectrum of poly-L-lysine·HCl in the solid state recorded with a Cary 81 spectrometer. Power at the sample is 100 mW, the spectral slit width is 7 cm^{-1}, the time constant is 10 sec, and the scan speed is 0.1 cm^{-1}/sec. The relative humidity is about 50% at 21°.

At pH 4, PLL assumes a random-coil conformation.[22] The amide III line of the 4% aqueous solution at 1248 cm^{-1} is moderately strong, while the 1665 amide I line can be more easily seen when the polymer is dissolved in D_2O (Figure 4). In contrast to the sharp amide I in PLA, PBLG, and PLV, the broad line in PLL, with a half-width of approximately 40 cm^{-1} compared to 20 cm^{-1} in PLA, is indicative of a wider range of torsional dihedral angles for the polypeptide backbone. A similar result has been reported by Peticolas, *et al.*,[11] for PLL and by Lord and Yu[23] and Koenig and Frushour[24] for the random-coil conformation of poly-L-glutamic acid. In the solid PLL, which exists predominantly in the α-helical form at 21° and about 50% relative humidity,[22,25] the intensity of the 1295 line increases considerably compared to that in aqueous solution, and the amide I line at 1655 cm^{-1} develops a shoulder at 1665 cm^{-1} (Figure 5). A weak remnant of an amide III line at 1248 cm^{-1} and the 1665 shoulder show that there is still some residual random-coil conformation in the polypeptide sample. When PLL is in a fully α-helical form, as at pH 11.8 and 4°, the amide III frequency at 1311 cm^{-1} is overlapped by the C_α–H bending mode. The latter line still remains, though with less intensity, when the polymer is N-deuterated or converted into the β form. Koenig and Sutton[8] have also reported the Raman spectrum of PLL in the solid phase, but the location of the amide III line in their spectrum is obscured by its coalescence with a side-chain frequency at 1340 cm^{-1}.

The amide III frequencies of the three conformations may be summarized as follows: α-helix, 1265–1300 cm^{-1} (medium); β-antiparallel-pleated sheet, 1229–1235 (strong) and 1289–1295 cm^{-1} (weak); and random coil, 1243–1253 cm^{-1} (medium strong). These frequencies have been verified in the Raman spectra of several proteins.[2-5,23] It is usually difficult to see the 1290-cm^{-1} line of the β form because it is considerably weaker than that at ~1230 cm^{-1} and there is considerable overlap with weak side-chain frequencies in this region of the spectra of proteins. In the β form of glucagon[26] and PLL,[11] the amide I line is sharp and intense at approximately 1670 cm^{-1}, which is about 4 cm^{-1} higher than that of PLV reported in this paper. However, as has been observed above for PLA and PBLG, the amide I frequency is only about one-third as sensitive to conformational change as the amide III, so this small difference might arise from other factors than hydrogen bonding.

Caution is clearly necessary in applying these characteristic frequencies to the spectra of proteins, because the symmetry in ordered homopolypeptides may restrict both the number and position of the lines in the spectrum. These restrictions are removed in proteins, where there are a variety of side chains, and thus in principle no restriction is imposed on the spectroscopic activity of the various modes of the amide III vibration. Nevertheless, interpretation of protein spectra is considerably aided by an understanding of the structural implications of these characteristic frequencies. In particular, if the integrated intensities and shapes of the amide III lines can be properly calibrated for model compounds, quantitative assessment of the amounts of α-helical, β-pleated-sheet, and random-coil conformations in proteins should be possible. We are pursuing this possibility with the Raman spectra of a number of globular proteins whose crystal structures are known.

Acknowledgment. We wish to thank Dr. M. Panar for the gift of purified PBLG samples. We are indebted to the Jarrell-Ash Division of Fisher Scientific Co. and Dr. G. J. Thomas, Jr., for the use of their ininstruments. This work was generously supported by the National Institutes of Health and the National Science Foundation under the respective Grants GM-15310, GP-13473, and GP-35677.

(22) (a) K. Rosenheck and P. Doty, *Proc. Nat. Acad. Sci. U. S.*, 47, 1775 (1961); (b) U. Shmueli and W. Traub, *J. Mol. Biol.*, 12, 205 (1965).
(23) R. C. Lord and N. T. Yu, *J. Mol. Biol.*, **51**, 203 (1970).
(24) J. L. Koenig and B. Frushour, *Biopolymers*, **11**, 1871 (1972).
(25) E. R. Blout and H. Lenormant, *Nature (London)*, **179**, 960 (1957).
(26) N. T. Yu and C. S. Liu, *J. Amer. Chem. Soc.*, **94**, 5127 (1972).