Raman Spectroscopy, the Sleeping Giant in Structural Biology, Awakes*

Paul R. Carey‡

From the Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106

Principally through the efforts of crystallographers, we are being presented with an ever expanding atomic view of the biological world. Although this brings into focus many questions regarding the mysteries of function, techniques are needed that facilitate the transition in our understanding from structure to function. Raman spectroscopy is one of these; because the Raman effect involves an intimate interplay between atomic positions, electron distribution, and intermolecular forces, it sits at the bridgehead between structure and function. Thus, the Raman technique can answer questions that lie at the heart of issues such as ligand macromolecule recognition and enzymatic catalysis. Raman spectroscopy involves analyzing the scattered photons from a laser beam focused into the sample solution (1). The inelastic scattered photons (the Raman spectrum) provide information on molecular vibrations that, in turn, yield data on molecular conformation and environment. At its most effective, Raman spectroscopy can provide exquisite detail from an important site in a much larger macromolecular complex. Although Raman was first applied to the definition of biological molecules in the 1920s (2), the giant has remained dormant due to the difficulties both in obtaining high quality data and in interpreting those data. Considerable advances have been made in these areas in the past few years, and the giant is stirring!

A major goal of this review is to provide biochemists with enough information to determine whether the Raman technique could provide structural insights into their systems. The specific issues addressed are which type of systems are amenable to study and what information could be obtained. Practically, present-day sample requirements are for 20 µl of clear solution, where the target molecule is in the 100–300 µm range. Because the number of vibrational modes of a molecule is 3n−6, where n is the number of atoms, the complete Raman spectrum of a macromolecule is exceedingly complex. Thus, Raman is most suited to systems where it is possible to focus upon a small region of interest, e.g. a ligand-receptor or enzyme-substrate binding site. Historically, this condition was achieved by using resonance Raman spectroscopy (1) to obtain the intensity-enhanced spectra from chromophores at specific sites in macromolecules. Recent technical advances mean that similar information can now be gleaned from non-chromophoric systems, markedly broadening the application of the technique.

The information obtained can be very detailed, exceeding the level of resolution found in x-ray or NMR analyses (3–5). In addition to providing structural data, the Raman spectrum can also reveal changes in the distribution of electrons in a bound ligand and details of active site-ligand interactions, such as hydrogen bonding strengths.

Raman spectroscopy is beginning to fulfill its potential to contribute to structural biology because the three roadblocks that impeded its application to biological systems have been all but removed. These were low sensitivity, interference from fluorescence background, and problems with data interpretation. Sensitivity has increased several orders of magnitude with the corresponding decrease in concentration requirements because of advances in optical filters and photon detectors (6). Fluorescence interference is now minimized by using deep-red excitation in the 650–800 nm range, made possible by the advent of photon detectors with high efficiency in this region (7). Problems with interpreting Raman spectra have receded with the availability of "friendly" software packages (8) and ever increasing computational power that enable us to calculate, ab initio, the Raman spectra of molecules (of the size of many ligands or co-factors found at biological sites). Interpretation is further strengthened by a comparison of the calculated and experimental shifts in Raman peak positions when a molecule is substituted with stable isotopes. Recently, this approach has been used to characterize hydrogen bonding in a complex of adenosine deaminase with a transition state analogue (9) and to discriminate between different protonation states of dihydrofolate binding to dihydrofolate reductase (10).

What Is Raman and What Does It Tell You?

The Raman spectrum of 5-methyl thienylacrylic acid (5-MTA) is shown in Fig. 1; the 5-MTA entity has been used extensively as a probe of protease active sites (3, 11, 12). The spectrum was obtained by focusing a laser beam into a solution in methanol and by analyzing the scattered light 90° to the direction of the beam using a Raman spectrometer. A small percentage of the scattered photons exchange energy with the vibrational energy levels (or, crudely, the "vibrations") of the molecules in solution. Thus, by analyzing the scattered photons information on the vibrational motions of atoms in molecules is obtained. These motions are a function of molecular conformation, of the distribution of electrons in the chemical bonds, and of the molecular environment. Thus, interpretation of the Raman spectrum provides information on all these factors. This is the underlying principle behind using Raman spectroscopy for defining the detailed chemistry of molecules at biological sites (1). With present day computational power, commercially available software packages allow us to undertake high level quantum mechanical calculations on molecules such as 5-MTA acid and to predict the stable conformational states for this molecule as well as the infrared and Raman active vibrations (13). Such calculations put interpretation of the data on a sure footing. They also reveal the complex nature of vibrational spectra; many peaks are due to vibrational motions that include contributions from many atoms in the molecule. However, some vibrations are more or less localized in molecular groupings and three such are indicated in Fig. 1 (the C=O and C=C stretching vibrations and the breathing-type motion of the thienyl ring). We will see below how the carbonyl peak can provide detailed chemical information on the chemistry of this group in serine protease active sites. In addition, the C=C stretch and ring modes can be used to follow the redistribution of π-electrons for 5-MTA in cysteine protease active sites (11), and marker bands in the 1000–1200 cm⁻¹ region give the conformation, cis or trans, about the =C–C=O single bond (13, 14).

The absorption spectrum of the 5-MTA chromophore shows a maximum in the near UV at 324 nm. If a laser wavelength far from this electronic transition, e.g. near 650 nm, is used to generate the Raman spectrum the resulting spectrum is relatively weak. If, however, we use an excitation wavelength near 330 nm the coincident between this and the absorption band leads to large (10³ or more) intensity enhancement. This is the resonance Raman (RR) effect, and because of its high intensity it has been the method most used for obtaining vibrational spectra at biological sites. RR plays a key role in obtaining data from natural chromophoric sites such as occur in heme (15–17) and metalloproteins (18). Moreover, time-

* This minireview will be reprinted in the 1999 Minireview Compendium, which will be available in December, 1999. Research in this laboratory is supported by National Institutes of Health Grants GM 54072 and DK 53053.‡ To whom correspondence should be addressed: Dept. of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106. Tel.: 216-368-0031; Fax: 216-368-4544; E-mail: carey@biochemistry.CWRU.edu.

The abbreviations used are: 5-MTA, 5-methyl thienylacrylic acid; RR, resonance Raman.
resolved RR is a powerful means of following changes at a chromophoric site in a rapidly evolving system such as cytochrome oxidase (19) or a peptide in the early stages of folding (20). However, for stable or slowly evolving systems, increases in sensitivity now allow us to obtain Raman difference spectra from specific sites under non-resonance conditions. The RR approach remains the only method available to probe short-lived species (with half-lives of less than a second) and picosecond range), and RR studies on some reactive acyl enzymes, prior to deacylation, at high pH in a rapid mixing experiment in the early stages of folding (20). However, for stable or slowly evolving systems, increases in sensitivity now allow us to obtain Raman difference spectra from specific sites under non-resonance conditions.

The work discussed above on acyl serine proteases utilized near UV lasers operating near 350 nm to generate RR spectra of the bound acyl groups. The spectra were recorded using double or triple monochromators (to separate the “Raman” from the interfering “Rayleigh” photons) and detected by a single photomultiplier or, later, an optical multichannel analyzer. Encouraged by the relationship cited above, attempts were made to extend the studies to other acyl cysteine proteases, e.g. R-C≡C-(C≡O)-S—papain, and to other α,β-unsaturated thiol esters such as hexadienyl-CoA binding to enoyl-CoA hydratase (29). Again these acyl groups have absorption features near 350 nm and are candidates for RR studies. However, satisfactory spectra could not be obtained from these systems because α,β-unsaturated thiol esters are highly photolabile, and the samples underwent uncontrolled photoisomerization and decomposition in the laser beam used to generate the RR data. Technical innovations came to the rescue: a combination of efficient holography-based optical filters to block the Rayleigh photons and high quantum efficiency charge coupled device photon detectors increased the sensitivity of Raman instrumentation dramatically (6). In practical terms this meant that it was no longer necessary to use the resonance condition. For example, high quality Raman data were obtainable using excitation near 500 or 650 nm. This is far removed from any absorption peak, and thus the samples are not photochemically modified by the laser beam (30).

To observe the Raman spectrum of the bound ligand, Raman difference spectroscopy was employed where the spectrum of the enzyme is subtracted from the spectrum of the enzyme-ligand complex (31, 32). The resultant difference spectrum contains features due to the bound ligand with the possibility of some protein modes appearing if there are conformational changes occurring upon ligand binding. Ligands such as the 5-MTA seen in Fig. 1 are strong Raman scatterers (they give rise to relatively intense Raman signals even under non-resonance conditions) because of the extended and polarizable π-electron system. For this reason the ligand modes dominate the difference spectrum. When experiments involve ligands that scatter less intensely, protein features appear in the difference spectrum in both the positive and negative directions, and these are a source of information on changes in protein structure, although the detailed interpretation of many of these features is still in its infancy (33, 34).

The Raman difference spectra for acyl cysteine proteases, supported by absorption spectral data, provided different insights from those uncovered for the serine analogs (11, 35). Principally, these involve so-called π-electron polarization; in the cysteine protease active sites there are strong electrostatic forces that bring about a major rearrangement of the π-electrons in the acyl group. An α-helix dipole, terminating at the active site cysteine, is likely one of the important factors causing polarization of the acyl group’s electrons, and it was proposed that the dipole, with its positive pole pointing toward the acyl C=O, functions to stabilize negative charge build-up in the transition state (11).

Ideas on active site-induced electron polarization have been further refined by Raman difference studies on acyl enzymes involving the semi-synthetic enzymes, thiol and selenol subtilisins (36). The Raman data for these 5-MTA acyl enzymes showed that the acyl group experiences no polarization in the active sites. However, when the active sites are enlarged, e.g. in selenol subtilisin by replacing asparagine 155 by a glycine residue, a new conformational state is detected, and this second conformation is polarized. Thus, the acyl group is able to switch from a region of null electric field to one where strong electrostatic forces come into play. Using the x-ray-derived structure of selenol subtilisin combined with modeling studies (13, 36), it was possible to provide a molecular explanation for the polarized and non-polarized forms of the acyl groups observed by spectroscopy. In essence, the additional room in the active site, created by making the Gly-155 mutant, allows the acyl group to “flip” about its C=O single bond. Then, it goes from an environment where it experiences minimal electrostatic forces to one where there is a negatively charged side chain near the thiolyl ring and strong electropositive effects at the carbonyl because of hydrogen bonds and an α-helix dipole. Thus, optimum
polarization appears to be achieved by a combination of electron “push” and “pull,” as shown in the schematic in Fig. 2.

The above discussion focuses on the use of the S-MTA group to demonstrate many of the details on structure and chemistry that can be elicited from Raman data. This molecule was one of the chromophoric acyl groups developed as RR probes of active sites in early studies. It is not part of a natural substrate for the enzymes discussed. Now, however, with the latest generation of highly sensitive Raman spectrometers and the use of red or deep red excitation to avoid fluorescence interference, we are able to consider a plethora of “natural” enzyme-ligand systems. One such involves the enzyme 4-chlorobenzoyl-CoA dehalogenase that carries out the transformation as shown in Reaction 1.

Initial attempts to undertake RR studies of dehalogenase-product complexes were thwarted again by the photolability of the CoA thiol esters. However, using a Raman difference spectrometer constructed in 1992 (6), red-excited Raman difference data for the product complex were obtained (37). These showed that the active site of dehalogenase can bring about a complete reorganization of the benzoyl group’s electrons and, in turn, provided insight into how the difficult chemical step of replacing the –Cl atom on the ring by an –OH is achieved. The determinants for this strong polarization in the active site of dehalogenase (38) are very similar to those found for the cysteine proteases and shown in Fig. 2. At the benzoyl C=O there is strong electron pull exerted by two H-bonds and an α-helix dipole combined with electron push by an aspartate side chain near the benzoyl’s para position.

Even with the 1992 vintage high throughput Raman spectrometer, studies on dehalogenase complexes were hampered by concentration requirements. The latter were close to 1 mM and were usually thwarted by the protein coming out of solution during concentration procedures. However, in 1997 we were able to modify a commercial spectrophotograph to optimize its performance for dilute aqueous solutions (7). With this instrument concentration requirements were lowered to 100–300 μM. Collecting data for dehalogenase complexes reached a success rate of 100%, and high quality spectra were obtained for several dozen complexes with different substrate analogs and protein-engineered forms of the enzyme (39). One of these complexes had the surprising property of evolving with time (40). It involved the substrate binding to dehalogenase where the catalytically vital group Asp-145 had been mutated to asparagine. The enzyme was expected to be unreactive, and yet a series of spectroscopic changes were detected over a period of several minutes. Fig. 3 shows Raman difference spectra recorded approximately 1, 1.5, 2, and 4.5 min after adding the substrate to the Asn-145 variant. The data collection time for each spectrum was only 30 s, but the spectral fingerprints of several species can be identified. Initially (in the top trace) the peaks at 1586 and 1216 cm⁻¹ show that there is the expected population of the 4-chloro-substrate bound to the active site of the Asn-145 enzyme. However, this population decays rapidly with time. There is also evidence at early times for peaks at 1543 and 1490 cm⁻¹, and detailed analysis (39) shows that these are because of the ionized (4-O⁻) form of the product bound to the Asn-145 enzyme. After 5 min the spectra cease evolving, and the signature (Fig. 3, bottom trace) is that of non-ionized product in the active site of the wild-type enzyme! The combined Raman and absorption data provide the following explanation for the spectral changes. Initially, the Asp-145 mutant enzyme contains a trace of the wild-type enzyme brought about by spontaneous deamidation of the Asn-145 side chain. The wild-type enzyme catalyzes the formation of product, which binds in its ionized form to the large population of Asn-145 molecules. Thus, early in the time sequence there are populations of substrate and product binding to the dominant Asn-145 population. However, the ionized product catalyzes the deamidation of D145N to give the wild-type enzyme. Thus, in a few minutes all the substrate is converted to product and all the mutant enzyme is converted to the wild type, and we see the interesting phenomenon of product catalyzing changes to the enzyme. In general these studies show: 1) that Raman difference spectra were able to identify species in a complex evolving reaction and to aid in identifying the reaction mechanism and 2) that transient species (with lifetimes of tens of seconds or longer) could be identified at the sub-100 μM level.

**Limitations and Prospects**

Although there are outstanding examples where the Raman approach provides important information on large macromolecular structures such as nucleic acid-protein complexes (41–43) it may find most use in defining small regions of large complexes. In that sense it complements techniques such as x-ray crystallography that provide the “big picture.” Another limitation is that, occasionally, samples are encountered that have high background signals that are difficult to remove. This is an area of active investigation since the source of the background is not understood fully. The target molecules most suited for Raman difference spectroscopy are those that have relatively intense normal (non-resonance) Raman signals, which usually means that they have extended π-electron
systems because these are polarizable and give rise to strong Raman scattering. Saturated systems such as carbohydrates are less amenable for Raman analysis.

Even with the above limitations the prospects are good that Raman spectroscopy will make an increasing impact in structural biology. In the author’s own area of interest, involving enzyme complexes, the last few years have seen the Raman technique go from being applicable to a very few systems to being a means of addressing mechanistic questions for a wide array of enzymes. It is now possible to follow changes in co-factor chemistry in detail. For example, flavoproteins have been difficult subjects for Raman investigations because of their intrinsic fluorescence. However, using red or deep red laser excitation, high quality non-resonance Raman data can now be obtained from the isoalloxazine ring system of flavins (32), and these promise to reveal a wealth of information. There are as yet few studies reported involving receptor-ligand complexes; these should appear in the near future. Similarly, the potential for Raman to provide, fairly rapidly, molecular information for a family of ligands binding to a receptor or enzyme target has yet to be exploited. In another application, Raman microscopes permit us to obtain the Raman spectra of microscopic objects under controlled conditions (44). High quality data can be obtained for protein crystals, and “Raman crystallography” has the potential to provide detailed information on complexes in the crystalline as well as the liquid phases and can thus provide a bridge between the results of the crystallographers and solution studies. Improvements in theory and computational power are leading theorists to calculate increasingly sophisticated models of macromolecular binding sites (45, 46), and Raman may have a role in providing benchmarks against which to test the results of calculations.

Finally, will Raman spectroscopy become an important tool in many biochemical laboratories as for example CD now is? Perhaps it will not be as widespread, but it may become an indispensable tool for scientists interested in the chemistry of many classes of small molecule-big molecule interactions. The tide is running in its favor, the cost of instrumentation is falling, experiments are becoming easier to undertake, and the results are becoming easier to interpret in a quantitative fashion.

Acknowledgments—I am indebted to my students and collaborators for contributions to the work discussed in this review and to colleagues for sharing some of their most recent advances.

REFERENCES

1. Carey, P. R. (1982) Biochemical Applications of Raman and Resonance Raman Spectroscopies, Academic Press, New York
2. Edsall, J. T. (1936) J. Chem. Phys. 4, 1–6
3. Carey, P. R., and Tonge, P. J. (1995) Acc. Chem. Res. 28, 8–13
4. Wang, J. H., Xiao, D. G., Deng, H., Webb, M. R., and Callender, R. (1998) Biochemistry 37, 1110–1116
5. Deng, H., Wang, J., Callender, R. H., Grammer, J. C., and Yount, R. G. (1998) Biochemistry 37, 10972–10979
6. Kim, M., Owen, H., and Carey, P. R. (1993) Appl. Spectrosc. 47, 1780–1783
7. Deng, J., Dinakarpandian, D., and Carey, P. R. (1998) Appl. Spectrosc. 52, 1117–1122
8. Frisch, M. J., et al., and Pople, J. A. (1995) Gaussian 94, Gaussian Inc., Pittsburgh, PA
9. Deng, H., Kurtz, L. C., Rudolf, F. B., and Callender, R. (1998) Biochemistry 37, 4968–4976
10. Deng, H., and Callender, R. (1998) J. Am. Chem. Soc. 120, 7730–7737
11. Doran, J. D., and Carey, P. R. (1996) Biochemistry 35, 12485–12502
12. Dinakarpandian, D., Shemyo, B., Pusztai-Carey, M., Malcolm, B. A., and Carey, P. R. (1997) Biochemistry 36, 4943–4948
13. Dinakarpandian, D., and Carey, P. R. (1999) Biospectroscopy 5, in press
14. O’Connor, M. J., Dunlap, R. E., Odom, J. D., Hilvert, D., Pusztai-Carey, M., Shemyo, B., and Carey, P. R. (1996) J. Am. Chem. Soc. 118, 239–246
15. Kitagawa, T., and Mizutani, Y. (1994) Coordination Chemistry Rev. 135/136, 685–735
16. Blackwood, M. E., Rush, T. S., III, Ronesberg, F., Schultz, P. G., and Spiro, T. G. (1998) Biochemistry 37, 779–782
17. Hu, S., Smith, K. M., and Spiro, T. G. (1996) J. Am. Chem. Soc. 118, 12638–12646
18. Dong, S., and Spiro, T. G. (1998) J. Am. Chem. Soc. 120, 10434–10440
19. Kitagawa, T., and Ogura, T. (1997) Prog. Inorg. Chem. 45, 431–479
20. Lednev, I. K., Karnoup, A. S., Sparrow, M. C., and Asher, S. A. (1999) J. Am. Chem. Soc. 121, 4076–4077
21. MacClement, B. A., Carriere, R. G., Phelps, D. J., and Carey, P. R. (1981) Biochemistry 20, 3438–3447
22. Phelps, D. J., Schneider, H., and Carey, P. R. (1981) Biochemistry 20, 3447–3454
23. Carey, P. R., and Phelps, D. J. (1983) Can. J. Chem. 61, 2590–2595
24. Tonge, P. J., and Carey, P. R. (1989) Biochemistry 28, 6701–6709
25. Tonge, P. J., and Carey, P. R. (1990) Biochemistry 29, 10723–10727
26. Herath, G., Illenyi, J., Pusztay, L., and Simon, R. (1987) Acta Chem. Hung. 124, 819–822
27. Tonge, P. J., and Carey, P. R. (1992) Biochemistry 31, 9122–9125
28. Tonge, P. J., Fausto, R., and Carey, P. R. (1996) J. Mol. Struct. 379, 135–142
29. Tonge, P. J., Anderson, V. E., Fausto, R., Kim, M., Pusztai-Carey, M., and Carey, P. R. (1995) Biospectroscopy 1, 387–394
30. Tonge, P. J., Carey, P. R., Callender, R., Deng, H., Ekiel, I., and Mahandiram, D. R. (1995) J. Am. Chem. Soc. 117, 8757–8762
31. Callender, R., and Deng, H. (1994) Annu. Rev. Biophys. Biomol. Struct. 23, 215–245
32. Kim, M., and Carey, P. R. (1993) J. Am. Chem. Soc. 115, 7015–7016
33. Clarkson, J., Jaffe, E. K., Petrovich, R. M., and Carey, P. R. (1997) J. Am. Chem. Soc. 119, 11556–11557
34. Clarkson, J., Palley, B. A., and Carey, P. R. (1997) Biochemistry 36, 9511–9516
35. Doran, J. D., Tonge, P. J., Mort, J. S., and Carey, P. R. (1996) Biochemistry 35, 12487–12494
36. Dinakarpandian, D., Shemyo, B. C., Hilvert, D., McRee, D., McTigue, M., and Carey, P. R. (1999) Biochemistry 38, 6650–6667
37. Taylor, K., Liu, R. Q., Liang, P. H., Price, J., Dunaway-Mariano, D., Tonge, P. J., Clarkson, J., and Carey, P. R. (1995) Biochemistry 34, 13881–13888
38. Benning, M. M., Taylor, K. L., Liu, R., Yang, G., Xiang, H., Wesenberg, G., Dunaway-Mariano, D., and Holden, H. M. (1996) Biochemistry 35, 8103–8109
39. Dong, J., Xiang, H., Luo, L., Dunaway-Mariano, D., and Carey, P. R. (1999) Biochemistry 38, 4207–4213
40. Laporte, L., and Thomas, G. J., Jr. (1998) Biochemistry 37, 1327–1335
41. Laporte, L., Benevides, J. M., and Thomas, G. J., Jr. (1999) Biochemistry 38, 582–588
42. Thomas, G. J., Jr. (1999) Annu. Rev. Biophys. Biomol. Struct. 28, 1–27
43. Puppels, G. J., de Mul, F. P. M., Otto, C., Greve, J., Robert-Nicoud, M., Arnt-Jovin, D. J., and Jevin, T. (1990) Nature 347, 301–303
44. Cunningham, M. A., and Bash, P. A. (1997) Biochimie (Paris) 79, 687–689
45. Warshel, A. (1998) J. Biol. Chem. 273, 27035–27038