

# NONRESONANCE RAMAN DIFFERENCE SPECTROSCOPY: A General Probe of Protein Structure, Ligand Binding, Enzymatic Catalysis, and the Structures of Other Biomacromolecules

*Robert Callender and Hua Deng*

Department of Physics, City College of the City University of New York,  
New York, NY 10031

KEY WORDS: vibrational spectroscopy, emerging technique,  
enzymes, active sites, energy-vibrational frequency  
correlations

---

## CONTENTS

PERSPECTIVES AND OVERVIEW .....	216
GENERAL CONSIDERATIONS OF RAMAN DIFFERENCE SPECTROSCOPY .....	218
<i>Raman Spectroscopy</i> .....	218
<i>Anticipated Signal Size and Sources of Error in Raman Difference     Spectroscopy</i> .....	219
<i>The Difference Raman Spectrometer Used in Our Studies</i> .....	221
STRUCTURES AND INTERACTIONS OF LIGANDS WITH PROTEINS PROBED BY RAMAN DIFFERENCE SPECTROSCOPY .....	228
<i>Covalent Bond Formation</i> .....	229
<i>Hydrogen Bonding (and Weak Electrostatic Interactions)</i> .....	230
<i>Ligand-Induced Changes in Protein Structure</i> .....	235
ENZYMATIC CATALYSIS .....	236
PROPERTIES OF INDIVIDUAL RESIDUES AND THE EFFECTS OF MUTATION ON PROTEIN STRUCTURE .....	240
CONCLUSION .....	242

215

## PERSPECTIVES AND OVERVIEW

Both infrared and Raman spectroscopies measure the vibrational frequency of a group of bonded atoms. The masses of the atoms and the force constants of the various bonds between the atoms determine this frequency, from which bond orders and bond lengths can be surmised (cf 29). The interactions that take place between molecular groups, such as hydrogen bonding, affect the force constants of certain bonds and, hence, vibrational frequencies. Thus, one can probe the energies of hydrogen bonds directly by examining shifts in the frequencies of certain bands from so-called Badger-Bauer relationships (2). This method for determining these parameters is very accurate; it is generally more accurate than probes of protein structure such as X-ray crystallographic and multidimensional NMR studies. On the other hand, while vibrational spectroscopy measures on a very fine scale, one cannot determine the full conformation or configuration of a protein from its vibrational spectrum—at least not at present. Hence, the information that these probes provide is quite synergistic.

Despite the obvious value vibrational spectroscopy could have in the determination of biomacromolecular structure, spectral crowding has greatly hampered its use. Many vibrational modes contribute to the spectrum of a protein at each frequency because the bandwidth of a given mode is relatively broad, making the spectrum very difficult to interpret. Selective measurements are usually necessary to obtain understandable results. Hence, resonance Raman spectroscopy is often performed using chromophores inside proteins. Because of resonance enhancement, the spectrum of the chromophore dominates the observed spectrum (cf 13, 49, 50). Also, Fourier transform infrared (FTIR) difference spectroscopy is widely used in systems that contain a photolabile chromophore, such as e.g. bacteriorhodopsin or the photosynthetic reaction center. In this technique, light is used to switch protein films from one state to another. This preserves the optical geometry to achieve a high degree of subtraction fidelity, and only those protein vibrational modes affected by the light are observed in the difference spectrum (cf 6–8, 41, 45). FTIR difference spectroscopy is also useful in studying proteins containing electrochemically switchable redox systems, such as cytochromes, hemoglobin, and the photosynthetic reaction center, so that one can control the state of the protein (reduced or oxidized) with the applied potential (cf 41). Although these methods are powerful and extremely useful in understanding the (very important) protein systems that fulfill the necessary experimental constraints for resonance Raman or light (redox)-induced

FTIR difference spectroscopy, they are not general methods for studying the structures of proteins or protein-ligand interactions as only certain systems and bonds can be studied.

In the past several years, very accurate nonresonance Raman difference spectroscopy, called simply Raman difference spectroscopy in this review, has emerged as a technically feasible, general method of probing protein structure (cf 10, 14, 19, 34, 42, 61). In such experiments, proteins are tagged in some way and the difference spectrum between the protein and its modified version measured. For example, one could measure a protein spectrum and the spectrum of a protein complexed with ligand. Subtraction of the two yields the spectrum of the bound ligand. Alternatively, an atom within a bond of interest could be labeled with a stable isotope ( $^2\text{H}$ ,  $^{13}\text{C}$ ,  $^{15}\text{N}$ ,  $^{18}\text{O}$ , etc), which shifts the frequency of the modes that involve the motion of the labeled atom. Subtraction of labeled and unlabeled protein spectra yields an isotopically edited difference spectrum that shows only those modes involving the labeled atom. These and other protocols, described below, provide sufficient selectivity and so yield an interpretable spectrum.

The analysis below suggests that discerning a particular band from the protein background requires a protein Raman signal with a signal-to-noise ratio in excess of 300/1; the Raman spectra of proteins can now be obtained in an hour with a signal-to-noise ratio of 1000/1. Moreover, water ( $\text{H}_2\text{O}$  or  $\text{D}_2\text{O}$ ) generally has a very small Raman cross section, allowing the study of samples in their biological environment, and virtually the entire spectrum of a protein or protein-ligand complex can be measured. One can also obtain the spectrum of membrane-bound proteins (60). Thus, Raman difference spectroscopy can be used on most protein systems, and results are available in a relatively short period of time. Following similar protocols, investigators have recently successfully applied FTIR difference techniques to small proteins. This work was performed on samples in  $\text{D}_2\text{O}$  in order to move the intense and otherwise masking water absorption band from about  $1640\text{ cm}^{-1}$ , the spectral region of interest for these experiments, down to  $1200\text{ cm}^{-1}$  in order to study certain protein-ligand interactions (3, 47, 54, 55, 58, 59). This procedure may also develop into a general tool, particularly if some technique can be devised to overcome the problem of the strong water ( $\text{H}_2\text{O}$  or  $\text{D}_2\text{O}$ ) absorption.

For the past few years, our laboratory has been developing techniques and protocols to perform Raman difference spectroscopy in order to better understand key aspects of protein structure in certain enzyme systems that interest us. The information we wanted was unavailable through other techniques, such as NMR and diffraction meth-

ods. These enzymes, it turned out, could also not be studied with resonance Raman or FTIR difference spectroscopy, for the reasons outlined above. These studies showed the analysis potential for many other issues of protein structure of Raman difference spectroscopy; we have partially investigated some of these. This review is written for those who are interested in how proteins function and therefore are interested in their structures. For many problems, vibrational spectroscopy is of real value, and Raman difference spectroscopy offers a method of measurement that can be used on most systems.

Virtually all work to date developing and using Raman difference spectroscopy of large molecules has been performed in our laboratory, and our work has essentially focused on proteins, so this review concerns only proteins. However, the methodology surely could be extended to virtually any biomacromolecule and its assemblies. Indeed, the spectra of individual components of DNA assemblies were recently obtained using Raman difference spectroscopy (52).

## GENERAL CONSIDERATIONS OF RAMAN DIFFERENCE SPECTROSCOPY

### *Raman Spectroscopy*

In Raman spectroscopy, when light of a certain frequency (say  $\nu_L$  since the light is usually produced by a laser) irradiates a sample, a (small) portion is scattered from the sample, and the frequency shifts by an amount corresponding to the frequency of a particular vibrational mode,  $\nu_O$ . Thus, photons are emitted from the sample with a frequency of  $\nu_L - \nu_O$ . Because the frequency of the laser is known, the frequency of the vibrational mode can be determined. Figure 1 shows a Raman spectrometer set up. The universally used unit of the frequency is wavenumbers ( $\text{cm}^{-1}$ ), which is the reciprocal of the wavelength of  $\nu_O$  expressed in centimeters. The plotted spectrum is the intensity on the y-axis (usually in arbitrary units but proportional to the number of detected photons) versus frequency. Each peak in this plot corresponds to the frequency of a vibrational mode.

Fluorescence is a general problem facing the execution of Raman spectroscopy. However, in the approach taken here, laser light in the visible (from the blue to the far red) is used to excite Raman scattering. This is known as nonresonance or classical Raman spectroscopy because the protein samples under study, and their ligands, generally do not have visible-spectrum absorption bands. Thus, in principle, there

is no fluorescence background because no absorption takes place; this chief advantage of the present approach is required for nonresonance Raman spectroscopy to function as a general probe. In practice, most biological samples contain trace amounts of impurities, and these may yield fluorescence. Of the 20 or so proteins we have studied over the past several years, about half of them were so contaminated. On the other hand, we have nearly always found purification routes (generally involving standard biochemical approaches) that have yielded samples sufficiently free of these impurities.

The characteristic time scale associated with vibrational spectroscopy is on the order of the vibrational motion ( $\sim 10^{-13}$  s) or faster so that signal size and its characteristics are not affected by slow motions of the protein, such as tumbling or rotational motions, that affect NMR spectra. Thus large proteins are easy to study. In Figures 2 and 3 (see below), the Raman spectrum of a single phosphate group inside an  $\sim 60$ -kDa protein is shown as is the spectrum of NADH bound to an  $\sim 160$ -kDa protein. Certainly larger larger proteins may be studied as well.

In our work, we adapted a conventional Raman spectrometer to perform difference spectroscopy. The essential feature of the measurement is that the two parent spectra must be very accurate so that small differences between them show up without error in the difference spectrum. We first estimate the necessary accuracy of the spectrum, which is governed by essentially two parameters. One is the intensity of the parent spectra—both spectra must be large enough so that shot noise is small. The second is the spectral mismatch between the two spectra, because subtraction of two spectra that are not in register produces subtraction artifacts.

### *Anticipated Signal Size and Sources of Error in Raman Difference Spectroscopy*

A typical protein contains a few hundred amino acid residues and has a weight of about 20–80 kDa. In Raman difference spectroscopy, most of the signal arising from the protein is subtracted out, and the question is how much signal-to-noise is necessary to detect the scattered light arising from a single vibrational mode from the protein background. The signal-to-noise ratios must be very high, but within the capabilities of present equipment. For example, a 40-kDa protein has approximately 4000 atoms, or 12,000 normal modes. Should about half these yield a signal, then 6000 bands comprise the spectrum. If we assume for this discussion that this signal is spread evenly over about 2000  $\text{cm}^{-1}$  (not strictly valid, see below) and each band is 10  $\text{cm}^{-1}$  in width, then some 30 normal modes contribute to the signal at each position

of the observed spectrum. This estimate suggests that one must have a protein signal-to-noise ratio in excess of 300/1 to discern a particular band from the protein background (30 times stronger) with a signal-to-noise ratio of 10/1.

There are two sources of error in a protein spectrum that can adversely affect the signal-to-noise of the difference spectrum: so-called shot noise and systematic factors. Because shot noise arises from the statistical nature of the scattering process, it may be decreased with increased detected signal (signal-to-noise scales as the square root of the number of detected photons). Hence, much effort in recent technological developments of Raman spectroscopy has been devoted to more efficient collection optics and detectors. With biological samples in solution, to obtain shot noise of less than a percent of the protein Raman signal requires the use of optically fast detectors such as intensified diode arrays or charge-coupled devices, optical multichannel detector (OMA) systems. These instruments detect an entire portion of the spectrum simultaneously, and the spectrum of a protein in solution, at a concentration of a fraction of a millimolar, can be determined in about an hour with a signal-to-noise of better than 1000/1. Earlier work (32, 46) exploring Raman difference spectroscopy described Raman systems employing photomultiplier detectors. These spectrometer systems can yield an accurate difference spectrum, but take more than 100 times longer to do the measurement.

Systematic error factors comprise the most important experimental problem, and we have found the most critical determinant leading to an inaccurate subtraction of two spectra, say A and B, is that the signal from run A hits the OMA detector with a small spectral displacement with respect to the signal from run B. Consequently, the two spectra appear to be shifted in frequency, and identical signals present in the two runs do not subtract to zero but yield a difference spectrum resembling a derivative. The apparent frequency difference can be very small and yet yield a substantial difference intensity (e.g. 46, 61). The magnitude of the systematic error can be reduced to an effective shift in frequency between the two samples. For example, the maximum in the derivative signal, which we will call  $\Delta I_{\max}$ , can be calculated if the shape of the band is known. Assuming a Gaussian band shape with peak intensity of  $I(0)$  and a full width at half maximum  $\Gamma$ , for small misalignments (61):

$$\Delta I_{\max}/I(0) = 1.4(\delta/\Gamma). \quad 1.$$

where  $\delta$  is the apparent frequency shift due to effective spectrometer misalignment. A Lorentzian line shape will change the numerical factor

to 1.9. This equation shows that the apparent difference signal becomes significant whenever  $\delta$  becomes large or  $\Gamma$  becomes small. Thus, how well the protein background subtracts out depends on the bandwidth of the protein band. A protein amide-I band is typically  $50\text{ cm}^{-1}$  broad so that the apparent frequency shift,  $\delta$ , must be kept to less than  $0.1\text{ cm}^{-1}$  if we wish to maintain a  $\Delta I_{\max}/I(0)$  of 0.3% or better. Most proteins also contain sharp bands, such as the phenylalanine mode at  $1004\text{ cm}^{-1}$ , which has a bandwidth of only  $10\text{ cm}^{-1}$ . In this case,  $\delta$  must be kept to less than  $0.02\text{ cm}^{-1}$  to accurately subtract out this band. These are quite tight requirements for currently available spectrometer systems, but they are feasible, as shown below. That sharp bands are more difficult to subtract out than broad bands has been illustrated (61).

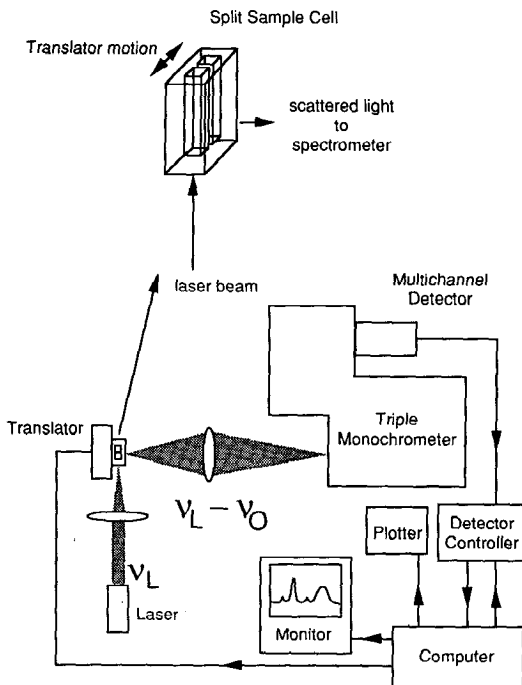
### *The Difference Raman Spectrometer Used in Our Studies*

Our Raman spectra are measured using a conventional Raman spectrometer employing an optical multichannel analyzer (OMA) system as depicted in Figure 1 (10, 61). Typically, 100 mW of visible light from either an argon ion or krypton ion laser is used to excite Raman scattering, and a typical experiment takes one to two hours. We can detect about  $1200\text{ cm}^{-1}$  of the Raman spectrum simultaneously. The parent spectra that make up the difference spectrum are obtained using a specially made split-cell cuvette, which is mounted on a translator stage-stepping motor combination. Each side of the split cell has an inside dimension of  $3 \times 2.5\text{ mm}$  and is loaded with  $30\text{ }\mu\text{l}$  of sample in a typical experiment. Sample concentrations are about 1 mM; thus, approximately 2.4 mg of a 40-kDa protein are presently required.

We believe that advances in spectrometer technology, in terms of light collection and detector efficiency as well as sample cell configuration and involving no new principles, can be developed to reduce the amount of protein needed by as much as 100-fold. In fact, Peticolas and his coworkers have employed a Raman microscope system that requires only  $2.5\text{ }\mu\text{l}$  of sample volume (42), and Carey and coworkers have demonstrated the measurement feasibility at a sample concentration of  $100\text{ }\mu\text{M}$  (34).

Several precautions must be taken to minimize systematic errors. The entire spectrometer system, including the exciting laser, is mounted on a vibration-free table. Ambient room temperature is controlled to within  $\pm 1.5^\circ\text{C}$  and relative humidity to within  $\pm 3\%$ . Sample positioning is especially critical.

The protein solution is loaded into one side of the split cell cuvette,



*Figure 1* A schematic of a Raman difference instrument. Laser light is incident on a specially fabricated split cell from the bottom. The split cell has been designed to make the light paths in both halves as equal as possible in order to reduce subtraction artifacts. The Raman scattered light is collected by a monochromator and multichannel detector combination. Scattered light is collected from one side of the split cell for 10–60 min, and this spectrum is stored in the computer. Then the cell is translated and scattered light is collected from the second side. The two spectra are then subtracted in the computer (see text) to form the difference spectrum.

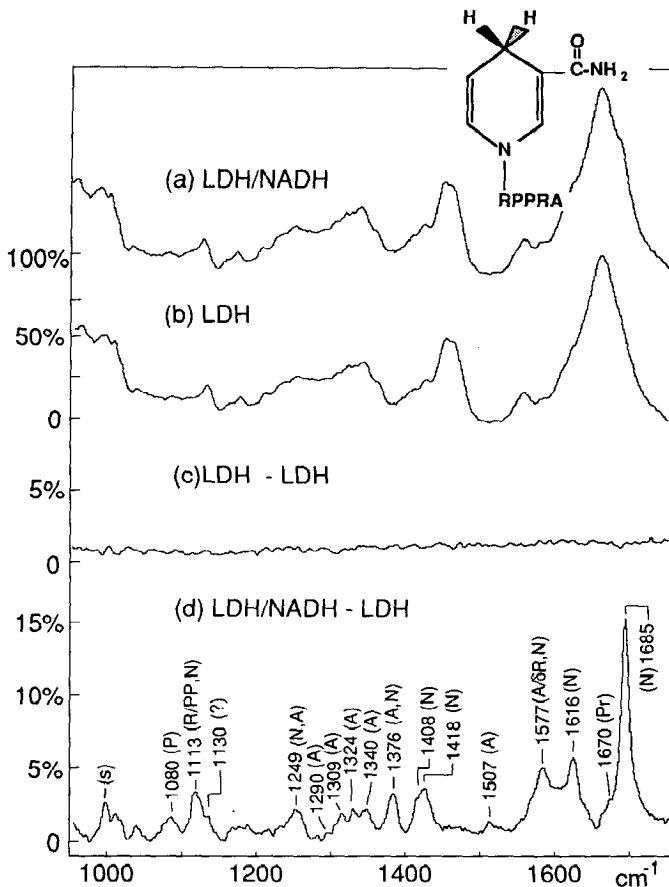
e.g. the A side, while the altered protein in solution is loaded into the other. After measuring the Raman spectra of the two samples sequentially, one then calculates the Raman difference spectrum of the bound substrate from the two spectra stored in the laboratory computer. Matching the signal levels arising from A and B is very difficult. For example, small mismatches in alignment and/or sample concentration on the order of 1% are nearly always present. Therefore, the difference spectrum that nulls out the protein background signal is  $A - xB$ , where  $x$  is a numerical factor close but not identical to one. The parameter  $x$  is determined by trial subtractions that vary this parameter until the background peaks of the protein, which are generally relatively broad

compared to those of the ligand, subtract out. Chen et al (14) discuss this procedure in more detail.

If we define subtraction fidelity as the ratio of the maximum false signal in the difference spectrum divided by the maximum signal in the parent spectra [ $\Delta I_{\max}/I(0)$ ], the instrument depicted in Figure 1 has a fidelity of about 0.1%. Essentially, systematic factors determine this value, and it is measured by loading up both sides of the split cuvette in a test run with identical samples. Another procedure for assessing subtraction fidelity during a run involves taking interleaved spectra in an ABAB . . . AB sequence, adding every other A spectrum and subtracting this from the sum of the remaining A spectra. In such a subtraction, the resulting A - A or B - B difference should result only in simple shot noise.

**AN EXAMPLE** One can perform many studies on bound ligands by measuring the Raman spectrum of the protein and the protein-ligand complex and subtracting these two. This type of difference spectrum provides the spectrum of the bound ligand as well as any signals that result from protein changes that are induced by the binding of ligand. While making solutions of the protein-ligand complex, one should determine whether or not the solution contains any free ligand. If so, the difference spectrum will also contain peaks caused by the free ligand. In all the work described here, the experimental conditions are such that the concentrations of free ligand were low enough to yield negligible signals.

Figure 2 shows early results (23). Spectrum *a* shows the Raman spectrum of lactate dehydrogenase (LDH) with bound NADH (i.e. reduced nicotinamide adenine dinucleotide). The nicotinamide group appears in Figure 2*a* for reference. Figure 2*b* shows the Raman spectrum of LDH itself. LDH is a tetramer with a molecular weight of about 160 kDa, containing four independently acting units. The strong broad protein band near  $1660\text{ cm}^{-1}$  is the amide I band and arises from polypeptide backbone amide C=O stretching motions. The amide III region lies between  $1230$  and  $1300\text{ cm}^{-1}$  and results from polypeptide backbone N-H in-plane bending and C-N stretching motions. The bands at  $1340$  and  $1450\text{ cm}^{-1}$  are the  $\gamma\text{-CH}_2$  and  $\delta\text{-CH}_2$  bands, respectively. The apparent difference between the LDH protein spectrum and the LDH/NADH binary complex spectrum is very small. Visually observing any difference in plots of the parent spectra is often impossible, but this example was chosen partly because a weak shoulder is barely observed near  $1685\text{ cm}^{-1}$  in Figure 2*a* that is absent in Figure 2*b*. Figure 2*d* shows the difference spectrum between the LDH/NADH binary com-

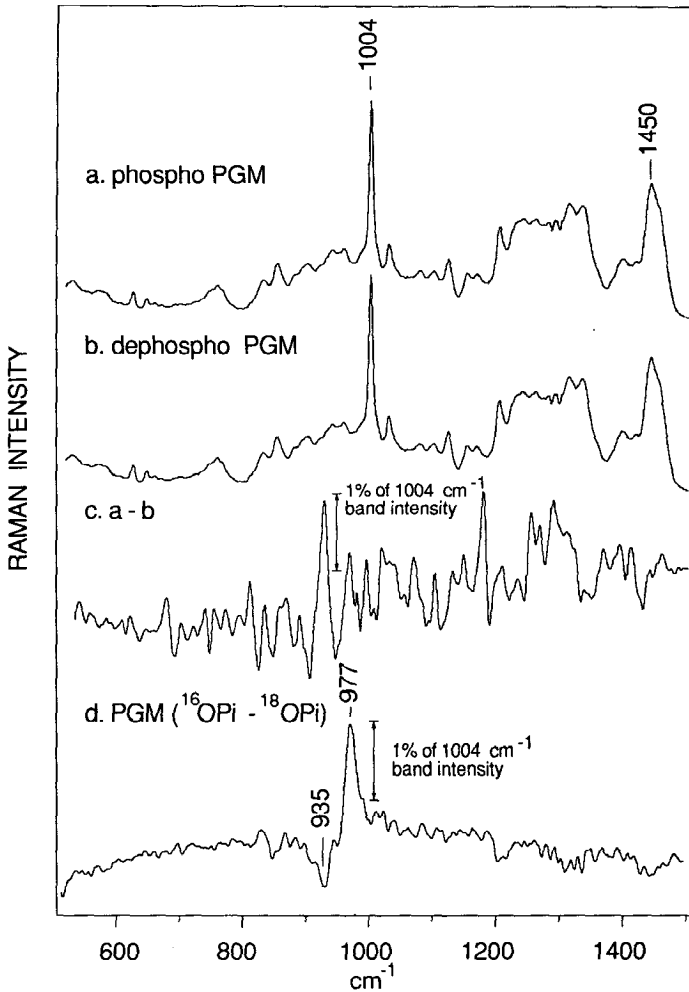


**Figure 2** Raman spectra of (a) LDH/NADH binary complex and (b) LDH, at 4°C in 0.1 M phosphate buffer, pH 7.2. (c) Results of subtracting the spectrum from b of an otherwise identical run as in b, except that the second position of the split cell shown in Figure 1 containing LDH was moved into the laser interaction area. (d) Results obtained by calculating the difference between a and b. The scale in c or d is five times that in a or b, as shown on the y axis. Data are taken with the spectrometer system shown in Figure 1 using 100 mW of 488-nm irradiation from an argon ion laser and a resolution of 8 cm<sup>-1</sup>. The molecular structure of the nicotinamide group of NADH is given in spectrum a for reference. Assignments of the peaks in d (23): A, adenine; N, nicotinamide; P, phosphate; Pr, protein; PP, pyrophosphate; R(δR), ribose; S, solvent; ?, unknown.

plex spectrum and that of LDH. The number of bands now observed in the difference spectrum, with very good signal-to-noise, is startling. The Raman bands in the difference spectrum are mostly from bound NADH, although some protein bands show up owing to the effect of NADH on the protein upon binding (see Figure 2 caption for assignment). Spectrum *c* shows the result of the control measurement of subtracting the spectra obtained from the A and B side of the cuvette, where both sides contain LDH. This difference spectrum is dominated by shot noise, which is about 0.5% of the protein's strongest (amide-I) peak in this example.

**ISOTOPE EDITING** In general, the protein structure and conformation should change upon ligand binding. These alterations will be evident in the Raman difference spectrum, and signals from the protein should be present in a protein ligand–minus–protein difference spectrum. Figure 2*d* shows protein signals, but their size is generally smaller than those of the bound ligand. As a (very rough) rule of thumb, we have found that the Raman spectrum of a protein minus that of a perturbed protein (perturbed by bound ligand or mutated residue) is about 1% of that of the protein itself. This observation is consistent with the difference FTIR spectroscopic studies of proteins (e.g. 6, 55). Thus, in many cases, the signal from bound ligand will be similar in size, or smaller, than that arising from changes in protein structure. Moreover, some proteins are not stable without bound ligand. The strategy of isotope editing (40, 58) of the difference spectrum obviates these problems; here, isotopically labeled ligands are prepared, so that the frequency of those modes that involve motions of the isotope are shifted. Thus, a subtraction of the labeled ligand from the unlabeled one yields a difference spectrum of positive and negative peaks of just those modes, and all others cancel. In an experiment using labeled protein–ligand minus unlabeled protein–ligand, the protein bands subtract out.

In a recent study (20), the protein bands observed in a protein–ligand–minus–protein difference spectrum were completely dominant, and isotope editing was employed to remove these features. Figure 3 shows the Raman spectrum of phospho (*a*) and dephospho-phosphoglucosmutase (*b*) in the P=O stretch region. Phosphoglucosmutase (PGM) catalyzes the interconversion of glucose 1-phosphate and glucose 6-phosphate. The active form of the enzyme contains a phosphorylated serine residue, and this phosphate group can be removed. The phosphate stretch bands for phosphate bound to PGM, and for bound glucose 1-phosphate and glucose 6-phosphate, were measured to study the catalytic mechanism of PGM, as discussed below. The difference spec-



*Figure 3* Raman spectra of phosphoglucosmutase at pH = 7.4 at a temperature of 4°C. The approximate time for data accumulation for the protein spectra was 2 h; a 514.5-nm laser was used. (a) The spectrum of the phospho enzyme; (b) the spectrum of the dephospho form to the enzyme; (c) the difference spectrum between *a* and *b*; (d) the difference between two phospho forms of the enzyme, one containing <sup>16</sup>O-labeled phosphate and the other containing <sup>18</sup>O-labeled phosphate.

trum between phospho and dephospho enzyme spectra is surprisingly complicated (Figure 3c; discussed further below), and the phosphate stretch band cannot be assigned. The  $\text{P}\text{-}^{16}\text{O}$  stretch shifts downward  $35\text{ cm}^{-1}$  for  $\text{P}\text{-}^{18}\text{O}$ , so  $^{18}\text{O}$ -labeled phosphates were used to produce an edited spectrum. Figure 3d shows the difference Raman spectrum of PGM with  $^{16}\text{O}$ -labeled minus  $^{18}\text{O}$ -labeled phosphate, and all the protein peaks that are masking the phosphate stretch band in spectrum c now subtract out. The positive peak in spectrum d represents the  $\text{P}\text{-}^{16}\text{O}$  stretch, and the negative peak the  $\text{P}\text{-}^{18}\text{O}$  stretch. The background noise is simple shot noise (at a level of about 0.1% of the main bands of PGM). This experiment demonstrates the feasibility of measuring small molecules or small molecular fragments bound to an  $\sim 60\text{-kDa}$  protein with a signal-to-noise ratio of about 10/1.

**A COMPUTATIONAL METHOD OF REDUCING SYSTEMATIC ERRORS** As discussed above, a slight misalignment of the two parent spectra on the detector gives rise to artifacts in the difference spectrum that have a derivative-like shape. In particular, this difference shows up when the two spectra are taken at different times, which makes the pooling of spectra, in order to reduce shot noise, difficult. One can eliminate or reduce these artifacts by shifting the frequency origin of one data set by a small amount before subtracting the two data sets (this may be done after the experiment because all data are stored in the computer). The amount of the shift can be calculated from the intensity and width of a band in the primary spectra and the unshifted difference spectrum that can be reasonably assumed to be, or is known to be, one that should subtract out in the difference spectrum, but does not.

An example of such a band is the prominent, but quite uninteresting, sharp protein band at  $1004\text{ cm}^{-1}$  in the spectrum of PGM (Figure 3a), which arises from a phenylalanine ring mode. Its bandwidth in the primary spectra is measured, as is the magnitude of the peak-valley distance in the unshifted difference spectrum. These parameters are used in Equation 1 to calculate the shift required to null out this signal. This procedure was used recently to null out remnants of the phenylalanine  $1004\text{ cm}^{-1}$  protein peak in some of the [ $^{16}\text{O}$ - $^{18}\text{O}$ ] PGM-phosphate difference studies (Figure 3d) (20). The shifts were typically less than  $0.1\text{ cm}^{-1}$  (the data in Figure 3d is from unshifted parent spectra). Neither the positions nor the widths of the  $^{16}\text{O}$  and  $^{18}\text{O}$  phosphate stretches in the difference spectra were affected by the procedure, and additional artifacts in the difference spectrum from other protein bands in the parent spectra were also cleanly nulled. A complete description of this is to be published (J Burgner & R Callender, unpublished study).

COMPARISON OF DIFFERENCE RAMAN TO RESONANCE RAMAN SPECTROSCOPY Because of its sensitivity and selectivity, resonance Raman spectroscopy is often the first choice in the study of ligand-protein or substrate-enzyme interactions by vibrational spectroscopy when the absorption maximum of the ligand or substrate is at a visible or UV wavelength that is separate from that of the protein (54). However, this method has two problems that can limit its usefulness. One is when the bound ligand is photolabile. Raman difference spectroscopy can be advantageous in this case because it can avoid unwanted photo-reactions associated with the resonance Raman technique. For example, studies by Carey and coworkers (54, 54a, 55) on the acyl carbonyl group in [3-(5-methyl-2-thienyl)acryloyl]chymotrypsin compare the results of the two techniques and also the results of FTIR. A similar avoidance of light-induced photoreactions in resonance Raman studies was accomplished in Raman difference studies of aspartate aminotransferase (19).

A second problem with resonance Raman spectroscopy is the occasional presence of large masking fluorescence backgrounds. Fluorescence backgrounds are not present, in principle, when the frequency of the laser excitation is not in resonance, as occurs in Raman difference spectroscopy. A recent study illustrates this by comparing resonance Raman versus nonresonance Raman spectroscopies (33).

## STRUCTURES AND INTERACTIONS OF LIGANDS WITH PROTEINS PROBED BY RAMAN DIFFERENCE SPECTROSCOPY

How ligands bind to proteins is of central importance to biology. Upon ligand binding, any number of structural changes may take place. Both the ligand and/or the protein may change their structure; hydrogen bonds and other electrostatic interactions may form between the ligand and the protein at the binding site; and possible covalent bonds may form. The now standard crystallographic and NMR (reviewed in 11) approaches to studying this problem provide atomic-resolution pictures placing the positions of atoms within the complex. However, the resolution of these pictures, at about 1–3 Å, while possibly sufficient to place a particular atom to within 0.1–0.3 Å, is not sufficient to quantitate small but important changes in key bonds or the size of weak forces such as hydrogen bonding, nor is it sufficient to probe whether or not many small changes take place that, in the sum, are important. In this context, vibrational spectroscopy is useful in that it provides very-high-resolution pictures of the structure of groups of atoms and

bonds and of how these structures respond to noncovalent interactions. The following discussion does not review particular protein systems but rather groups the discussion into categories of structural information relevant to the binding of ligands to proteins, including the formation of covalent bonds, hydrogen bonding, and the effects of the ligand binding on the protein. In principle, one could also determine the molecular geometry of the bound ligand from vibrational spectroscopy. This is well documented (12, 48, 49), and we do not discuss it further here.

### *Covalent Bond Formation*

A new covalent bond may form when a ligand binds to a protein. This new bond, characterized by a specific normal mode involving motions of the newly bonded atoms, will likely produce a new peak in the vibrational spectrum of the complex. The appearance of such a peak, when properly assigned, demonstrates the formation of the new bond.

For example, a Schiff base ( $-C=N-$ ) linkage forms between the  $\epsilon$ -amino group of active-site Lys258 of *E. coli* aspartate aminotransferase (AATase) and the carbonyl moiety of enzyme bound cofactor, pyridoxal 5'-phosphate (PLP). Recent Raman difference studies (28) examined the spectrum of the high and low pH forms of this complex, and Raman bands characteristic of Schiff base and protonated Schiff base were determined, respectively. A Y225F mutant form of the enzyme, in which the near-UV absorption of the bound PLP at high pH resembles that of an aldehyde rather than a Schiff base, was also studied. Evidence from isotope-edited Raman difference spectra proved conclusively that the near-UV spectrum is anomalous and that PLP is bound to Y225F at high pH as a Schiff base. A strong peak at  $1630\text{ cm}^{-1}$  was observed for PLP bound to Y225F, which is characteristic of a Schiff base bond and not an aldehyde. Moreover, this band redshifted by  $18\text{ cm}^{-1}$  in enzyme labeled with  $^{15}\text{N}$  at Lys258 (28).

An inhibitor,  $\alpha$ -methyl-L-aspartate ( $\alpha\text{MeAsp}$ ), complexed with enzyme-PLP was also studied because it is believed to be a Michaelis complex of the enzyme. Four forms of the complex are possible because the inhibitor may, in principle, form a Schiff base with Lys258 (so-called internal Schiff base) or with PLP (the external Schiff base) and may be protonated or unprotonated. The Raman spectrum of  $\alpha\text{MeAsp}$  bound to AATase complex contains a  $1630\text{ cm}^{-1}$  band, which was assigned to the unprotonated internal Schiff base C=N stretch mode based on a downshift of  $18\text{ cm}^{-1}$  with  $^{15}\text{N}$  labeling of the Lys258  $\epsilon$  nitrogen. This study also showed that the protonated external Schiff base form coexisted in this complex, a conclusion supported by the

insensitivity of the observed  $\nu_{\text{C}=\text{NH}^+}$  of the  $\alpha\text{MeAsp}$ -enzyme complex at  $1655\text{ cm}^{-1}$  to lysine  $^{15}\text{N}$  substitution. No significant amount of other two possible Schiff base forms, namely the protonated internal or unprotonated external forms, were found. Interestingly, the Raman study, which of course was performed on protein solutions, showed that the solution structure of the *E. coli* AATase- $\alpha\text{MeAsp}$  complex differs significantly from the reported crystal structures in terms of the equilibrium concentrations found amongst the four forms (18).

### *Hydrogen Bonding (and Weak Electrostatic Interactions)*

The energies of hydrogen bonds and weak electrostatic interactions are on the order of 2–10 kcal/mol or more and play a major role in the structure of proteins and nucleic acids and their interactions with ligands (cf 30). H-bonding is generally considered a major determinant in the specificity of interactions between a protein and ligand, and these interactions are important to enzymatic catalysis (see below), although they have not been well characterized in proteins.

**H-BOND ENERGIES FROM VIBRATIONAL FREQUENCIES** Clearly, hydrogen bonding affects the electronic distributions within bonds of localized modes (and hence force constants). Thus, the frequencies of particular modes may be sensitive to the strengths of H-bonding. Badger & Bauer (2) suggested quite some time ago that the enthalpy of formation of a hydrogen bond,  $\Delta H$ , is linearly related to the vibrational frequency shift,  $\Delta\nu$ , of the O-H stretch frequency of an alcohol. A stronger hydrogen bond weakens the O-H bond and the stretch force constant; this weakening significantly decreases the observed stretching frequency. Numerous investigations of such thermodynamic correlations have shown that the relationship is often very direct and simple (cf 31). The general methodology of this body of work is to determine the association constants, enthalpies, and binding entropies of donor-acceptor pairs in organic solvents. Badger-Bauer relationships have been found for O-H, C=O, N-H, and other groups (reviewed in 17, 31, 56). The stated relationships are generally between  $\Delta H$  and  $\Delta\nu$ , but  $\Delta G$  versus  $\Delta\nu$  relationships have also been derived (cf 31). The exact relationship depends on the nature of the group and nearby chemical constituents. Quantitative use of such relationships for a specific ligand then generally involves some sort of calibrating set of measurements. In general, the empirically derived relationships only extend over a fairly small range in  $\Delta H$  (<10 kcal/mol). Theoretical development of this problem will greatly help define the accuracy and applicability of

such relationships over wider ranges and in conditions other than those found in solution.

For example, the relationship between carbonyl stretch and hydrogen bond energy was recently investigated with *ab initio* calculations of H<sub>2</sub>CO interacting with various cations and water (38). The shift in C=O stretching frequency in H<sub>2</sub>CO fits a linear correlation with the computed interaction energy,  $\Delta E$ , having a slope of 0.5 kcal/mol for each wavenumber (cm<sup>-1</sup>) of frequency shift—which agrees remarkably well with experiment (51). This study also addressed several important issues that the empirically observed correlations cannot. First, the linear relationship was found to hold up to an interaction energy of 27 kcal/mol, at which point the character of the donor-acceptor pair becomes covalent in nature. The empirical solution studies are limited to 5–10 kcal/mol in  $\Delta E$  because model systems generally do not involve large hydrogen bond energies. Another important finding is that the linear correlation was found to be essentially independent of the nature of the cation and the distance between cation and the carbonyl oxygen (until the two were so close that the bond became more nearly covalent in character). The  $\Delta E$  versus  $\Delta\nu$  relationship differed among calculations in which the cation angle of approach to the carbonyl oxygen was varied; consequently, complete analyses of the experimental data will have to take this variation into account.

**EXAMPLES OF H-BOND PROTEIN-LIGAND INTERACTIONS** Raman difference studies on two systems provide a basis of examining how well Raman difference spectroscopy tells us about hydrogen bonding of small molecules to proteins. The first system is the binding of substrates and coenzymes to the NAD dehydrogenases; the second is the binding of guanine nucleotides to a selected set of so-called G-proteins.

The NAD-linked dehydrogenases form a large class of oxidation-reduction enzymes needed for a variety of metabolic functions. Some of these have been studied with Raman difference spectroscopy: liver alcohol dehydrogenase (LADH), lactate dehydrogenase (LDH), malate dehydrogenase (cytoplasmic sMDH and mitochondrial mMDH), and dihydrofolate reductase (DHFR) (9, 10, 14, 15, 21–23, 62–64). These enzymes catalyze the reversible direct stereochemical transfer of a hydride ion, H<sup>-</sup>, from either the pro-R or pro-S face of NADH (but not both) to the carbonyl of aldehyde or ketone carbon of substrate to form alcohol and NAD<sup>+</sup> (see, e.g. 26). The hydride transfer proceeds with a high degree of stereochemical fidelity. For example, the hydride transfer occurs from NADH's pro-R face with an error rate of <1/10<sup>8</sup> in LDH (1).

At the NADH protein-binding site (see Figure 2), hydrogen bonds are formed with the carboxamide [ $-(C=O)NH_2$ ] moiety of the nicotinamide headgroup. In Raman studies, this group is characterized by observation of the C=O stretch and the  $-NH_2$  rock mode. The  $-NH_2$  rock mode is assigned to the  $1113\text{ cm}^{-1}$  peak in Figure 2*d* for NADH bound to LDH. This band has shifted upward  $30\text{ cm}^{-1}$  from its in-solution position at  $1083\text{ cm}^{-1}$  (14, 23, 62). In the same way, the C=O stretch of keto substrates bound to LDH and LADH shift upon binding (9, 22). Table 1 summarizes the observed frequency shifts, from which the interaction enthalpies are calculated from the Badger-Bauer relationships described above for the C=O moieties. No accurate relationship has been derived for the  $-NH_2$  groups. The reported hydrogen bond enthalpies are with respect to water.

It is sometimes useful to know the value of the hydrogen bond with respect to a nonhydrogen environment. A keto group forms a hydrogen bond with an interaction enthalpy estimated at about 4.5 kcal/mol from spectroscopic studies employing Badger-Bauer relationships (23). Thus, one must add approximately this amount to the values in Table

**Table 1** Frequency shifts of certain vibrational bands and deduced binding energies of various coordinates<sup>a</sup>

Group	Wavenumbers ( $\text{cm}^{-1}$ )		Reference
	$\Delta\nu$ (in situ, solution)	$-\Delta H$ (kcal/mol)	
C=O stretch of bound substrates			
DABA in LADH	-94	-14	9
Pyruvate adduct in LDH	-35	-14	22
C=O stretch of NAD's carboxamide group			
LDH	-10	-2.8	15, 21, 23
sMDH	-9	-2.5	15
mMDH	-9	-2.5	15
DHFR	+5	+1.4	64
$-NH_2$ rock of NAD's carboxamide group			
LDH	+30	$\sim -3$	15, 21, 23
LADH	+30	$\sim -3$	14
sMDH	+35	$\sim -3$	15
mMDH	-4	0	15
DHFR	+35	$\sim -3$	64
$-COO^-$ stretch of pyruvate's carboxylate			
LDH	-5		22

<sup>a</sup> Data are for the coenzymes NADH and  $NAD^+$  and substrates of the NAD dehydrogenases, binding to the specified proteins as derived from Badger-Bauer relationships. The derived binding  $\Delta H$  is expressed with respect to water. See text and the listed references for details.

1 to determine the absolute value of the hydrogen bond to the C=O groups. The energies obtained from this analysis of the vibrational data are very useful in understanding some important properties of these complexes.

For example, the stereospecific nature of the hydride transfer step may be a result of relative positioning of the substrate and the coenzyme. For enzymes that are pro-S, only the pro-S hydrogen of NADH is positioned close to the carbonyl carbon of substrate, and the opposite is true for pro-R enzymes (about half of the 300 or so known dehydrogenases are pro-R and half are pro-S). The positioning apparently results from enzymic control of the conformation of the glycosidic bond between the nicotinamide ring and ribose group of NADH. Structural studies show that generally pro-R enzymes bind the nicotinamide ring in the *anti* conformation about the glycosidic bond, while pro-S binds the ring in the *syn* conformation. Assuming this is the case, the energy barrier to rotation about this bond, which is normally very small in solution, must adopt large values at the binding site. For LDH, a stereochemical transfer error rate of  $<1/10^8$  suggests the barrier is greater or equal to 10.4 kcal/mol (1). The essential source of the energy barrier presumably arises from hydrogen-bond breaking between the carboxamide moiety and protein at the binding site, caused by rotation of the glycosidic bond from the normal *anti* conformation that NAD adopts when bound to LDH (15, 23). Assuming that enthalpic terms are entirely responsible for the energy difference between the *anti* and *syn* conformations, the value of 10.4 kcal/mol results from the addition of 3 kcal/mol for the hydrogen bond to  $-\text{NH}_2$  and 7.3 kcal/mol for the hydrogen bond to C=O of the carboxamide moiety (equal to the 2.8 kcal reported in Table 1, which is with respect to water, plus the 4.5 kcal/mol for breaking the water hydrogen bond to the keto group) (Table 1). Alternatively, LeReau & Anderson estimated a minimum energy difference of 8.2 kcal/mol in experiments on PAAD<sup>+</sup> in which the  $-\text{NH}_2$  was replaced with a hydrogen to test the importance of the  $-\text{NH}_2$  group of NAD (37). This value is satisfyingly close to the 7.3 kcal/mol difference in hydrogen bond energy estimated for the hydrogen bond to the carbonyl group alone.

One can use frequency shifts of appropriate spectroscopic marker bands and empirical relationships to relate the markers to interaction energies in order to probe the magnitude of various components of the binding enthalpy between protein and ligand. Such studies are quite important and can be applied to, for example, the rational design of drugs. One experimental study examined the binding of pyruvate to the LDH-NADH binary complex using the numbers given in Table 1

(24). The net total binding enthalpy is the result of numerous factors: the shift of pyruvate from water to enzyme and of structural water at the active site to solution, differences in coenzyme interactions with protein induced by substrate binding, and changes in the internal interactions within the enzyme produced by subtle or not-so-subtle protein conformational changes upon complex formation. The contribution to the net enthalpy of interaction by pyruvate should arise primarily from differences in interactions involving its carboxylate and carbonyl moieties. The value of about  $-17$  kcal/mol for the overall binding enthalpy for pyruvate to LDH/NADH and the value of  $-14$  kcal/mol derived for its carbonyl moiety (Table 1) is remarkably close. Thus, even though ligand binding may cause protein and other conformational changes, the net binding enthalpy of these changes seems to reflect the sum of several large, mostly compensating effects, which is consistent with recent theoretical studies (27). The hydrogen bond strengths reported in the Raman data for the C=O groups of the substrates bound to LADH and LDH are unusually large for such bonds in proteins (cf 30) and are quite provocative. Given the assumptions inherent with the use of the empirical relationships, the value of 14 kcal/mol for the C=O of pyruvate interacting with LDH has an error of perhaps  $\pm 3$  kcal/mol (24). Still the lower end of this range represents a hydrogen bond rarely reported inside proteins.

Another study along these lines looks at how the 6-keto and the 2-amino groups, the polar binding handles, of the purine ring of GDP interact with the EF-Tu and p21 proteins, which belong to the G-protein family (57). This protein family is involved in a variety of key cellular processes such as signal transduction, protein transport, and protein biosynthesis, and their activity is regulated by whether they bind GDP (inactive) or GTP (active). Because no Badger-Bauer-like calibrating empirical studies for the C=O stretch and the  $-\text{NH}_2$  deformation mode of GDP have been done, a quantitative relationship between frequency shift and  $\Delta H$  is not known for these groups. However, one could glean some insight into the strengths of the interactions from the shifts themselves, assuming that at least a rough correlation does exist—a reasonable approximation in view of all the earlier studies.

The amine deformation mode of GDP, observed at  $1646\text{ cm}^{-1}$  in solution, is shifted up by 8 and  $16\text{ cm}^{-1}$  upon nucleotide binding to EF-Tu and p21, respectively, and the ketone stretch shifts down 15 and  $10\text{ cm}^{-1}$  for GDP in EF-Tu and p21, respectively. The directions of the shifts indicate that the interaction between the keto and the amine groups are stronger for the two groups with either EF-Tu or p21 than with water; this is certainly reasonable because both proteins bind GDP

very strongly ( $K_{ds}$  tighter than  $10^{-9}$  M) and because the structure of the binding pockets is such that appropriate proton donors and acceptors are present to bind to the polar groups.

The hydrogen-bond strengths do not exhibit a consistent pattern in the binding of the carboxamide moiety of NAD to the NAD-dehydrogenases that have been studied (summarized in Table 1), and in the binding of GDP to the two G-proteins, despite the fact that the binding pockets for NADH among the dehydrogenases and for GDP by the G-proteins are much the same. For the NAD proteins, the H-bonds generally arise from a protein backbone C=O group that pairs with the  $-NH_2$  moiety of NAD and a structural water that pairs with the C=O group. These observations indicate that the C=O group should bind more strongly to the proteins than to water, which is generally the case (15, 21). However, the C=O group of NAD actually binds more weakly to DHFR and inhibitor than to water (Table 1) (64). Similarly, interactions with the 6-keto and 2-amino positions are conserved in EF-Tu and p21 as well as among other members of the GTPase super family, as part of a NKXD motif. The 2-amino group of GDP is known to be only 3 Å or less away from an invariant Asp residue (Asp138 in EF-Tu and Asp119 in p21), while the 6-keto group is hydrogen bonded to the main-chain NH of a conserved alanine residue (Ala174 in EF-Tu and Ala146 in p21). However, the relative strengths of the hydrogen bonds formed with these groups and the two proteins are very different. If we assume that at least an approximate Badger-Bauer relationship holds between interaction energy and frequency shift of the marker bands, the hydrogen bonds can differ by up to a factor of two.

### *Ligand-Induced Changes in Protein Structure*

How is the structure of a protein affected by the binding of a ligand? The standard method of studying this problem involves crystallographic or multidimensional NMR studies comparing the protein and the protein-ligand complex. The data shown in Figure 3c suggests that vibrational spectroscopy can also be helpful for this problem because of its high resolution. Figure 3c shows the difference spectrum of phospho-PGM minus dephospho-PGM, i.e. of two protein spectra from an ~60-kDa protein that differ only by the presence or absence of a single phosphate group (see above). The difference spectrum is surprisingly robust, but it has not yet been interpreted.

In general, an analysis would proceed as follows. The amide-I ( $\sim 1660$ – $1690$   $cm^{-1}$ ) and amide-III ( $1230$ – $1280$   $cm^{-1}$ ) regions of the protein spectrum, which involve molecular motions of the polypeptide linkage, are sensitive to secondary structure. In addition, the char-

acteristic signatures of the various residues in the spectrum provide useful structural information: the ionization states of histidines, carboxyls, etc; whether or not sulfide groups have formed disulfide linkages and, if not, the degree to which an S-H group is hydrogen bonded; the nature of how tyrosines have interacted with their environment; and so forth (see e.g. 12, 48). That the difference spectrum contains substantial structure essentially everywhere suggests, qualitatively and quite speculatively, that the presence of the single phosphate group has caused subtle changes throughout the protein. The size of the difference signal suggests that perhaps scores of residues find themselves in an altered environment. Crystallographic studies of PGM show minor changes between the two forms (44). However, the phosphate group forms an important bridging group between domains one and two of this four-domain protein. Evidence indicates that dephospho-PGM in the crystal state is stabilized by crystal-packing forces that are absent for PGM in solution (W Ray, private communication).

## ENZYMATIC CATALYSIS

Enzymes perform two functions. One is to selectively bind the substrates that they catalyze. This feature is partly responsible for their ability to catalyze a particular chemical reaction (or a small set of reactions). Such chemical specificity is determined in part by the particular binding interactions between enzyme and substrate, and we presented examples above of how vibrational spectroscopy can be of help in studying these forces. Enzymes also lower the reaction barrier. This often involves bond distortion affecting the vibrational frequencies of the key internal modes that make up the reaction coordinates. For two enzymic problems, Raman difference spectroscopy has furthered our knowledge of enzymes' catalytic mechanisms. Before we discuss these examples, it is to be noted that resonance Raman spectroscopy has provided a wealth of information that has been invaluable in understanding some enzymes and for some internal coordinates; this body of work is very well reviewed (12, 49, 54). We emphasize that non-resonance Raman difference spectroscopy can extend this body of work because it is not tied to particular internal coordinates associated with chromophoric substrates. Rather, the entire set of coordinates that make up the reaction coordinate can be studied, one by one, as can any substrate. This feature of the technique should allow the study of most enzymic systems.

For example, the catalytic binding site of lactate dehydrogenase involves two domains. One binds the coenzyme, either NADH or NAD<sup>+</sup>

(oxidized nicotinamide adenine dinucleotide), and the other binds a substrate, either pyruvate or lactate. NADH provides hydride ion,  $H^-$ , which attacks the aldehyde moiety of pyruvate to form the secondary alcohol, lactate, and  $NAD^+$ . The imidazole ring of His195 interacts with substrate and performs several important functions. The N3 nitrogen donates a proton to the carbonyl of pyruvate and accepts a proton from the alcohol of lactate, thus acting as a general acid-base catalyst. Its interaction with substrate (discussed above)—certainly with the carbonyl group of pyruvate—helps position the substrate properly for its interaction with the C4 hydrogen of NADH. This interaction also polarizes pyruvate's  $C=O$  bond, promoting nucleophilic attack of  $H^-$ .

Figure 4 shows the reaction coordinate from pyruvate and NADH to the transition state and then to lactate and  $NAD^+$ . During reduction of pyruvate, the reaction scheme involves three important internal coordinates: the C4-H bond of the NADH coenzyme, the  $C=O$  bond of pyruvate, and the hydrogen bond between the  $C=O$  group and the imidazole ring. The properties of the C4-H bond have been studied by preparing NADH deuterated stereospecifically at the pro-R or the pro-S positions. This results in a localized C4-D stretch band that is unique in the difference spectrum. The frequency of the pro-R C4-H bond stretch of NADH does not change during formation of a binary LDH/NADH complex (25). This frequency is also not affected, apart from band narrowing, when oxamate, a substrate analogue, binds to form a ternary (LDH/NADH/oxamate) complex (16). This band narrowing was attributed to a stiffening of the binding site, which decreased the number of accessible states available to NADH and is essentially an entropic effect. Thus, the dominating ground-state enthalpic interaction in the reaction coordinate appears to be the hydrogen bond between

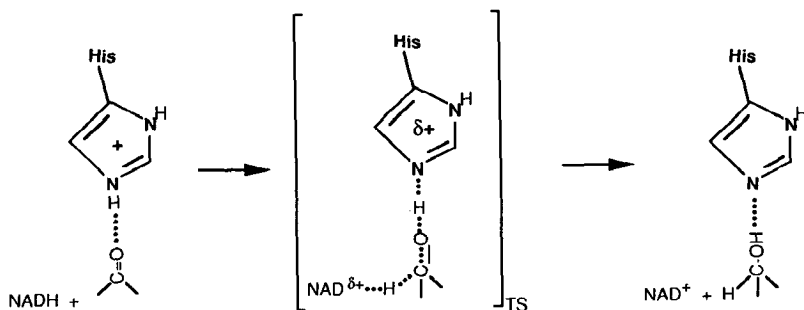


Figure 4 A simplified schematic of the reaction coordinate for the LDH-catalyzed reaction.

C=O and the imidazole ring, as suggested by the polarization of the C=O bond of pyruvate when it binds to LDH (22).

As discussed above, one can characterize the strength of the H-bond as well as the C=O bond order and bond length with the Badger-Bauer relationships, and other empirical relationships [as developed, for example, in some recent studies of acyl-proteases (53–55, 59)], by using the C=O stretch frequency as an internal probe. Keto (aldehyde) groups inside enzymes have received much attention in vibrational studies partly for technical reasons. For instance, the C=O stretch sometimes lies in an otherwise spectroscopically silent region (higher in frequency than the amide-I protein band near  $1660\text{ cm}^{-1}$ ). Thus, FTIR measurements are feasible and have shown substantial bond polarization of a substrate keto group in triosephosphate isomerase (4, 35), in yeast aldolase (5), in phospholipase  $A_2$  (47), and in citrate synthase (36). Recently, Wharton and his coworkers measured the C=O stretch of a C=O containing substrate bound to chymotrypsin and phospholipase  $A_2$  by using FTIR difference spectroscopy and employing the isotope-editing techniques discussed above to remove the protein IR absorbance (47, 55, 58, 59). The stretch frequency of this particular keto group lies within the amide-I protein peak. Also, Carey and his coworkers (cf 13, 54) and Peticolas and coworkers (cf 42) have performed extensive resonance Raman studies on acyl-chymotrypsins (and other proteases) made with chromophoric substrates.

In the case of lactate dehydrogenase, the C=O stretch of bound pyruvate lies within the strong amide-I protein band but could be measured employing isotope-edited Raman difference spectroscopy (22, 24). The observation of substantial bond polarization of the bound keto group implied that the total enthalpy of the essential reactant groups is stabilized by  $14 (\pm 3)$  kcal/mol from the C=O interaction with His195 (and other groups at the active site) relative to its interaction with water (Table 1). Measurements of mutant lactate dehydrogenases that affect the  $k_{\text{cat}}$  of this enzyme showed a definite correlation of rate of the hydride transfer step with the C=O stretch frequency (24). This correlation was rationalized by supposing that the enthalpy of the transition state is stabilized to a greater extent than the ground state so that the electrostatic interaction lowers the net reaction barrier. This explanation seems plausible because the carbonyl moiety is substantially more polarized in the transition state, with more negative character on the oxygen. Also, the proton located on the imidazole in the ground state is probably closer to the oxygen in the transition state, as in Figure 5.

A numerical analysis of the correlation suggested that about  $10^{5.5}$ -fold of lactate dehydrogenase's  $10^9$ -fold rate enhancement arises from the  $C=O \cdots His195^+$  interaction. Thus, taken all together, vibrational studies on this protein have suggested that the catalytic mechanism arises from entropic factors (bringing the reactants together and holding them in a configuration appropriate for reacting) and bond polarization of the reacting  $C=O$  coordinate. Note the high degree of precision in these results; the shift in frequency of  $35\text{ cm}^{-1}$  for the  $C=O$  stretch observed when pyruvate binds to lactate dehydrogenase represents a change in bond length of the  $C=O$  group of only  $\sim 0.02\text{ \AA}$  (53).

The enzymic mechanism of phosphoglucosyltransferase (PGM) was recently studied using Raman difference spectroscopy (20, 43). This enzyme catalyzes the interconversion of glucose 1-phosphate and glucose 6-phosphate so that the bond-breaking step is the hydrolysis of the phosphate ester bond. Bound vanadate, when substituted for phosphate, likely behaves as a transition-state analogue. Thus, studies comparing the properties of phosphate to vanadate when bound to PGM compare the effect of the enzyme on the ground and transition states, respectively, upon binding.

The phosphate  $P=O$  stretch frequency of bound phosphate groups was measured using isotope editing techniques and  $^{18}O$  labeled samples (Figure 3). The frequencies of the dianionic phosphate  $P=O$  stretch from either bound glucose 1-phosphate or glucose 6-phosphate did not differ from their values in solution, suggesting that a binding-induced distortion/polarization of the enzymic phosphate group in the ground state enzyme-substrate complex is not responsible for PGM's catalytic mechanism.

On the other hand, the frequency of the  $V=O$  bonds within a vanadate group bound at the same site in the transition analogue complex involving glucose 1-phosphate 6-vanadate is much lower than for a normal vanadate in solution (20). A series of model studies of vanadate in solution (43) suggested the formation of a weak fifth bond between the central vanadium and some internal fifth-ligand atom of PGM, very likely the  $O$  of Ser116. The addition of the fifth ligand draws electrons out of the three  $V=O$  bonds, and this lowers the frequency of the  $V=O$  stretch. Broadly speaking, the hydrolysis of a phosphate ester may proceed by either a dissociative ( $S_N1$ ) or an associative ( $S_N2$ ) mechanism. In the dissociative case, the ester bond breaks, which promotes nucleophilic attack. In the associative case, the phosphorus atom undergoes attack, which expels the ester group. The inference of a fifth bond to the phosphorus atom in the transition state supports an as-

sociative mechanism rather than a dissociative one (20). This finding contrasts with those of solution studies in which virtually all reactions involving hydrolysis of dianionic phosphate proceed via a dissociative mechanism.

## PROPERTIES OF INDIVIDUAL RESIDUES AND THE EFFECTS OF MUTATION ON PROTEIN STRUCTURE

The previous two sections have emphasized what we can learn about the structures of ligands bound to their protein-binding sites from vibrational spectroscopy. Conceptually, Raman difference spectroscopy could report on the vibrational spectrum of an individual amino acid residue within a protein. This may be accomplished by forming the difference spectrum between a protein and its analogue: where a specific residue (or particular group of residues) is isotopically labeled; where a residue is mutated; or where its pH is altered so that a residue or group of residues are in a different ionization state. One can determine the pKs of ionizable residues or determine the effects of environmental factors imposed by the protein on a particular residue (as in conformations of particular bonds, H-bonding properties, and so forth) as well as investigate how the mutation affects other residues.

In one study (62a), investigators titrated the various histidine residues of human transferrin to obtain the Raman difference spectra at different pHs (3.0–9.0) with respect to pH 8.9. About  $12 \pm 2$  of the 18 residues titrated over this range. A titration curve with a  $pK_a$  of  $6.08 \pm 0.01$  fit the data of histidine in solution, and  $6.56 \pm 0.02$  was the average value of the 12 histidine residues inside transferrin. The study indicated that the technique had enough sensitivity to monitor a single histidine residue in an 80-kDa molecule and to determine the titration curve of one residue in a 40-kDa protein. Another study determined the  $pK_a$  values for thiol-thiolate equilibria of thioredoxin (39).

In an unpublished study, we obtained the Raman difference spectrum between a wild-type protein and a mutant, in which a serine replaced a carboxyl group. We are performing these measurements to determine the  $pK_a$  of the carboxyl because these groups are often at the active sites of enzymes and their ionization state is difficult to determine otherwise. Also, we wish to investigate how mutation affects the structure of the protein. Figure 5a shows the difference spectrum between wild-type dihydrofolate reductase (DHFR) and its D27S mutant, in which serine replaces Asp27, in 6 M urea (Y Chen, J Kraut & R Callender, unpublished data). This experiment should yield only the dif-

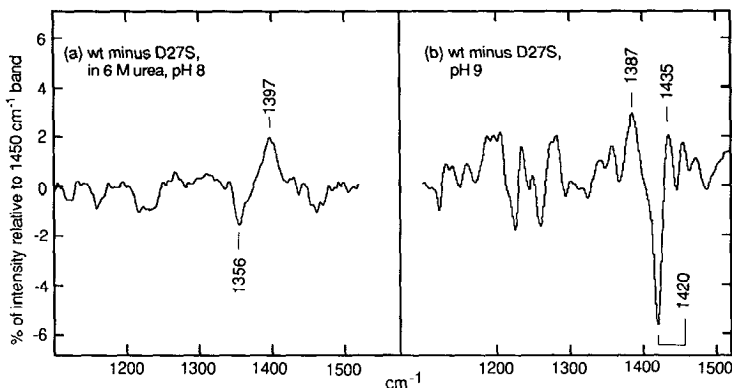


Figure 5 Difference spectra between wild-type (wt) dihydrofolate reductase (DHFR) and its D27S mutant in (a) 6 M urea and 10 mM Tris buffer, pH 8.1 and (b) 25 mM acetic acid, 15 mM phosphate, 15 mM Tris, and 200 mM KCl, pH 9.0.

ference spectrum of the Raman spectra of an aspartate with a serine without any other residue signatures because the concentration of urea is sufficient to unfold most of the protein (the spectrum of urea has been subtracted). The positive peak at  $1397\text{ cm}^{-1}$  and the negative peak at  $1356\text{ cm}^{-1}$  are then tentatively assigned to the carboxylate  $\text{COO}^-$  symmetric stretch mode of Asp27 and a vibrational mode of Ser27, respectively, based on the observation that bands at similar (but not identical) positions are found in the solution aspartate and serine spectra. The scale of the y-axis refers to a percentage of the intensity of the protein  $1450\text{ cm}^{-1}$  band; and the intensity of the observed bands in the difference spectrum are about right for the putative carboxylate and serine bands.

Figure 5b shows the difference Raman spectrum of wild-type DHFR and its D27S mutant in a buffer without urea. Which bands can be assigned to Asp27 or Ser27 in this spectrum is not obvious. Apparently, the network of interactions among protein residues is disturbed upon mutation of Asp27 to serine, and many Raman bands from various amino acids, which are sensitive to the environment, then become prominent in the difference spectrum. The character in this difference spectrum is similar to that found for PGM between the phospho and dephospho forms (see above; Figure 3c). Although an extensive analysis (which would follow the general outline sketched above for PGM) has not been done, the sharp and intense band at  $1420\text{ cm}^{-1}$ , which is about three times more intense than the carboxylate  $\text{COO}^-$  symmetric stretch band, probably arises from the ring N-H bending motion of one of DHFR's five protein tryptophan residues. The shift in frequency of

this band from its normal position at  $1435\text{ cm}^{-1}$  in water suggests that the ring nitrogen of this tryptophan in the wild-type DHFR is strongly hydrogen bonded, and its environment becomes much more hydrophobic in the D27S mutant. Further studies with specific mutated DHFR tryptophan residues should reveal which one is disturbed in the D27S mutant.

## CONCLUSION

The above discussion clearly demonstrates the value of vibrational studies in answering important questions of the structures and functions of proteins and other biological macromolecules, as do many other investigations using resonance Raman and FTIR difference spectroscopies and other forms of vibrational spectroscopy. Their use has been limited because of the difficulty of assigning bands in a complicated protein spectrum. Because Raman difference spectroscopy deals directly with this problem and may be used on essentially any protein system, many problems of protein structure are now open to study. Moreover, the instrumentation is relatively inexpensive. Currently, an instrument can be purchased for about the price of a couple of high performance liquid chromatographs. Measurements may be performed rather quickly, on relatively small quantities of sample, and on protein solutions.

Some problems are particularly suited for Raman difference spectroscopy. One concerns characterization of the interactions between a protein and a ligand, indeed a very large area of protein research and one that is not well understood. Another has to do with investigations of the distortions of particular bonds, such as those that take place in enzymatic catalysis.

As suggested above, Raman difference spectroscopy may be useful for the study of how ligand binding or mutation changes the structure of a protein. The resolution of vibrational spectroscopy is very high compared with other techniques, and high-resolution pictures are needed in the study of these relatively poorly characterized, but very important, problems.

## ACKNOWLEDGMENTS

The development of the techniques discussed here, and their applications, have been the enterprise of several colleagues. We acknowledge the efforts of Drs. TK Yue, D Chen, J van Beck, J Ball, J Burgner, D Mannor, G Weng, J Zheng, D Sloan, C Martin and students, Y-Q Chen, and D Xiao. Support from the National Institutes of Health

(grants GM35183 and EYO3142) and the National Science Foundation (grant MCB-8912322) is gratefully acknowledged.

Any Annual Review chapter, as well as any article cited in an Annual Review chapter, may be purchased from the Annual Reviews Preprints and Reprints service.

1-800-347-8007; 415-259-5017; email: arpr@class.org

### Literature Cited

1. Anderson VE, LaReau RD. 1988. Hydride transfer catalyzed by lactate dehydrogenase displays absolute stereospecificity at the C4 of the nicotinamide ring. *J. Am. Chem. Soc.* 110:3695-97
2. Badger RM, Bauer SH. 1937. Spectroscopic studies of the hydrogen bond. II. The shift of the O-H vibrational frequency in the formation of the hydrogen bond. *J. Chem. Phys.* 5:839-55
3. Baenziger JE, Miller KW, Rothschild KJ. 1992. Incorporation of the nicotinic receptor into planar multilamellar films: characterization by fluorescence and Fourier transform infrared difference spectroscopy. *Biophys. J.* 61:983-92
4. Belasco JG, Knowles JR. 1980. Direct observation of substrate distortion by triosephosphate isomerase using fourier transform infrared spectroscopy. *Biochemistry* 19:472-77
5. Belasco JG, Knowles JR. 1983. Polarization of substrate carbonyl group by yeast aldolase: investigation by fourier transform infrared spectroscopy. *Biochemistry* 22:122-29
6. Braiman MS, Rothschild KJ. 1988. Fourier transform infrared techniques for probing membrane protein structure. *Annu. Rev. Biophys. Biophys. Chem.* 17:541-70
7. Breton J, Nabedryk E, Parson WW. 1992. A new infrared electronic transition of the oxidized primary electron donor in bacterial reaction centers: a way to assess resonance interactions between the bacteriochlorophylls. *Biochemistry* 31:7503-10
8. Breton J, Vermeglio A, ed. 1992. *The Photosynthetic Bacterial Reaction Center*, Vol. 2, *Structure, Spectroscopy, and Dynamics*. New York: Plenum
9. Callender R, Chen D, Lugtenburg J, Martin C, Ree KW, et al. 1988. The molecular properties of p-dimethylamino benzaldehyde bound to liver alcohol dehydrogenase: a Raman spectroscopic study. *Biochemistry* 27:3672-81
10. Callender R, Deng H, Sloan D, Burgner J, Yue TK. 1989. Raman difference spectroscopy and the energetics of enzymatic catalysis. *Proc. Int. Soc. Opt. Eng.* 1057:154-60
11. Campbell AP, Sykes BD. 1993. The two-dimensional transferred nuclear overhauser effect: theory and practice. *Annu. Rev. Biophys. Biomol. Struct.* 22:99-122
12. Carey PR. 1982. *Biochemical Applications of Raman and Resonance Raman Spectroscopy*. New York: Academic
13. Carey PR, Tonge PJ. 1990. Chemistry of enzyme-substrate complexes revealed by resonance Raman spectroscopy. *Chem. Soc. Rev.* 19:293-316
14. Chen D, Yue KT, Martin C, Rhee KW, Sloan D, Callender R. 1987. Classical Raman spectroscopic studies of NADH and NAD<sup>+</sup> bound to liver alcohol dehydrogenase by difference techniques. *Biochemistry* 26:4776-84
15. Deng H, Burgner J, Callender R. 1991. Raman spectroscopic studies of NAD coenzymes bound to malate dehydrogenases by difference techniques. *Biochemistry* 30:8804-11
16. Deng H, Burgner J, Callender R. 1992. Raman spectroscopic studies of the effects of substrate binding on coenzymes bound to lactate dehydrogenase. *J. Am. Chem. Soc.* 114:7997-8003
17. Deng H, Callender R. 1993. Enzymatic catalysis and molecular recognition: the energetics of ligand binding to proteins as studies by vibrational spectroscopy. *Comments Mol. Cell. Biophys.* 8:137-54
18. Deng H, Goldberg JM, Kirsch JF, Callender R. 1993. Elucidation of the solution structure of the *Escherichia coli* aspartate aminotransferase- $\alpha$ -methyl-L-aspartate complex by isotope edited Raman difference spectroscopy. *J. Am. Chem. Soc.* 115:8869-70
19. Deng H, Manor D, Weng G, Chen C-X, Balogh-Nair V, Callender R. 1993. Difference Raman spectroscopic studies of ligand-protein interactions. *Proc. Int. Soc. Opt. Eng.* 1890:114-22
20. Deng H, Ray WJ, Burgner JW, Callen-

- der R. 1993. A comparison of vibrational frequencies of critical bonds in ground state complexes and in a vanadate-based transition state complex of muscle phosphoglucomutase. Mechanistic implications. *Biochemistry*. 32: 12984-92
21. Deng H, Zheng J, Burgner J, Callender R. 1989. Hydrogen bonding and reaction specificity in lactate dehydrogenase studied by Raman spectroscopy. *J. Phys. Chem.* 93:4710-13
22. Deng H, Zheng J, Burgner J, Callender R. 1989. Molecular properties of pyruvate bound to lactate dehydrogenase: a Raman spectroscopic study. *Proc. Natl. Acad. Sci. USA* 86:4484-88
23. Deng H, Zheng J, Burgner J, Sloan D, Callender R. 1989. Classical Raman spectroscopic studies of NADH and NAD<sup>+</sup> bound to lactate dehydrogenase by difference techniques. *Biochemistry* 28:1525-33
24. Deng H, Zheng J, Clarke A, Holbrook JJ, Callender R, Burgner JW. 1994. The source of catalysis in the lactate dehydrogenase system, part II: ground state interactions in the enzyme-substrate complex. *Biochemistry*. In press
25. Deng H, Zheng J, Sloan D, Burgner J, Callender R. 1992. A vibrational analysis of the catalytically important C4-H bonds of NADH bound to lactate or malate dehydrogenase: ground state effects. *Biochemistry* 31:5085-92
26. Fersht A. 1985. *Enzyme Structure and Mechanism*. New York: Freeman
27. Gao J, Kuczera K, Tidor B, Karplus M. 1989. Hidden thermodynamics of mutant proteins: a molecular dynamics analysis. *Science* 244:1069-72
28. Goldberg JM, Zheng J, Deng H, Chen YQ, Callender R, Kirsch JF. 1993. The structure of the complex between pyridoxal 5'-phosphate and the tyrosine-225 to phenylalanine mutant of *Escherichia coli* aspartate aminotransferase determined by isotope edited classical Raman difference spectroscopy. *Biochemistry* 32:8092-97
29. Gordy W. 1946. A relation between bond force constants, bond orders, bond lengths and the electronegativities of the bonded atoms. *J. Chem. Phys.* 14:305-20
30. Jeffrey GA, Saenger W. 1991. *Hydrogen Bonding in Biological Structures*. Berlin: Springer-Verlag
31. Joesten M, Schaad LJ. 1974. *Hydrogen Bonding*. New York: Marcel Dekker
32. Kiefer W. 1973. Raman difference spectroscopy with a rotating cell. *Appl. Spectrosc.* 27:253-57
33. Kim M, Carey PR. 1993. Observation of a carbonyl feature for riboflavin bound to riboflavin-binding protein in the red-excited Raman spectrum. *J. Am. Chem. Soc.* 115:7015-16
34. Kim M, Owen H, Carey PR. 1993. A high-performance Raman spectroscopic system based on a single spectrograph, CCD, notch filters and a Kr laser ranging from the Near-IR to Near-UV regions. *Appl. Spectrosc.* In press
35. Komives EA, Change LC, Lolis E, Tilton RF, Petsko GA, Knowles JR. 1991. Electrophilic catalysis in triosephosphate isomerase: the role of histidine-95. *Biochemistry* 30:3011-19
36. Kurz LC, Drysdale GR. 1987. Evidence from Fourier transform infrared spectroscopy for polarization of the carbonyl of oxaloacetate in the active site of citrate synthase. *Biochemistry* 26:2623-27
37. LaReau RD, Anderson VE. 1992. An inquiry into the source of stereospecificity of lactate dehydrogenase using substrate analogues and molecular modeling. *Biochemistry* 31:4174-80
38. Latajka Z, Scheiner S. 1990. Correlation between interaction energy and shift of the carbonyl stretching frequency. *Chem. Phys. Lett.* 174:179-84
39. Li H, Hanson C, Fuchs JA, Woodward C, Thomas JGJ. 1993. Determination of the pKa values of active-center cysteines, cysteines-32 and -35, in *Escherichia coli* thioredoxin by Raman spectroscopy. *Biochemistry* 32:5800-8
40. Manor D, Weng G, Deng H, Cosloy S, Chen CX, et al. 1991. An isotope edited classical Raman difference spectroscopic study of the interactions of guanine nucleotides with elongation factor Tu and H-ras p21. *Biochemistry* 30: 10914-20
41. Mäntele W. 1993. Reaction-induced infrared difference spectroscopy for the study of protein function and reaction mechanisms. *Trends Biochem. Sci.* 18: 197-202
42. Peticolas WL, Bajdor K, Patapoff TW, Wilson KJ. 1987. New methods of studying enzyme-substrate interactions using ultra violet resonance Raman and microscopic Raman difference technique. In *Studies in Physics and Theoretical Chemistry*, ed. J Stepanek, P Anzenbacher, B Sedlacek, pp. 249-70. Amsterdam: Elsevier
43. Ray JWJ, Burgner IJW, Deng H, Callender R. 1993. A comparison of internal chemical bonding in solutions of simple phosphates and vanadates. *Biochemistry*. 32:12977-83
44. Ray WJ, Post CB, Liu Y, Rhyu GI.

1993. Structural changes at the metal binding site during the phosphoglucosyltransferase reaction. *Biochemistry* 32:48–57
45. Rothschild KJ. 1992. FTIR difference spectroscopy of bacteriorhodopsin: toward a molecular model. *J. Bioenerg. Biomembr.* 24:147–67
46. Rousseau DL. 1981. Raman difference spectroscopy as a probe of biological molecules. *J. Raman Spectrosc.* 10:94–99
47. Slaich PK, Primrose WU, Robinson D, Drabble K, Wharton CW, et al. 1992. The binding of amide substrate analogues to phospholipase A2. *Biochem. J.* 288:167–73
48. Spiro TG, ed. 1987. *Biological Applications of Raman Spectroscopy*, Vol. 1, *Raman Spectra and the Conformations of Biological Molecules*. New York: Wiley & Sons
49. Spiro TG, ed. 1987. *Biological Applications of Raman Spectroscopy*, Vol. 2, *Resonance Raman Spectra of Polyenes and Aromatics*. New York: Wiley & Sons
50. Spiro TG, ed. 1988. *Biological Applications of Raman Spectroscopy*, Vol. 3, *Resonance Raman Spectra of Heme and Metalloproteins*. New York: Wiley & Sons
51. Thijs R, Zeegers-Huyskens T. 1984. Infrared and Raman studies of hydrogen bonded complexes involving acetone, acetophenone, and benzophenone-1. Thermodynamic constants and frequency shifts of the  $\nu_{\text{OH}}$  and  $\nu_{\text{C=O}}$  stretching vibrations. *Spectrochim. Acta* 40A:307–13
52. Thomas GJ, Tsuboi M. 1993. Raman spectroscopy of nucleic acids and their complexes. *Adv. Biophys. Chem.* 3:1–69
53. Tonge PJ, Carey PR. 1990. Length of the acyl carbonyl bond in acyl-serine proteases correlates with reactivity. *Biochemistry* 29:10723–27
54. Tonge PJ, Carey PR. 1993. Raman, resonance Raman and FTIR spectroscopic studies of enzyme-substrate complexes. In *Biomolecular Spectroscopy*, ed. RJH Clarke, RE Hester, pp. 129–61. New York: Wiley & Sons
- 54a. Tonge PJ, Carey PR, Callender RH, Deng H, Ekrel I, Muhandiram R. 1993. Characterization of *trans* and *cis* 5-methylthienylacryloyl-chymotrypsin using Raman difference spectroscopy, NMR, and kinetics: carbonyl environment and reactivity. *J. Am. Chem. Soc.* 115:8757–62
55. Tonge PJ, Pusztai M, White AJ, Wharton CW, Carey PR. 1991. Resonance Raman and Fourier transform infrared spectroscopic studies of the acyl carbonyl group in [3-(5-methyl-2-thienyl)acryloyl]chymotrypsin: evidence for artifacts in the spectra obtained by both techniques. *Biochemistry* 30:4790–95
56. Vinogradov SN, Linnel RH. 1971. *Hydrogen Bonding*. New York: Van Nostrand Reinhold
57. Weng G, Chen CX, Chen Z, Balogh-Nair V, Callender R, Manor D. 1994. The hydrogen bonding of G proteins with the guanine ring moiety of guanine nucleotides. *Protein Sci.* In press
58. White AJ, Drabble K, Ward S, Wharton CW. 1992. Analysis and elimination of protein perturbation in infrared difference spectra of acyl-chymotrypsin ester carbonyl groups by using isotopic substitution. *Biochem. J.* 287:317–23
59. White AJ, Wharton CW. 1990. Hydrogen-bonding in enzymatic catalysis. *Biochem. J.* 270:627–37
60. Yager P, Gaber BP. 1987. Membranes. See Ref. 48, pp. 81–133
61. Yue KT, Deng H, Callender R. 1989. Raman difference spectroscopy in measurements of molecules and molecular groups inside proteins. *J. Raman Spectrosc.* 20:541–46
62. Yue KT, Martin CL, Chen D, Nelson P, Sloan DL, Callender R. 1986. Raman spectroscopy of oxidized and reduced nicotinamide adenine dinucleotides. *Biochemistry* 25:4941–47
- 62a. Yue KT, Minghe L, Zheng J, Callender R. 1991. The determination of the  $\text{pK}_a$  of histidine in proteins by Raman difference spectroscopy. *Biochim. Biophys. Acta* 1078:296–302
63. Yue KT, Yang JP, Charlotte M, Lee SK, Sloan D, Callender R. 1984. A Raman study of reduced nicotinamide adenine dinucleotide bound to liver alcohol dehydrogenase. *Biochemistry* 23:6480–83
64. Zheng J, Chen YQ, Callender R. 1993. A study of the binding of NADP coenzymes to dihydrofolate reductase by Raman difference spectroscopy. *Eur. J. Biochem.* 215:9–16