Protein Influences on Porphyrin Structure in Cytochrome *c*: Evidence from Raman Difference Spectroscopy†

John A. Shelnutt,‡ Denis L. Rousseau,* Judy K. Dethmers,§ and Emanuel Margoliash

ABSTRACT: To probe the details of protein heme interactions, we have developed a Raman difference spectroscopic technique which allows reliable detection of very small, ≈0.01 cm^{-1}, frequency differences. When this technique is applied to heme proteins, structural differences in the protein which perturb the porphyrin macrocycle may be examined by obtaining Raman difference data on the porphyrin vibrational modes which are strongly enhanced in the Raman spectrum produced with visible laser excitation. We report here Raman difference spectroscopic data on cytochromes *c* from 24 species. The differences in the Raman spectrum of the porphyrin between the cytochromes *c* of any two species are small, confirming that all of the cytochromes we have examined have the same "cytochrome fold". However, many small (0.02–2 cm^{-1}) but systematic differences were detected which indicate structural differences among these proteins. These differences could be classified into three different groups and interpreted in terms of different types of structural variations resulting from specific differences in the amino acid sequences. First, direct interactions between near-heme residues and the porphyrin influence the electron density in the π orbitals of the porphyrin macrocycle. Second, variation in the residue at position 92, far removed from the heme, affects the frequency of the core-size marker line at 1584 cm^{-1}. Third, the conformation near cysteine 14 affects the shape of the Raman mode which is sensitive to the pyrrole ring substituents (≈1313 cm^{-1}). From these data we conclude that there are several ways in which the protein amino acid sequence may regulate the oxidation–reduction potential and several ways in which the sequence can modify the binding site between cytochrome *c* and its redox partners.

Nearly every known electron transport system in both eukaryotic and prokaryotic species contains forms of cytochrome *c*. The molecule was first discovered by Keilin (1930), and because it is a small, stable, water-soluble protein, it has been the subject of innumerable investigations (Ferguson-Miller et al., 1979). Consequently, its basic structural features are now very well characterized. Its single prosthetic group, a low-spin heme, is covalently bound to the protein by two thioether bonds, and the heme iron is bound to methionine residue 80 and the imidazole of residue 18. Amino acid sequences of over 80 species are now known (Ferguson-Miller et al., 1979). X-ray crystallographic structures from three cytochromes *c* have been obtained (Dickerson & Timkovich, 1975) and reveal a common tertiary structure, designated the "cytochrome fold" (Margoliash, 1972). However, in spite of the wealth of information now available, there is as yet no clear understanding of the pathway by which the protein carries out its basic function—the transfer of an electron from a donor to an acceptor.

Determination of the molecular basis for the mechanism of electron transfer in cytochromes *c* has been hindered by the complexity of its mitochondrial oxidoreductases coupled with the fact that they are membrane bound. As such, crystallographic studies have been limited to the isolated protein which can be crystallized and most spectroscopic studies have also been done on pure samples to be free of the added complexity of the other proteins. Because of the complexity of the oxidoreductase systems for cytochrome *c*, many workers have found it useful to do comparative studies of cytochromes and examine the difference in properties as a function of the amino acid sequence. In such crystallographic and spectroscopic studies on cytochromes *c* from various species, in some cases either only minor species dependent changes are observed or changes are so complex that interpretation is very difficult. However, in certain cases correlations between spectroscopic parameters and specific amino acid sequence variations have been detected. For example, in EPR[1] studies of the electron distribution around the iron atom, Brautigan et al. (1977) found that ferricytochromes *c* of several species could be separated into two classes. It was argued that the classes resulted from the protonation or deprotonation of N-1 of the imidazole ligand. Functional implications of this variability could not be determined. In an NMR study of a series of cytochromes *c*, Cookson et al. (1977) were able to conclude that the protein fold in solution is the same as that in crystals. They also found that the conformations of mitochrondial cytochromes *c* and bacterial cytochromes c_2 are the same, and a correlation between redox potential and the methionine ligand resonances was observed (Cookson et al., 1978; Moore & Williams, 1977). They attributed the correlation to a change in the iron–sulfur bond lengths. In a recent series of papers, Moore & Williams (1980a–f) studied the NMR spectra of eukaryotic cytochromes under a variety of conditions and made comparisons among species. The found small but clear differences in structure when polar groups are substituted for nonpolar groups, and they noted that substitution of several amino acids may result in widespread structural modifications which are more extensive than would be expected from each substitution taken individually (Moore & Williams, 1980f). In a recent study of the pK_a and transition temperature of the 695-nm absorption band in ferricytochromes *c*, considerable variation in the stability of the native conformation of the heme crevice was inferred (Osheroff et al., 1980). It was concluded

† From Bell Laboratories, Murray Hill, New Jersey 07974 (J.A.S. and D.L.R.), and the Department of Biochemistry and Molecular Biology, Northwestern University, Evanston, Illinois 60201 (J.K.D. and E.M.).
Received February 10, 1981; revised manuscript received May 27, 1981. The portion of the work carried out at Northwestern University was supported by National Institutes of Health Grants GM-19121 and HL-11119.
‡ Present address: Sandia Laboratories, Albuquerque, NM 87115.
§ Present address: Department of Biochemistry, Cornell University Medical College, New York, NY 10021.

[1] Abbreviations used: RDS, Raman difference spectroscopy; CDNP, 4-carboxy-2,6-dinitrophenyl; EPR, electron paramagnetic resonance; NMR, nuclear magnetic resonance.

0006-2960/81/0420-6485$01.25/0

that several sequence positions influence the strength of the two surface interactions which constitute the bonds favoring the closed conformation of the heme crevice.

The technique of resonance Raman scattering has also been applied to heme proteins and was found to be sensitive to oxidation state, spin state, and anomalous structure (Spiro & Strekas, 1972; Strekas & Spiro, 1974; Rakshit & Spiro, 1974; Spiro & Burke, 1976). Normal coordinate analyses of the porphyrin macrocycle have also been reported (Stein et al., 1975; Kitagawa et al., 1976; Abe et al., 1976, 1978; Sunder & Bernstein, 1976; Susi & Ard, 1977). With such a foundation, it is evident that the resonance Raman technique can probe several important heme protein properties. Specifically, these are as follows: (1) The intensity dependence of the lines of the Raman spectrum with excitation frequency yields information about the electronic states (Shelnutt, 1980) including the vibrational frequencies of excited states. Vibronic coupling between Q and B states of the heme can be studied and effects of nonadiabatic coupling and Jahn-Teller distortion in heme proteins can be determined. (2) Excited-state lifetimes may be directly measured from widths of resonance Raman excitation profiles (Friedman et al., 1977). (3) Large frequency shifts (20 cm^{-1}) of certain ground-state vibrational modes occur between the oxidized and reduced forms and between high- and low-spin states of heme proteins (Spiro & Strekes, 1972). These serve as probes of heme structure and allow ready identification of those proteins that conform to a given class and those that have anomalous properties. (4) In a study of a homologous series of model compounds (Spaulding et al., 1974), a vibrational mode frequency vs. porphyrinato core-size correlation has been made, demonstrating the specificity and sensitivity of resonance Raman scattering as a structural probe. However, until recently, resonance Raman scattering was found to be insensitive to the more subtle perturbations of the properties of the prosthetic group by the protein. In this paper we describe a Raman scattering technique, Raman difference spectroscopy (RDS)[1] (Rousseau, 1981), in which the spectra of two samples are obtained simultaneously. By this method frequency differences of less than 10^{-2} of the width of a given Raman line have been accurately measured. In the present study we have used this technique to examine the cytochromes *c* from several species. Frequency differences as small as 0.02 cm^{-1} have been interpreted in terms of the influence of variations in amino acid sequences on the properties of the heme. The larger body of data reported here has upheld previous conclusions based on our preliminary data (Shelnutt et al., 1979a) that large frequency differences result from comparisons of two cytochromes *c* with different near-heme environments. Further, we have observed one series of differences that results from apparent protein conformational changes induced by amino acid sequence differences away from the heme and a series of line shape or frequency differences resulting from sequence variations at positions adjacent to a cysteine covalent linkage.

Experimental Procedures

Materials. Horse, tuna, pigeon, dog, chicken, pig, cow, sheep, rabbit, and *Candida krusei* cytochromes *c* were purchased from Sigma Chemical Co. Spider monkey, human, donkey, cow, snapping turtle, mouse, rabbit, guanaco, hippopotamus, Pekin duck, screwworm fly, tobacco hornworm moth, *Samia cynthia* moth, and horse cytochromes *c* were prepared by the procedure of Margoliash & Walasek (1976), as modified by Brautigan et al. (1978a). Human tissue was obtained at autopsy. Before spectra were taken, all preparations were filtered through a column of Sephadex G-50 superfine (Pharmacia) to completely separate polymeric material (Brautigan et al., 1978a), which caused a characteristic background fluorescence and a decrease in sensitivity of the difference spectra resulting from the increase in noise level. In general, the spectra were obtained with ferrocytochrome *c* (minimal dithionite or ascorbate reduced) at a concentration of about 2×10^{-4} M, buffered in 10 or 25 mM Tris–acetate, pH 7.8. Spectra were obtained in Pyrex or quartz cells at room temperature, taking from 2 to 12 h for completion. At the end of most runs there were no changes in the Raman spectra that could be attributed to the formation of ferricytochrome *c*. On several occasions optical spectra were taken at the end of the run to further demonstrate that no autooxidation had occurred. Oxidation, if it occurs, is indicated in the Raman spectrum by the appearance of a line at 1640 cm^{-1}. With the CDNP derivatives, some oxidation was noted. However, the amounts were sufficiently small so as to affect the difference spectrum only slightly. Gel filtration before and after a run confirmed that no polymeric material was present.

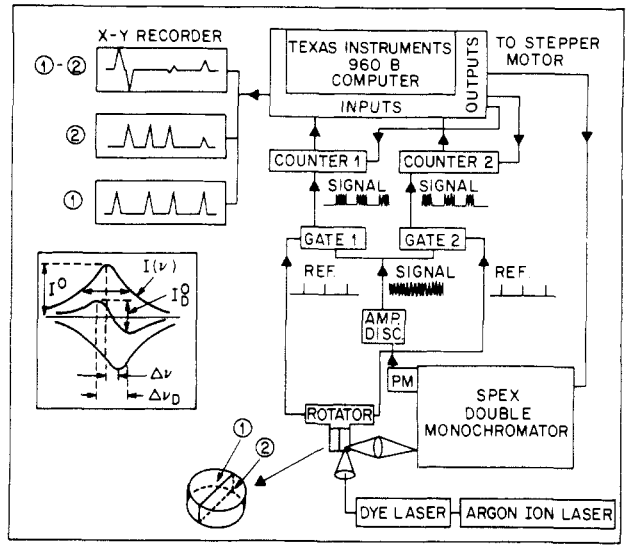

FIGURE 1: Experimental arrangement for obtaining Raman difference spectra. The incident laser beam is focused on the split rotating cell with different samples in side 1 and side 2. The scattered light is focused on the entrance slit of a Spex double monochromator, dispersed, and detected by the photomultiplier (PM). The signal from the photomultiplier goes to the amplifier discriminator (Amp. Disc.) and then to the two linear gates. A reference signal from the rotator controls the gates such that gate 1 only passes the signal from side 1 of the rotating cell and gate 2 only passes that from side 2. The separated signals are accumulated in their respective counters for a preset time and then stored in separate memory locations in the computer. A signal from the computer to the stepper motor advances the spectrometer and data accumulation resumes. When the spectra are completed, each Raman spectrum or difference spectrum may be displayed and plotted on the X–Y recorder and the complete data set may be stored on cassettes. The inset shows the shape obtained in a difference spectrum and the parameters used to calculate the frequency separations.

Raman Difference Spectra. Raman difference spectra were obtained by a modification of the technique originally reported by Kiefer (1973, 1977). Samples were placed in a cylindrical cell with a partition along a diameter such that the two samples were alternatively illuminated by the laser as the cell was rotated (see Figure 1). Thus both samples could be examined for each frequency setting of the monochromator. After dispersion by the monochromator, the signal was amplified and sent to the inputs of two linear gates. The shutters of the gates were controlled by a synchronous signal from the rotator, and the gate delay times and window widths were adjusted such that only scattered light from side 1 of the cell passed

through gate 1 and only scattered light from side 2 passed through gate 2. The signals were accumulated in their respective counters for a present time prior to storage in the 960B Texas Instruments computer with 64K words of memory. Following transfer of information from the counters to the computer, the double monochromator was advanced by a preset frequency increment. Typically the cell was rotated at 30 rps and the monochromator was advanced in steps ranging from 0.1 to 0.5 cm^{-1}. The dwell time at each step was 0.5 s. The spectral slit width was 2 cm^{-1}. Ten to fifty scans were performed to signal average the data, and up to 4096 points could be collected for each of the two spectra. The data were removed from the computer by plotting on an X–Y recorder and were stored permanently on cassettes.

The apparatus described here is similar to that described by Kiefer (1973, 1977) and others (Gardiner et al., 1975) except for electronics which allow for the full capabilities of photon counting and the utilization of a computer for the analysis of the data. Another computerized Raman difference spectrometer has been reported (Christman et al., 1976), but it has not been applied to the study of very small frequency differences. For Raman difference spectra of heme proteins, the computer is a necessary part of the system. It not only allows signal averaging but also, most importantly, it provides the capability of a posteriori selection of the balance between signal 1 and signal 2. This is an essential feature of the apparatus since it is not feasible to a priori balance the two signals of every Raman line if one has a fluorescence background associated with it or if the Raman modes have different relative intensities in the two samples being compared. Such relative intensity differences occur when the incident frequency is in resonance with optical transitions that vary slightly.

The line widths of the Raman lines in heme proteins are typically in the 5–15-cm^{-1} range. For such broad lines, spectra of different samples that are obtained in separate scans are difficult to compare accurately. With conventional instrumentation, confidence in the difference in Raman frequencies of a line in the two spectra is only about 1 or 2 cm^{-1}. In contrast, if the spectra are obtained simultaneously in a split cell by the RDS technique described here, sensitivity to frequency differences smaller than $^1/_{100}$ of the line width can be achieved.

When difference spectra of lines that differ by a small amount are recorded, a "derivative" type of shape is obtained as shown in the inset in Figure 1. When the frequency separation between the lines is large, the difference spectrum is intense. When the frequency difference between the two lines is smaller, the intensity in the difference spectrum is correspondingly reduced, but the peak to peak frequency separation in the difference spectrum reaches a limiting value. For Lorentzian lines this value is $\Gamma/3^{1/2}$ where Γ is the full width at half-height of the Lorentzian line.

Quantitative frequency shifts may be readily calculated from the difference spectrum (Gardiner et al., 1975; Laane & Kiefer, 1980; Rousseau, 1981). Let us define a difference function, $I_D(\nu)$, as the difference between two lines of the same shape and intensity but shifted in frequency by $\Delta\nu$; i.e.

$$I_D(\nu) = I(\nu + \Delta\nu) - I(\nu) \quad (1)$$

For very small $\Delta\nu$, $I_D(\nu)$ may be approximated by a Taylor series expansion and with only the first-order term included so that

$$I_D(\nu) = [\partial I(\nu)/\partial\nu]\Delta\nu \quad (2)$$

If we assume $I(\nu)$ is a Lorentzian, the frequency separation, $\Delta\nu_D$, between the peak and the valley in the difference spectrum may be readily calculated by setting the derivative in eq 2 equal to zero, and $\Delta\nu_D$ turns out to be $\Gamma/3^{1/2}$. A relationship between the frequency separation, $\Delta\nu$, of the lines and the peak to peak intensity, I_D^0, in the difference spectrum may then be calculated by evaluating eq 2 at $\pm\Gamma/[2(3^{1/2})]$. We obtain

$$\frac{\Delta\nu}{\Gamma} = 0.38\left|\frac{I_D^0}{I^0}\right| \quad (3)$$

where we have also used the relationship between the peak intensity I^0 of the Lorentzian line and Γ. Equation 3 is only valid when $\Delta\nu \ll \Gamma$. As $\Delta\nu$ becomes larger, higher order terms in the expansion must be included in order to relate $\Delta\nu/\Gamma$ to I_D^0/I^0 or eq 1 must be evaluated numerically. It should be noted that in the limit of small differences selection of a Gaussian line rather than a Lorentzian line only changes the coefficient in eq 3 from 0.38 to 0.35. Additional extensive discussions of the Raman difference technique (Rousseau, 1981) and the data analysis (Laane & Kiefer, 1980) were recently reported.

From the above discussion, a qualitative picture of the difference spectrum behavior should be clear: As the difference between two spectral lines becomes smaller, the frequency spacing of the maximum and minimum reach a limiting value but the intensity continues to drop. The lower limit of a frequency difference that is detectable is then *not* limited by spectral resolution but instead by the signal to noise ratio. Assuming that the noise limitation is shot noise so that the noise level equals the square root of the number of counts, then a 10^4 count line should have a difference sensitivity limit of $\sim 10^{-2}\Gamma$. In order to attain such levels, it is necessary to isolate carefully the scattering system from room vibrations, stabilize the intensity of the laser output, and properly construct the optical cell and its holder to prevent wobble or movement of the scattering column. Also, the spectrometer step size must be small enough to allow accurate determination of the peak and valley in the difference spectrum. The entrance slit must be set to give high enough resolution so as not to broaden the Raman lines, since Γ goes into the calculation of the frequency differences. When signal averaging, the additional precaution of controlling the spectrometer temperature must be taken so that each scan starts at approximately ($\leq\pm 1$ cm^{-1}) the same frequency. In this way very small frequency differences may be detected since errors due to thermal drift and irreproducibility of the monochromator are eliminated by obtaining data for both samples at each setting of the spectrometer. The spectral resolution and the spectrometer frequency stepping interval which we used in this study have been shown from model calculations to introduce no larger than a 6% error in the frequency differences.

Since the completed construction of the Raman difference apparatus (March 1978), many tests have been run to ensure that no artifactual differences occur. The error limits for frequency differences in isolated lines result from the noise limitation in the difference spectrum. However, if there are closely spaced Raman lines in a spectrum, the accuracy is limited by our ability to distinguish between the contributions of each of the lines to the spectrum.

In our preliminary RDS study of cytochromes (Shelnutt et al., 1979a), we reported an error limit for the detection of frequency differences in moderate strength lines of 0.2 cm^{-1}. The data for that study were obtained by advancing the monochromator in 0.5-cm^{-1} steps. In the current study the monochromator was advanced in 0.1- or 0.2-cm^{-1} steps. This allows greater precision in determining the magnitude of the signal in the difference spectrum from which our frequency

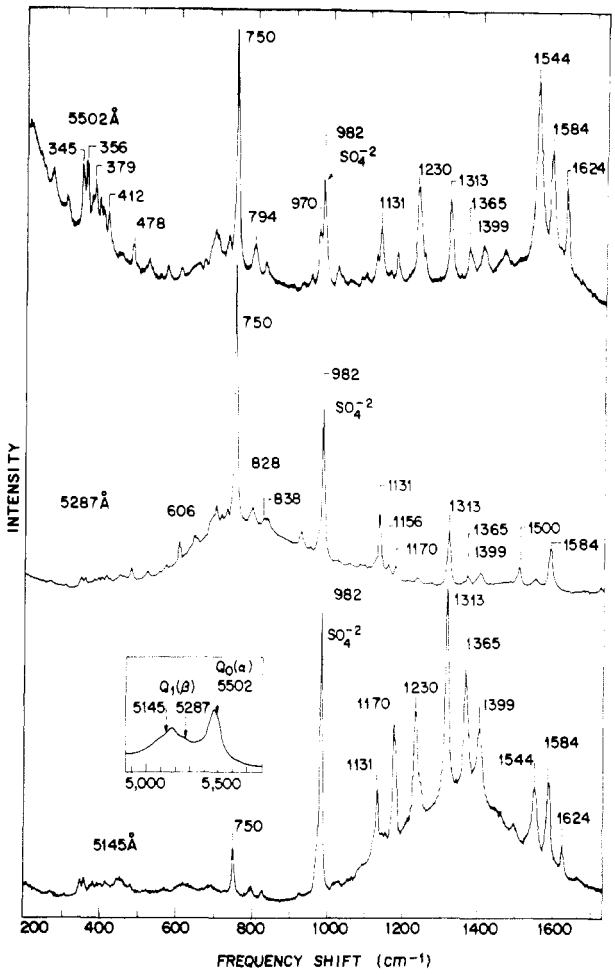

FIGURE 2: Resonance Raman spectra of horse ferrocytochrome *c* for three difference excitation wavelengths. The frequencies of the strongest lines are labeled. The line at 982 cm^{-1} results from SO_4^{2-} which was added to the samples as a relative intensity reference. The relationship between the excitation wavelengths and the visible absorption spectrum is shown in the inset. The excitation wavelength of the top spectrum, 5502 Å, lies near the peak of the α band. The excitation wavelength of the middle spectrum, 5287 Å, is shifted from the peak of the α band by about 1350 cm^{-1}. No polarization scrambler was used.

differences are determined. Because of this and by longer signal averaging, we have been able to reduce our errors in strong lines to 0.02 cm^{-1} and in the moderate strength and weaker lines to 0.05 cm^{-1}.

Results

Resonance Raman Spectra of Cytochrome c. When laser excitation is with visible light, the Raman spectrum of a heme protein shows only the vibrational lines associated with the heme. This selectivity is a result of enhancement of the Raman scattering of the chromophore due to resonance with its lowest allowed $\pi \rightarrow \pi^*$ transition. Resonance Raman spectra of ferrocytochrome *c* from horse for the three excitation frequencies in the region of the visible absorption bands are shown in Figure 2. The positions of these excitation frequencies with respect to the optical transition are indicated in the inset. In each spectrum a standard amount (2 M) of SO_4^{2-} ion was added to allow comparison of the relative intensities of the resonance Raman lines obtained at different excitation frequencies. There are no optical absorption bands from SO_4^{2-} in this region, so its intensity should display only a slowly varying and monotonic ν^4 frequency dependence. The frequencies of the more prominent Raman scattering modes are labeled in each spectrum.

The top spectrum in Figure 2 was recorded near the peak of the α band. The spectrum is resplendent with lines ranging from those at relatively low frequency (345 cm^{-1}) to those at high frequency (1624 cm^{-1}). The lines at 1131, 1313, 1399, and 1584 cm^{-1} are anomalously polarized (A_{2g} modes) (Spiro & Stekas, 1972), and those at 750, 970, 1170, 1230, 1365, 1544, and 1624 cm^{-1} are polarized (A_{1g} modes) or depolarized (B_{1g} or B_{2g} modes). The depolarization ratios ($I_\parallel/I_\perp$) of the other lines in the spectrum have not been determined.

At the 5287-Å excitation wavelength, corresponding to a shift in frequency from the peak of the α band by $\approx$700 cm^{-1}, the strongest line in the spectrum is the 750-cm^{-1} line. Further decreasing the wavelength to 5145 Å, a 1350-cm^{-1} shift from the peak of the α band, results in relatively strong lines in the 1100–1600-cm^{-1} region. There is a broad fluorescence centered near the zero frequency shift in the spectrum excited with 5502-Å light, at about 750 cm^{-1} in the spectrum excited with 5287 Å, and at 1350 cm^{-1} in the spectrum excited with 5145 Å.

RDS Spectra. We have obtained RDS data for the cytochromes *c* from 17 vertebrates (horse, donkey, cow, pig, sheep, mouse, rabbit, guanaco, dog, hippopotamus, duck, pigeon, chicken, spider monkey, turtle, human, and tuna), 3 insects (screwworm fly, tobacco hornworm moth, and *Samia cynthia* moth), and 2 yeasts (baker's yeast *iso*-1 and *iso*-2, and *Candida krusei*) with the 5287-Å excitation wavelength. These data were obtained by comparing the cytochrome *c* of each species to that of horse. Typical RDS data for the 750-cm^{-1} line of a selection of cytochromes *c* from various species are shown in Figure 3, and a survey of the 900–1700-cm^{-1} region for the protein from several species is shown in Figure 4. The data in these figures were obtained by subtracting the spectrum of each protein from that of the horse protein except for the pigeon spectrum in which the subtraction was reversed.

The Raman frequency differences for each protein relative to the Raman frequencies of the horse protein were calculated from the difference spectra, and the values are shown in Table I. The data for the yeasts are from Shelnutt et al. (1979a). However, the RDS data for all of the other species studied previously have been repeated with improved signal to noise to obtain improved accuracy. The errors are estimated to be $\pm$0.02 cm^{-1} for the strong line at 750 cm^{-1} and $\pm$0.05 cm^{-1} for the other lines. The frequency differences for 12 species not reported earlier are included in Table I. In addition, frequency differences relative to that of the native horse protein are given for the horse protein modified with CDNP at either lysine-13 or lysine-72. Some oxidation was noted with these samples. However, estimates of the possible error could be made from entirely oxidized vs. reduced RDS spectra, and these errors were determined to be too small to affect significantly the difference spectra. We also report frequency differences for the horse and tuna proteins above pH 11. In addition to the data reported in Table I for some samples, RDS data were obtained with other laser frequencies (4579 and 5145 Å) with which other Raman lines were enhanced. These results were consistent with those obtained at 5287-Å excitation. For example, frequency differences of about +0.1 cm^{-1} were detected in the 1365-cm^{-1} line in the 72-CDNP derivative and the cytochromes *c* from pig and hippopotamus. Frequency differences of +0.55 and +0.60 were detected in the 1365- and 1500-cm^{-1} lines in cytochrome *c* of human, consistent with the large differences in the other modes.

In addition to the frequency differences listed in Table I, a line shape difference relative to the horse protein was observed in the 1313-cm^{-1} line for cytochromes *c* from the birds

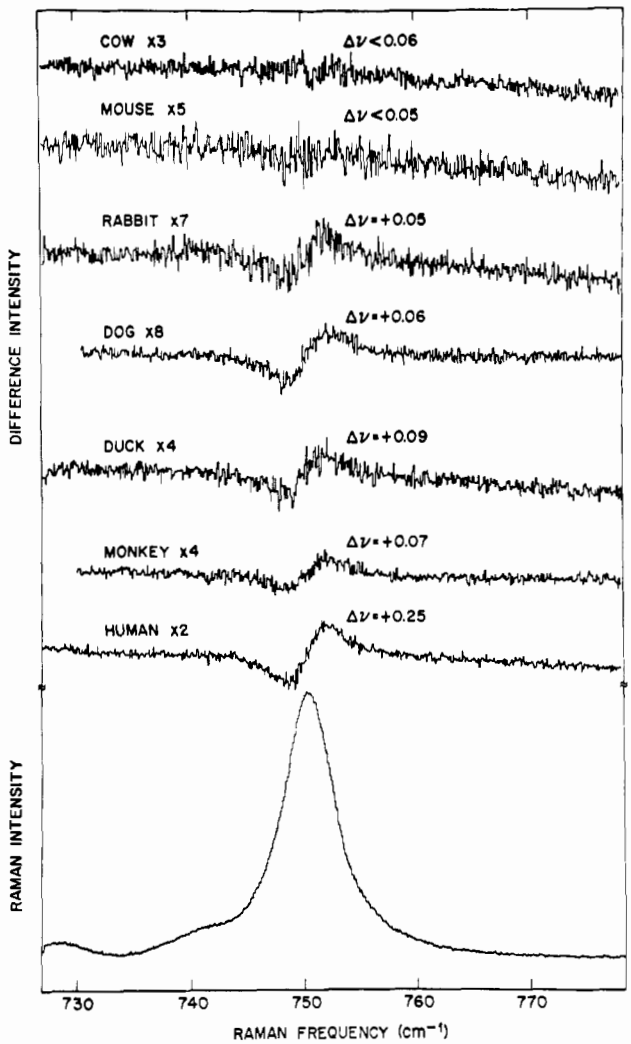

FIGURE 3: Raman difference spectra of the 750-cm^{-1} line of cytochrome *c* for several species. In all the difference spectra the spectrum of the cytochrome *c* from the labeled species is subtracted from that of horse. The multiplicative factor indicates the increase in scale sensitivity of the difference spectra compared to that of the Raman spectrum of horse cytochrome *c* on the bottom. These data were obtained with 5287-Å excitation.

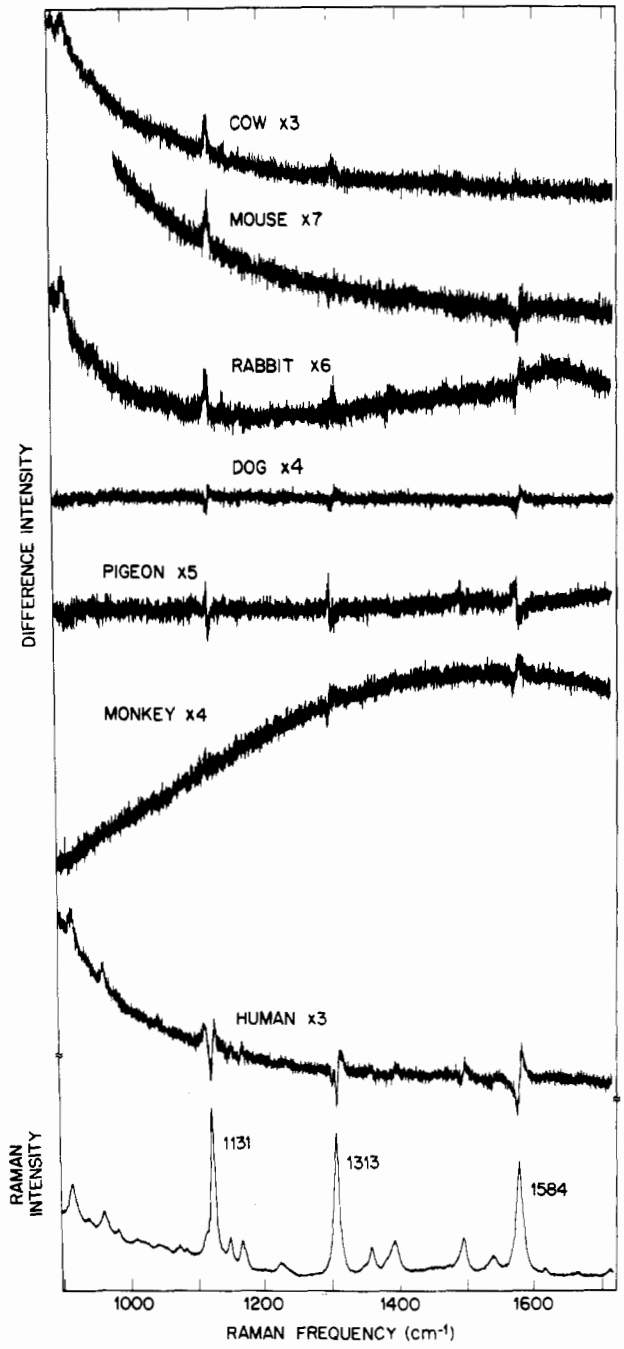

FIGURE 4: Raman difference spectra of the high frequency region (900–1700 cm^{-1}) of cytochromes *c* of several species. With the exception of pigeon, in which the subtraction is reversed, the spectrum of the cytochrome *c* of the labeled species is subtracted from that of horse. The scale factor in the difference spectrum is relative to that of the Raman spectrum of horse cytochrome *c* at the bottom.

and the primates. The difference spectra for these proteins are shown in Figure 5. In this figure we also include the donkey vs. horse difference spectrum of cytochrome *c* in which no difference (<0.07 cm^{-1}) was detected and the hippopotamus vs. horse cytochrome *c* difference spectrum in which a +0.09-cm^{-1} difference was detected. In contrast, the line shape differences between the cytochrome *c* from horse and those from the birds and the primates are very complex. The difference spectra of the cytochromes *c* from these species have negative deflections near 1307 and 1313 cm^{-1} and positive deflections near 1311 and 1317 cm^{-1}. This difference in shape may also be seen in the Raman spectra at the bottom in which the lines from human cytochrome *c* and that of horse are compared.

Discussion

Resonance Raman Scattering from Cytochrome c. Most aspects of the resonance Raman spectrum of cytochrome *c* are now well understood (Felton & Yu, 1978; Rousseau et al., 1979). With visible excitation, the vibrational modes of the porphyrin macrocycle are enhanced, and therefore only these modes are observed in the resonance Raman spectrum. The relative intensities of the modes depend on the excitation frequency. For excitation into the α band, all of the in-plane symmetry allowed vibrational modes may be enhanced. Consequently, a rich spectrum of lines is observed in Figure 2 for the 5502-Å excitation wavelength. If the excitation frequency is in the region of the vibrational side band of the porphyrin macrocycle (the β-band absorption), then those vibrational modes responsible for the optical absorption are the ones that are most enhanced. As the laser is tuned through the β band, the resonance Raman spectrum will continue to change, being first dominated by low-frequency modes when the laser wavelength is just above the zero vibrational level of the state Q_0 to being dominated by high frequency modes when the laser wavelength is on the high-energy side of the β band. This is illustrated by the data in Figure 2 in which

Table I: Raman Frequencies of Horse Ferrocytochrome *c* (in cm^{-1}) and Frequency Differences (Compared to Those of Horse Cytochrome *c*) of Several Other Species[d]

species	Raman frequencies and differences			
horse (pH 7.8)	750	1131	1313	1584
horse (pH 11.6)	<0.03	<0.1	−0.7	<0.1
CDNP-lysine-13 (pH 8.1)	<0.04	<0.1	−0.15	+0.2[b]
CDNP-lysine-13 (pH 5.7)	<0.04	<0.1	<0.2	+0.12
CDNP-lysine-72	+0.03	<0.08	<0.10	+0.17
donkey	<0.02	<0.08	<0.07	<0.08
cow	<0.06	<0.09	<0.10	<0.15
pig	+0.02	<0.07	+0.07[c]	<0.12
sheep	<0.06	<0.01	+0.1[c]	<0.2
mouse	<0.05	<0.04	<0.06	+0.15
rabbit	+0.05	<0.06	<0.08	+0.18
guanaco	+0.07	+0.07	+0.10	+0.20
dog	+0.06	+0.08	+0.07	+0.21
hippopotamus	+0.02	+0.09	+0.09	+0.12
duck	+0.09	+0.06	*a*	+0.24
pigeon	+0.08	+0.11	*a*	+0.21
chicken	+0.03	+0.06	*a*	+0.10
turtle	+0.09	+0.07	+0.17	+0.31
monkey	+0.07	<0.05	*a*	+0.25
human	+0.25	+0.26	*a*	+0.73
tuna	+0.12	+0.18	+0.52	−0.3
screwworm fly	+0.05	+0.26	+0.10	−0.35
tobacco hornworm moth	+0.31	+0.2	+0.29	+0.63
Samia cynthia moth	+0.26	+0.2	+0.5	+0.7
baker's yeast *iso*-1	<0.1	+0.3	+0.9	+0.5
baker's yeast *iso*-2	+0.1	<0.2	+1.0	+0.7
C. krusei	+0.3	+0.6	+1.6	+0.7
tuna (pH 7.8) − tuna (pH 11.2)	−0.03	<0.04	−0.14	<0.08

[a] Frequency differences could not be determined for these species because of accompanying line shape changes. See the text and Figure 5. [b] Some oxidation occurred during data gathering in this sample. [c] The apparent differences in the 1313-cm^{-1} line of the cytochromes *c* from pig and sheep could result from line shape differences. [d] Positive differences indicate that the Raman line of the cytochrome *c* of the indicated species is at a higher frequency than that of horse. When no difference was detected, the limit was conservatively set by analyzing the noise level. All data were obtained with 5287-Å excitation.

with excitation at 5287 Å (700 cm^{-1} to the blue of Q_0) the strongest line in the spectrum is the 750-cm^{-1} line. With the 5145-Å excitation wavelength (1350 cm^{-1} blue of Q_0), the higher frequency modes are most strongly enhanced. This type of excitation frequency dependence is expected if the vibronic coupling is weak. Strong vibronic coupling in some metalloporphyrins gives rise to a more complicated dependence on excitation frequency (Shelnutt & O'Shea, 1978).

The broad fluorescence especially evident in the 5145-Å spectrum results from dephasing processes (Friedman & Rousseau, 1978). Although the Raman lines are very intense with 5145-Å excitation, the fluorescence contributes to the noise level in the RDS spectra. We found, therefore, that we could obtain better signal to noise in the RDS spectrum with excitation at 5287 Å than at 5145 Å. Although the signal was even larger at 5502-Å excitation, difference spectra here are noisier because of the use a dye laser as the source. We therefore obtained most of the RDS data reported here with 5287-Å excitation.

The strong electronic transitions from groups in heme proteins, other than the porphyrin macrocycle, are well into the ultraviolet, so their Raman spectra are not present when the spectrum is obtained with visible frequency excitation. Under these conditions the Raman spectrum contains only the vibrational modes of the heme group. No modes from the amino acids of the surrounding protein appear, and lines from the buffer appear in the spectrum only very weakly.

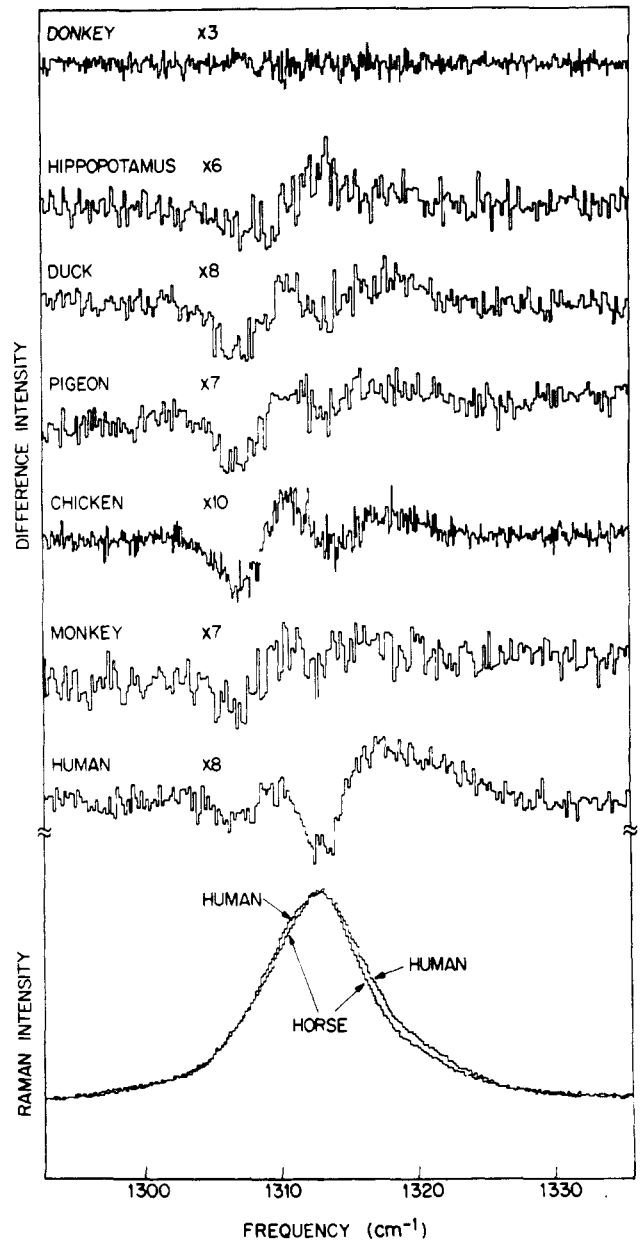

FIGURE 5: Raman difference spectra of the 1313-cm^{-1} Raman line for several species. The spectrum of the cytochrome *c* of the labeled species is subtracted from that of horse, and the scale factors are relative to the Raman spectrum of cytochrome *c* from horse. For cytochrome *c* from donkey no difference was detected. For cytochrome *c* from hippopotamus, the difference is 0.09 cm^{-1}. For the cytochromes *c* from the birds and the primates, a characteristic shape difference was detected. For illustration of the magnitude of this difference, the Raman spectra of the cytochromes *c* of human and horse are compared at the bottom.

Although direct observation of the amino acid residues is not possible with resonance Raman scattering experiments using visible or near-ultraviolet excitation, the influence of the protein may be detected by its perturbations on the porphyrin resonance Raman spectrum. The interaction of the protein with the heme affects the electronic structure and thereby the atomic arrangement of the ground and excited states. These effects are detected in the Raman line shifts and the Raman excitation spectra. Specifically, the intensity of the Raman reemission as a function of excitation frequency (excitation profile) will directly depend on the properties of the excited-state electronic structure which may be varied by the protein (Shelnutt, 1980). Similarly, the polarization properties of the Raman lines are a critical function of the symmetry of the

heme environment. Finally, changes in the ground-state electron density distribution within the porphyrin ring will modify the force constants and the equilibrium positions of the nuclei and hence shift porphyrin ring vibrational frequencies (Shelnutt et al., 1979a,b; Rousseau et al., 1982).

Several normal coordinate analyses (Stein et al., 1975; Kitagawa et al., 1976; Abe et al., 1976, 1978; Sunder & Bernstein, 1976) have been reported for the porphyrin macrocycle of model compounds and have been refined by comparison with vibrational spectroscopic data. It was found that the normal modes are very complex combinations of the symmetry coordinates. Most normal modes have a potential energy distribution with contributions from bonds between the atoms of the pyrrole rings, meso carbon atoms, substituent carbon atoms, and meso hydrogen atoms. Since the results reported in this study deal with small frequency variations between proteins from different species, our concern is with the forces which give rise to variations in the normal mode frequencies rather than their particular potential energy distributions.

The behavior of the frequencies of the Raman modes upon change of oxidation state, degree of back-donation, spin state, and core size has been well studied (Spiro & Strekas, 1972; Spiro & Burke, 1976; Felton & Yu, 1978). Large variations (≈ 20 cm^{-1}) have been found in some lines (the lines at 1365, 1500, 1544, and 1624 cm^{-1} in the ferrous porphyrin) upon changing the iron oxidation state. On the other hand, upon going from the low to high spin state, the low-spin lines at 1500, 1544, 1584, and 1624 cm^{-1} change frequency. When the particular iron orbitals that are changing as the oxidation and spin states change are considered, a qualitative understanding may be obtained for the origin of the behavior of the vibrational modes of the porphyrin ring. From such considerations and from the many investigations of variations of Raman frequencies associated with changes in oxidation states, the 1366-cm^{-1} line is the most reliable indicator of oxidation state since it appears to be nearly independent of spin state. In addition, this line as well as those at 1131 and 1313 cm^{-1} has been observed to change in a systematic way as the degree of back-donation from the iron to the porphyrin is changed by modifying the axial ligands (Spiro & Burke, 1976). Increased back-donation occurs when there is an increase in the π donation to the iron from the axial ligand. Due to mixing between the iron d(π) orbitals and the porphyrin $e_g(\pi^*)$ orbitals, increased π density at the iron results in a concomitant increase in the $e_g(\pi^*)$ orbitals of the porphyrin. Because these orbitals are antibonding, increased electron density within the $e_g(\pi^*)$ orbitals causes most vibrational modes to decrease in frequency. This is consistent with Raman shifts observed for this vibration in the mono- and dianion porphyrins in which electrons are introduced into the π^* orbital electrochemically (Ksenofontov et al., 1976). Moreover, Shelnutt (1981) has detected systematic changes in the Raman frequencies upon complex formation in copper uroporphyrins.

Influence of Near-Heme Residues on the Heme. In this investigation the cytochromes *c* from 24 species were examined (see Table I). We found that even in spectra of cytochromes *c* from closely related species where only a few differences in the amino acid sequence exist (see Table II), there are significant differences in the Raman frequencies. With respect to cytochrome *c* of horse, all of the porphyrin frequencies of the cytochromes *c* of all the species we examined in this work increased in frequency with a few exceptions which will be discussed below. In a preliminary report of this work in which cytochrome c_2 of *Rhodospirillum rubrum* was also examined (Shelnutt et al., 1979a), we classified a series of proteins into three groups: group I in which the frequencies of the modes of the porphyrin were the same as or only slightly different from those of the horse protein; group II in which the frequencies were higher than those of horse cytochrome *c*; group III containing only *R. rubrum*, in which the frequencies were generally lower than those of horse cytochrome *c*. These RDS data were interpreted by considering sequence changes in the near heme environment. None of the cytochromes *c* of the species in group I have changes in heme-contact residues. All of the cytochromes *c* of species in group II have heme-contact residue differences, compared to the horse protein, including a tyrosine at position 46 which in the group I proteins is occupied by phenylalanine. The data are consistent with the concept that changes in the porphyrin vibrational modes were induced by a direct interaction between the porphyrin and the residue at position 46. The general increase in frequency of most of the high-frequency Raman lines was interpreted in terms of changes in the electron density in the antibonding π orbitals of the ring, due to the previously discussed correlation between π electron density and Raman frequencies. We concluded that the proteins of group II, in which the Raman frequencies are generally higher than those in group I, must have lower electron density in the porphyrin $e_g(\pi^*)$ orbitals. This electron density change could arise from charge transfer, van der Waals interactions, or electrostatic interactions between the heme and near-heme residues.

The present data, in which the proteins from many more species are examined and higher sensitivity was obtained, confirm the importance of the near-heme residues. Among the 15 vertebrates (horse, donkey, cow, pig, sheep, mouse, rabbit, guanaco, dog, hippopotamus, duck, pigeon, chicken, turtle, and spider monkey) with no changes in near-heme residues, the frequency differences in the modes sensitive to electron density are all very small (less than 0.2 cm^{-1}), and in the 750-cm^{-1} mode the differences are less than 0.1 cm^{-1}. (The larger differences in the 1584-cm^{-1} line will be discussed below.) In the group of cytochromes *c* including those from human, tuna, the insects, and the yeasts, the differences in the 750-, 1131-, and 1313-cm^{-1} lines are mostly larger. Because the cytochromes *c* from every species of this latter group have at least one heme contact residue different from the horse protein (see Table II), this larger body of data upholds our previous conclusions about the importance of near heme amino acid sequence differences.

In the yeasts, human, and tuna fish proteins there is also an *aromatic* contact residue difference at position 46. In Figure 6 the relationship between tyrosin-46 and the heme for tuna cytochrome *c* is illustrated. In this space-filling model, based on the crystallographic coordinates, it is apparent that tyrosine-46 is in van der Waals contact with the heme. It is therefore not surprising that a change in this heme-contact residue from phenylalanine, which is nonpolar, to tyrosine, which is polar, would influence the heme properties.

In NMR comparisons of cytochromes *c* from tuna and horse, Moore & Williams (1980e) have detected three regions in which there are significant structural differences. Two of these are away from the heme (residues 33 and 57) and the third is in the region of residue 46. They found the phenylalanine to tyrosine change at this position influences the mobility of residue 48 which is hydrogen bonded to one of the propionic acids of the heme. Structural changes inferred from the reactivity of bacterial cytochrome oxidase with a series of cytochromes *c* have also been attributed to the phenylalanine to tyrosine change at position 46 (Yamanaka & Fukumori,

Table II: Amino Acid Sequences of Cytochromes *c* from the Species for Which Raman Data Were Obtained[g]

	Seq. Diffs	-0---------1	GDVEKGKKIF[10]	VQKCAQCHTV[20]	EKGGKHKTGP[30]	NLHGLFGRKT[40]	GQAPGFTYTD[50]
Horse[a]	-						
Donkey[a]	1	----------	----------	----------	----------	----------	------S---
Cow[a] = Pig[a]	3	----------	----------	----------	----------	----------	------S---
Mouse[b]	6	----------	----------	----------	----------	----------	---A--S---
Rabbit[a]	6	----------	----------	----------	----------	----------	---V--S---
Guanaco[a]	5	----------	----------	----------	----------	----------	---V--S---
Dog[a]	6	----------	----------	----------	----------	----------	---S------
Hippopotamus[c]	6	----------	----------	----------	----------	----------	--S---S---
Duck[a]	10	----------	----------	----S-----	----------	----------	---E--S---
Pigeon[a]	11	----------	--I-------	----S-----	----------	----------	---E--S---
Chicken[a]	11	----------	--I-------	----S-----	----------	----------	---E--S---
Turtle[a]	11	----------	----------	----------	----------	--N--I----	---E--S--E
Spider Monkey[d]	12	----------	-------R--	IM--S-----	----------	----------	---S-----E
Human[a]	12	----------	----------	IM--S-----	----------	----------	-----YS--A
Tuna[a]	19	----------	---A----T-	----------	-N-----V--	--W-------	---E-YS---
Screw worm fly[a]	22	------GVPA	----------	--R-------	-A-----V--	----------	---A--A--N
Tobacco horn worm moth[a]	28	------GVPA	-NADN-----	--R-------	-A-----V--	----F-----	------S-SN
Samia Cynthia moth[a]	29	------GVPA	-NA-N-----	--R-------	-A-----V--	----FY----	------S-SN
Baker's Yeast ISO 1[a]	46	-----TEFKA	-SAK--ATL-	KTR-I-----	----P--V--	----I---HS	---E-YS---
Baker's Yeast ISO 2[a]	55	-AKESTGFKP	-SAK--ATL-	KTR-Q----I	-E--PN-V--	----I-S-HS	--VV-YS---
C. Krusei[a]	51	----PAPFFQ	-SAK--ATL-	KTR------I	-A--P--V--	----I-S-HS	---E-YS---
Aromatic Residues	..	------*---	---------*	---------*	----------	--*--*----	-----*-*--
Packed Against Heme[e]	..	----------	--------*-	---*---*--	--------*-	-*--*-----	-----*----
Interior Residues[e]	..	----------	-----*---*	---*---*--	---*----**	-*-***---*	**---*-*--
Covalent to Heme	..	----------	----------	---*--**--	----------	----------	----------
H-bond to Heme[f]	..	----------	----------	----------	----------	----------	I------IO-

	[51]ANKNKGITWK[60]	EETLMEYLEN[70]	PKKYIPGTKM[80]	IFAGIKKKTE[90]	REDLIAYLKK[100]	ATNE
Horse						
Donkey	----------	----------	----------	----------	----------	----
Cow	---------G	----------	----------	--------G-	----------	----
Mouse	---------G	-D--------	----------	--------G-	-A--------	----
Rabbit	---------G	-D--------	----------	--------D-	-A--------	----
Guanaco	---------G	----------	----------	--------G-	-A--------	----
Dog	---------G	----------	----------	-------TG-	-A--------	--K-
Hippopotamus	---------G	----------	----------	--------G-	-A-------Q	----
Duck	---------G	-D--------	----------	--------S-	-A-------D	--AK
Pigeon	---------G	-D--------	----------	--------A-	-A-------Q	--AK
Chicken	---------G	-D--------	----------	--------S-	-V-------D	--SK
Turtle	---------G	----------	----------	--------A-	-A-------D	--SK
Spider Monkey	-------I-G	-D--------	----------	--V-----E-	-A--------	----
Human	-------I-G	-D--------	----------	--V-----E-	-A--------	----
Tuna	---S---V-N	ND--------	----------	--------G-	-Q--V----S	--S*
Screw worm fly	---A-----Q	DD--F-----	----------	----L--PN-	-G-------S	--K*
Horn worm moth	---A-----Q	DD--F-----	----------	V---L--AN-	-A-------Q	--K*
Samia Cynthia moth	---A-----G	DD--F-----	----------	V---L--AN-	-A-------E	S-K*
ISO-1	--IK-NVL-D	-NNMS---T-	-X--------	A-G-L--EKD	-N---T----	-CE*
ISO-2	--I--NVK-D	-DSMS---T-	-X--------	A---L--EKD	-N---T-MT-	-AK*
C. Krusei	---PA-VE-A	-P-MSD----	-X--------	A-G-L--AKD	-N--VT-MLE	-SK*
Aromatic	-*-------*-	------*---	---*------	-*--------	-------*---	----
Packed	----------	---*--**--	*--------*	*---*-----	---**--*--	----
Interior	-*---*--*-	---*--**--	*---*--*-*	---**-----	---**--*--	-*--
Covalent	----------	----------	---------*	----------	----------	-*--
H-bond	-I------I-	----------	-------OO-	----------	----------	----

[a] Dayhoff & Eck (1972) and Borden & Margoliash (1976). [b] Carlson et al. (1977). [c] Thompson et al. (1978). [d] E. Margoliash (unpublished results). [e] Dickerson et al. (1971). [f] Takano et al. (1977). [g] We have adopted the standard one-letter notation for the amino acids (Dayhoff & Eck, 1972).

1978). The yeast and tuna proteins have additional heme contact residue differences not involving aromatic amino acids. The insect cytochromes *c* also have sequence differences in the heme-contact residues. These include a single residue difference involving an aromatic amino acid at position 35 (leucine → phenylalanine), *except* for the screwworm fly

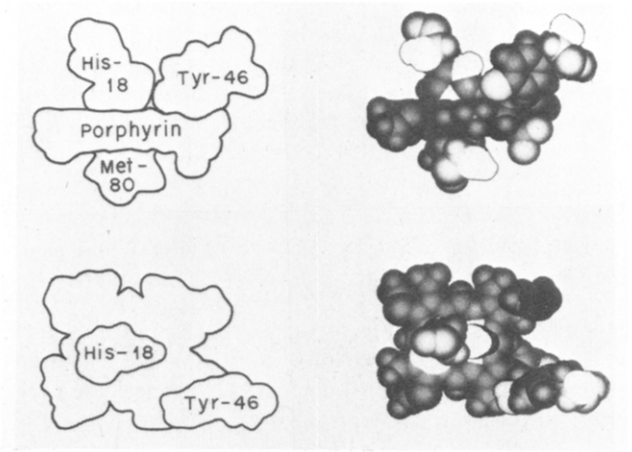

FIGURE 6: Space-filling model computer generated from atomic coordinates illustrating van der Waals contact between the porphyrin macrocycle and tyrosine-46 of cytochrome *c* from tuna. (This was kindly supplied by Richard J. Feldman of National Institutes of Health.) The outline on the left is included to allow rapid identification of the residues.

protein, which has leucine as does the horse protein at this position. The frequency differences for the screwworm fly protein are significantly smaller than those of the other proteins of the insects. This underscores the importance of the interaction between the heme and the aromatic residues in contact with it. However, the frequency differences are larger than those of the proteins of group I, consistent with the (nonaromatic) heme-contact sequence difference at residue 85 (isoleucine → leucine). Although a quantitative assessment of the relative importance of these changes cannot be made at present, the π electron system of the aromatic residue (phenylalanine) would be expected to interact with the heme. In the yeast proteins the frequency differences are substantially larger. However, in these species in which about half of the residues differ from the horse protein and in which there are from five to seven heme-contact residue changes, it is not possible to isolate the effect of a single residue. Speculation on the importance of some of the residues in the yeast proteins was made previously (Shelnutt et al., 1979a).

The results reported here establish the importance of near-heme residues in governing the electronic properties of the heme. Because most lines of a given protein change in the same direction with respect to the horse protein and because the Raman lines that change frequency are established indicators of π electron density, the near-heme residues apparently modify the porphyrin π electron density. It is the oxidation and reduction of the iron atom that are important in cytochrome *c* function. Because of back-donation the π electronic properties of the porphyrin macrocycle might be expected to exert some control on the redox potential and therefore one would expect some dependence of the Raman probe of π electron density on the redox potential. On the other hand, in model compounds the redox potentials are known to depend on the total ($\sigma + \pi$) electron density on the porphyrin ring (Zerner & Gouterman, 1966), so the importance of the contribution of the π electron density difference among the proteins is not obvious. However, the range of redox potentials in these materials is too small to test this assumption meaningful. A study of a homologous series of cytochromes c_2 with a wide range of redox potentials is currently under way.

Influence of Differences in Protein Conformation on the Heme. The line at 1584 cm^{-1} exhibits substantial frequency differences for the cytochromes *c* of many species. The direction of the frequency differences in this line do not correlate with the direction of the changes in other lines in the spectrum. This is not entirely unexpected as in various heme proteins and model compounds the line at 1584 cm^{-1} is known to behave in a unique way (Spaulding et al., 1975; Felton & Yu, 1978; Lanir et al., 1979; Spiro et al., 1979). This line is insensitive to oxidation state but shifts dramatically with change in spin state. Furthermore, in a study of a large number of model compounds, the frequency of this line has been shown to correlate with that of the porphinate core size [heme center-to-nitrogen (pyrrole) distance].

By examining the amino acid sequences of the species given in Table II, one can see that the residue at position 92 determines whether or not a shift is observed in the 1584-cm^{-1} line as follows. Since we see no changes, with respect to cytochrome *c* of horse, in the line in the proteins from donkey, cow (or sheep or pig) changes in this line cannot result from the changes of threonine to serine at position 47, lysine to glycine at position 60, or threonine to glycine at position 89. The presence of a shift in mouse cytochrome *c* therefore implicates the proline to glycine change at position 44, the glutamic acid to aspartic acid change at position 62, or the glutamic acid to alanine change at position 92. It should be noted that the frequency shifts in this line in the mouse, rabbit, guanaco, dog, hippopotamus, duck, pigeon, turtle, and monkey proteins are all in the 0.12–0.25-cm^{-1} range. The cytochromes *c* of all of the species have alanine at position 92, suggesting that it gives rise to the shift in the 1584-cm^{-1} line. The cytochromes *c* of dog and hippopotamus both lack the change at position 44, thereby excluding that position as influencing the 1584-cm^{-1} line significantly. The cytochromes *c* of dog, guanaco, hippopotamus, and turtle all lack the change at position 62. It is apparent then that the presence of alanine at position 92 causes a shift in the 1584-cm^{-1} line when comparing such proteins to those with the glutamic acid at this position.

Additional evidence supporting the influence of position 92 on the frequency of the 1584-cm^{-1} line comes from the data on cytochrome *c* from chicken, tuna, and screwworm fly. The shift in this line in the chicken protein is 0.10 cm^{-1} compared to 0.24 and 0.21 cm^{-1} for that of duck and pigeon, respectively. The sequence of chicken cytochrome *c* differs from that of duck only at positions 3 (isoleucine → valine), 92 (valine → alanine), and 103 (serine → alanine), and the chicken protein differs from that of pigeon only at positions 89 (serine → alanine), 92 (valine → alanine), 100 (aspartic acid → glutamic acid), and 103 (serine → alanine). Since the residues at position 89 and position 100 are the same in cytochrome *c* of duck as in that of chicken, we can rule them out as causes of the difference in the frequency shift in the 1584-cm^{-1} line for the chicken protein compared to those shifts measured for the cytochromes *c* of the other two birds. Similarly, since the residue in the pigeon protein at position 3 is the same as that in chicken cytochrome *c*, we can exclude the influence of the residue change at this position and we are left with position 92 and position 103. However, cytochrome *c* from turtle, which also has a large frequency shift in the 1584-cm^{-1} line, has serine at position 103 as does chicken cytochrome *c*, so we may exclude this position as well. Only position 92 remains as a possibility for giving rise to the observed 1584-cm^{-1} line frequency differences. This interpretation is consistent with the results from tuna cytochrome *c* in which there is another different residue at position 92 (glutamine) and the 1584-cm^{-1} line shifts in the opposite direction than it does in most of the other proteins which we examined. The negative shift of the 1584-cm^{-1} line in screwworm fly cytochrome *c* could be a result

Table III: Conformational Parameters and Frequency Shifts for Proteins with Different Amino Acids at Residue 92

residue at 92	frequency shift	conformational parameters
alanine (dog)	+0.2	*nonpolar*, strong α-helix former, indifferent to β-sheet formation, poor β-turn former
valine (chicken)	+0.1	*nonpolar*, strong β-sheet former, weak α-helix former, strong β-turn breaker
glutamic acid (horse)	0.0	*negatively charged*, strong α-helix former, strong β-sheet breaker, β-turn indifferent
glutamine (tuna)	−0.3	*polar*, weak α-helix former, weak β-sheet former, and weak β-turn former
glycine (screw-worm fly)	−0.3	*polar*, strong β-turn former, strong α-helix breaker, weak β-sheet breaker

of position 92 which is occupied by glycine in this protein. However, in this case we are unable to exclude contributions from other residue changes (for example, the changes at positions 47, 88, and 100) to the behavior of the 1584-cm^{-1} line.

From the above analysis it appears that differences in amino acid sequence at position 92 are at least partially responsible for the frequency differences in the 1584-cm^{-1} Raman line of ctyochrome *c*. These differences apparently do not result from simple charge differences since at other positions, for example, at residue 60, changes in the charge of residues do not correlate with changes in any of the heme Raman frequencies that we have measured. On the other hand, owing to the wide range of properties of the residue at position 92, protein conformational differences would be expected.

There has been much effort devoted to the study of the properties of the amino acids and their sequence in polypeptides in order to determine under what conditions it becomes favorable to form various protein conformations (Sternberg & Thornton, 1978; Chou & Fasman, 1978). Although these analyses are for the most part empirical, it is instructive to compare our frequency differences with the predicted structural properties of the residues that occupy position 92 in the cytochromes *c* of the species we have examined. This comparison is shown in Table III in which the amino acids at position 92 are tabulated on the left, the frequency differences in the 1584-cm^{-1} line with respect to cytochrome *c* from horse are listed in the center, and the properties of the residues are listed at the right. These properties are those reported by Chou & Fasman (1978), who established conformational parameters for forming the α helix, β sheet, and β turns for amino acids by analyzing reported structures of 29 proteins and determining the frequency with which each amino acid appeared in various structural segments.

The residues from positions 88 to 103 are known to have an α-helical structure in crystals of tuna ferrocytochrome *c* (Takano et al., 1977). However, the residue at position 92 in cytochrome *c* of tuna (glutamine) is only a weak α-helical former (Chou & Fasman, 1978). There is about a 0.5-cm^{-1} difference in the 1584-cm^{-1} line between cytochrome *c* of tuna and that of those cytochromes *c* with alanine at position 92, a strong α-helix former. This suggests that the α-helical character in this region is maintained in cytochrome *c* of tuna as well as in those of species with alanine at position 92. By similar considerations of the proteins with other residues at position 92, it is difficult to see any obvious categorization according to conformation differences. However, since the range of residues at this position varies from nonpolar to polar to negatively charged, residue–solvent interactions may cause the entire α-helical section to twist somewhat in order to establish more favorable hydrophobic interactions depending on the properties of the residue 92 and the nearby residues. Such twisting of the entire α-helical section could result in specific tertiary conformational differences in the heme pocket which we have detected in the heme vibrational modes.

What is the mechanism by which residue 92, which is far removed from the heme, can interact with it to change the "core size" marker line? It is possible that a change in the tertiary conformation at the beginning of the long helical section from 88 to 103 due to a change in the attitude of the helix may, in turn, influence the conformation of the "random coil" section from 88 to methionine-80. Two residues in this section of the protein interact directly with the heme. Methionine-80 is the sixth Fe ligand and phenylanine-82 is very near the heme and is directly in the crevice. It is conceivable that a difference in conformation of the peptide chain between residue 92 and methionine-80 might affect the interaction of the methionine sulfur with the iron atom and thereby change the core-size marker line frequency by influencing the interaction between the iron and the porphyrin ring. It is interesting to note that Cookson et al. (1978) have observed a correlation between the shift in the methionine methyl resonances in the NMR spectra of reduced cytochromes *c* and the redox potential. They interpreted this to be a result of changes in the iron–sulfur bond length and the ability of the sulfur to donate electrons to stabilize the iron(III) state. More experiments to determine if the core size sensitive Raman mode or other Raman modes are also sensitive to redox potential.

The interpretation presented here of the origin of the differences in the 1584-cm^{-1} line of cytochrome *c* is consistent with the amino acid sequence analysis. However, we cannot preclude other more complicated origins for the differences we have detected due to cooperative structural changes resulting from concerted effects of several amino acid substitutions. On the other hand, our conclusions are generally consistent with those of Moore & Williams (1980f) based on NMR data, who found that changes in charged amino acids gave rise to small structural rearrangements in many regions of the protein.

By studying monocarboxydinitrophenyl derivatives (CDNP) of the lysine residues (Brautigan et al., 1978b,c), Margoliash and co-workers (Ferguson-Miller et al., 1978; Kang et al., 1978; Speck et al., 1979a,b; König et al., 1980) have been able to identify the binding domain between horse cytochrome *c* and its oxidoreductase partners. They found that for mammalian mitochondrial oxidase and reductase as well as for yeast peroxidase, liver sulfite oxidase, and purified cytochrome c_1, interaction domains are roughly centered on the point at which the dipole axis of cytochrome *c* crosses the surface of the protein near phenylalanine-82. On the basis of these results, they proposed that the distribution of charged and hydrophobic side chains within and around the binding domain and the magnitude and direction of the dipole moment of the protein serve to direct the interaction between cytochrome *c* and its reaction partners and thereby form the productive cytochrome *c*–cytochrome oxidase complex. Therefore the Raman differences in the 1584-cm^{-1} line that we have interpreted as resulting from conformational changes in this region of the binding domain may have functional significance. Any conformational change in that region would be expected to modify the distribution of charged and hydrophobic groups in the binding domain. Control of the conformation in the binding domain by a residue not directly at the binding site could be

a means of fine tuning the interaction region, thereby matching each cytochrome *c* with its oxidoreductases.

In an examination of the spectral features and the physicochemical properties of mono-CDNP horse cytochromes *c*, modified at a variety of different lysyl residues, Brautigan et al. (1978c) concluded that the chemical modification did not alter the overall protein conformation nor the hemochrome structure. In a comparison of the NMR spectrum of these derivatives and native cytochrome *c*, Brautigan et al. (1978c) could detect no differences in the heme resonances. They concluded that the CDNP moiety neither intercalates into the protein structure nor does it occlude the heme edge. We have examined two of the derivatives (CDNP-lysine-13 and CDNP-lysine-72 cytochromes *c*) in which the modification is closest to the heme to determine if there are some subtle perturbations at the heme to which the NMR spectrum was not sensitive. For both derivatives the 1584-cm^{-1} line has a positive shift with respect to the native horse protein by about 0.2 cm^{-1}. In tuna ferrocytochrome *c* lysine-13 forms a salt bridge linkage across the heme crevice to the side chain of glutamic acid-90 near the beginning of the α-helical section from 88 to 103. Therefore, one could account for the difference between CDNP-lysine-13 cytochrome *c* and cytochrome *c* of horse by a rupture of the lysine-13 to glutamic acid-90 salt bridge due to the presence of the negative CDNP group. The reason for the shift with CDNP-lysine-72 is less clear, but lysine-72 is near enough to the section of random coil chain in the binding domain (especially residues 80–85) for the CDNP group to influence its conformation. The influence on the reaction kinetics of the changes at the heme induced by the CDNP modification remains to be explored. Also, the spectra from other CDNP derivatives must be examined to determine if the changes observed in the CDNP-lysine-13 and CDNP-lysine-72 cytochromes *c* are unique to these two derivatives.

Sequence Differences Affecting the Covalent Protein–Heme Linkage. The cytochromes *c* of the birds and the primates all displayed a characteristic difference spectrum with a complex line shape change in the 1313-cm^{-1} line (see Figure 5). It has been noted that this line is very sensitive to the pyrrole ring substituents (Adar, 1975, 1977). Protoheme, with vinyl substituents, has anomalously polarized lines at 1306 and 1342 cm^{-1} whereas saturated mesoheme with the $-CH_2CH_3$ groups instead of the $-CHCH_2$ groups has only one anomalously polarized line at 1313 cm^{-1}. Adar has interpreted this as resulting from a resonance interaction between the vinyl side groups and the deuteroheme macrocycle. The line shape changes of the 1313-cm^{-1} line of the cytochromes *c* are therefore suggestive of possible changes at the cysteine covalent linkage sites. Indeed, examination of the amino acid sequences of the cytochromes *c* from the birds and primates (Table II) reveals that the only common difference between their cytochromes *c* and that of horse is the alanine to serine change at position 15 adjacent to the covalently linked cysteine at 14. It appears that this nonpolar to polar residue change may result in some conformational changes which are reflected in this heme vibrational mode due to coupling through the thioether linkages.

CDNP-lysine-13 at high pH (8.1) and the unmodified horse protein at high pH (>11) exhibit the only instances in which the 1313-cm^{-1} line is at lower frequency than that of the native horse protein. The 1313-cm^{-1} line of tuna cytochrome *c* also shifts to lower frequency, relative to that of the native protein at pH 7, when the pH is raised to 11.6. In ferricytochrome *c* at high pH in both the 13-CDNP derivative (pH 8.1–8.2) and the native horse protein (pH 9) the two heme crevice bonds are thought to open as evidenced by the titration behavior of the 695-nm band (Osheroff et al., 1980). One of these bonds has been proposed to result from a salt bridge between the ϵ-amino group of lysine-13 and the γ-carboxyl group of glutamic acid-90, and the other has been proposed to result from a hydrogen bond between the ϵ-amino group of lysine-79 and the backbone carbonyl of residue 47. Breaking these bonds at pH 9 in ferricytochrome *c*, allowing the heme crevice to open, is associated with the disappearance of the 695-nm absorbance due to the replacement of methionine-80 with some other group as the sixth iron ligand. No changes in optical bands have been detected in ferrocytochrome *c* at the pH ~9. The p*K* for cleavage of the second heme crevice bond in the ferrous form of the lysine-13 derivative has not been determined. Titration behavior by NMR of native cytochromes *c* has been reported though (Moore & Williams, 1980c). It was observed that at pH values >9.5 there were a number of resonances that underwent small shifts in directions consistent with a loosening of the protein structure. The p*K* of the ϵ-amino group of lysine in solution is 10.5 (Lehninger, 1975). It is therefore tempting to speculate that one or both heme crevice bonds may rupture above pH 9.5 without the concerted change in axial ligation that occurs in the ferric state. Lysine-13 is adjacent to the cysteine-14 linkage to the heme. In view of the results from the cytochromes *c* of the birds and the primates, if the heme crevice bond involving lysine-13 is ruptured under our conditions, it is not surprising that we would detect a change in the 1310-cm^{-1} line that is sensitive to the local structure at cysteine-14.

In the normal coordinate analysis of nickel(II) octaethylporphyrin, Abe et al. (1978) assigned the 1308-cm^{-1} line to ν_{21}, an A_{2g} normal mode. Their analysis gives more than a 50% contribution from a bending motion involving the proton on the methine bridge carbon and some α-carbon–β-carbon stretching motion. In a normal coordinate analysis of copper(II) octamethylporphyrin, Sunder & Bernstein (1976) found a more complicated potential energy distribution for the 1306-cm^{-1} line involving substituent carbon stretching and deformation motions. A conformational change due to titration of the lysine-13 to glutamic acid-19 salt bridge or the different residues at position 15 could affect the interaction of the meso hydrogen atoms or some other peripheral substituents with the protein.

Conclusions

The Raman difference technique described here has sufficient sensitivity to detect structural differences in the protein induced by amino acid substitutions when the resulting structural change influences heme vibrational mode frequencies. In the comparisons between eukaryotic cytochromes *c*, many differences have been detected, but the magnitudes of the differences are generally small (typically < 1 cm^{-1}). These data are therefore consistent with other observations that in cytochromes *c* from very diverse species spanning a wide range of eukaryotic taxonomy, the basic "cytochrome fold" is identical. In an effort to organize the large body of data that we have collected, we have isolated three types of protein structural variations which may be detected at the heme: (I) The first is direct interaction between near-heme residues and the porphyrin macrocycle. These Raman differences result from changes in the amino acid sequence in positions that have residues in van der Waals contact with the heme. These near heme residue changes seem to modify or redistribute electron density delocalized in the antibonding orbitals of the porphyrin moiety. (II) Species-specific differences in the amino acid

sequence at residue 92, a position far removed from the heme, bring about a conformation change that propagates to the heme. This effect is manifest in a change in the core size marker line at 1584 cm^{-1} and thereby makes it a structural probe of conformational differences in this region of the protein surface. (III) The conformation near cysteine-14 appears to affect the mode at 1313 cm^{-1} that is sensitive to pyrrole substituents. This effect is especially evident in comparisons between cytochromes *c* with alanine at position 15 and those with serine at this position.

Moore & Williams (1980f) have used NMR to examine cytochromes *c* from horse, donkey, cow, dog, rabbit, chicken, pigeon, and tuna. They found that the chemical shift values of the isoleucine-57 resonances were the most sensitive to structural changes. From the variations in the resonances of this residue, they classified the cytochromes *c* into three groups relative to that of horse: donkey in which there were no changes; cow, dog, rabbit, chicken, and pigeon in which there were small changes; tuna in which there were large changes. The Raman difference data classify the cytochromes *c* into the same groupings, although it is not clear that the change in position 46 believed to be the origin of the large difference in the heme which we detect is also the origin of the difference Moore & Williams (1980f) detect in the isoleucine-57 resonances.

A major biological function of cytochrome *c* is to transfer an electron from cytochrome reductase to cytochrome oxidase. To fulfill this function, the cytochrome *c* must be able to efficiently bind to and possibly release from its redox partners and also have the proper oxidation–reduction potential. Control of the structure of the binding domain and control of the redox potential is achieved by variations in the amino acid sequence which in turn gives rise to changes in the secondary and tertiary structure. The changes which we have detected suggest that the amino acid sequence changes may influence the functional properties in a variety of ways. First, near-heme residues may modulate heme electron densities and thereby control the value of the redox potential. Second, the changes in the core size marker line at 1584 cm^{-1} suggest a change in the iron–methionine bond strength. Protein-induced changes in the properties of the bond may consequently influence the redox properties of the cytochrome *c*. The importance of the properties of this bond was emphasized by Moore & Williams (1980d), who cited the change in the length of the bond upon oxidation of ferrocytochrome *c* as the "trigger" for inducing conformational changes in the ferric protein. Finally, in addition to binding domain changes by amino acid substitutions within the binding domain, data reported here suggest changes in the binding domain induced by amino acid sequence changes outside of the binding domain. This could serve as a fine-tuning mechanism. It is hoped that these experiments will be a starting point for many more investigations into the biological implications of protein–heme interactions. It must be determined if the sequence-dependent changes we have detected are solely reporting structural variations in the protein or if they represent energetically important changes in redox potentials. This question will only be resolved after more experiments are done in which chemical modifications of the amino acid residues near the heme are examined, binding experiments are performed, and a study of heme modes are made for a homologous series of cytochromes with a wide range of redox potentials.

Acknowledgments

We thank J. M. Friedman and M. R. Ondrias of Bell Laboratories for helpful discussions. We thank R. J. Feldman of National Institutes of Health for supplying the computer-generated space-filling model used in Figure 6.

References

Abe, M., Kitagawa, T., & Kyogoku, Y. (1976) *Chem. Lett.*, 249.

Abe, M., Kitagawa, T., & Kyogoku, Y. (1978) *J. Chem. Phys. 69*, 4526.

Adar, F. (1975) *Arch. Biochem. Biophys. 170*, 644.

Adar, F. (1977) *Arch. Biochem. Biophys. 181*, 5.

Borden, D., & Margoliash, E. (1976) in *Handbook of Biochemistry and Molecular Biology* (Fasman, G. D., Ed.) 3rd ed., Vol. 3, pp 268–281, The Chemical Rubber Co., Cleveland, OH.

Brautigan, D. L., Feinberg, B. A., Hoffman, B. M., Margoliash, E., Peisach, J., & Blumberg, W. E. (1977) *J. Biol. Chem. 252*, 574.

Brautigan, D. L., Ferguson-Miller, S., & Margolish, E. (1978a) *Methods Enzymol. 53*, 128.

Brautigan, D. L., Ferguson-Miller, S., & Margoliash, E. (1978b) *J. Biol. Chem. 253*, 130.

Brautigan, D. L., Ferguson-Miller, S., Tarr, G. E., & Margoliash, E. (1978c) *J. Biol. Chem. 253*, 140.

Carlson, S. S., Mross, G. A., Wilson, A. C., Mead, R. T., Wolin, L. D., Bowers, S. F., Foley, N. T., Muijsers, A. O., & Margoliash, E. (1977) *Biochemistry 16*, 1437.

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251.

Christman, R. W., English, J. C., & Tobias, R. S. (1976) *Appl. Spectrosc. 30*, 168.

Cookson, D. J., Moore, G. R., Pitt, R. C., Williams, R. J. P., Campbell, I. D., Ambler, R. P., Bruschi, M., & Gall, J. L. (1978) *Eur. J. Biochem. 83*, 261.

Dayhoff, M. O., & Eck, R. V. (1972) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington, DC.

Dickerson, R. E., & Timkovich, R. (1975) *Enzymes, 3rd Ed. 11*, 397–547.

Dickerson, R. E., Takano, T., Eisenberg, D., Kallai, O. B., Samson, L., Cooper, A., & Margoliash, E. (1971) *J. Biol. Chem. 246*, 1511.

Felton, R. H., & Yu, N.-T. (1978) in *The Porphyrins* (Dolphin, D., Ed.) Vol. 3, Chapter VIII, Academic Press, New York.

Ferguson-Miller, S., Brautigan, D. L., & Margoliash, E. (1978) *J. Biol. Chem. 253*, 149.

Ferguson-Miller, S., Brautigan, D. L., & Margoliash, E. (1979) in *The Porphyrins* (Dolphin, D., Ed.) Vol. 7, pp 149–240, Academic Press, New York.

Friedman, J. M., & Rousseau, D. L. (1978) *Chem. Phys. Lett. 55*, 488.

Friedman, J. M., Rousseau, D. L., & Adar, F. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 2607.

Gardiner, D. J., Girling, R. B., & Hester, R. E. (1975) *J. Chem. Soc., Faraday Trans. 2 71*, 709.

Kang, C. H., Brautigan, D. L., Osheroff, N., & Margoliash, E. (1978) *J. Biol. Chem. 253*, 6502.

Keilin, D. (1930) *Proc. R. Soc. London, Ser. B 106*, 418.

Kiefer, W. (1973) *Appl. Spectrosc. 27*, 253.

Kiefer, W. (1977) *Adv. Infrared Raman Spectrosc. 3*, 1–42.

Kitagawa, T., Abe, M., Kyogoku, Y., Ogoshi, H., Watonabe, E., & Yoshida, Z. (1976) *J. Phys. Chem. 80*, 1181.

König, W. B., Osheroff, N., Wilms, J., Muijsers, A. O., Bekker, H. L., & Margoliash, E. (1980) *FEBS Lett. 111*, 395.

Ksenofontova, N. M., Maslov, V. G., Siderov, A. N., & Bobovich, Y. S. (1976) *Opt. Spectrosc. 40*, 462.
Laane, J., & Kiefer, W. (1980) *J. Chem. Phys. 722*, 5305.
Lanir, A., Yu, N.-T., & Felton, R. H. (1979) *Biochemistry 18*, 1656.
Lehninger, A. L. (1975) *Biochemistry*, p 81, Worth, New York.
Margoliash, E. (1972) *Harvey Lect. 66*, 117.
Margoliash, E., & Walasek, O. F. (1976) *Methods Enzymol. 10*, 339.
Moore, G. R., & Williams, R. J. P. (1977) *FEBS Lett. 79*, 229.
Moore, G. R., & Williams, R. J. P. (1980a) *Eur. J. Biochem. 103*, 493.
Moore, G. R., & Williams, R. J. P. (1980b) *Eur. J. Biochem. 103*, 503.
Moore, G. R., & Williams, R. J. P. (1980c) *Eur. J. Biochem. 103*, 513.
Moore, G. R., & Williams, R. J. P. (1980d) *Eur. J. Biochem. 103*, 523.
Moore, G. R., & Williams, R. J. P. (1980e) *Eur. J. Biochem. 103*, 533.
Moore, G. R., & Williams, R. J. P. (1980f) *Eur. J. Biochem. 103*, 543.
Osheroff, N., Borden, D., Koppenol, W. H., & Margoliash, E. (1980) *J. Biol. Chem. 255*, 1689.
Rakshit, G., & Spiro, T. C. (1974) *Biochemistry 13*, 5317.
Rousseau, D. L. (1981) *J. Raman Spectrosc. 10*, 94.
Rousseau, D. L., Friedman, J. M., & Williams, P. F. (1979) *Top. Curr. Phys. 11*, 203–252.
Rousseau, D. L., Shelnutt, J. A., Ondrias, M. R., Friedman, J. M., Henry, E. R., & Simon, S. R. (1982) in *Proceedings of Symposium on Interaction Between Iron and Proteins in Oxygen and Electron Transport* (Ho, C., Ed.) pp 223–229, Academic Press, New York.
Shelnutt, J. A. (1980) *J. Chem. Phys. 72*, 3948.
Shelnutt, J. A. (1981) *J. Am. Chem. Soc. 103*, 4275.
Shelnutt, J. A., & O'Shea, D. C. (1978) *J. Chem. Phys. 69*, 5361.
Shelnutt, J. A., Rousseau, D. L., Dethmers, J. K., & Margoliash, E. (1979a) *Proc. Natl. Acad. Sci. U.S.A. 76*, 3865.
Shelnutt, J. A., Rousseau, D. L., Friedman, J. M., & Simon, S. R. (1979b) *Proc. Natl. Acad. Sci. U.S.A. 76*, 4409.
Spaulding, L. D., Chang, C. C., Yu, N.-T., & Felton, R. H. (1975) *J. Am. Chem. Soc. 97*, 2517.
Speck, S. H., Ferguson-Miller, S., Osheroff, N., & Margoliash, E. (1979a) *Proc. Natl. Acad. Sci. U.S.A. 76*, 155.
Speck, S. H., Koppenol, W. H., Osheroff, N., Dethmers, J. K., Kang, C. H., Margoliash, E., & Ferguson-Miller, S. (1979b) in *Membrane Bioenergetics* (Lee, C. P., Schatz, J., & Ernster, L., Eds.) pp 31–43, Addison-Wesley Publishing Co., Reading, MA.
Spiro, T. G., & Strekas, T. C. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 2622.
Spiro, T. G., & Burke, J. M. (1976) *J. Am. Chem. Soc. 98*, 5482.
Spiro, T. G., Stong, J. D., & Stein, P. (1979) *J. Am. Chem. Soc. 101*, 2648.
Stein, P., Burke, J. M., & Spiro, T. G. (1975) *J. Am. Chem. Soc. 97*, 2304.
Sternberg, M. J. E., & Thornton, J. M. (1978) *Nature (London) 271*, 15.
Strekas, T. C., & Spiro, T. G. (1974) *Biochim. Biophys. Acta 351*, 237.
Sunder, S., & Bernstein, H. J. (1976) *J. Raman Spectrosc. 5*, 351.
Susi, H., & Ard, J. S. (1977) *Spectrochim. Acta, Part A 33A*, 561.
Takano, T., Trus, B. L., Mandel, N., Mandel, G., Kallai, O. B., Swanson, R., & Dickerson, R. E. (1977) *J. Biol. Chem. 252*, 776.
Thompson, R. B., Borden, D., Tarr, G. E., & Margoliash, E. (1978) *J. Biol. Chem. 253*, 8957.
Yamanaka, T., & Fukumori, Y. (1978) in *Evolution of Protein Molecules* (Matsubara, H., & Yamanska, T., Eds.) pp 387–399, Japan Scientific Societies Press, Tokyo.
Zerner, M., & Gouterman, M. (1966) *Theor. Chim. Acta 4*, 44.