Schachman, H. K. (1957), *Methods Enzymol. 4*, 32.
Scheidegger, J. J. (1955), *Int. Arch. Allerg. 7*, 103.
Solomon, A., and McLaughlin, C. L. (1969), *J. Biol. Chem. 244*, 3393.
Terry, W. D., and Ohms, J. (1970), *Proc. Nat. Acad. Sci. U. S. 66*, 558.
Timasheff, S. N., Susi, H., and Steverns, L. (1967), *J. Biol. Chem. 242*, 5467.
Zimmerman, B., and Grey, H. M. (1972), *Biochemistry 11*, 78.

Derivative Spectroscopy Applied to Tyrosyl Chromophores. Studies on Ribonuclease, Lima Bean Inhibitors, Insulin, and Pancreatic Trypsin Inhibitor†

John F. Brandts* and Lawrence J. Kaplan‡

ABSTRACT: The spectroscopic properties of the tyrosyl chromophores in several tryptophan-free proteins (ribonuclease, lima bean inhibitors, insulin, and pancreatic trypsin inhibitor) have been examined using derivative spectroscopy. These are compared to a model chromophore, *N*-acetyl-L-tyrosine ethyl ester (AcTyrOEt). In the simplest case of a lima bean inhibitor having only a single tyrosyl chromophore, presumably largely exposed to solvent, the derivative bands are considerably sharper and of larger amplitude than would be expected if it resided in a *normal* aqueous environment. In the case of other proteins containing more than one chromophore, the situation is more complicated due to the possibility of spectral heterogeneity resulting from different local environments of the individual chromophores. When this spectral heterogeneity is partially removed for ribonuclease by selective acetylation, it can also be seen that the spectrum for both the three "accessible" and the three "inaccessible" chromophores is again much sharper than anticipated from the corresponding spectra of AcTyrOEt in water or in other liquid hydrogen-bonding solvents. Although covered up to a large extent by broadening from heterogeneity, it is probable that the same situation exists for the other proteins which were examined. Possible explanations for the relatively low degree of environmental broadening of the spectra of protein chromophores are discussed, including the possibility that any "solid-like" character of the protein or the surrounding solvent will minimize configurational fluctuations about native protein chromophores and the possibility that the chromophores are not completely hydrogen bonded. Either of these factors could lead to reduced spectral broadening. Also, the solvent perturbation spectra (25% ethylene glycol) of ribonuclease and its acetylated derivatives have been measured and compared to the equivalent spectrum of AcTyrOEt. It is found that the three accessible chromophores give rise to a perturbation spectrum which is somewhat at variance with the model spectrum. More significantly, there is a moderately large perturbation spectrum for the three inaccessible chromophores which bears little resemblance to the model spectrum and which could arise from factors not explicitly taken into account in the normal analysis of solvent perturbation data.

Ultraviolet spectroscopy is the most commonly used technique for studying protein conformational transitions and for investigating the local environment of the aromatic chromophores of tyrosine and tryptophan residues. In spite of the large dependence on this technique, very little is actually known about the spectroscopic properties of *individual protein chromophores* since most proteins contain a number of chromophores of different type, as well as multiple chromophores of the same type. Even for multiple chromophores of the same type (*e.g.*, the six tyrosyl chromophores in ribonuclease), it is expected that the spectra of the individual chromophores will not be identical due to nonidentity of their local environments within the native protein. If the spectrum of each chromophore within a protein is slightly different (*i.e.*, shifted along the wavelength scale) from the other chromophores of the same type, then the observable or composite spectrum of the protein will be broadened due to this nonsuperposition of the individual spectra which contribute to the composite spectrum. *Spectral heterogeneity* of this type can give rise to protein spectra which do not accurately reflect the spectroscopic properties of the individual chromophores, even when dealing with chromophores of only a single type.

Environmental heterogeneity in native proteins is of interest in itself, but spectroscopic complications arising from this must be overcome before any definitive information on the spectra of the individual chromophores can be obtained. As mentioned above, there is at present no information in the literature on the spectra of single chromophores in native proteins and it has commonly been assumed that their spectroscopic properties can be adequately reproduced from the spectra of suitable model chromophores in hydrogen-bonding solvents such as water, alcohols, and other organic solvents. The purpose of this work is to examine more carefully the $\pi \rightarrow \pi^*$ transition of the phenolic chromophore in both model compounds and proteins in order to reach more definite conclusions concerning spectral heterogeneity in proteins and the effects of local environment on individual chromophore spectra. The large breadth of the absorption curve associated with tyrosyl chromophores in proteins has generally made it

† From the Department of Chemistry, University of Massachusetts, Amherst, Massachusetts 01002. *Received March 6, 1972.* Supported by National Institutes of Health Grant GM11071 and by a Sloan Fellowship to J. B.

‡ Present address: Department of Chemistry, Williams College, Williamstown, Mass.

difficult to obtain any direct information on individual chromophores. Laskowski (1970) has already discussed this problem at some length. For example, the half-width of the 277-nm band of ribonuclease is about 25 nm. The displacement of the spectrum of any one of the six tyrosyl residues from the spectrum of any of the others will probably be 5 nm or less, since this seems to be approximately the maximum shift to be expected from various environmental factors such as burial and exposure, hydrogen-bonding configuration, charge perturbations, etc. When the spectral displacement of individual chromophores is considerably smaller than the composite half-width, the problem of information extraction on individual chromophores is difficult.

The other problem which makes it difficult to obtain useful information from protein spectra is that the spectral contributions from different types of chromophores show considerable overlap. Even in the absence of tryptophan, it is difficult to separate the phenylalanine and cystine contributions from that due to tyrosine in the 250–310-nm region, and this must be done before any detailed analysis of the tyrosyl contribution can be made.

Both of these problems, large bandwidths and spectral overlap of chemically different chromophores, can be partially overcome by working with derivative spectra rather than with conventional spectra. Although this is an artificial way of creating sharp "bands," it nevertheless provides the necessary analytical factors which permit one to draw some conclusions concerning spectra of the individual chromophores which comprise a composite protein spectrum. As will be seen, the bandwidths are small and of the same order of magnitude as the shifts to be expected from environmental heterogeneity, so that spectral broadening in first-derivative spectra of proteins should be much more apparent. For proteins containing tyrosyl and no tryptophyl residues, the spectral contamination from phenylalanine and cystine can also be handled more conveniently in derivative spectra than in conventional spectra.

This paper presents the results of a study of model chromophore spectra in various liquid solvents, using the derivative method. Following this, an attempt is made to correlate the model chromophore data with data on four proteins which contain tyrosine and no tryptophan.

Experimental Section

Materials. The ribonuclease (Worthington RASE) was used without further purification. The amount of high molecular weight material in this preparation was found to be only about 1% by chromatography on Sephadex G-50. Lima bean inhibitor (Worthington) was further purified on DEAE-Sephadex (A25) as described in the text. Pancreatic trypsin inhibitor (Worthington), ribonuclease S (Sigma), and insulin (Worthington) were used without further purification, although insulin solutions were chromatographed on Sephadex G-25 prior to use to remove any heavy metal ions. The acetylated derivatives of ribonuclease were made according to the procedure of Riordan *et al.* (1965).

The acetyl-L-tyrosine ethyl ester (Mann chromatographically homogeneous) and *p*-cresol (Baker Grade) were used without further purification. Where possible, the organic solvents were spectroscopic grade. In other cases (formamide, ethyl acetate, *n*-heptane) the solvents were the best grade available and nearly transparent above 260 nm.

Methods. In the early stages of this work, the derivative spectra were measured directly in the first-derivative mode of the Perkin-Elmer Model 356 double beam–double wavelength spectrophotometer, using a $\Delta\lambda$ setting of 0.5–1.0 nm and a spectral bandwidth of about 0.2 mm. It was discovered however that the "first-derivative spectra" obtained directly in this way showed significant deviations from the hand-calculated spectra obtained from the normal absorption curve of the same instrument operating in the conventional double beam mode, even when the bandwidths and $\Delta\lambda$ values were the same. It was found that the derivative curves calculated from Cary 14 tracings were experimentally indistinguishable from the hand-calculated curves obtained on the Perkin-Elmer 356, and all spectra included here were obtained by this indirect method on one of these two instruments.

The spectra were all recorded at the slowest scan rate of the instruments, and it was ascertained that there was no significant pen lag. The chart speed was adjusted so that the vertical markings on the chart paper coincided with the desired $\Delta\lambda$ intervals to be used for calculating the "derivative," *i.e.*, $\Delta A/\Delta\lambda$, where A is the absorbance. The derivative spectra of AcTyrOEt[1] were examined as a function of the wavelength interval $\Delta\lambda$. The spectra changed very little as $\Delta\lambda$ was reduced below 1 nm, so that a close approximation to the true derivative spectrum is obtained at $\Delta\lambda = 1$ nm and this value or a smaller value of $\Delta\lambda$ was used for all spectra. Also, derivative spectra were examined as a function of the slit width of both instruments since distortion from large spectral bandwidths might be important. Such distortions were found to be very small for slit widths below 0.2 mm on the Cary instrument and 0.1 mm on the Perkin-Elmer instrument and spectra were taken with slit widths at or below these values. In the case of *p*-cresol in *n*-heptane where vibrational bands were narrow, lower slit widths were used.

Due to the low concentrations of AcTyrOEt in the solutions studied, the direct determination of concentrations was not practical. To determine concentrations and thereby extinction coefficients, a stock solution of concentrated AcTyrOEt (*ca.* 0.1 M) in methanol was prepared. A volumetric aliquot of this stock solution was then diluted immediately (a slow change in the spectrum of this AcTyrOEt solution occurred over long periods of time) into 200 parts of the desired solvent, repeating this for the various solvents, and the spectrum then recorded. The concentration in each solution could then be determined from the known dilution and from the concentration of AcTyrOEt in the stock solution. The concentration in the stock solution was determined by dry weight. A constant weight of the residual ester was attained after the methanol was allowed to evaporate off for several hours at 40°. No further change in weight occurred over long periods at either 40 or at 50°. It was assumed that the extinction coefficients for the AcTyrOEt solutions containing 0.5% methanol were the same as the extinction coefficients in the pure solvents in the absence of any methanol.

The extinction coefficients for the *p*-cresol solutions were determined less accurately. A small quantity of *p*-cresol was weighed out and diluted with a large amount of solvent. The concentration was calculated assuming that the *p*-cresol contained no moisture.

The concentrations of protein solutions were determined spectrophotometrically, using extinction coefficients determined in this laboratory by dry weight analysis in most cases. The extinction coefficients (molar units) were 9800 at 277 nm for RNase (mol wt 13,680), 6000 at 277 nm for Ac_3RNase (mol wt 13,700), 3050 at 270 nm for Ac_6RNase (mol wt

[1] Abbreviations used are: AcTyrOEt, *N*-acetyl-L-tyrosine ethyl ester; LBI, lima bean inhibitors; PTI, pancreatic trypsin inhibitor.

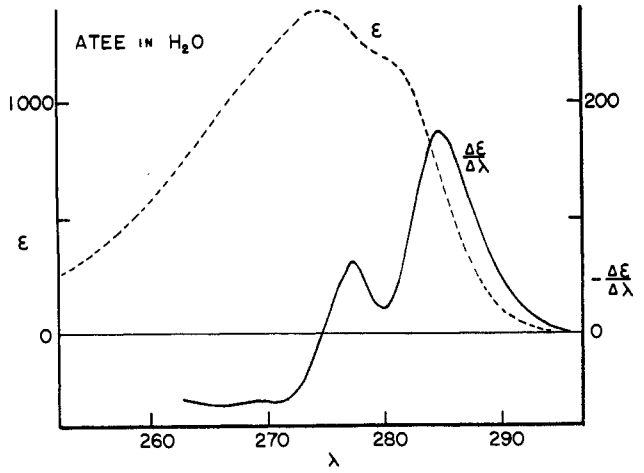

FIGURE 1: Absorption spectrum of acetyl-L-tyrosine ethyl ester (ATEE) in water. The dashed curve and the left-hand ordinate show the normal absorption curve while the solid curve and right-hand ordinate are for the derivative spectrum. The derivative spectrum was calculated from the normal spectrum as described in the Experimental Section. The continuous curve for the derivative spectrum was drawn from calculated points spaced at 1-nm intervals. Extinction coefficients are for a 1 M solution in a 1-cm cell.

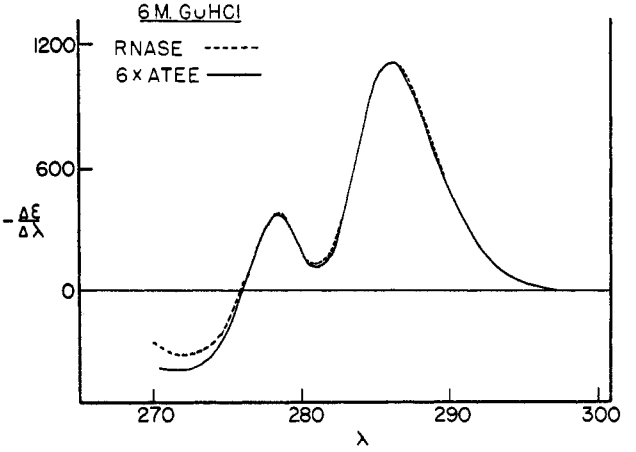

FIGURE 2: Comparison of the derivative spectrum of acetyl-L-tyrosine ethyl ester (ATEE) with the corresponding spectrum of reduced ribonuclease in 6 M guanidine hydrochloride at pH 2.5. The spectrum of acetyl-L-tyrosine ethyl ester has been multiplied by six and shifted to the red by 0.7 nm. The estimated contribution from the cystine residues has been subtracted from the experimental ribonuclease spectrum in order to obtain the spectrum shown here. See text for details.

13,860), 4150 at 276 nm for LBI component VI (mol wt 9900), 3180 at 276 nm for LBI component VII (mol wt 9500), 5430 at 280 nm for pancreatic trypsin inhibitor (mol wt 6510), and 6100 at 277 nm for insulin (mol wt 5800).

Solvent perturbation spectra were obtained in much the same way, using tandem cells. Solutions were prepared by adding small volumes of a concentrated stock solution of the protein or model compound to identical volumes of the different solvents to be examined, using a microburet. All samples were run at least in duplicate with very high internal consistency in all cases.

Results and Discussion

Model Chromophores. In trying to reproduce the spectra of protein chromophores by the use of model systems, there are two factors to consider. (1) A model chromophore must be chosen to approximate as closely as possible the spectroscopic properties of the protein chromophore of interest. (2) The local environment of the protein chromophore must be adequately approximated by some appropriate "solvent" system. Discrepancies between model spectra and protein spectra could result from a breakdown of either of these approximations, although it is invariably the second which causes the greatest problems.

THE MODEL CHROMOPHORE. The model chromophore which has been used for most of these studies is *N*-acetyl-L-tyrosine ethyl ester (AcTyrOEt). The normal absorption curve of AcTyrOEt in water (slightly acid) is shown in Figure 1. The maximum extinction coefficient of 1390 is attained at a wavelength of 274.6 nm. Also shown in Figure 1 is the first-derivative spectrum under the same conditions, where negative values of $\Delta\epsilon/\Delta\lambda$ are plotted along the abscissa. The prominent features of the latter include a major negative band (I) at 284.7 and a minor negative band (II) at 277.2 nm. There is also a positive band at wavelengths below the crossover point at 274.6 but this will not be of much importance in the present study due to the fact that the inherent broadness makes it relatively unsuitable for analytical purposes, and also because in most proteins this wavelength region below 275 nm contains relatively large contributions from phenylalanine absorption.

The major band for AcTyrOEt in water is characterized by a peak height $[-(\Delta\epsilon/\Delta\lambda)_{max}]$ of 175 and a half-width $(\Delta\lambda_{1/2})$ of 6.2 nm. Due to its extreme narrowness, it is the minor band which will be shown to be the most sensitive indicator of spectral heterogeneity in proteins. Since the concept of a bandwidth is rather vague in this case, the degree of resolution of the minor peak will be measured by the ratio, R, of heights of the two peaks, *i.e.*, $R = h_I/h_{II}$, where the heights h_I and h_{II} are measured from the extremum separating the two bands. In this case, R has a value of 4.7. Factors which produce band sharpening will tend to decrease the value of R, while broadening from spectral heterogeneity in proteins will lead to large values of R as the minor band becomes less well resolved.

In order to demonstrate the suitability of AcTyrOEt as a model for tyrosyl chromophores in proteins, the two must be compared under conditions where they are known to have identical local environments. The best way to achieve this situation is to compare the two under conditions where the protein is "completely" unfolded, so that all chromophores in the protein are equally exposed to solvent and heterogeneity is absent. The dashed curve in Figure 2 shows the derivative spectrum of fully reduced ribonuclease in 6 M guanidine hydrochloride. The solid curve is the spectrum of AcTyrOEt in the same solvent, after multiplication by six to adjust for the number of tyrosyl residues in ribonuclease. In order to obtain the best superposition of the two spectra, it is necessary to first shift the AcTyrOEt spectrum to the red by a very small amount, 0.7 nm. With this correction, the two spectra are experimentally indistinguishable above 275 nm where the phenylalanyl chromophores of RNase make no significant contribution. Small contributions to the ribonuclease spectrum from the four disulfide bonds have been subtracted out, as described later.

We conclude from this that AcTyrOEt is a very good model for the protein chromophores *under conditions where the model chromophore and the protein chromophore are in similar environments.* The small 0.7-nm shift necessary to achieve superposition could result from a difference in inductive effects in

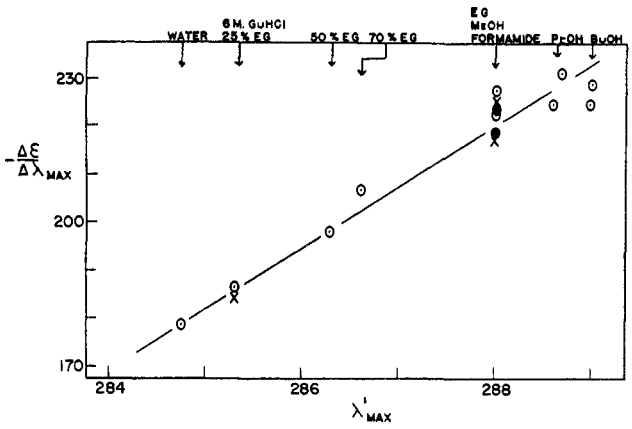

FIGURE 3: The maximum amplitude of the major peak in the derivative spectrum of AcTyrOEt in various solvents, plotted as a function of the wavelength where the maximum amplitude occurs.

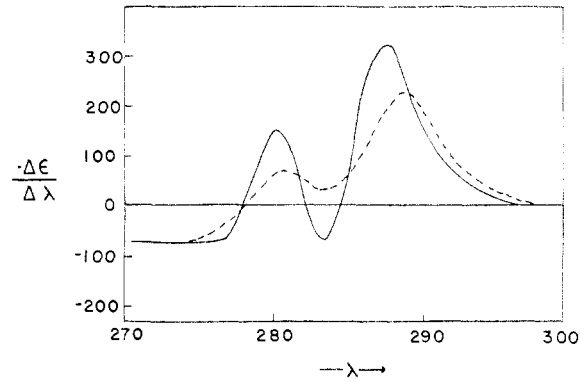

FIGURE 4: The derivative spectrum of AcTyrOET in ethyl acetate. The derivative spectrum of AcTyrOET in isopropyl alcohol is shown also as the dashed curve.

the model, as opposed to a long polypeptide chain, or from a small residual shielding of the chromophores from solvent by the polypeptide backbone. Inductive effects are known to cause relatively large spectral displacements for $\pi \rightarrow \pi^*$ transitions of phenol-like chromophores. For example, the spectrum of *p*-cresol is very similar in shape to that of AcTyrOEt but displaced to the blue by 2.2 nm. In some of the studies to be described below, it will be necessary to use *p*-cresol rather than AcTyrOEt as the model chromophore because of low solubility of the latter.

EFFECTS OF LOCAL ENVIRONMENT. The overall shape of the spectrum of AcTyrOEt in other solvents which possess the potential for both donating and accepting protons in hydrogen-bond formation with the chromophore is very much the same as shown in Figure 1 for water. The most prominent difference occurs in positioning of the spectra along the wavelength scale. As the aqueous character of the solvent is reduced, there is a moderately strong red shift and this is accompanied by an increase in extinction coefficient (ϵ_{max}) at the peak position. These two effects are illustrated in the data of Table I, where the value of ϵ_{max} is compared with the wavelength at which the maximum occurs (λ_{max}) for AcTyrOEt in water, in aqueous mixtures (6 M guanidine hydrochloride, 25, 50, and 70% ethylene glycol), and in pure organic solvents (methanol, 1-propanol, 2-propanol, 1-butanol, formamide, and ethylene glycol). A nearly linear relationship exists between ϵ_{max} and λ_{max} for all of these diverse solvent systems, encompassing a change of about 25% in the value of ϵ_{max}.

The situation is very similar when comparing the derivative spectra of AcTyrOEt in these same solvent systems. Looking first at the major band, there is a linear relationship between the value of the first derivative at the peak position, $-(\Delta\epsilon/\Delta\lambda)_{max}$, and the wavelength where this peak occurs, λ_{max}', in the various solvents. This is shown graphically in Figure 3 (see also Table I). The total change in peak height over the range of solvents studied is about 30%.

There are several generalizations that are consistent with the data in Table I. Considering only the pure solvent systems (water, methanol, propanol, butanol, ethylene glycol, and formamide), the positional parameters λ_{max} and λ_{max}' are relatively constant (±0.5 nm) among the diverse organic solvents, but in all of the organic solvents they are very different (*ca.* 3.0–3.5 nm) from the water value. The differences between water and hydrogen-bonding organic solvents are not nearly so great in terms of the parameters which measure the degree of resolution or sharpness of the peaks in the derivative spectrum of AcTyrOEt. The half-width of the major band in water (6.2 nm) is only slightly greater than the average value in the five organic solvents (5.75 nm). The value of *R* in water (4.7) is at the low end of the range (4.7–5.9) found for these organic solvents, indicating that the minor peak in the derivative spectrum is at least as well resolved in water. *Thus, the major differences between the spectra of AcTyrOEt in water and in these organic solvents are in terms of wavelength positioning and intensities and not band shapes.*

The constancy of the various parameters among the organic solvents discussed above is very good considering the fact that these solvents encompass large differences in dipole moment, "hydrophobicity," density, and viscosity, and that they include solvents with both oxygen and nitrogen as the hydrogen bonding atom. This is perhaps partly fortuitous, but does tend to suggest that rather similar spectra will occur for AcTyrOEt in most organic solvents capable of both donating and accepting protons in hydrogen bond formation with the –OH moiety of the phenolic chromophore.

In organic solvents where complete hydrogen bonding is not possible, the observed derivative spectrum of AcTyrOEt is markedly different. Shown in Figure 4 is the derivative spectrum in ethyl acetate, a solvent which can accept but not donate protons for hydrogen-bond formation with the chro-

TABLE I: Characteristic Spectroscopic Parameters for Acetyl-L-tyrosine Ethyl Ester in Various Solvent Systems at 25°.

Solvent	ϵ_{max}	λ_{max}	$-(\Delta\epsilon/\Delta\lambda)_{max}$	λ_{max}'	$\Delta\lambda_{1/2}$	*R*
Water	1390	274.6	175	284.8	6.2	4.5
6 M guanidine-HCl	1450	275.3	185	285.3	6.1	4.0
25% ethylene glycol	1450	275.1	185	285.3	6.2	4.2
50% ethylene glycol	1540	275.8	198	286.3	6.0	4.6
70% ethylene glycol	1585	276.5	207	286.6	6.1	4.8
Ethylene glycol	1708	277.4	226	288.0	5.8	5.9
Methanol	1665	277.4	228	288.0	5.6	4.7
Formamide	1675	277.6	225	288.0	6.0	4.7
1-Propanol	1710	277.8	225	288.6	5.8	5.2
2-Propanol	1720	277.7	233	288.6	5.6	5.3
1-Butanol	1725	278.0	230	288.9	5.6	5.0
Ethyl acetate	1815	277.9	310	287.5	4.5	1.7

mophore. The spectrum in propanol is also shown for comparison. There are large differences in peak amplitudes in these two solvents (see Table I) and the half-width of the major peak is only 4.5 nm in ethyl acetate, well below the value found in other organic solvents and in water. Another obvious difference is the high degree of resolution of the minor peak in ethyl acetate (R = 1.75) and the fact that the extremum between the two peaks occurs considerably below the zero axis. The maxima in the normal absorption curves in the two solvents, however, occur at almost identical wavelengths of 277.8 nm where the derivative curves intercept the zero axis in Figure 4. Due to the sharpness of the spectrum in ethyl acetate, however, the two derivative bands occur more to the blue than the corresponding bands in propanol.

The differences noted above are accentuated when the chromophore is examined in *n*-heptane where no hydrogen bonding to solvent will occur. Due to the low solubility of AcTyrOEt in this solvent, the spectra shown in Figure 5 are for *p*-cresol. In pure *n*-heptane, the derivative bands are extremely narrow and of very large amplitude. In addition, the vibrational fine structure is very prominent below 280 nm while it is totally undetectable in the other organic solvents examined, including ethyl acetate.

Chignell and Gratzer (1968) have examined spectral shifts of *p*-cresol as various amounts of hydrogen-bonding solvents are added to a solution of the chromophore in a nonpolar solvent (isooctane). When the hydrogen-bonding solvent is isopropyl alcohol, for example, relatively large shifts occur from 0 to about 10% isopropyl alcohol followed by much smaller shifts from 10 to 100% isopropyl alcohol. Their analysis of the data indicates that hydrogen bonding of the -OH moiety of the chromophore is nearly complete at 10% propanol. Shown in Figure 5 is the derivative spectrum of *p*-cresol in 10% isopropyl alcohol–90% *n*-heptane. This spectrum in 10% isopropyl alcohol is extremely similar to the spectrum in 100% isopropyl alcohol, shown also in Figure 5, and very different from that in 100% *n*-heptane. It seems likely, as concluded by Chignell and Gratzer, that the -OH end of the chromophore is preferentially interacting with the alcohol in the 10% solution while the aromatic ring is probably still solvated predominantly by the major heptane component. If this is true, the data in Figure 5 suggest that *the shape of the p-cresol spectrum is controlled predominantly by the interactions of the OH end of the chromophore and that the ring portion plays only a minor role.*

One final experimental observation is worth mentioning. All of the above spectra were taken on very dilute solutions where chromophore–chromophore interactions are of little or no importance. On the other hand, the effective concentration of chromophores for a tyrosyl-rich protein such as ribonuclease is 0.6 M and it is known from X-ray work that some chromophores are in close proximity in the crystalline enzyme. It is conceivable that such spatial factors might markedly alter the spectra of the chromophores involved. To investigate this possibility, the derivative spectrum of *p*-cresol in isopropyl alcohol was examined at concentrations from 0.0005 up to 4 M, using pathlengths from a centimeter down to several microns. No detectable change in shape or positioning of the spectrum occurred up to concentrations of 1 M. Between 1 and 4 M, the spectrum shifted blue by about 1 nm and a slight broadening occurred. On the basis of these observations, it seems probable that *large* spectral distortions due to close proximity of chromophores in proteins will not generally occur. It will also be shown in the following section that one of the unusual aspects of protein chromophores is their

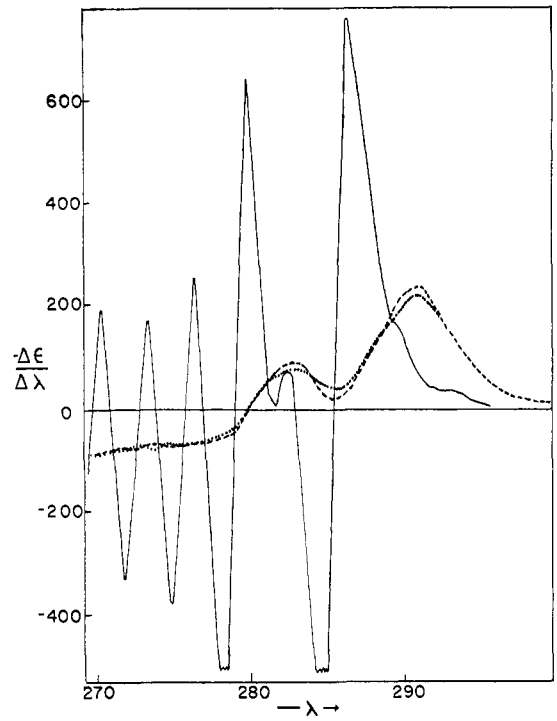

FIGURE 5: The derivative spectrum of *p*-cresol in *n*-heptane. Also shown are the spectrum of *p*-cresol in 10% isopropyl alcohol–90% *n*-heptane (dashed curve) and the spectrum of *p*-cresol in 100% isopropyl alcohol (dotted curve).

extremely *sharp* spectrum, so that the proximity effects noted above appear to be in the wrong direction to account for this.

Proteins. CORRECTIONS FOR CYSTINE RESIDUES. Since the proteins chosen contain no tryptophan, the high wavelength transition of the phenolic chromophores will be contaminated only by phenylalanyl and cystyl chromophores. Rather than attempting to subtract out the phenylalanyl contribution, it will simply be avoided by restricting attention to wavelengths greater than *ca.* 275 nm where it does not interfere appreciably. However, corrections for the disulfide absorption will be made since its spectrum, with a maximum at about 245 nm for L-cystine in water, is very broad and extends throughout the region where the tyrosyl chromophores absorb. Due to both the low oscillator strength and the broadness, however, the correction for disulfide absorption in the derivative spectrum of most proteins is relatively minor. The derivative spectrum of L-cystine in water is shown in Figure 6. The extremum in the derivative spectrum occurs at about 270 nm and the high-wavelength tail extends even beyond 320 nm. In the region between 280 and 290 nm, where tryosyl chromophores exhibit a major peak in the derivative spectrum with an amplitude of *ca.* 200, the value of the first derivative for L-cystine is only about 5 and relatively independent of wavelength. Under circumstances where the content of tyrosine and cystine are nearly equivalent in a protein, the disulfide corrections will then be very small relative to the tyrosyl contributions so that we do not have to worry too much about whether disulfide bonds are buried, exposed, adjacent to charges, etc., or precisely what the dihedral angle might be. An adequate correction can usually be made without considering these secondary effects and the procedure here will be to use the spectrum shown in Figure 6 as being adequately representative of all cystine residues in all proteins. In particularly unfavorable cases where the cystine:tyrosine ratio is very high, as for lima

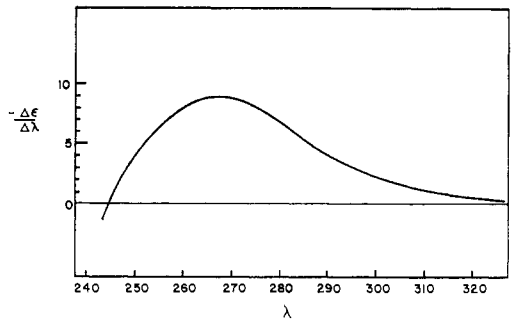

FIGURE 6: The derivative spectrum for L-cystine in water.

bean inhibitor, this method of correction may lead to significant errors.

LIMA BEAN INHIBITOR, A MODEL PROTEIN. The best way to obtain information on the spectrum of a single tyrosyl chromophore in a native protein is to examine a protein which has only one chromophore and thereby avoid the difficult problems associated with separating contributions from multiple-chromophore proteins. Lima bean inhibitor (LBI) is a good choice in this respect since some variants of this protein contain but a single tyrosyl and no tryptophyl chromophores. Life is somewhat complicated however by the extremely high content of disulfide bonds (6–8) in these small proteins (mol wt 8,500–10,000).

The chromatogram of commercial LBI used in this study is shown in Figure 7. Seven well-defined peaks were observed upon elution from the DEAE-Sephadex column. Although the trypsin activity of the various fractions was not checked, all exhibited ultraviolet (uv) absorption spectra indicating a high cystine:tyrosine ratio and the absence of tryptophan. Previous studies have shown four (Jones *et al.*, 1963) or five (Ikeda *et al.*, 1968) well-resolved components on DEAE-cellulose and six components on DEAE-Sephadex (Haynes and Feeney, 1967). In all cases, the second to last component off of the column was the only component shown by amino acid analysis to contain more than one tyrosine per mole; it has two along with 16 half-cystines and a mol wt of 9900. Component VI of Figure 7 is clearly identifiable as this component since the A_{275}/A_{250} ratio is much larger than for the other components. This then also permits component VII to

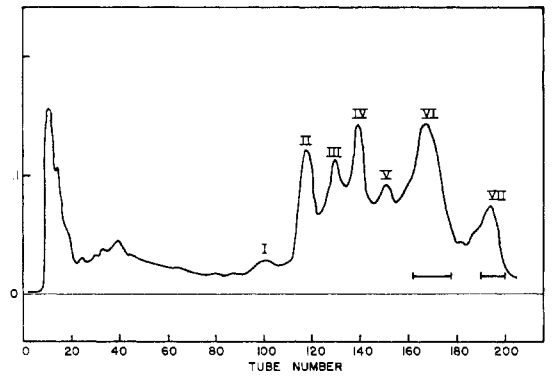

FIGURE 7: Elution pattern of lima bean inhibitor on DEAE-Sephadex at room temperature. The column bed was 2.5 × 40 cm and the flow rate was 30–35 ml/hr with a fraction size of 10 ml. Starting buffer was 0.05 M Tris–0.1 M NaCl, pH 8.07, in a 1-l. mixing chamber. A logarithmic gradient was used with an eluting buffer 0.05 M Tris–0.33 M NaCl, pH 8.07. The fractions which were pooled to obtain components VI and VII are indicated.

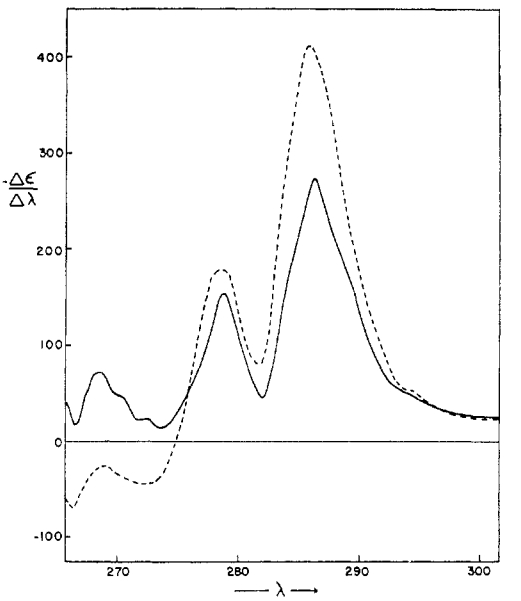

FIGURE 8: The derivative spectra of lima bean inhibitor components VI (dashed curve) and VII (solid curve), 25°.

be identified with component 4 of Jones *et al.* and component 6 of Haynes and Feeney. It has a single tyrosine, 14 half-cystines, and a mol wt of about 9500.

The derivative spectra of components VI and VII are shown in Figure 8. The high disulfide content is evident in the long-wavelength tail which extends even beyond 320 nm. The higher tyrosine content of component VI is also apparent from the increased magnitude of the peak near 286 nm. After correction for the disulfide contributions to the derivative spectra, the curves shown in Figure 8 are obtained. The wavelength-corrected spectrum for AcTyrOEt in water is also included for comparison. The fact that the crossover point for the LBI chromophore is very close to that expected for an aqueous environment suggests that this chromophore is well exposed to solvent. Nevertheless, it is also clear that the spectrum for the single tyrosyl chromophore in LBI VII differs substantially from that of the model chromophore in several important respects. The amplitude of the major peak is about 30% greater and the half-width is smaller for the LBI chromophore (5.1 nm) than for AcTyrOEt (6.2 nm). Even more significantly, the ratio R of the heights of the major peak to the minor peak for LBI (2.2) is only about half as large as for the model chromophore (4.5), which clearly illustrates the much greater resolution of the minor peak in the protein spectrum.

Since component VI contains two tyrosines per mole, it is expected that the corrected derivative spectrum will be broadened by heterogeneity if the two chromophores are not spectrally equivalent. This seems to be the case as can be seen from its spectrum in Figure 9, which has been divided by two to normalize for the tyrosine content. The band amplitudes are considerably smaller and the resolution of the small band considerably poorer than for component VII which contains only one tyrosine. Actually, the spectrum of component VI is very similar in shape to the model AcTyrOEt spectrum but we feel this is accidental and results from the fact that broadening from heterogeneity acts to cover up the inherently sharper spectra of the individual chromophores in component VI, which individually are probably quite similar to the single chromophore in VII.

Inspection of the derivative spectra for the LBI components

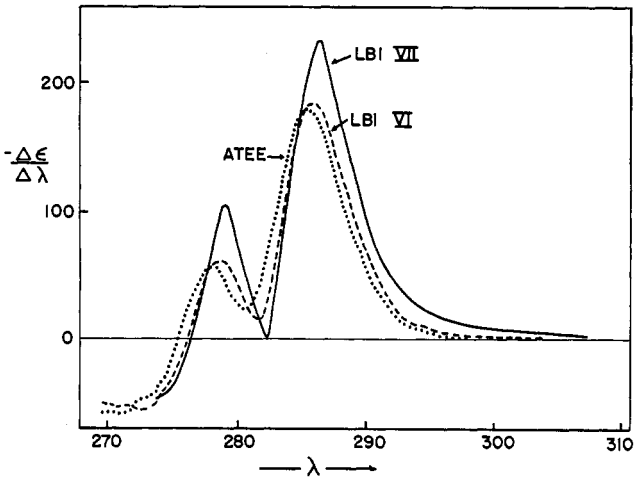

FIGURE 9: The derivative spectra of lima bean inhibitor components VI (dashed curve) and VII (solid curve), after having subtracted the estimated contributions from the disulfide chromophores. The corrected spectrum of component VI has been divided by two to normalize with respect to tyrosine content. The wavelength-corrected spectrum of acetyl-L-tyrosine ethyl ester (ATEE) in water (dotted curve) is shown for comparison.

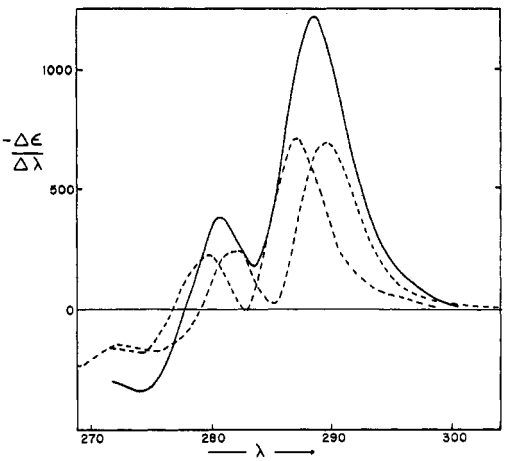

FIGURE 10: Solid curve, the derivative spectrum for native ribonuclease after corrections for disulfide contributions, 25°. Dashed curves, the corrected derivative spectra for the three accessible tyrosyl chromophores in native ribonuclease (blue-shifted) and for the three inaccessible tyrosyl chromophores.

shows that the corrections for the disulfide contribution leave something to be desired, particularly for component VII which has the higher cystine: tyrosine ratio. There remains a long-wavelength tail which is not seen for AcTyrOEt nor for other proteins with more normal disulfide content, so that some contamination is undoubtedly present even in the corrected spectra. Due to the inherent broadness of disulfide bands, the removal of this contamination, were it possible, would probably act to accentuate the differences between the single LBI chromophore of component VII and AcTyrOEt in water.

RIBONUCLEASE, A MORE COMPLEX PROTEIN. Since no additional proteins with but a single tyrosyl residue are available to us, the generality of the conclusion regarding LBI must be established on more complex proteins. Ribonuclease has turned out to be a reasonably good choice, in spite of the fact that it contains a large number of tyrosyl chromophores.

The derivative spectrum of native RNase is shown in Figure 10, after very small corrections for the four cystines. This spectrum is quite different from the spectrum of LBI component VII shown in Figure 9, since the derivative bands for RNase are significantly broader and of lower amplitude (per chromophore). In fact, the RNase spectrum is very similar in shape to the spectrum of AcTyrOEt in liquid solvents capable of hydrogen bonding. For example, the half-width of the major peak is 5.9 nm which is intermediate between the value found for AcTyrOEt in water (6.3 nm) and in *n*-butyl alcohol (5.6). Likewise, the value of R (5.1) and the amplitude of the major peak (200 per chromophore) are intermediate with respect to AcTyrOEt in water and in various organic solvents. The crossover point in the derivative spectrum of RNase (277.5 nm) is 2.2 nm to the red of the water reference point and suggests a moderately high degree of burial.

These data are then consistent with the idea that the composite spectrum of ribonuclease is the sum of six identical spectra of spectroscopically equivalent tyrosine residues, *if we are willing to accept the idea that each chromophore absorbs similarly to AcTyrOEt in hydrogen-bonding solvents.* There is, however, good reason to question the validity of this interpretation. Much evidence in the literature suggests that three of the chromophores are of very different reactivity than the other three. Also, the results on LBI suggest that protein chromophores are really not equivalent to model chromophores in hydrogen-bonding solvents. Consequently, alternative interpretations seem more attractive and it can be demonstrated in a fairly convincing way that the above picture is incorrect.[2] It is possible, using *N*-acetylimidizole, to chemically modify the three reactive phenolic groups without affecting the three inaccessible chromophores (Riordan *et al.*, 1965). The three reactive chromophores have been identified as 73, 76, and 115 (Burstein and Patchornik, 1972). Comparison of the CD spectra suggests that Ac_3RNase does not differ substantially in conformation from the native protein (Riordan *et al.*, 1965), so that the three unmodified chromophores in Ac_3RNase should absorb almost identically with the same three chromophores in native RNase. Fortunately, the absorption of the acetylated tyrosyl chromophores contributes very little to the Ac_3RNase spectrum in the wavelength region above 275 nm so that the derivative spectrum of the *unmodified chromophores* can be obtained relatively accurately from the spectrum of Ac_3RNase (after disulfide corrections) in the following way

$$\text{spectrum of three unmodified chromophores} = \text{spectrum of } Ac_3\text{RNase} - 1/2(\text{spectrum of } Ac_6\text{RNase}) \quad (1)$$

The second term on the right assumes that the spectrum of the three acetylated tyrosines in Ac_3RNase is equivalent to one-half of the spectrum of Ac_6RNase (obtained by complete acetylation in 8 M urea followed by removal of the urea and acetylimidizole). It is likely that this assumption is not strictly valid. The circular dichroic (CD) spectrum of Ac_6RNase in the peptide region is quite different from either the native protein or Ac_3RNase (Riordan *et al.*, 1965) and it is probably partially unfolded. Nevertheless, the second term in eq 1 is

[2] At the time an abstract of this work was submitted (Brandts and Kaplan, 1971), it was felt by us that the six tyrosines in RNase were nearly spectrally equivalent. This view was later corrected in the actual presentation of the work at the 62nd Annual Meeting of the Federation of American Societies for Experimental Biology, but the erroneous conclusion in the abstract has caused some confusion which should be corrected.

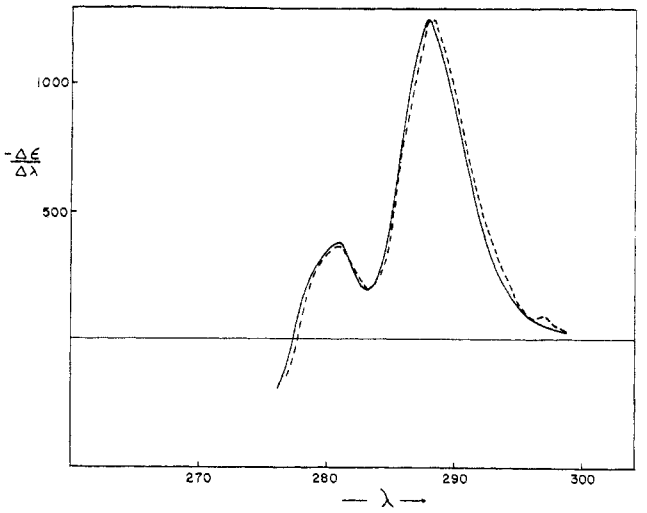

FIGURE 11: The corrected derivative spectra for ribonuclease (dashed curve) and ribonuclease S (solid curve), 25°.

small compared to the first (at its maximum, it is only 8.5% of the first term) so that relatively large errors can be tolerated.

After having obtained the spectrum of the three unreactive chromophores as described above, the spectrum of the three reactive chromophores (in the *native, nonacetylated* state) may be obtained by subtracting the spectrum of the three unreactive groups from the spectrum of native RNase. The resulting spectrum for the three reactive chromophores as well as the spectrum for the three unreactive chromophores are shown in Figure 10. In terms of band shapes and amplitudes, the two spectra are very similar. The R values are identical (3.0), although the extremum between the two peaks occurs slightly below the zero axis in the spectrum of the reactive groups. The amplitude of the major peak is slightly larger for the reactive chromophores (236 *vs.* 231 per chromophore) and the bandwidth is slightly smaller (5.0 *vs.* 5.2). Since for the model chromophore AcTyrOEt a red shift is always accompanied by an *increase* in amplitude and a small *decrease* in half-width (Table I), these results suggest that the residual spectral heterogeneity may be greater among the three unreactive chromophores.

For each of the two groups of chromophores in ribonuclease, the derivative spectrum is similar in shape and normalized amplitudes to that of the single tyrosyl chromophore in LBI VII and very different from the spectra of AcTyrOEt in water and in hydrogen-bonding organic solvents. Even for the three reactive chromophores, generally regarded as being highly exposed to solvent, the amplitude of the major peak is 35% greater, the half-width of the major peak is over 20% smaller, and the value of R is 35% smaller than the values which would be predicted from the spectrum of AcTyrOEt in water. This shows that the spectrum of these three chromophores is considerably less broadened than would be the case if they were solvated normally and completely by water. Even this comparison underestimates the actual differences, since there will still exist some spectral broadening and amplitude reduction from heterogeneity among these three chromophores. Much the same picture holds true for the spectrum of the unreactive chromophores, but to a somewhat lesser extent.

From ultraviolet spectra and CD spectra in water–glycerol mixtures at 77°K, Horwitz *et al.* (1970) (see also Horwitz and Strickland, 1971) have concluded that native ribonuclease contains three "types" of tyrososyl chromophores, as judged by the location of the center of their 0–0 electronic transition obtained by curve fitting. Three residues have this characteristic wavelength at 283.5 nm, two residues at 286 nm, and a single residue at 288.5 nm in a 1:1 ethylene glycol–water mixture at the low temperature. From an analysis of the CD spectrum at room temperature in water, they conclude that the characteristic wavelengths are nearly the same and they have tentatively assigned tyrosine-73, -76, and -115 to the group absorbing at lowest wavelength, tyrosines-92 and -97 to the intermediate group, with tyrosine-25 being the chromophore absorbing 2.5 nm further to the red. In this study, we find only two "types" of chromophores, the grouping being tyrosines-73, -76, and -115 in the low-wavelength triad and -25, -92, and -97 in the high-wavelength triad. As mentioned above, however, the band shapes and amplitudes do imply that the chromophores in the high-wavelength group may be more dissimilar than those in the low-wavelength group and in this sense perhaps our results can be construed as being in qualitative agreement with the above picture.

Horwitz and Strickland (1971) have also examined RNase S at both low temperature and room temperature. From the comparative results with RNase, they have suggested that the spectrum of tyrosine-25 shifts to the blue by 2.5 nm upon action of subtilisin so that RNase S contains only two "types" of chromophores. If a shift of this magnitude does occur, it should be readily apparent in the derivative spectra of RNase and RNase S at room temperature. These spectra are shown in Figure 11 for solutions at pH 6.0 in 0.01 M acetate buffer. Small differences are apparent since the spectrum of RNase S appears to be shifted further to the blue by *ca.* 0.2 nm relative to the spectrum of the native enzyme. Although this is in the proper direction to be consistent with the picture of Horwitz and Strickland (1971), it is difficult to rationalize our results with a wavelength shift of 2.5 nm for a single chromophore. The difference in the derivative spectra of RNase and RNase S is much more consistent with a blue shift of 1 nm or less for a single chromophore (or, of course, a smaller shift for two or more chromophores), and seems inconsistent with the larger shift suggested by Horwitz and Strickland.

At pH 3.4, native RNase is thermally unfolded at 65°. The derivative spectrum under these conditions is shown in Figure 12. This spectrum is considerably broader than that of native RNase at room temperature, also shown in Figure 12. Substantial broadening is to be expected simply due to the higher temperature. However, careful comparison of the protein spectrum to AcTyrOEt spectra obtained at 65° in water and several organic solvents suggests that all of the broadening cannot be accounted for by temperature alone. The crossover point in the derivative spectrum of thermally denatured RNase (275.8 nm) occurs slightly to the red of the water reference point (275.6 at 65°) and the maximum amplitude (156 per chromophore) is slightly smaller than for AcTyrOEt in water (160) at the same temperature. More significantly, the larger half-width for the major band (7.3 nm for RNase *vs.* 7.0 nm for AcTyrOEt in water) and the considerably larger value of R (12.5 *vs.* 7.2) both suggest that thermally denatured RNase does not contain six identical and fully solvated chromophores, in contrast to what was found for reduced RNase in 6 M guanidine hydrochloride. This suggests that the thermally denatured form contains at least one chromophore which is rather different from the others.

INSULIN AND PANCREATIC TRYPSIN INHIBITOR. These two proteins have similar molecular weights and both contain 4 mol of tyrosine and three disulfide bonds. The corrected derivative

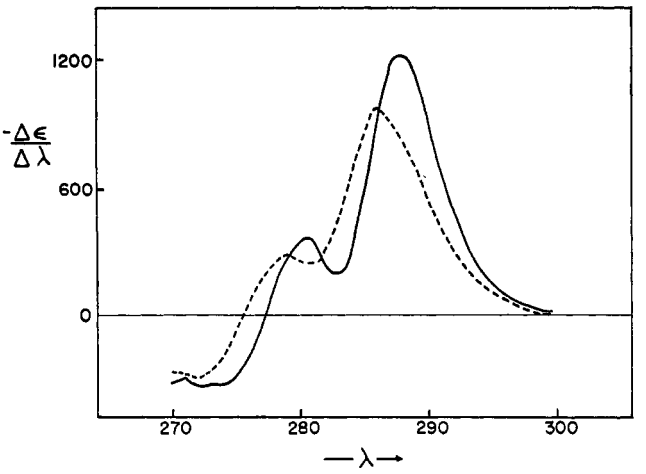

FIGURE 12: The corrected derivative spectra for thermally denatured ribonuclease (65°, pH 3.4, dashed curve) and native ribonuclease (25°, pH 5.5, solid curve).

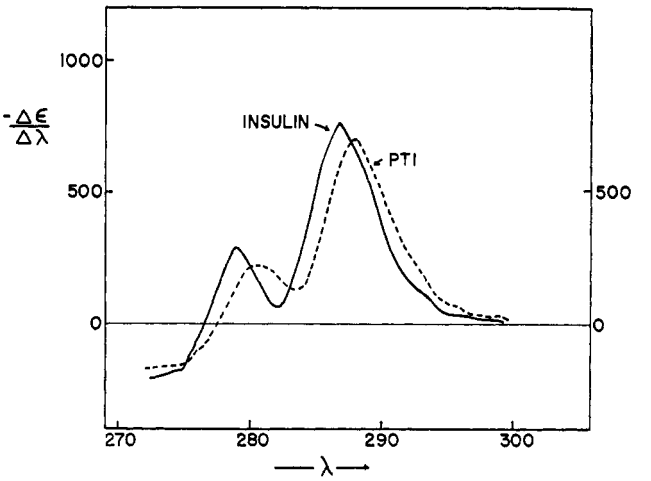

FIGURE 13: The corrected derivative spectra for insulin (solid curve) and pancreatic trypsin inhibitor (dashed curve), 25°, pH 7.0.

spectra for both insulin (Zn free) and pancreatic trypsin inhibitor are shown in Figure 13. In spite of the similarity in size and tyrosine content, the derivative spectra for the two proteins are quite different. In addition to the fact that the PTI spectrum is further to the red by more than 1 nm, it is also considerably broader and less well resolved than the insulin spectrum. This is apparent in the larger half-width for the major peak (5.8 *vs.* 5.5), the smaller amplitude (175 per chromophore *vs.* 190 for insulin), and particularly in the *R* value (6.1) which is almost twice as large as the value for insulin (3.2). It was shown from studies on AcTyrOEt that the spectrum of a single chromophore should show increases in amplitude and decreases in half-width as it is shifted to the red by changes in environment, so that the trends observed here run counter to the intrinsic effects expected for individual chromophores. The greater broadness of the PTI spectrum is undoubtedly due to the fact that the tyrosyl chromophores in this protein are more different from one another, with regard to wavelength positioning, than are the chromophores in insulin, *i.e.*, PTI has a greater degree of spectral heterogeneity.

SOLVENT PERTURBATION STUDIES ON RIBONUCLEASE. Solvent perturbation spectra are closely related to derivative spectra since a perturbation which gives only a pure shift in the absorption curve will give rise to a difference spectrum which, for small wavelength shifts, will have the same shape as the derivative spectrum in the absence of the perturbant. Shown in Figure 14A is the solvent perturbation spectrum of AcTyrOEt (multiplied by six so it can be directly compared to RNase) produced by the addition of 25% ethylene glycol (solid curve). This can be compared with the corresponding derivative curve in water (dashed curve). Small differences in shape can be distinguished, since the solvent perturbation spectrum has broader peaks, less separation between peaks, and is of considerably lower amplitude below the wavelength (270 nm) where it intercepts the zero axis, relative to the amplitude above that wavelength. Because of these differences, other factors besides spectral shift must contribute to some extent to the solvent perturbation spectrum and these factors would include changes in oscillator strength and changes in the broadness of the absorption curve produced by the addition of perturbant. Nevertheless, it appears from the comparative spectra in Figure 14A that the perturbation spectrum of AcTyrOEt produced by 25% ethylene glycol does result *primarily* from shift effects.

The solid curve in Figure 14B shows the perturbation spectrum of native ribonuclease, again using 25% ethylene glycol. This is also compared to the derivative spectrum in the absence of perturbant. Several differences are apparent, relative to the situation for AcTyrOEt discussed above. To begin with, the largest maximum in the solvent perturbation spectrum is of considerably smaller amplitude relative to the corresponding peak in the derivative spectrum. This can be easily seen in Figures 14A and 14B since the perturbation spectra and derivative spectra for ribonuclease and for six AcTyrOEt chromophores are plotted on the same scales. A second difference is that the peak positions in the perturbation spectrum of RNase are blue shifted relative to the corresponding peaks in the RNase derivative spectrum. This immediately suggests that the six tyrosyl chromophores are different and that those which absorb more to the blue are perturbed to the greatest extent by ethylene glycol. This is qualitatively consistent with the usual assumption (Herskovits and Laskowski, 1962) regarding the selectivity of the solvent perturbation method for exposed groups, and with our previous conclusion regarding spectral heterogeneity in native ribonuclease. Thirdly, there are large differences in the shape of the solvent perturbation spectrum

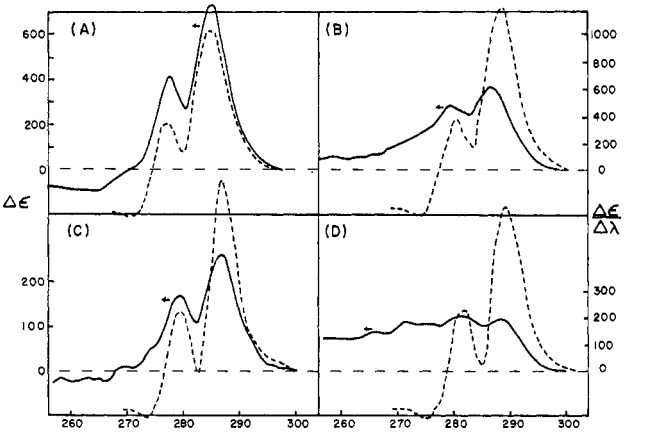

FIGURE 14: The derivative spectra (dashed curves) and the 25% ethylene glycol solvent perturbation spectra (solid curves) for acetyl-L-tyrosine ethyl ester (A), native ribonuclease at pH 5.5 (B), the three accessible tyrosines in native ribonuclease (C), and the three inaccessible tyrosines in native ribonuclease (D).

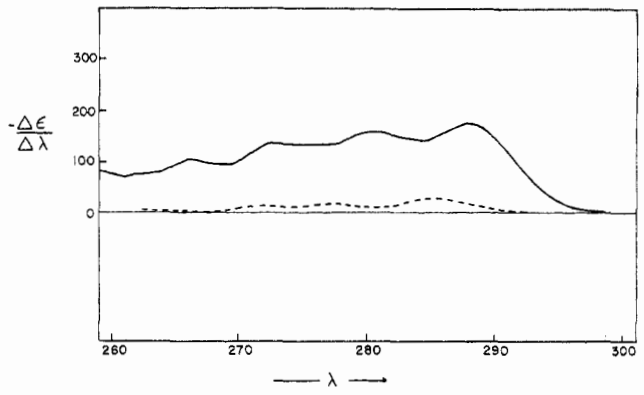

FIGURE 15: Solvent perturbation spectra for Ac_3RNase, pH 5.5 (solid curve), and for *N,O*-diacetyltyrosine (dashed curve), using 25% ethylene glycol.

of ribonuclease, as compared to the "expected" shape based on the AcTyrOEt perturbation spectrum. The two peaks at 279 and 287 nm are less well resolved than the corresponding peaks in the AcTyrOEt perturbation spectrum. Of perhaps greater significance is the fact that the perturbation spectrum for native RNase is positive at all wavelengths shown in Figure 14B while the corresponding spectrum for AcTyrOEt is negative below about 270 nm. The difference between the RNase and model compound spectra in the 260–270-nm region amounts to about 150 in the extinction coefficient. It does not seem possible that this difference is due to perturbation of the four disulfide chromophores by the ethylene glycol since the maximum contribution expected, even assuming 100% exposure, would be only about 20 as estimated from the perturbation spectrum of cystine in 20% ethylene glycol (Donovan, 1969). Likewise, the expected contributions from phenylalanine are small and highly oscillatory in nature (Donovan, 1969) and could not reasonably be expected to account for a significant portion of the large differences between the perturbation spectrum of ribonuclease and AcTyrOEt in the low-wavelength region.

The most reasonable assumption regarding the difference in shape of the protein and model perturbation spectra is that it arises because the contributions of the tyrosyl chromophores in native RNase are not exactly what is expected from straightforward analogy to AcTyrOEt. It is important to ascertain whether the anomalous contribution arises from perturbation of the most accessible chromophores or from perturbation of the most inaccessible chromophores in the native protein, or perhaps from both. To pursue this point, we have measured the solvent perturbation spectrum of the Ac_3RNase derivative where the three accessible chromophores have been acetylated so that their contribution to the total perturbation spectrum will be considerably reduced. The perturbation spectrum (25% ethylene glycol) for Ac_3RNase is shown in Figure 15 and it is seen that this spectrum is in fact considerably smaller in amplitude than that for the unmodified protein in the 275–290-nm region where the large peaks occur in the spectrum of unmodified protein. However, the perturbation spectrum of the acetylated derivative is actually larger in amplitude at wavelengths below 270 nm.

Although the contributions from the three accessible chromophores can be expected to be considerably smaller in Ac_3-RNase than in native RNase, it is still necessary to attempt to subtract their contribution from the perturbation spectrum of Ac_3RNase in order to arrive at some estimate of the absolute contribution from the three inaccessible chromophores. There is no highly accurate way of doing this, so we have used what must be regarded as an approximate method which, due to the small absolute contribution of an acetylated chromophore, will probably be suitable at a semiquantitative level. The solvent perturbation spectrum of *N,O*-diacetyltyrosine is shown in Figure 15, again using 25% ethylene glycol. This might be regarded as a first approximation to the expected contribution from a fully exposed acetylated tyrosyl chromophore in a protein. Since the acetyl group is likely to protrude further into the solvent than the OH group for a structurally equivalent chromophore near the protein surface, we will assume that the three acetylated chromophores in Ac_3RNase have a relatively high average exposure of 70%. Also, it is known that the three accessible chromophores in the native protein absorb *ca.* 2 nm further to the red than does the model chromophore AcTyrOEt in water. It will be assumed that this is also true for the three acetylated chromophores in Ac_3RNase relative to the diacetyltyrosine chromophore, so that the solvent perturbation spectrum of the model chromophore will be shifted by this amount before subtracting it (*i.e.*, before subtracting 70% of it for each of three chromophores) from the perturbation spectrum of Ac_3RNase.

The resulting solvent perturbation spectrum for the three inaccessible tyrosines of RNase is shown in Figure 14D, along with the corresponding derivative spectrum discussed earlier. By subtracting this spectrum from that of the native protein, the solvent perturbation spectrum of the three accessible chromophores may be obtained, and this is shown in Figure 14C.

The perturbation spectrum of the three accessible chromophores shows a much closer resemblance in shape to the corresponding perturbation spectrum of AcTyrOEt (Figure 14A) than does that of native RNase. Nevertheless, the peaks in the perturbation spectrum are still slightly broader *in spite of the fact that the derivative spectrum is much sharper* than for the model chromophore. There is, however, a good correspondance in the location of the peaks in the perturbation spectrum and the derivative spectrum for the accessible chromophores, which would be expected if the method of separation of contributions that we have used is valid.

The direct comparison of the maximum amplitude in the perturbation spectrum of the accessible chromophores with the corresponding amplitude for three AcTyrOEt chromophores in the usual way leads to an estimate of 70% exposure. However, this does not take cognizance of the fact that the derivative spectrum of the protein chromophores is nearly 35% larger in maximum amplitude than for the model chromophores. Since the perturbation spectrum is largely determined by spectral shift and therefore should be nearly proportional to the magnitude of the derivative spectrum (for a given shift), this would reduce the estimate to about 52% for these three chromophores.

It is not possible to estimate the degree of exposure of the three unreactive chromophores from their solvent perturbation spectrum (Figure 14D) or even to decide if they are exposed to any significant extent at all. There are two small extrema in the perturbation spectrum at 281.5 and 288.5 nm which correspond closely with the extrema in the derivative spectrum suggesting a small contribution to the perturbation spectrum from spectral shifts brought about by the addition of perturbant. However, the overall shape of this perturbation spectrum is very different from that of AcTyrOEt and it seems pointless to obtain an estimate of the degree of exposure by comparing two spectra of very dissimilar shape at some se-

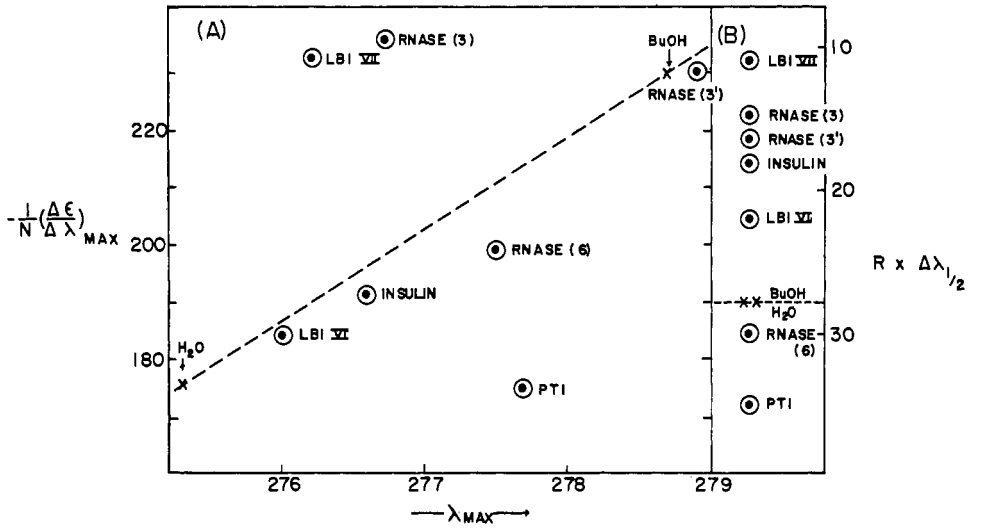

FIGURE 16: Spectroscopic parameters for the native proteins examined in this study. See text for details.

lected wavelength. Because of the large difference in shape, it seems entirely possible that the origins of the two spectra are different, *i.e.*, perhaps the solvent perturbation spectrum of the three inaccessible chromophores does not arise from *direct* perturbation of the protein chromophores by the added perturbant.

These results do tend to confirm the basic assumption inherent in the use of the solvent perturbation technique, *i.e.*, accessible chromophores will be selectively perturbed. However, they also pose some problems which caution against placing too much reliance on the precise estimates of exposure which are derived from the method. These problems include (1) the lack of correspondence in the derivative spectra of individual protein chromophores and model chromophores in water, and (2) the presence of a broad perturbation spectrum for the chemically inaccessible chromophores, which bears little resemblance to the perturbation spectra of model chromophores. This latter perturbation spectrum contributes about one-third of the total amplitude for native RNase at the wavelength where estimates of exposure are generally made.

It has frequently been observed in other solvent perturbation studies that the protein perturbation spectrum is broader and more poorly defined than the corresponding model perturbation spectrum, so that the results discussed above for ribonuclease appear to have some generality. This in the past sometimes been attributed to spectral heterogeneity; *i.e.*, the perturbation spectra for the individual protein chromophores do not superimpose and this leads to the observed broadening. This explanation cannot account for the present results. Effects from spectral heterogeneity are reflected in the composite derivative spectrum. Even if all chromophores are affected to the same extent by the addition of perturbant, then the derivative spectrum and perturbation spectrum for the protein should bear the same relationship to one another as do the equivalent spectra for the model chromophore. If the solvent perturbation is at all selective toward those groups that are exposed, as it is expected to be, this will have the effect of sharpening the protein perturbation spectrum relative to the derivative spectrum. However, exactly the opposite effect is seen for native ribonuclease. The only explanation for this is if the perturbation spectrum of at least some of the individual chromophores is substantially broader than the model chromophore perturbation spectrum, this occurring in spite of the fact that the derivative spectra of the individual protein chromophores are in fact sharper than those of model chromophores in liquid solvents.

Further Discussion

Results on all of the proteins examined in this study are summarized in Figure 16. In Figure 16A, the abscissa, λ_{max}, is the wavelength where the derivative crosses the zero axis and thereby corresponds to the wavelength of maximum extinction in the normal spectrum. This parameter is probably most closely related to the average degree of accessibility of the chromophores to solvent. The ordinate of Figure 16A shows the maximum value of the derivative function at the wavelength where the major peak occurs, normalized with respect to the number of tyrosyl chromophores contributing to the spectrum, N. In the case where there is no spectral heterogeneity (LBI VII), this will be equal to the maximum amplitude for each individual chromophore. When heterogeneity does occur, it will act to reduce $1/N(\Delta\epsilon/\Delta\lambda)_{max}$. The dashed line in Figure 16A summarizes the results obtained on AcTyrOEt in a number of different hydrogen-bonding solvents, although only the points obtained in water and butanol are actually plotted. Since these latter data have been corrected for the inherent differences between AcTyrOEt and protein chromophores, individual protein chromophores would be expected to fall on this line if there is no spectral heterogeneity and if the local environments can be approximated by liquid hydrogen-bonding solvents of the type investigated here. Spectral heterogeneity in multiple-chromophore proteins will always tend to make the protein data lie below the dashed curve.

Although none of the proteins show a value of λ_{max} which would suggest that the chromophores are *completely* exposed to solvent, the results do show that the chromophores of LBI VI, LBI VII, insulin, and RNase(3) (*i.e.*, the three reactive chromophores in native RNase) are well exposed on average. The chromophores of native RNase(6) and pancreatic trypsin inhibitor are apparently exposed to an intermediate extent while those of RNase(3') (the three unreactive chromophores of native RNase) are clearly the least accessible to solvent. In fact, the position of the absorption maximum for the three unreactive chromophores is even slightly to the red of the wave-

TABLE II: Summary of Properties of Tyrosyl Chromophores in Native Proteins, Based on Spectroscopic Parameters Shown in Figure 16.

Protein	No. of Chromophores	Av Accessibility to Solvent	App Deg of Heterogeneity
LBI VII	1	High	Low
LBI VI	2	High	Intermediate
Insulin	4	High	Low–intermediate
Rnase(3)	3	High	Low
Rnase(6)	6	Intermediate	High
PTI	4	Intermediate	Very high
Rnase(3′)	3	Low	Low–intermediate

length expected for protein chromophores in a liquid butanol environment.

Two of the proteins shown in Figure 16A, LBI VII and RNase(3), are found to have amplitudes considerably larger than the values expected for chromophores located in local environments similar to liquid hydrogen-bonding solvents. Since this is contrary to any effects expected from heterogeneity, it necessarily follows that the local environment of these chromophores is not adequately approximated by hydrogen-bonding solvents. This is probably true for all of the proteins examined, but it is not so apparent in certain of these proteins because spectral heterogeneity is high and tends to reduce the maximum amplitude accordingly. That this is in fact the case is suggested by data shown in Figure 16B. The relative sharpness of the protein spectra is reflected in the half-width of the major derivative peak, $\Delta\lambda_{1/2}$, and in the ratio of the maximum amplitude of the major to the minor peak, R. Both of these parameters can only increase in magnitude when spectral heterogeneity is present in multiple chromophore proteins, relative to the values characteristic of the individual chromophores in such proteins. We will use the product of these two parameters, $R \times \Delta\lambda_{1/2}$, to summarize data on proteins and these values are shown in Figure 16B, along with the values for AcTyrOEt in water and in butanol (which are identical). In terms of this parameter, all of the proteins except RNase(6) and pancreatic trypsin inhibitor have lower values than expected if local protein environment could be adequately approximated by solvents such as water and butanol. This occurs in spite of the spectral heterogeneity which exists in these proteins and again emphasizes that liquid hydrogen-bonding solvents are poor approximations for the local environment of chromophores in native proteins.

Except in the case of the well-exposed chromophore of LBI VII, we are not able to know precisely the spectrum of any individual chromophores in these proteins due to the complications of spectral heterogeneity. Nevertheless, certain qualitative conclusions about average properties of the chromophores can be reached on the basis of the data shown in Figures 16A and 16B. These conclusions are stated in Table II for each of the proteins examined. It is encouraging that the two proteins with an intermediate degree of exposure (RNase(6) and pancreatic trypsin inhibitor) also show the largest apparent degree of heterogeneity, since this would be intuitively expected. The conclusions reached on the basis of spectral characteristics are also in excellent agreement with independent information from the literature, which can be briefly summarized. For example, it is known that the tyrosyl residues in unfractionated LBI are both acetylated (Gorbunoff, 1970) and iodinated (Steiner, 1966) with relative ease in the absence of denaturing conditions, suggesting a high degree of exposure. Likewise, in the case of insulin it is known that three of the four tyrosines ionize normally (Inada, 1961) and that all four are readily iodinated (Covelli and Wolff, 1967) and acylated with either *N*-acetylimidizole (Riordan *et al.*, 1965) or cyanuric fluoride (Kurihara *et al.*, 1963). All are oxidized rapidly by tyrosinase (Cory and Frieden, 1967) as well. However, only two are readily nitrated (Sokolovsky *et al.*, 1966). Perhaps most important, the latest refinement of the crystal structure suggests that all four tyrosines are hydrogen bonded to water (Blundell *et al.*, 1971). The conclusion from this study—that the four tyrosines in insulin are well exposed on the average, and that none of the chromophores are grossly different from the others—seems to be in general agreement with these data.

There is also independent information available on the four tyrosines of pancreatic trypsin inhibitor. It is known that two of these titrate normally and two abnormally (Scholtan and Rosenkranz, 1966). It has been reported that residues 10 and 21 are readily nitrated while 35 and 23 cannot be nitrated (Meloun *et al.*, 1968). The crystal structure (Huber *et al.*, 1972) confirms that 10 and 21 are situated in relatively open crevices while 35 and 23 are buried and hydrogen bonded to main chain carbonyls. Our results, suggesting a moderately high degree of burial on the average with large differences between certain of the chromophores, are in excellent agreement with this information.

There is probably more information available on the tyrosine chromophores of native ribonuclease than on any other protein. From solution studies, there is a strong suggestion that there are three available and three unavailable tyrosines. This grouping is found in studies on ionization (Tanford and Hauenstein, 1956), acylation with cyanuric fluoride (Gorbunoff, 1967), nitration (Sokolovsky *et al.*, 1966), iodination at neutral pH (Woody *et al.*, 1966), and acetylation with *N*-acetylimidizole (Riordan *et al.*, 1965). In the case of iodination (Woody *et al.*, 1966) and reaction with *N*-acetylimidizole (Burstein and Patchornik, 1972), the three reactive residues have been definitely identified as tyrosines-73, -76, and -115. These results suggest intermediate exposure and a high degree of spectral heterogeneity, in keeping with our expectation in Table II.

Some of the heterogeneity among the tyrosines in native ribonuclease was removed in this study by estimating the separate contributions of the three reactive and the three unreactive chromophores. As would be expected, the spectra of each of these triads were considerably less broadened than the spectrum of the six tyrosines in native ribonuclease. Also, the three reactive chromophores were found to absorb much further to the blue than the three unreactive, in keeping with what is known about the effect of local environment on spectral positioning. In examining this in a little more detail, however, some questions arise. Lee and Richards (1971) have estimated the static accessibilities of the OH group of the tyrosine side chains in crystalline ribonuclease S (which should be very similar to those for ribonuclease A). The accessibilities which they calculate (measured relative to the accessibility in tripeptides of the type Ala-Tyr-Ala and Gly-Tyr-Gly) for each of the chromophores are shown in Table III.

If it is assumed, for the moment, that there exists a linear relationship between λ_{max} and the degree of contact with

solvent (*i.e.*, OH accessibility) then the average degree of accessibility might be estimated from the spectroscopic data by the following equation

$$\% \text{ accessibility} = \frac{\lambda_{max} - \lambda_{max}(0\%)}{\lambda_{max}(100\%) - \lambda_{max}(0\%)} \times 100$$

where λ_{max} is the experimental value and $\lambda_{max}(100\%)$ and $\lambda_{max}(0\%)$ are the reference values for the most accessible and the least accessible cases. A reasonable estimate of $\lambda_{max}(100\%)$ can be made from the value observed for fully reduced ribonuclease in 6 M guanidine hydrochloride, after correcting this value to pure water from the known shift experienced by AcTyrOEt in going from 6 M guanidine hydrochloride to water, leading to the value $\lambda_{max}(100\%) = 275.3$ nm. It is more difficult to come up with a reasonable estimate for the fully buried and inaccessible chromophore. About the best that can be done is to use Lee and Richards' estimate of 21% accessibility for the three unreactive chromophores in ribonuclease and solve the above equation for $\lambda_{max}(0\%)$. This leads to a value of 280.0 for this parameter. This seems like a reasonable estimate and suggests that the protein interior is a somewhat more extreme environment for chromophores than is butanol, for example, since it would be expected that a protein chromophore in pure butanol would absorb at 278.7 nm.

If these same values of $\lambda_{max}(100\%)$ and $\lambda_{max}(0\%)$ are used to estimate the average accessibility for the reactive triad of ribonuclease, an estimate of 71% is obtained, which is considerably higher than the 21% for the unreactive triad. It is also apparently in substantial disagreement with the estimate of 46% made from Lee and Richard's data. However, there may be a good reason for this disagreement. In the crystalline enzyme, the oxygen atoms of tyrosines-73 and -115 protrude away from the protein surface into the aqueous phase. The low estimates of accessibility (28 and 35%) are largely due to the fact that the oxygen atoms of 73 and 115 are within hydrogen-bonding distance of each other. There is a possibility that the hydrogen bond between these two tyrosines will be broken when dissolved in solution since they are largely surrounded by solvent. Even if the bond does persist, the spectrum of each of these tyrosines is not going to be too different from a chromophore with much larger exposure to solvent, since the OH from the nearby chromophore will probably perturb much the same as a water molecule. Consequently, the spectroscopic estimate of 70% seems to be in reasonable accord with the crystalline structure.

The broadening of the spectrum of a simple chromophore in liquid solution, as opposed to the sharp spectrum generally observed in gas phase or in crystal, is largely attributable to structural fluctuations in the liquid solutions (Jaffe and Orchin, 1962; Bayliss and McRae, 1954). Such fluctuations are particularly productive of energy broadening when the chromophore and solvent are strongly interacting, as for polar chromophores in polar solvents where hydrogen bonding is prevalent. This effect can be clearly seen when comparing the spectra of AcTyrOEt in hexane, ethyl acetate, and propanol where the broadening increases progressively as the solvent goes from nonbonding to partially bonding to completely bonding, respectively. In general, broadening will be decreased when (a) the local environment interacts less strongly with the chromophore, or (b) the local environment becomes less mobile so that fluctuations are damped.

We have observed that native protein chromophores exhibit much sharper spectra than would be expected if their local environment were similar to liquid water or to liquid organic solvents with full hydrogen-bonding potential. This seems to be true for readily accessible chromophores (LBI VI, the three reactive chromophores of ribonuclease, and the four chromophores of insulin) as well as for inaccessible chromophores (the three unreactive chromophores of ribonuclease). From the above discussion, this would imply one of two possibilities (or both). First, it is possible that the OH groups of these chromophores in native proteins do not achieve the same degree of hydrogen bonding as is achieved by AcTyrOEt or by fully unfolded proteins in water so that some or all of the individual chromophores might display spectra more characteristically found in solvents such as ethyl acetate. This is a particularly strong possibility for chromophores which are largely inaccessible to solvent. For those proteins whose crystal structure is precisely known, it does appear that each buried tyrosine has at least one hydrogen-bonding partner nearby and that this is usually a proton acceptor such as a backbone carbonyl. It is not clear, however, that all such buried tyrosines are able to overcome the geometrical restrictions which would arise in trying to provide a full complement of hydrogen-bonding partners. Indeed, it has been suggested that hydrogen bonds where tyrosine acts as the proton acceptor are not nearly so strong as those where it acts as the proton donor (Nemethy and Ray, 1973) due to the high acidity of the oxygen resulting from attachment to an aromatic ring. If this is true, there may be no thermodynamic necessity for providing each buried tyrosine with a full complement of hydrogen-bonding partners. Although this offers a plausible explanation for the decreased broadness observed for proteins with largely inaccessible tyrosyls (such as RNase-(3′)) it is not apparent why this would be an important factor for those other proteins containing chromophores where the OH group is largely exposed, unless the degree of exposure is less than sufficient to allow completely free access of water molecules.

Secondly, reduced spectral broadening would occur if motion is restricted in the vicinity of the chromophore. This effect will be operative even though the chromophore has a full complement of bonding partners. Although the spatial arrangement of groups within a native protein may not be completely fixed, the number of thermodynamically significant configurations will undoubtedly be considerably smaller than is present in simple liquids. If structural fluctuations are minimized in the interior of native proteins, it is reasonable to expect a reduced degree of spectral broadening. It has also frequently been suggested, but never really substantiated, that the water of solvation about proteins is more ordered and less mobile than bulk water. If this is true, then broadening from geometrical fluctuations should be reduced even for those chromophores which have their OH moiety "completely exposed."

TABLE III: Static Accessibilities of the Tyrosyl OH to Water Molecules in Crystalline Ribonuclease S (Estimated from the Data of Lee and Richards (1971)).

Less Reactive Chromophores		Reactive Chromophores	
Amino Acid	%	Amino Acid	%
Tyr-25	0	Tyr-73	35
Tyr-92	64	Tyr-76	74
Tyr-97	0	Tyr-115	28
Av accessibility = 21%		Av accessibility = 46%	

Earlier work from this laboratory (Brandts *et al.*, 1970) showed that the changes in volume and compressibility during the denaturation of ribonuclease do not seem consistent with the "expected" changes based on predictions from model transfer reactions which also rely on water and organic solvents to approximate local protein environments. Klapper (1971) has suggested a possible explanation for this discrepancy. On the basis of available volumetric data, he has shown that the packing of protein groups in native proteins is very efficient and leads to a surprisingly low void volume. He concludes that native proteins in solution must therefore be considered to be solid or semisolid in character, rather than liquid, and that this accounts for the inability of the usual reference liquids to correctly predict local environment. These arguments are in accord with the second alternative discussed above. The high degree of resolution obtainable from X-ray diffraction data on crystalline proteins might also be taken as evidence to support the notion that structural fluctuations in native proteins are relatively minor.

It is, however, difficult to be certain as to the causes of the low degree of spectral broadening of tyrosyl chromophores in native proteins and the above discussion is only intended to point to two of the more obvious possibilities. In spite of the problems in interpretation, the present results do show with certainty that the popular methods of estimating spectroscopic properties of protein chromophores are not very suitable at a quantitative level. Significantly large errors are present in the expectation values based on analogy with small model chromophores or unfolded proteins in liquid solvents, the exact magnitude of the errors being uncertain but probably in excess of 30% for effects based on shift-induced difference spectra. This discrepancy arises solely from the fact that the derivative spectrum for protein chromophores has larger amplitude than for model chromophores in the usual reference solvents so that an equivalent spectral shift will produce a larger difference spectrum for protein chromophores than would be predicted from model studies. Other factors may tend to enhance or reduce the errors concerned with amplitude effects. For example, even though the derivative spectrum of the three accessible chromophores in ribonuclease is considerably sharper than for the model chromophore AcTyrOEt in water, the solvent perturbation spectrum is actually broader. In the case of the three inaccessible chromophores, the perturbation spectrum is so broad and poorly defined that it bears little resemblance to the perturbation spectrum of model chromophores in liquid solvents. Factors such as these caution against placing too much reliance on the usual methods of quantitating spectroscopic information until a better understanding of the spectra of individual protein chromophores is achieved.

References

Bayliss, N., and McRae, E. (1954), *J. Phys. Chem. 58*, 1002.

Blundell, T., Cutfield, J., Cutfield, S., Dodson, E., Dodson, G., Hodgkin, D., Mercola, D., and Vijayan, M. (1971), *Nature (London) 231*, 506.

Brandts, J., Oliveira, R., and Westort, C. (1970), *Biochemistry 9*, 1038.

Brandts, J. F., and Kaplan, L. J. (1971), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 30*, 1108.

Burstein, Y., and Patchornik, A. (1972), *Biochemistry 11*, 2939.

Chignell, D., and Gratzer, W. (1968), *J. Phys. Chem. 72*, 2934.

Cory, J., and Frieden, E. (1967), *Biochemistry 6*, 116.

Covelli, I., and Wolff, J. (1967), *J. Biol. Chem. 242*, 881.

Donovan, J., (1969), *in* Physical Principles and Techniques of Protein Chemistry, Part A, Leach, S. J., Ed., New York, N. Y., Academic.

Gorbunoff, M. (1967), *Biochemistry 6*, 1606.

Gorbunoff, M. (1970), *Biochim. Biophys. Acta 221*, 314.

Haynes, R., and Feeney, R. (1967), *J. Biol. Chem. 242*, 5378.

Herskovits, T., and Laskowski, M., Jr. (1962), *J. Biol. Chem. 237*, 2481, 3418.

Horwitz, J., and Strickland, E. (1971), *J. Biol. Chem. 246*, 3749.

Horwitz, J., Strickland, E., and Billups, C. (1970), *J. Amer. Chem. Soc. 92*, 2119.

Huber, R., Kukla, D., Ruhlmann, A., and Steigemann, W. (1972), *Cold Spring Harbor Symp. Quant. Biol. 36*, 141.

Ikeda, K., Hamaguchi, K., Yamamoto, M., and Ikenaka, T. (1968), *J. Biochem. (Tokyo) 63*, 521.

Inada, Y. (1961), *J. Biochem. (Tokyo) 49*, 217.

Jaffe, H., and Orchin, M. (1962), Theory and Applications of Ultraviolet Spectra, New York, N. Y., Wiley.

Jones, G., Moore, S., and Stein, W. (1963), *Biochemistry 2*, 66.

Klapper, M. (1971), *Biochim. Biophys. Acta 229*, 557.

Kurihara, K., Horinishi, H., and Shibata, K. (1963), *Biochim. Biophys. Acta 74*, 678.

Laskowski, M., Jr. (1970), *in* Spectroscopic Approaches to Biomolecular Conformation, Chicago, Ill., American Medical Association.

Lee, B., and Richards, F. (1971), *J. Mol. Biol. 55*, 379.

Meloun, B., Fric, I., and Sorm, F. (1968), *Eur. J. Biochem. 4*, 112.

Nemethy, G., and Ray, A. (1973), *J. Phys. Chem. 77*, 64.

Riordan, J., Wacker, W., and Vallee, B. (1965), *Biochemistry 4*, 1758.

Scholtan, W., and Rosenkranz, H. (1966), *Makromol. Chem. 99*, 254.

Sokolovsky, M., Riordan, J., and Valee, B. (1966), *Biochemistry 5*, 3582.

Steiner, R. (1966), *Biochemistry 5*, 1964.

Tanford, C., and Hauenstein, J. (1956), *J. Amer. Chem. Soc. 78*, 5287.

Woody, R., Friedman, M., and Scheraga, H. (1966), *Biochemistry 5*, 2034.