Roitt, J. M., Campbell, P. N., & Poniach, D. (1958) *Biochem. J. 69*, 248–297.
Sokolovsky, M., Riordan, J. F., & Vallee, B. L. (1966) *Biochemistry 5*, 3582–3589.
Spackman, D. M., Stein, W. H., & Moore, S. (1958) *Anal. Chem. 30*, 1190–1206.
Strickland, E. H. (1974) *CRC Crit. Rev. Biochem. 2*, 113–175.
Strong, P. N., Goerke, J., Oberg, S. G., & Kelley, R. B. (1976) *Proc. Natl. Acad. Sci. U.S.A. 73*, 178–182.
Tanford, C. (1961) *Physical Chemistry of Macromolecules*, p 556, Wiley, New York.
Tanford, C., Hauenstein, J. D., & Rands, D. G. (1955) *J. Am. Chem. Soc. 77*, 6409–6413.
Thomas, C. B., Maurice, J. G., & Sandra, L. W. (1968) *J. Am. Chem. Soc. 96*, 1612–1619.
Tsernoglou, D., & Petsko, G. A. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 971–974.
Tsernoglou, D., Petsko, G. A., & Hudson, R. A. (1978) *Mol. Pharmacol. 14*, 710–716.
Tu, A. T. (1973) *Annu. Rev. Biochem. 42*, 235–258.
Walkinshaw, M. D., Saenger, W., & Maelicke, A. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 2400–2404.
Woody, R. W. (1974) in *Peptides, Polypeptides and Proteins (Rehovot Symposium)* (Blout, E., Bovey, F. A., Goodman, M., & Lotan, N., Eds.) pp 338–350, Wiley, New York.
Yang, C. C. (1974) *Toxicon 12*, 1–43.

Quantitation of Aromatic Residues in Proteins: Model Compounds for Second-Derivative Spectroscopy†

Rodney L. Levine* and M. Marcia Federici‡

ABSTRACT: The ultraviolet absorption spectrum of proteins in 6 M guanidine is approximately that of the sum of the spectra of the constituent aromatic amino acids, phenylalanine, tyrosine, and tryptophan, plus contributions from light scattering and disulfides. A multicomponent analysis of the spectrum would theoretically permit simultaneous quantitation of each aromatic amino acid in the protein. In practice, this has not been possible, because of the similarities of the spectra of the amino acids, large differences in molar absorptivity, variable absorption by the disulfides, light scattering, and wavelength shifts which occur when the amino acids are incorporated into proteins. We describe a method for the simultaneous quantitation of the aromatic amino acids in purified proteins. We used second-derivative ultraviolet spectroscopy coupled with a statistically weighted multicomponent analysis. Use of the second derivative virtually eliminated interference from light scattering and from cystine. Empirical selection of model compounds obviated the problem of wavelength shifts. The models are *N*-acetylphenylalanine ethyl ester in 6 M guanidine for phenylalanine, *N*-acetyltyrosine ethyl ester in 55% methanol for tyrosine, and mellitin in 6 M guanidine for tryptophan. This method permits accurate, rapid quantitation of phenylalanine, tyrosine, and tryptophan in intact, denatured proteins.

The ultraviolet spectrum of proteins approximates that of its constituent aromatic amino acids (Wetlaufer, 1962). The near-ultraviolet spectrum of a denatured protein can be closely, but not exactly, mimicked by an equimolar mixture of its constituent aromatic amino acids plus cystine (Edelhoch, 1967). The differences between the observed and "synthesized" spectra remain of interest. Exact reconstruction of protein spectra from simpler models is of particular interest because one might then derive rapid spectroscopic methods for quantitation of the aromatic residues within proteins. Edelhoch (1967) introduced a simple method for determination of the tryptophan content of proteins from their absorbance spectra, although the analysis gives erroneous values for some proteins (Hugli & Moore, 1972; Levine, 1982). Tyrosine values are less reliable, and phenylalanine cannot be estimated.

Derivative spectroscopy might provide a more powerful technique for quantitative analysis of amino acid residues in proteins (Balestrieri et al., 1978, 1980; Federici & Levine, 1980). The potential of increased resolution and precision was pointed out by Giese & French (1955). Others subsequently provided theoretical and experimental contributions to derivative spectroscopy (Olson & Alway, 1960; Butler & Hopkins, 1970; Grum et al., 1972; Brandts & Kaplan, 1973; Shibata, 1976; O'Haver & Green, 1976; Hawthorne & Thorngate, 1978). With the addition of microprocessors or computers to commercial spectrophotometers, the calculation of derivative spectra became a simple task. The first derivative ($dA/d\lambda$) is the rate of change of absorbancy with wavelength; the second derivative ($dA^2/d\lambda^2$) is the velocity of that change. A derivative spectrum thus enhances small dips, peaks, and shoulders compared to the direct spectrum.

Balestrieri et al. (1978) took advantage of this enhancement to develop a technique for quantitating phenylalanine with proteins. They added known amounts of *N*-acetylphenylalanine ethyl ester to the protein solution and were able to determine the content of phenylalanine in the protein by regression analysis. The second derivative obeys Beer's law. If the second derivative spectrum of a protein could be dissected to isolate the contributions of the three aromatic amino acids, then it would be possible to quantitate all three amino acids in that protein. We used simple compounds to provide models whose spectra match those of the aromatic residues. Using second-derivative spectra, we were able to accurately quantitate

† From the Laboratory of Biochemistry, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, Maryland 20205. *Received December 10, 1981.*
‡ Present address: Monoclonal Antibodies, Inc., Mountain View, CA 94043.

the phenylalanine, tyrosine, and tryptophan content of intact, denatured proteins.

Materials and Methods

The *N*-acetyl ethyl esters of phenylalanine, tyrosine, and tryptophan were used as received from Sigma Chemical Co. Guanidine hydrochloride ("ultrapure") was from Bethesda Research Laboratories, Gaithersburg, MD. The concentrated stock was 8.0 M guanidine–26.7 mM potassium phosphate, pH 6.35 (make by dissolving 764 g of guanidine hydrochloride in about 350 mL of water with warming; add 3.63 g of KH_2PO_4; adjust the pH with KOH; bring to 1 L). When 3 volumes of stock are diluted with 1 volume of protein solution, one has 6 M guanidine–20 mM potassium phosphate, pH 6.5 (Edelhoch, 1967). Lysozyme was obtained from Sigma, Worthington, and Boehringer Mannheim. The latter was further purified as described (Sophianopoulos et al., 1962). S-Carboxymethylated lysozyme was a gift from Dr. T. Martensen, NIH, who prepared it with iodoacetate. Rabbit muscle aldolase and ribonuclease were gifts from Dr. A. Ginsburg, NIH. The purity of this ribonuclease has been reported (Ginsburg & Carroll, 1965). Bovine and human albumin were from Sigma. They were further purified by acid–charcoal treatment (Chen, 1967) and Cibacron Blue chromatography (Bio-Rad Laboratories, 1977). Monomer-free, highly purified reference bovine serum albumin, 70.45 g/L, was a gift from Dr. D. Reader of the National Bureau of Standards. Bovine liver catalase from Boehringer Mannheim was used after buffer exchange on Sephadex G-25 into 50 mM tris(hydroxymethyl)aminomethane (Tris)–50 mM KCl, pH 7.40. Insulin and its A chain, parvalbumin, and mellitin were from Sigma. The parvalbumin was purified by using DEAE-cellulose (Blum et al., 1977). Mellitin was used as supplied; in 6 M guanidine, the ratio of absorbances of the major and minor peaks was 1.17, exactly as reported by Edelhoch (1967) for *N*-acetyltryptophanamide.

Protein molecular weights and residue content were taken from sequence data (Dayhoff, 1978). The concentrations were determined spectrophotometrically as listed: insulin, ϵ_M = 5220 at 280 nm in 35 mM HCl (Praissman & Rupley, 1968); insulin A chain, ϵ_M = 2610 (one-half that of insulin); ribonuclease, ϵ_M = 9390 at 276 nm in 6 M guanidine (Sela et al., 1957); aldolase, ϵ_M = 35 980 at 280 nm in water (Donovan, 1964; Sia & Horecker, 1968); bovine albumin, ϵ_M = 44 300 at 278 nm in water (Foster & Sterman, 1956); human albumin, ϵ_M = 35 200 at 279 nm in water (Foster & Sterman, 1956); catalase, ϵ_M = 297 000 at 404 nm in 50 mM Tris–50 mM KCl, pH 7.4 (Deisseroth & Dounce, 1967); lysozyme, ϵ_M = 37 600 at 280 nm in 100 mM KCl (Imoto et al., 1972).

The multicomponent analysis of aromatic residues of proteins required estimation of ϵ_M of the residues (Table I). The peak ϵ_M of phenylalanine is not very sensitive to solvent effects [e.g., see Gratzer (1970)]. Although the peaks shift slightly to the red, we found the ϵ_M of the *N*-acetylphenylalanine ethyl ester to be the same in 6 M guanidine as in water (data not shown). Thus, we use ϵ_M = 110.6 at 248 nm for the phenylalanine model (Mihalyi, 1968). Of course, tyrosine is quite sensitive to solvent effects [e.g., see Solli & Herskovits (1973)]. Harrison & Garratt (1969) noted that in acid solution the ϵ_M (corrected for scattered light) of insulin is identical with that of tyrosine in acid. Further, we found less than 1% difference in the ϵ_M of insulin in acid or in 6 M guanidine. Thus, for tyrosine in proteins, we use ϵ_M = 1330 at 276 nm (Mihalyi, 1968). The ϵ_M from *N*-acetyltryptophan ethyl ester is also sensitive to solvent effects (Solli & Herskovits, 1973). However, in 6 M guanidine, the ϵ_M values of various amino-blocked tryptophan models are similar (Edelhoch, 1967). Thus, we used ϵ_M = 5620 at 281 nm for tryptophan.

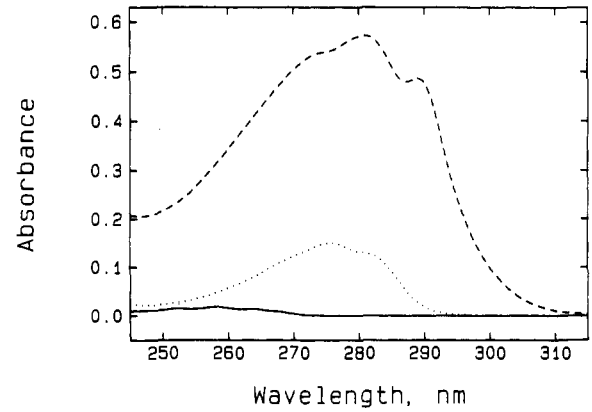

FIGURE 1: Absorbance spectra for 100 μM *N*-acetyl ethyl esters of phenylalanine (—), tyrosine (···), and tryptophan (- - -). Compare with the second derivatives in Figure 2. The solvent was 6 M guanidine–20 mM potassium phosphate, pH 6.5.

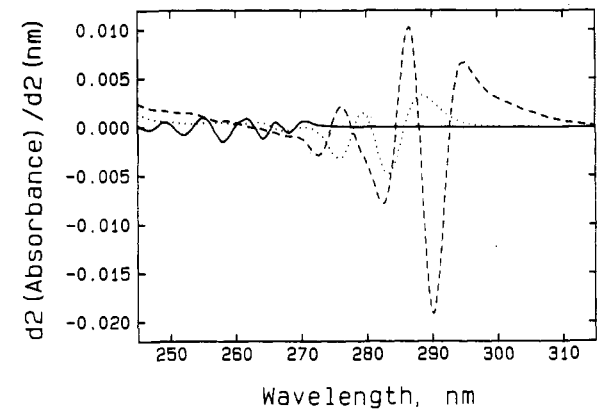

FIGURE 2: Second-derivative spectra from Figure 1: 100 μM *N*-acetyl ethyl esters of phenylalanine (—), tyrosine (···), and tryptophan (- - -).

Spectrophotometry. Spectra were taken with a Hewlett-Packard Model 8450 A spectrophotometer. This is an unusual spectrophotometer. It employs an array of photodiodes as its detector. This permits parallel (simultaneous) acquisition of absorbances at 401 specific wavelengths. The entire spectrum from 200 to 800 nm can be measured twice, the results averaged, and the spectrum displayed on the cathode ray screen in about 1 s. We used 10-s measurements, which provides the average of 20 individual spectra.

The spectrophotometer also includes a microprocessor which allows rapid calculation of derivatives of the absorbance spectrum. The microprocessor also implements a multicomponent analysis method (Schleifer & Willis, 1980) which takes advantage of estimates of precision available with the averaged spectra (see Discussion).

Actual spectra were obtained by placing 0.75 mL of 8 M guanidine solution in semimicro, masked cuvettes in the reference and sample positions. Then the balance was performed (i.e., automatic base-line correction). After the balance was confirmed, 0.25 mL of buffer was added to the reference cuvette and 0.25 mL of protein solution to the sample cuvette. The solutions were mixed by inversion or by stirring. Spectra were followed until they became constant, when denaturation was presumed complete. For many proteins, this takes less than 1 min. If desired, the spectrum may be stored on cassette tape or a floppy disk.

Results

Derivative Spectra of Model Compounds. The absorbance spectra of the *N*-acetyl ethyl esters of phenylalanine, tyrosine,

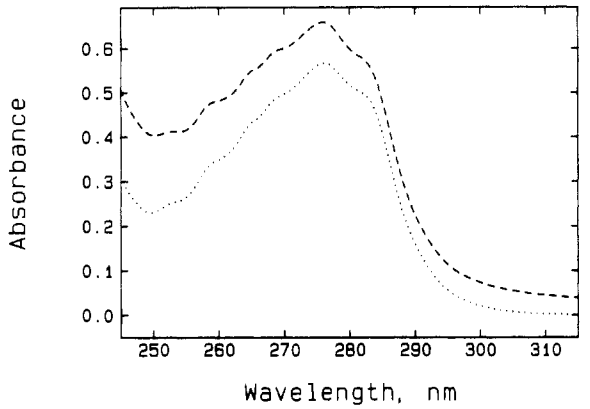

FIGURE 3: Absorbance spectrum of 18.4 μM human albumin in 6 M guanidine–20 mM potassium phosphate, pH 6.5. (---) Observed spectrum; (···) spectrum corrected for light scatter. The apparent absorbance from light scattering was determined by fitting a linear regression to a plot of log wavelength vs. log absorbance from 320 to 350 nm (a region in which only scatter is assumed to give absorbance). The equation then permits calculation of the light-scatter component at any wavelength. The equation is log absorbance = 14.91 − 6.534 × log wavelength, with a correlation coefficient of 0.9999.

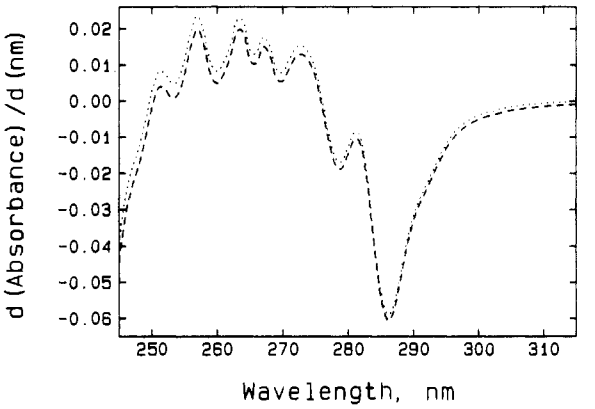

FIGURE 4: First-derivative spectrum of 18.4 μM human albumin, from Figure 3. (---) Observed spectrum; (···) spectrum corrected for light scatter.

and tryptophan are plotted in Figure 1 and their second derivatives in Figure 2. In the absorbance spectra, one sees that the dramatic differences in molar absorptivities will make multicomponent analysis of proteins difficult or impossible, particularly for phenylalanine. However, the second derivative of phenylalanine is easily distinguished from those of tyrosine and tryptophan (240–270 nm). Note that the second derivatives of tyrosine and tryptophan are relatively flat, but are not zero in the 240–270-nm range [compare Ichikawa & Terada (1977)]. Consequently, the contribution of all three aromatic amino acids must be considered in developing an accurate method of multicomponent analysis of protein spectra. From 275 to 310 nm, the second derivative of phenylalanine is zero while those of tyrosine and tryptophan show distinct qualitative differences. In principle, then, it should be possible to analyze the second-derivative spectra of proteins and quantitate the aromatic amino acids within proteins.

Light Scattering. The UV absorption spectrum of a protein is the sum of the true absorption plus a contribution from scattering. The scattering component can be quite large, and it increases exponentially with decreasing wavelength (Figure 3). This practically eliminates the possibility of direct spectral quantitation of phenylalanine. Since scattering is a smoothly varying function of wavelength, each successive derivative will reduce the contribution of scatter to the spectrum. Figure 4 shows that the first derivative greatly reduces, but does not

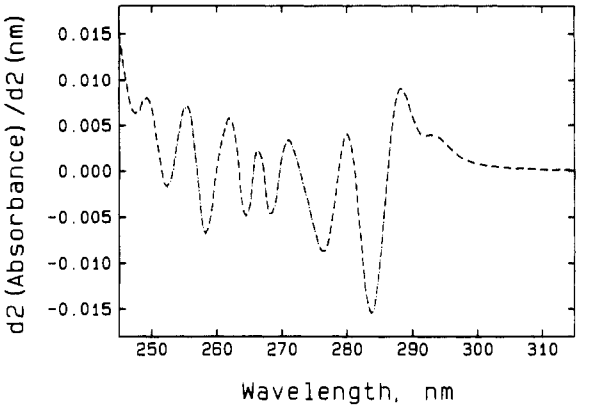

FIGURE 5: Second-derivative spectrum of 18.4 μM human albumin, from Figure 3. (---) Observed spectrum; (···) spectrum corrected for light scatter. The spectra are essentially identical.

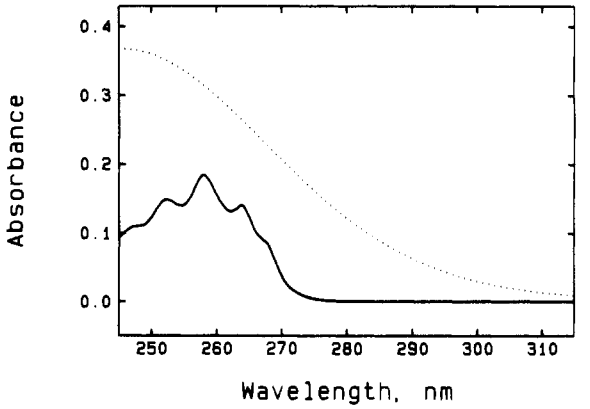

FIGURE 6: Absorbance spectra for 1 mM *N*-acetylphenylalanine ethyl ester (—) and cystine (···). Compare with the second derivatives in Figure 7. The solvent was 6 M guanidine–20 mM potassium phosphate, pH 6.5.

eliminate, the scatter component. However, the second derivative effectively eliminates the scatter (Figure 5).

Figure 5 also provides an example of a second-derivative spectrum of a typical protein. Although such a spectrum seems strange at first, we found that we readily learned to recognize the components contributed by the constituent aromatic residues. The characteristic second-derivative spectrum of phenylalanine is obvious from 245 to 260 nm. Human serum albumin has only one tryptophan per subunit, and its spectral contribution is difficult to appreciate in the direct absorbance spectrum (Figure 3). However, the shoulder around 295 nm is obvious in the second-derivative spectrum (Figure 5).

Disulfide Bonds. Disulfide bridges also contribute to the spectrum of proteins. Like light scattering, this absorption can seriously interfere with quantitation of the aromatic amino acids (Figure 6). This problem might be dealt with by including in the analysis a model compound for the disulfide bridge. This proved unnecessary because the second derivative also effectively eliminates the spectral contribution of the disulfide (Figure 7).

Selection of the Model Compounds. Model compounds such as the *N*-acetyl ethyl esters of amino acids show molar absorptivities close to those of their corresponding residues in proteins (Edelhoch, 1967; Solli & Herskovits, 1973). However, the spectral match is not perfect, primarily because of a red shift of the residues within proteins (Wetlaufer, 1962; Edelhoch, 1967; Tanford, 1968; Brandts & Kaplan, 1973; Rafferty et al., 1980). This shift seriously interfered with quantitative, multicomponent analysis of protein spectra (data not shown). Solvent perturbation might permit shifting of the amino acid spectrum to match that within proteins (Solli & Herskovits,

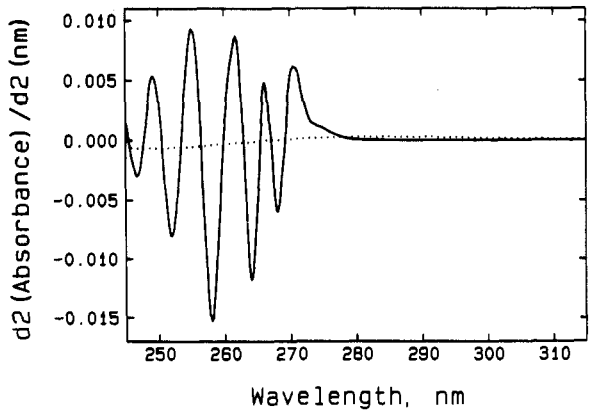

FIGURE 7: Second-derivative spectra from Figure 6; 1 mM *N*-acetylphenylalanine ethyl ester (—) and cystine (···).

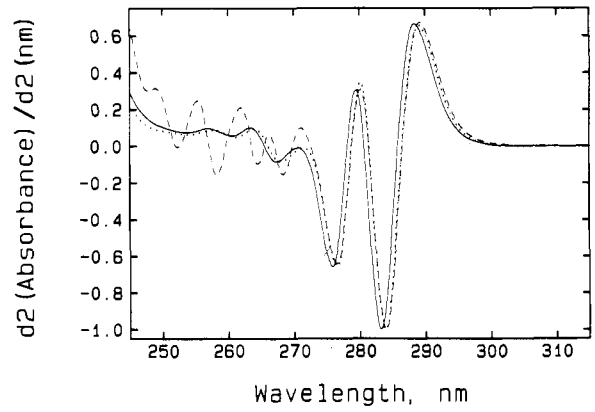

FIGURE 8: Second-derivative spectra of insulin in 6 M guanidine (---) and of *N*-acetyltyrosine ethyl ester in 6 M guanidine (—) or in 55% methanol (···). The spectra were normalized to a trough value of −1.0 to permit ready comparison.

1973). An empirical search revealed that 55% methanol did shift the spectrum of *N*-acetyltyrosine ethyl ester to provide an excellent match with that of insulin (Figure 8). (Insulin lacks tryptophan residues, thus permitting easy evaluation of the spectral match.) Ribonuclease also lacks tryptophan, and its normalized second derivative superimposed on that of insulin, in the region above the phenylalanine signal (>275 nm). *N*-Acetyltyrosinamide gave a spectrum identical with that of the *N*-acetyl ethyl ester.

Phenylalanine shows little solvent perturbation (Gratzer, 1970). Not surprisingly, the spectrum of *N*-acetylphenylalanine ethyl ester in 6 M guanidine provided an excellent model for phenylalanine residues. (The model was compared to the spectrum of rabbit parvalbumin, a protein with nine phenylalanine but no tyrosine or tryptophan residues.)

This leaves only the tryptophan model to be selected. Tryptophan demonstrates a red shift upon incorporation into protein, although the shift is not as great as that for tyrosine. Despite empirical experimentation with methanol and ethanol, solvent perturbation did not provide an adequate match for the tryptophan residue. Although the amino acid's spectral peaks could be red shifted, the resultant spectrum was distorted when compared to that of mellitin, a peptide containing one tryptophan and no phenylalanine or tyrosine residues (Figure 9). Not surprisingly, multicomponent analysis of proteins with the shifted model spectrum gave poor quantitative results (data not shown). The limited solubility of polytryptophan prevented our obtaining an adequate second-derivative spectrum.

Although the amino acids did not provide an adequate model, we have an acceptable model in the mellitin itself. Thus, we used the second-derivative spectrum of mellitin as

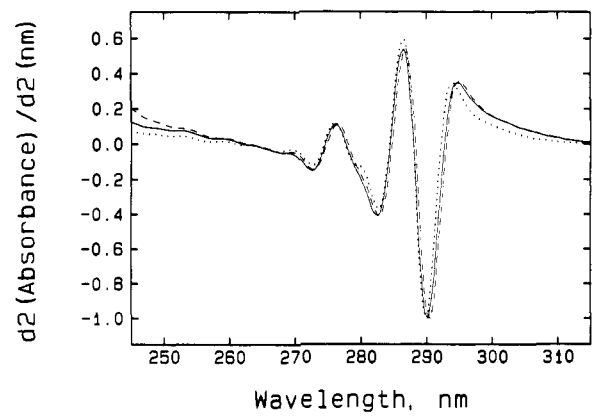

FIGURE 9: Second-derivative spectra of mellitin in 6 M guanidine (---) and of *N*-acetyltryptophan ethyl ester in 6 M guanidine (—) or in 55% methanol (···). The spectra were normalized to a trough value of −1.0 to permit ready comparison.

Table I: Model Compounds

residue	model	solvent	ϵ_M	λ (nm)
phenylalanine	*N*-acetyl-phenylalanine ethyl ester	6 M guanidine[a]	110.6	248
tyrosine	*N*-acetyl-tyrosine ethyl ester	55% methanol	1330	276
tryptophan	mellitin	6 M guanidine[a]	5620	281

[a] That is, 6 M guanidine hydrochloride–20 mM potassium phosphate, pH 6.5 (Edelhoch, 1967).

the tryptophan model. We could also have used selected proteins for the other aromatic residues (e.g., parvalbumin for phenylalanine). We chose not to do so since the amino acids are readily available in pure form and at reasonable cost.

Table I summarizes the three spectral models and their molar absorptivities. The spectra and their second derivatives were stored on cassette tape, permitting easy recall of the model spectra. When new spectra were obtained on solutions prepared 3 months later, they matched the original spectra perfectly. Thus, spectra of the standard model compounds need not be repeated with each multicomponent analysis.

Multicomponent Analysis of Protein Spectra. Using the second-derivative spectra of the three model compounds (Table I), we obtained good quantitative analysis of proteins by multicomponent analysis from 255 to 300 nm. However, better results were obtained when the analysis was run twice: first, from 245 to 265 nm with all three models to obtain only the phenylalanine concentration; second, from 280 to 300 nm with the tyrosine and tryptophan models to obtain their concentrations. This two-part analysis required only a few extra seconds but improved the phenylalanine results by isolating that part of the spectrum where its second derivative dominates.

The results of this analysis for several proteins are shown in Table II. All values agree to within 5% of the result expected from sequence data, except for the tyrosine content of bovine serum albumin. The available sequence data for this protein give 19 residues/molecule (Brown, 1977) while our method gave 20. We found 20 residues/molecule for the highly purified albumin from the National Bureau of Standards as well as several lots from Sigma. The value for the Sigma product did not change after charcoal treatment nor after affinity chromatography. The sequence data may be incorrect, as amino acid analysis also gives 20 residues of tyrosine (Spahr & Edsall, 1964). The residues expected at

Table II: Multicomponent Analysis of Proteins[a]

protein	concn (μM)	multicomponent analysis Phe (μM)	Tyr (μM)	Trp (μM)	residues Phe theory	Phe found	Tyr theory	Tyr found	Trp theory	Trp found	% error Phe	Tyr	Trp
insulin A chain	337	35.3	659.2	8.0	0	0.10	2	1.95	0	0.02		−2	
insulin	148	448.4	587.1	−3.6	3	3.04	4	3.97	0	−0.02	1	−1	
ribonuclease	37.7	115.7	229.5	−2.4	3	3.07	6	6.09	0	−0.06	2	2	
aldolase	24.0	166.0	285.2	75.03	7	6.91	12	11.86	3	3.12	−1	−1	4
bovine albumin	24.1	625.3	488.2	46.77	26	25.97	19[b]	20.27	2	1.94	0	7[b]	−3
human albumin	18.2	551.5	328.6	17.87	31	30.27	18	18.04	1	0.98	−2	0	−2
catalase	3.64	107.7	72.03	21.57	31	29.62	20	19.81	6	5.93	−4	−1	−1

[a] Analyses were performed as described under Materials and Methods. [b] Bovine albumin has 19 tyrosine residues according to sequence data (Brown, 1977). However, as discussed in the text, the correct value may be 20. If so, the analytical error is not 7% but 1%.

Table III: Effect of Bovine Albumin Concentration on the Multicomponent Analysis[a]

albumin (μM)	absorbance at 276 nm	residues Phe	Tyr	Trp	% error Phe	Tyr	Trp
53.75	2.243	25.3	20.4	2.05	−3	2	3
26.88	1.178	26.6	20.6	2.03	2	3	2
13.44	0.5882	26.5	20.3	1.99	2	2	−1
6.719	0.2929	26.7	20.1	1.95	3	1	−3
3.360	0.1465	26.5	19.9	1.91	2	0	−4
1.680	0.0777	25.5	19.3	1.83	−2	−4	−8
0.8399	0.0356	23.8	18.9	1.84	−8	−6	−8

[a] Serial dilutions of the National Bureau of Standards bovine albumin were prepared in 6 M guanidine hydrochloride–20 mM potassium phosphate, pH 6.5. The actual residues were taken as 26 for phenylalanine, 20 for tyrosine, and 2 for tryptophan. See the text for discussion of the tyrosine content.

positions 400–403 have not yet been identified in the bovine albumin, and a tyrosine might be present there (Brown, 1977). Note that the values for catalase agree with those expected from sequence data (Schroeder et al., 1969). Apparently, the heme chromophore did not interfere in the second-derivative analysis.

The current results, with less than 5% error, are probably close to the error limits imposed by the purity of the proteins or even their reported molar absorptivities. Thus, we have not attempted to improve on the results shown in Table II, e.g., by adjusting the molar absorptivities of the model compounds.

For clarity, we summarize the method of multicomponent analysis which evolved from these studies:

(1) Record the absorption spectra and second derivatives of the model compounds and quantify their concentrations with the molar absorptivities shown in Table I.

(2) Obtain the second-derivative spectrum of the protein in 6 M guanidine hydrochloride–20 mM potassium phosphate, pH 6.5.

(3) Perform the multicomponent analysis in two steps: (a) For the phenylalanine concentration, use all three models and analyze 245–265 nm. (b) For the tyrosine and tryptophan concentrations, use only those two models, analyzing 280–300 nm.

Effect of Protein Concentration. Table III gives the multicomponent analyses of serial dilutions of bovine serum albumin. The results are independent of protein concentration within the range giving a peak absorbance of about 0.1–1.0, as expected from the stray light and shot noise characteristics of the spectrophotometer (Willis, 1980). When the peak absorbance was outside this range, deviations from the theoretical increased, although not dramatically. Thus, the method has good accuracy over a wide range of absorbancy. For routine work, we typically use a protein concentration which gives a peak absorbance of about 0.2.

Table IV: Multicomponent Analysis of Lysozyme[a]

protein	concentration Phe (μM)	Tyr (μM)	Trp (μM)	residues Phe	Tyr	Trp
lysozyme	36.38	50.56	70.01	3.04	4.22	5.84
lysozyme + 10 mM dithiothreitol	37.30	42.67	70.90	3.11	3.56	5.92
carboxymethylated lysozyme	44.45	54.21	88.64	3.00	3.66	5.98
theoretical				3	3	6

[a] The concentration of lysozyme was 11.98 μM. The protein was incubated in 6 M guanidine–20 mM potasssium phosphate, pH 6.5, at 37 °C for 30 min under argon with and without dithiothreitol. Then analysis was performed. Since a precise ϵ_M value was not available for the carboxymethylated lysozyme, the number of residues was calculated by normalizing the phenylalanine residues to 3.

Lysozyme: Anomalous Results. As already noted, success in multicomponent analysis of a protein spectrum required models whose spectra match those of the aromatic residues in the denatured proteins. Most proteins are fully denatured in 6 M guanidine; that is, they behave as cross-linked random coils (Tanford, 1968). If a protein is not fully denatured in 6 M guanidine, then the multicomponent analysis would likely be inaccurate. Although optical rotatory dispersion measurements indicated that lysozyme was denatured in 6 M guanidine, numerous other studies showed that residual structure is present (Hamaguchi & Kurono, 1963; Tanford, 1968; Aune & Tanford, 1969; Coan et al., 1975; White, 1976).

As shown in Table IV, multicomponent analysis of lysozyme in 6 M guanidine was much less accurate than for the proteins in Table II. The tyrosine value is high while the tryptophan value is low. Results were the same using lysozyme from Worthington, Sigma, and Boehringer Mannheim and were unchanged by additional purification (see Materials and Methods). Secondary structure in lysozyme is due in part to its disulfide bonds (Coan et al., 1975; White, 1976). The multicomponent analysis was improved by addition of dithiothreitol or by carboxymethylation of the protein (Table IV). Significant deviation still persists, consistent with residual structure. Thus, quantitative multicomponent analysis of lysozyme requires conversion of the protein to the true random coil (Tanford, 1968).

Discussion

The ultraviolet absorption spectrum of a native protein approximates the sum of the spectra of its constituent, blocked amino acids (Wetlaufer, 1962). However, small but definite spectral changes occur upon incorporation of the amino acid into a polypeptide. In general, the absorptivity increases, and the peak absorption shifts to the red (Wetlaufer, 1962). Denaturation of proteins in 6 M guanidine reduced these

discrepancies, but a red shift still remained in the direct and derivative spectra (Edelhoch, 1967; Brandts & Kaplan, 1973; this paper). The reason for this persistent discrepancy remains speculative, but it may result from some residual order in the protein (Tanford, 1968).

In any case, Edelhoch (1967) demonstrated that the spectrum of a denatured protein could be closely approximated by the sum of its constituent blocked amino acids. Synthesizing the protein spectrum from its known composition is obviously simpler than dissection of the protein spectrum to yield the isolated spectra of its constituent amino acids. We found that the slight differences noted by Edelhoch introduced major errors in our attempts at "dissection" or multicomponent analysis. Solvent perturbation shifts the spectra of proteins as well as the spectra of our model compounds (Solli & Herskovits, 1973). Empirical trial showed that 55% methanol perturbed the tyrosine spectrum to provide an excellent model of the tyrosine residue in denatured proteins. However, the same derivative of tryptophan in methanol was not an adequate model for the tryptophan residue. We have no compelling explanation for this dichotomy. Perhaps there are subtle differences in the residual structure or environment around tyrosine and tryptophan residues in 6 M guanidine.

Having empirically chosen model spectra by using commercially available compounds, we attempted to resolve the spectra of denatured proteins into their component spectra. This multicomponent analysis succeeded only when we used the second derivative of the absorption spectrum. A derivative spectrum contains no additional information than is present in its parent spectrum. Indeed, it contains less information. Specifically, the contribution of light scattering and disulfide groups is effectively eliminated. With these extraneous signals removed, the contributions of the three aromatic residuals could be accurately assessed.

This assessment was quantitated with the multicomponent analysis program supplied with the Hewlett-Packard 8450A spectrophotometer. In essence, the computer utilizes the spectra of the three model components to synthesize a second-derivative spectrum which closely matches that observed for the protein. A simple method of multicomponent analysis solves a series of simultaneous equations, one equation per component. The Hewlett-Packard spectrophotometer does not utilize this method, but fits to the entire spectrum (Schleifer & Willis, 1980). This provides more data points than the number of components, and the analysis is thus said to be "overdetermined". Coupling this overdetermination with the high precision of the absorbance yields a remarkably accurate multicomponent analysis. The actual analysis proceeds in two steps. First, the computer calculates an initial fit based on the method of least squares. Second, the values are refined by the method of maximum likelihood (Beck & Arnold, 1977; Schleifer & Willis, 1980). In principle, this same approach could be used with other spectrophotometers which are interfaced to a computer. In practice, such an analysis could be quite time consuming, as one must average multiple scans to obtain improved precision. With the spectrophotometer employed here, we can obtain a full spectrum (200–800 nm) in about 0.5 s. We used 10-s measurements to permit averaging of about 20 separate spectra. The multicomponent analysis requires an additional 20 s so that the entire analytical procedure takes less than 1 min.

Multicomponent analysis of the second-derivative spectrum provides a simple, rapid, and accurate method for quantitation of aromatic residues in proteins. It should prove useful in following the chemical modification of these residues. The spectrophotometer can provide multicomponent analysis of up to 12 components, so that other species can also be studied. It may prove possible to quantitate other amino acids such as histidine by analysis of the far-ultraviolet spectrum.

Multicomponent analysis also permits rapid determination of the minimum molecular weight of a protein, or when an estimate of the molecular weight is already available, multicomponent analysis allows refinement of the subunit molecular weight. Both of these estimations are obtained simply by assuming that the number of aromatic residues per molecule must be integral. For example, assume that a solution of a protein at 0.10 mg/mL has a tryptophan concentration of 2.00 μM. The minimum molecular weight would be 50 000. If a previous estimate of molecular weight had been 90 000, then the refined estimate becomes 100 000 with two tryptophans per subunit.

As already noted, the high accuracy of this method requires that the model compounds closely mimic the spectra of their corresponding residues within the denatured proteins. Proteins which retain secondary structure in 6 M guanidine will likely give erroneous results when analyzed with the current method. Lysozyme provides an example. While one could add a sulfhydryl reagent such as dithiothreitol, this will add a significant absorbance in the wavelength region utilized for analysis. This additional absorbance should be essentially removed by the second derivative (Figure 6), but some loss of analytical accuracy may occur. In general, reduction of disulfides should not be required (Tanford, 1968), as shown by the results in Table II. Analytical errors may also result from UV-absorbing impurities in the protein preparation, because the multicomponent analysis routine simply fits a spectrum to match the observed absorption. However, we have also utilized this to our advantage to follow the purification of proteins of known composition.

Covalent modifications or prosthetic groups can also introduce errors in the analysis, if they have significant second-derivative signals. In the case of catalase, the heme group did not interfere with the analysis. When such interference is suspected, it might be eliminated by including its spectrum as an additional component in the multicomponent analysis. Then, the prosthetic group can also be quantitated. Bacterial glutamine synthetase undergoes a covalent modification involving attachment of an adenyl moiety to a tyrosine residue (Stadtman & Ginsburg, 1974). Our preliminary experiments show that inclusion of the spectrum of AMP in guanidine allows accurate determination of the aromatic amino acids and the adenyl group.

References

Aune, K. C., & Tanford, C. (1969) *Biochemistry 8*, 4579–4585.

Balestrieri, C., Colonna, G., Giovane, A., Irace, G., & Servillo, L. (1978) *Eur. J. Biochem. 90*, 433–440.

Balestrieri, C., Colonna, G., Giovane, A., Irace, G., & Servillo, L. (1980) *Anal. Biochem. 106*, 49–54.

Beck, J. V., & Arnold, K. J. (1977) in *Parameter Estimation in Engineering and Science*, pp 154–204, Wiley, New York.

Bio-Rad Laboratories (1977) *Technical Bulletin 1049*, Bio-Rad Laboratories, Richmond, CA.

Blum, H. E., Lehky, P., Kohler, L., Stein, E. A., & Fischer, E. H. (1977) *J. Biol. Chem. 252*, 2834–2838.

Brandts, J. F., & Kaplan, L. J. (1973) *Biochemistry 12*, 2011–2024.

Brown, J. R. (1977) in *Albumin Structure, Function and Uses* (Rosenoer, V. M., Oratz, M., & Rothschild, M. A., Eds.) pp 27–51, Pergamon Press, Oxford.

Butler, W. L., & Hopkins, D. W. (1970) *Photochem. Photobiol. 12*, 439–450.
Chen, R. F. (1967) *J. Biol. Chem. 242*, 173–181.
Coan, C. R., Hinman, L. M., & Deranleau, D. A. (1975) *Biochemistry 14*, 4421–4427.
Dayhoff, M. O. (1978) in *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 2, National Biomedical Research Foundation, Washington, DC.
Deisseroth, A., & Dounce, A. L. (1967) *Arch. Biochem. Biophys. 120*, 671–692.
Donovan, J. W. (1964) *Biochemistry 3*, 67–74.
Edelhoch, H. (1967) *Biochemistry 6*, 1948–1954.
Federici, M. M., & Levine, R. L. (1980) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 39*, 1805.
Foster, J. F., & Sterman, M. D. (1956) *J. Am. Chem. Soc. 78*, 3656–3659.
Giese, A. T., & French, C. S. (1955) *Appl. Spectrosc. 9*, 78–96.
Ginsburg, A., & Carroll, W. R. (1965) *Biochemistry 4*, 2159–2174.
Gratzer, W. B. (1970) in *Handbook of Biochemistry* (Sober, H. A., Ed.) 2nd ed., p B-18, CRC Press, Cleveland, OH.
Grum, F., Paine, D., & Zoeller, L. (1972) *Appl. Opt. 11*, 93–98.
Hamaguchi, K., & Kurono, A. (1963) *J. Biochem. (Tokyo) 54*, 111–122.
Harrison, D. M., & Garratt, C. J. (1969) *Biochem. J. 113*, 733–734.
Hawthorne, A. R., & Thorngate, J. H. (1978) *Appl. Opt. 17*, 724–729.
Hugli, T. E., & Moore, S. (1972) *J. Biol. Chem. 247*, 2828–2834.
Ichikawa, T., & Terada, H. (1977) *Biochim. Biophys. Acta 494*, 267–270.
Imoto, T., Johnson, L. N., North, A. C. T., Phillips, D. C., & Rupley, J. A. (1972) *Enzymes, 3rd Ed. 7*, 665–868.
Levine, R. L. (1982) *J. Chromatogr. 236*, 499–502.
Mihalyi, E. (1968) *J. Chem. Eng. Data 13*, 179–182.
O'Haver, T. C., & Green, G. L. (1976) *Anal. Chem. 48*, 312–318.
Olson, E. C., & Alway, D. C. (1960) *Anal. Chem. 32*, 370–373.
Praissman, M., & Rupley, J. A. (1968) *Biochemistry 7*, 2431–2445.
Rafferty, C. N., Muellenberg, C. G., & Shichi, H. (1980) *Biochemistry 19*, 2145–2151.
Schleifer, A., & Willis, B. G. (1980) *Hewlett-Packard J. 31* (*2*), 11–17.
Schroeder, W. A., Shelton, J. R., Shelton, J. B., Robberson, B., & Apell, G. (1969) *Arch. Biochem. Biophys. 131*, 653–655.
Sela, M., Anfinsen, C. B., & Harrington, W. F. (1957) *Biochim. Biophys. Acta 26*, 503–512.
Shibata, S. (1976) *Angew. Chem., Int. Ed. Engl. 15*, 673–679.
Sia, C. L., & Horecker, B. L. (1968) *Arch. Biochem. Biophys. 123*, 186–194.
Solli, N. J., & Herskovits, T. T. (1973) *Anal. Biochem. 54*, 370–378.
Sophianopoulos, A. J., Rhodes, C. K., Holcomb, D. N., & Van Holde, K. E. (1962) *J. Biol. Chem. 237*, 1107–1112.
Spahr, P. F., & Edsall, J. T. (1964) *J. Biol. Chem. 239*, 850–854.
Stadtman, E. R., & Ginsburg, A. (1974) *Enzymes, 3rd Ed. 10*, 755–807.
Tanford, C. (1968) *Adv. Protein Chem. 23*, 121–282.
Wetlaufer, D. B. (1962) *Adv. Protein Chem. 17*, 303–390.
White, F. H., Jr. (1976) *Biochemistry 15*, 2906–2912.
Willis, B. G. (1980) *Hewlett-Packard J. 31* (*2*), 3–9.