Tack, B. F., Harrison, R. A., Janatova, J., Thomas, M. L., & Prahl, J. W. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 5764-5768.

Unson, C. G., Erickson, B. W., & Hugli, T. E. (1979) in *Peptides: Structure and Biological Function* (Gross, E., & Meinhofer, J., Eds.) pp 459-462, Pierce Chemical, Rockford, IL.

Weigle, W. O., Morgan, E. L., Goodman, M. G., Chenoweth, D. C., & Hugli, T. E. (1982) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 41*, 3099-3103.

Azo Protein Analogues: Synthesis and Characterization of Arsanilazo and Sulfanilazo Derivatives of Tyrosine and Histidine†

Gary J. Pielak,‡ Mickey S. Urdea,§ Kozo Igi,‖ and J. Ivan Legg*

ABSTRACT: The synthesis, purification, and characterization of the products obtained upon reaction of N^{α}-acetyl-L-histidine and *N*-acetyl-L-tyrosine with either diazotized arsanilic or sulfanilic acid were undertaken. The products previously reported to be pure monoazo- and bisazo-L-tyrosine and -L-histidine [Tabachnick, M., & Sobotka, H. (1959) *J. Biol. Chem. 234*, 1726] are shown to consist of a number of chromophoric species. The visible spectra described here differ substantially from the earlier report. To date, the use of the visible spectral properties of these azo protein analogues to determine the extent of histidine and tyrosine modification of proteins with diazotized arsanilic or sulfanilic acid has been in error. From the model studies reported here, it is now possible to characterize the corresponding azo proteins. It is also possible to distinguish spectrally between the C-2 and C-4 azo-coupled products of the histidine imidazole. Characterization of azo proteins is essential to metal-azo protein incorporation studies under investigation.

Simple and systematic methods for the incorporation of substitution-inert metal ions into proteins are under investigation in this laboratory. A particularly promising method involves the production of an azo dye chelating agent on the protein with a diazonium salt followed by incorporation of Co(III). We have successfully applied this methodology to carboxypeptidase A azo coupled[1] at tyrosine-248 (Urdea & Legg, 1979a,b) and have reported preliminary results for several other proteins (Urdea et al., 1979). As a first step toward making this method generally applicable, we have undertaken a reinvestigation of protein azo coupling.

Typically, determination of the extent and specificity of diazonium salt incorporation has been based on the visible spectral properties (band position and molar absorptivities) of the azo proteins as compared to simple azo amino acids (Tabachnick & Sobotka, 1960).[2] We have found that for characterization of proteins modified with either diazotized arsanilic or sulfanilic acid, this technique is highly erroneous since the compounds originally reported to be homogeneous are mixtures of chromophoric species. These compounds have now been purified and characterized, and it is possible to distinguish spectrally between C-2 and C-4 azo-coupled histidine.

Experimental Procedures

Preparation of Arsanilazo Derivatives of N-Acetyl-L-tyrosine. The method of synthesis was similar to that described by Tabachnick & Sobotka (1959). Arsanilic acid (Aldrich Chemicals, 0.651 g) was dissolved in 50.0 mL of 0.30 N HCl at 0 °C. Sodium nitrite (0.311 g) and sodium bromide (0.062 g) were added, and the solution was stirred for 30 min. The diazonium salt was added over a 30-min period to 25.0 mL of 0.120 M *N*-acetyl-L-tyrosine (Vega) and 0.01 M borate, pH 9.0 at 0 °C, and 1 N NaOH was added as necessary to keep the pH between 9.0 and 9.5. The reaction was stirred for an additional 3.5 h, turning first yellow then yellowish brown.

Isolation of N-Acetylmono(arsanilazo)-L-tyrosine (NAMAAT).[3] The above solution was acidified to pH 1.8 with 1 N HCl. The yellow-orange precipitate was filtered and washed with a small portion of cold water. The product (ca. 150 mg) was dissolved in 1 mL of the 50/50 ISTEA solvent

† From the Department of Chemistry and the Program in Biochemistry and Biophysics, Washington State University, Pullman, Washington 99164. *Received March 28, 1983; revised manuscript received September 19, 1983.* This study was supported by the National Institutes of Health (GM 23081). The Nicolet 200-MHz spectrometer was acquired with the assistance of the Boeing Co. M.S.U. was supported in part by a C. Glenn King Fellowship, Washington State University. Studies reported were submitted by G.J.P. in partial fulfillment of the requirements for the Ph.D. degree of Washington State University.

‡ Present address: Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.

§ Present address: Chiron Corp., Emeryville, CA 94608.

‖ Present address: Callery Chemical Co., Callery, PA 16024.

[1] The terms "diazotization" and "diazotized" have been used in previous work to indicate both the formation and coupling of diazonium salts. However, in this paper, in accordance with current usage, these terms refer only to diazonium salts while "azo coupling" and "coupled" refer to the reaction of diazonium salts.

[2] The equations of Tabachnick & Sobotka (1960) later appeared (and were cited) in Riordan & Vallee (1972).

[3] Abbreviations: NAMAAT, *N*-acetylmono(arsanilazo)-L-tyrosine; NABAAT, *N*-acetylbis(arsanilazo)-L-tyrosine; NAMAAH(C-2), N^{α}-acetyl-2-mono(arsanilazo)-L-histidine; NAMAAH(C-4), N^{α}-acetyl-4-mono(arsanilazo)-L-histidine; NABAAH, N^{α}-acetylbis(arsanilazo)-L-histidine; NAMSAT, *N*-acetylmono(sulfanilazo)-L-tyrosine; NABSAT, *N*-acetylbis(sulfanilazo)-L-tyrosine; NAMSAH(C-2), N^{α}-acetyl-2-mono(sulfanilazo)-L-histidine; NAMSAH(C-4), N^{α}-acetyl-4-mono(sulfanilazo)-L-histidine; NABSAH, N^{α}-acetylbis(sulfanilazo)-L-histidine; ISTEA, distilled 2-propanol and triethylamine bicarbonate, pH 9.5; ^{1}H NMR, proton nuclear magnetic resonance; TLC, thin-layer chromatography.

0006-2960/84/0423-0589$01.50/0

described by Warner & Legg (1979), applied to a 20 × 20 cm Whatman PR1F preparative TLC plate, and eluted with 50/50 ISTEA (8–10-h elution time). The R_f values (Table I) vary somewhat with eluant composition and are therefore reported for the same batch of ISTEA. After being dried at room temperature, the silica containing the yellow product was removed and extracted with two 50-mL portions of water. The material was precipitated as described above and powdered with anhydrous ether. The recrystallization was repeated. From 450 mg of crude reaction material, 240 mg of pure NAMAAT was obtained. Anal. Calcd for $C_{17}H_{18}N_3O_7As_1$: C, 45.24; H, 4.03; N, 9.31. Found: C, 45.35; H, 4.27; N, 9.46.

Isolation of N-Acetylbis(arsanilazo)-L-tyrosine (NABAAT). NABAAT was synthesized as described for NAMAAT except the quantities of arsanilic acid and sodium nitrite were doubled. Since even with excess diazonium salt NABAAT represented less than 5% of the reaction mixture, interfering low R_f side products made its purification by preparative TLC impossible. Partial purification was effected by repeated fractional precipitation with silver nitrate. The reaction mixture was acidified to pH 2 with concentrated HNO_3, 0.4 M $AgNO_3$ was added dropwise, and AgCl was removed by centrifugation. This regimen was repeated until the color of the precipitate changed from white (AgCl) to black. An excess of $AgNO_3$ (50 mL of 0.4 M) was then added, and the black precipitate was collected by centrifugation. The pH of the supernatant was readjusted to 2 with 0.1 N HNO_3, and the solution was evaporated under a stream of air until more black precipitate formed. The precipitate was removed and the process repeated until the supernatant yielded no more precipitate.

To the combined precipitates were added 10 mL of water and sufficient 0.3 N NaOH to make the solution pH 8. This caused the azo compounds to redissolve while precipitating AgOH, which was removed by centrifugation. The pH was adjusted to 2, and the fractional precipitation with $AgNO_3$ was repeated. The final supernatants were also combined, concentrated, and subjected to $AgNO_3$ precipitation. The entire process resulted in enrichment of NABAAT as evidenced by TLC in 70/30 ISTEA. A total of 400 mg of black precipitate was obtained after about 30 fractional precipitations.

The product was divided into two equal portions, and each was dissolved in 15 mL of 0.25 M $NaHCO_3/Na_2CO_3$ buffer, applied separately to the top of a 5 × 55 cm G-10 (Pharmacia) column, and eluted with the same buffer. Ag_2CO_3 remained at the top of the column while the NABAAT was eluted as a purple band that was well resolved from other colored impurities.

Carbonate was removed from the pure NABAAT by batching the 90 mL of column effluent with the hydrogen-ion form of Dowex 50W-X8 (three 5-mL portions). The olive drab microcrystalline NABAAT obtained by aerial evaporation of the solution (ca. pH 3) was removed by centrifugation, washed with cold 1 N HCl and then cold water, powdered in diethyl ether, and dried in vacuo over P_2O_5. The pure NABAAT migrated as a single purple spot upon TLC in 50/50, 60/40, and 70/30 ISTEA. The yield was 65 mg. Anal. Calcd for $C_{23}H_{25}N_5O_{11}As_2$: C, 39.61; H, 3.61; N, 10.05. Found: C, 39.66; H, 3.44; N, 9.84.

Preparation of Arsanilazo Derivatives of N^{α}-Acetyl-L-histidine. Sodium bromide (0.248 g) and arsanilic acid (2.60 g) were dissolved in 60 mL of 0.5 N HCl, and $NaNO_2$ (0.828 g) was added over a period of 10 min at 0 °C. After being stirred for 50 min, the solution was made up to 200 mL and added to a solution of 2.58 g of N^{α}-acetyl-L-histidine monohydrate (Vega) in 200 mL of 0.01 M sodium borate over a 30-min period with constant stirring at 0 °C. The pH was kept between 9.0 and 9.5 with 1 N NaOH. After being stirred for 60 min, the solution (which had changed from yellow to orange red) was concentrated to ca. 30 mL with a rotary evaporator at 34 °C and then chromatographed with 0.01 M KOH in batches on Sephadex LH-20 (6 mL on a 5 × 47 cm column). A red band and then a yellow band constituting the major portion of the reaction mixture were collected.

Isolation of N^{α}-Acetyl-2- and N^{α}-Acetyl-4-mono(arsanilazo)-L-histidine [NAMAAH(C-2) and NAMAAH(C-4)]. To the combined yellow bands adjusted to pH 9.2 with 0.1 M HCl was added $CaCl_2 \cdot 2H_2O$ (5.2 g). Upon concentration to 50 mL with a rotary evaporator, a gelatinous substance gradually formed. This material (fraction 1) was filtered and washed with ethanol and ether. Two more fractions (2 and 3) were isolated in an analogous manner by further evaporation of the filtrate to 20 and 10 mL, respectively. A final fraction was isolated after addition of 5 mL of ethanol. Generally, fraction 1 consists of NAMAAH(C-2), and fraction 4 is predominantly NAMAAH(C-4). Fractions 2 and 3 usually contained both isomers, which could be separated as described above. The two isomers are formed in approximately equal quantities.

The NAMAAH(C-2) fraction was dissolved in a small amount of water and passed through a cation-exchange column (Dowex 50W-X4, Ba^{2+} form). The solid obtained by evaporation of the eluant to dryness was recrystallized from hot water. The orange-yellow crystals were washed with 50% ethanol, ethanol, and ether. Anal. Calcd for $C_{28}H_{26}O_{12}N_{10}As_2Ba_3 \cdot 8H_2O$: C, 24.01; H, 3.02; N, 10.00. Found: C, 24.49; H, 3.01; N, 10.05. The NAMAAH(C-4) fraction was recrystallized from water by the addition of ethanol, washed as described for NAMAAH(C-2), and dried in vacuo in a drying pistol (boiling ethanol). Anal. Calcd for $C_{28}H_{26}O_{12}N_{10}As_2Ca_3 \cdot 2C_2H_5OH \cdot 11H_2O$: C, 30.62; H, 4.82; N, 11.16. Found: C, 30.37; H, 4.58; N, 11.37. The ethanol content was quantified by ^1H NMR.

Isolation of N^{α}-Acetylbis(arsanilazo)-L-histidine (NABAAH). To the combined red bands was added 10.6 g of $CaCl_2 \cdot 2H_2O$. Aerial evaporation of the solution to 20 mL yielded 2.7 g of crude product, which was washed with ethanol and ether. The crude product (1.5 g) was dissolved in 300 mL of water, filtered, and passed through a Dowex 50W-X4 (Na^+ form) column. The eluant was concentrated to ca. 35 mL. Addition of $CaCl_2 \cdot 2H_2O$ (1 g in 5 mL of water) caused immediate precipitation. The precipitate was washed with 50% ethanol, ethanol, and ether. Anal. Calcd for $C_{20}H_{17}O_9N_7As_2Ca_2 \cdot 7H_2O$: C, 28.08; H, 3.65; N, 11.46. Found: C, 28.26; H, 3.57; N, 11.10.

Preparation and Isolation of N-Acetylmono(sulfanilazo)-L-tyrosine (NAMSAT). NAMSAT was synthesized as described for the corresponding arsanilazo derivative by employing 0.521 g of sulfanilic acid (Eastman Organic Chemicals). The diazonium salt of sulfanilic acid precipitates upon formation and must be stirred vigorously. The reaction mixture was rotary evaporated to dryness and then powdered with diethyl ether. A portion of the crude powder (700 mg of 2.6 g) dissolved in 5 mL of 0.15 M $(NH_4)HCO_3$, pH 8.0, was applied to a 40 × 5 cm column of cellulose anion exchanger (Whatman DE-52) equilibrated in the same buffer. Elution with the same buffer yielded at least six bands. The major red band was rotary evaporated and powdered with ether. TLC in 70/30 and 80/20 ISTEA showed one major yellow spot with small amounts of colored impurities. The

powder was dissolved in 5 mL of buffer and rechromatographed. The pure (by TLC) compound was dissolved in 6 mL of water and passed through a cation-exchange column (Dowex 50W-X8, Ba^{2+} form). The colored effluent was reduced to 20 mL by rotary evaporation. Upon addition to 20 mL of absolute ethanol, a gelatinous, burnt orange precipitate formed, which was collected by centrifugation and washed twice with ethanol. The precipitate was dissolved in 20 mL of warm water and filtered, and 20 mL of ethanol was added. The precipitate that formed when the solution was allowed to stand at 4 °C overnight was washed with 50% ethanol, 95% ethanol, and anhydrous ether and desiccated in vacuo over P_2O_5 overnight and further dried in vacuo in a drying pistol (boiling ethanol). TLC in 50/50, 60/40, 70/30, and 80/20 ISTEA established purity. The yield was 15 mg. Anal. Calcd for $C_{17}H_{15}H_3O_7SBa \cdot 2H_2O$: C, 35.28; H, 3.32; N, 7.26. Found: C, 35.40; H, 3.26; N, 7.04.

Preparation and Isolation of N-Acetylbis(sulfanilazo)-L-tyrosine (NABSAT). Sulfanilic acid (12.4 g) was slurried in 1 L of 70 mM HCl at 0 °C, $NaNO_2$ (4.98 g) and NaBr (0.25 g) were added, and the solution was stirred for 30 min. The solution was added over 40 min to a 200-mL solution of 0.173 M *N*-acetyl-L-tyrosine and 0.01 M borate, pH 9.0 at 0 °C. The pH was kept between 8.5 and 9.0 with 6 N NAOH. The solution was stirred for an additional 4 h at 4 °C and turned from orange to black. Both TLC and spectral analysis indicated NABSAT represented at least 50% of the azo products.

Approximately 10 mL of the reaction mixture was applied to a 35 × 2 cm column of the hydroxide-ion form of (aminoethyl)cellulose (Sigma). The column was eluted with 0.05 M $(NH_4)HCO_3$ until the purple band reached the bottom, and the eluant was changed to 0.1 M $(NH_4)HCO_3$. The band was collected, concentrated to 10 mL by rotary evaporation, applied to a 10 × 0.5 cm column of barium-ion form Dowex 50W-X8, and eluted with H_2O. The main yellow band was collected. Ethanol was added to the yellow solution (the color of the desired product changes dramatically with pH, vide infra) until it became cloudy, and after 5 h at −20 °C, a precipitate formed. The precipitate was collected by centrifugation, dissolved in a few milliliters of H_2O, and applied to a 55 × 5 cm column of Sephadex G-10. Upon elution with H_2O, a fast-eluting brown band was partially resolved from the desired, slower eluting yellow band. Pure (by TLC in 80/20 ISTEA) fractions were pooled and rotary evaporated to a few milliliters. Pure NABSAT was precipitated by addition of ethanol, collected by centrifugation, washed with ethanol and anhydrous ether, and dried in vacuo. TLC in 70/30 and 80/20 ISTEA established purity. Anal. Calcd for $C_{23}H_{19}N_5O_{10}S_2Ba_1 \cdot 6H_2O \cdot 0.1EtOH$: C, 31.82; H, 3.65; N, 8.00. Found: C, 32.09; H, 3.49; N, 7.97. The structure was confirmed and the ethanol content quantified by ^1H NMR.

Preparation of the Monosubstituted Sulfanilazo Derivatives of N^{α}-Acetyl-L-histidine. Sodium bromide (0.25 g) and sulfanilic acid (5.98 g) were added to 600 mL of 0.12 M HCl, and $NaNO_2$ (2.38 g) was slowly added at 0 °C with stirring. After 30 min, the stirred slurry was added to a stirred solution of 7.74 g (34.5 mmol) of N^{α}-acetyl-L-histidine sesquihydrate (Sigma) in 200 mL of 0.01 M borate, pH 9.0, over 40 min at 0 °C. The pH was maintained between 8.5 and 9.0 with 6 N NaOH. The reaction was stirred for 1 h at 0 °C and turned from yellow to red. N^{α}-Acetylbis(sulfanilazo)-L-histidine (NABSAH) was removed by batching the reaction solution with six equal portions of Whatman DE-52 (180 g wet wt total) that had been equilibrated at pH 2 with HCl. The solution was readjusted to pH 2 and the process repeated. This procedure removed over 90% of the NABSAH as evidenced by TLC in 70/30 ISTEA. The volume was reduced to 200 mL by rotary evaporation.

Isolation of N^{α}-Acetyl-2- and N^{α}-Acetyl-4-mono(sulfanilazo)-L-histidine [NAMSAH(C-2) and NAMSAH(C-4)]. A 3-L linear gradient from 0.05 M $(NH_4)HCO_3$, pH 7.9, to 0.25 M $(NH_4)_2CO_3$, pH 8.9, on a 14 × 5 cm Whatman DE-52 column, equilibrated in the starting eluant, was used to separate NAMSAH(C-2) from NAMSAH(C-4). About 10 mL of the NABSAH-depleted reaction mixture (described above) was adjusted to pH 7.9 with NaOH and chromatographed. The major fast-eluting yellow band was collected in two fractions. TLC in 70/30 ISTEA and visible absorption spectroscopy indicated that the front of the band was primarily NAMSAH(C-4) and the end was primarily NAMSAH(C-2). This procedure was repeated 6 times on the original reaction mixture; the appropriate eluants were pooled and rotary evaporated to dryness. The two pooled fractions were each dissolved in 10 mL of 0.05 M $(NH_4)HCO_3$ and rechromatographed as described above. This time the sample was completely resolved; the pale NAMSAH(C-4) eluted first and the dark yellow NAMSAH(C-2) second. The two 500-mL fractions were rotary evaporated to dryness.

The NAMSAH(C-4) fraction in 20 mL of water was batched with the hydrogen-ion form of Dowex 50W-X8 (15 g) to remove carbonate. When bubbling ceased, the solution was passed through a Ba^{2+} Dowex 50W-X8 column and rotary evaporated to dryness. The yellow glass (209 mg) was dissolved in D_2O (0.5 mL) and the identity of the isomer confirmed by ^1H NMR. The compound was precipitated with ethanol, redissolved in H_2O, and reprecipitated. The process was repeated twice to ensure removal of D_2O. The powder was washed with absolute ethanol and anhydrous ether and dried in a drying pistol (boiling ethanol). Anal. Calcd for $C_{14}H_{13}N_5O_6S_1(Ba_{0.70}H_{0.60}) \cdot 3.5H_2O$: C, 31.18; H, 3.85; N, 12.99; Ba, 17.8. Found: C, 31.37; H, 3.41; N, 12.81; Ba, 17.8. The water content was quantified by ^1H NMR.

The NAMSAH(C-2) fraction was treated in the same manner as NAMSAH(C-4). Anal. Calcd for $C_{14}N_{13}N_5O_6S_1(Ba_{0.95}H_{0.10}) \cdot 4.25H_2O \cdot 0.2EtOH$: C, 29.03; H, 3.87; N, 11.76; Ba, 21.9. Found: C, 29.05; H, 3.36; N, 11.72; Ba, 21.7. The water and ethanol content was quantified by ^1H NMR.

Preparation and Attempted Isolation of NABSAH. Synthesis was as for the mono(sulfanilazo) derivatives of N^{α}-acetyl-L-histidine except that twice the amount of sulfanilic acid and $NaNO_2$ was used in 920 mL of 0.12 M HCl. Chromatography on (aminoethyl)cellulose (Sigma) at pH 7.8 with water (to remove mono derivatives) followed with 0.1–0.2 M $(NH_4)HCO_3$ produced a major red band. Preparative TLC (Whatman PLK-5, 80/20 ISTEA) resolved the red band. Two-dimensional analytical TLC in 70/30 ISTEA confirmed that the red band was slowly decomposing. However, the visible spectrum of the product and its pH dependence were very similar to that of NABAAH.

Determination of pK_as of the Azo Amino Acids. The pK_as of the various azo derivatives were determined by spectrophotometric titration in water. The absorbance was measured as a function of pH at the wavelength found to give the maximum change in absorption. pH was measured with a pH stat designed in this laboratory (Warner et al., 1980) in conjunction with a Radiometer GK 2321C electrode.

Spectral Analysis. Visible absorption spectra were recorded on either a Cary Model 14 spectrophotometer, a Varian Superscan 3, or a Cary 219. ^1H NMR spectra were obtained

Table I: TLC Results for ISTEA Solvent System

compd	R_f in 50/50 ISTEA	compd	R_f in 70/30 ISTEA
N-acetyl-L-tyrosine	0.87[a]	*N*-acetyl-L-tyrosine	0.72[a]
N^{α}-acetyl-L-histidine	0.81[a]	N^{α}-acetyl-L-histidine	nd[b]
NAMAAT	0.71	NAMSAT	0.61
NAMAAH(C-2)	0.68	NAMSAH(C-2)	0.58
NAMAAH(C-4)	0.63	NAMSAH(C-4)	0.53
NABAAH	0.49	NABSAH	not stable
NABAAT	0.44	NABSAT	0.49
self-coupling product of diazotized arsanilic acid	0.31	self-coupling products of diazotized sulfanilic acid	0.36, 0.11

[a] Visualized with Pauly spray. All other compounds visible under room lights. [b] nd, not determined.

Table II: Summary of Absorption Spectral Properties of Azotyrosines

compd	pH (in H_2O)	ionization state	band position (nm) [molar absorptivity ($\times 10^{-4}$ M^{-1} cm^{-1})] maximum	shoulder/minimum	maximum	isosbestic point
NAMAAT[a]	7	phenol	326 (2.55)	(sh) 380 (1.04)		418 (0.738)
	11	phenolate	327 (1.68)	(min) 386 (0.410)	485 (1.20)	
mono(arsanilazo)(chloroacetyl)tyrosine[b]	6.2	phenol	325 (2.22)			417 (0.620)
	0.1 N NaOH	phenolate	328 (1.43)		490 (1.05)	
NAMSAT[a]	7	phenol	326 (2.16)	(sh) 380 (0.840)		419 (0.579)
	11	phenolate	328 (1.33)	(min) 383 (0.327)	510 (0.882)	
mono(sulfanilazo)(chloroacetyl)tyrosine[b]	6.2	phenol	325 (2.21)			~420
	0.1 N NaOH	phenolate	330 (1.40)		490 (1.15)	
NABAAT[a]	6	phenol	329 (3.92)	(sh) 420 (1.25)		467 (0.842)
	11	phenolate	317 (2.77)	(min) 410 (0.649)	548 (2.20)	
bis(arsanilazo)(chloroacetyl)tyrosine[b]	6.2	phenol	330 (3.56)	(sh) 420 (1.14)		~465
	0.1 N NaOH	phenolate	325 (2.60)		545 (1.75)	
NABSAT[a]	6	phenol	332 (3.79)	(sh) 420 (1.22)		470 (0.787)
	11	phenolate	318 (2.53)	(min) 410 (0.430)	554 (2.09)	

[a] From this study. [b] From Tabachnick & Sobotka (1959).

Table III: pK_a Values of Arsanilazo and Sulfanilazo Derivatives of Histidine and Tyrosine

compd	pK_a (in H_2O)	ionization
N-acetyltyrosine[a]	10.1	phenol, phenolate
NAMAAT	9.4	
NAMSAT	9.6	
NABAAT	7.8	
NABSAT	8.4	
N^{α}-acetylhistidine[a]	7.5	imidazolium, imidazole
NAMSAH(C-2)	3.6	
NAMSAH(C-4)	4.1	
N^{α}-acetylhistidine[a]	14.5	imidazole, imidazolate
NAMAAH(C-2)	11.4	
NAMSAH(C-2)	11.0	
NAMAAH(C-4)	11.9	
NAMSAH(C-4)	11.6	
NABAAH	8.8	

[a] From Perrin (1965).

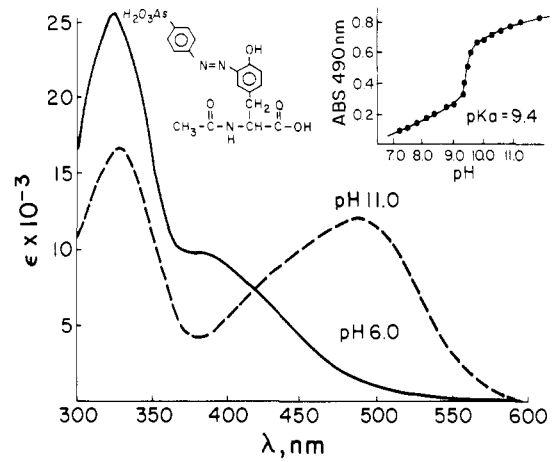

FIGURE 1: Visible spectral properties of NAMAAT. The inset shows the change in absorbance at 450 nm with pH used to determine the pK_a of the azophenol.

on either a JEOL JNM-MH-100 or a Nicolet NT-200 instrument. The sodium salt of 3-(trimethylsilyl)propanesulfonic acid was employed as an internal reference.

Chemical Analysis. Analyses for C, H, and N were usually performed in duplicate by either Canadian Microanalytical Services, Limited (Vancouver, B.C., Canada) or the C, H, and N Analytical Facility at the University of Idaho (Moscow, ID). Barium analysis was performed on a Perkin-Elmer 303 atomic absorption spectrophotometer equipped with a DCR1 time-averaging digital display by using a nitrous oxide–acetylene flame and a 2000 ppm potassium (as KCl) radiation buffer. ^1H NMR was used to quantify bound water and ethanol by the standard addition method. A calibration curve was made by adding known amounts of glass-distilled H_2O or absolute ethanol to 99.8% D_2O (Aldrich Chemicals).

Results

When *N*-acetyl-L-tyrosine or *N*-acetyl-L-histidine is coupled with either the diazonium salt of arsanilic acid or sulfanilic acid as described by Tabachnick & Sobotka (1959), both the mono- and bis-coupled products are obtained. The yield of bis product can be increased by doubling the quantity of diazonium salt. Pure products are isolated by fractional precipitation and various forms of chromatography. The TLC system of Warner & Legg (1979) was very important in establishing the purity of the products (Table I).

The absorption spectral data for the two ionization states of the azotyrosine derivatives reported by Tabachnick & So-

Table IV: ^{1}H NMR Spectral Data for the Histidine Derivatives

compd	chemical shift (ppm)				
		CH of imidazole			
	CH_3-	C-2	C-4	phenyl[d]	pH[e]
N^{α}-acetylhistidine[a]	2.03	8.61[c]	7.29		5.10
NAMAAH(C-2)[a]	1.98		7.18	7.92	9.05
NAMSAH(C-2)[a]	2.41		7.49	8.14	7.26
NAMAAH(C-4)[a]	1.73	7.73[c]		7.92	9.05
NAMSAH(C-4)[a]	2.19	8.92[c]		8.15	7.26
NABAAH[a]	1.75			7.79	8.37
peak integration	3	1	1	4	
N^{α}-acetyl-2-[(*p*-bromophenyl)azo]histidine methyl ester[b]			7.20		
N^{α}-acetyl-2-[(*p*-carbomethoxyphenyl)azo]histidine methyl ester[b]			7.30		
N^{α}-acetyl-4-[(*p*-bromophenyl)azo]histidine methyl ester[b]		7.86			
N^{α}-acetyl-4-[(*p*-carbomethoxyphenyl)azo]histidine methyl ester[b]		7.83			

[a] From this study. [b] From Nagai et al. (1973). Spectra obtained in Me_2SO-d_6. [c] Susceptible to rapid deuterium exchange. [d] The phenyl protons of the arsanilic acid group appear as a singlet under these conditions and as a quartet for the sulfanilic acid group. [e] No correction was made for the difference between the activities of hydrogen and deuterium ions.

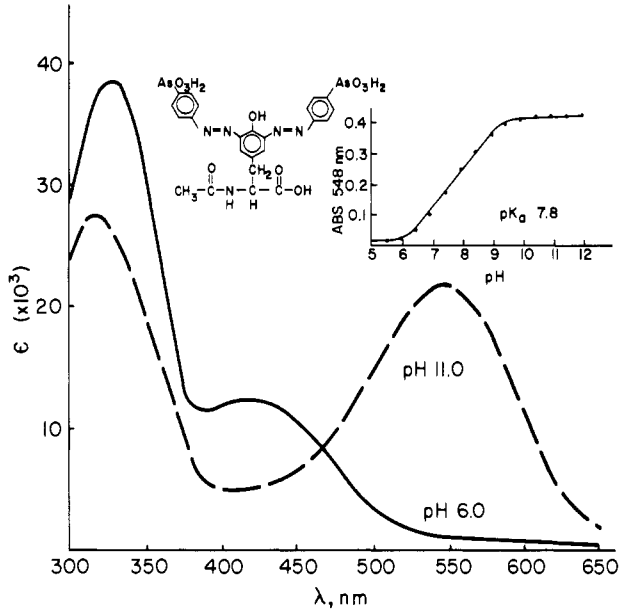

FIGURE 2: Visible spectral properties of NABAAT. The inset shows the change in absorbance at 550 nm with pH used to determine the pK_a of the azophenol.

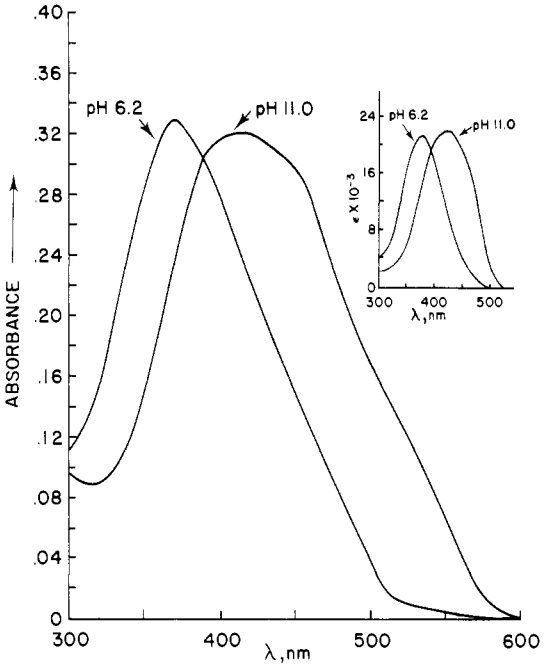

FIGURE 3: Visible absorption spectrum of the crude reaction mixture of diazotized arsanilic acid and N^{α}-acetylhistidine prior to chromatography. The inset shows the spectrum of mono(arsanilazo)histidine reported by Tabachnick & Sobotka (1959).

botka (1959) and obtained in this study are given in Table II. The corresponding pK_a values, as determined in this study, are given in Table III.

The pH dependence of the visible spectrum of the NAMAAT is shown in Figure 1. In water at pH 6.0, the azophenol spectrum is characterized by a 324-nm maximum and a 390-nm shoulder. With increasing pH, a peak builds in at 485 nm, representing the azophenolate form (pK_a = 9.4). Concomitantly, the 390-nm shoulder is lost, and the 324-nm peak is shifted to a lower intensity band at 328 nm. Similar but quantitatively different spectra and pH dependence are observed for NAMSAT.

A 329-nm band and a broad 412-nm absorption characterize the spectrum of NABAAT below pH 7.0 (Figure 2). At higher pH, the 329-nm band decreases in intensity and shifts to 317 nm. The 412-nm band is lost while a new maximum appears at 547 nm, indicative of the azophenolate (pK_a = 7.8). Again similar, but quantitatively different spectra and spectral changes are found for NABSAT.

The method of Tabachnick & Sobotka (1959) was used to synthesize the mono(arsanilazo) and mono(sulfanilazo) derivatives of histidine. The mixture obtained possessed essentially the same spectral properties as those reported in the earlier study (Figure 3 and inset). However, upon TLC of these materials (Table I), a number of compounds were detected.

The pure 2- and 4-azohistidine derivatives isolated in this study were distinguished by ^{1}H NMR spectroscopy (Table IV). At pH 5.1 (see footnote *e* of Table IV), the C-4 and C-2 methine resonances of N^{α}-acetyl-L-histidine are at 7.29 and 8.61 ppm, respectively, in close agreement with Sachs et al. (1971). The spectrum of NAMAAH(C-2) retains a high-field resonance at 7.18 ppm (C-4 methine), and that of NAMAAH(C-4) retains a low-field resonances at 7.73 ppm (C-2 methine). The same pattern is found for the corresponding sulfanilazo derivatives. Nagai et al. (1973) found similar spectral changes in the C-2 and C-4 methine resonances of C-2 and C-4 azo derivatives of N^{α}-acetylhistidine methyl ester (Table IV).

Unambiguous assignment was possible in the present study due to the differential deuteration rates of the C-2 and C-4 methines. The C-2 methine of histidine is deuterated by

Table V: Summary of Absorption Spectral Properties of Azohistidines

compd	pH (in H_2O)	ionization state	band position (nm) [molar absorptivity ($\times 10^{-4}$ M^{-1} cm^{-1})] maximum	isosbestic point
NAMAAH(C-2)[a]	6	imidazole	389 (2.54)	397 (2.36)
	13.7	imidazolate	440 (2.84)	
NAMAAH(C-4)[a]	6	imidazole	352 (2.75)	365 (2.48)
	13.0	imidazolate	386 (3.13)	
NAMAAH[b]	6.2	imidazole	370 (2.16)	
	0.1 N NaOH	imidazolate	420 (2.23)	
NAMSAH(C-2)[a]	1.5	imidazolium	364 (2.75)	382 (2.39)
	6	imidazole	386 (2.40)	400 (2.21)
	13.1	imidazolate	443 (2.88)	
NAMSAH(C-4)[a]	1.5	imidazolium	337 (2.32)	347 (2.02)
	6	imidazole	351 (2.03)	366 (1.71)
	13.0	imidazolate	388 (2.14)	
NAMSAH[b]	6.2	imidazole	370 (2.24)	
	0.1 N NaOH	imidazolate	423 (2.34)	
NABAAH[a]	6	imidazole	419 (3.18)	448 (2.84)
	11	imidazolate	498 (4.09)	

[a] From this study. [b] From Tabachnick & Sobotka (1959).

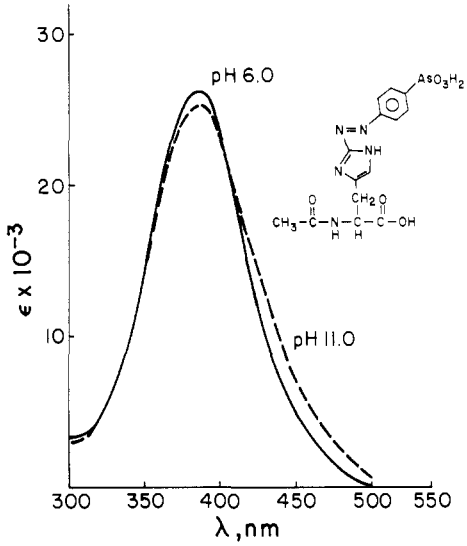

FIGURE 4: Visible spectral properties of NAMAAH(C-2).

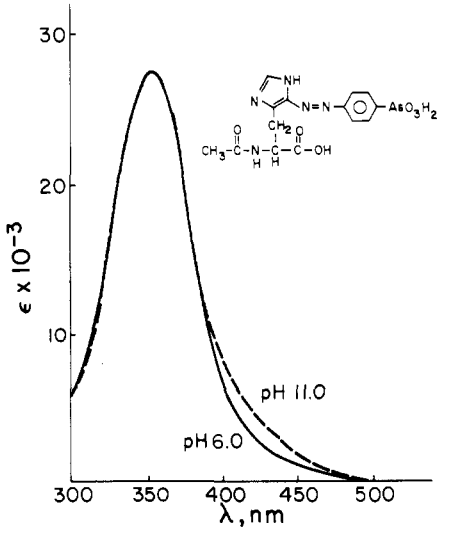

FIGURE 5: Visible spectral properties of NAMAAH(C-4).

heating under neutral or basic conditions, whereas the C-4 is much slower to exchange (Matsuo et al., 1972). A solution of NAMAAH(C-4) in D_2O was heated in 75 °C, pH 9.05. After 4 h, the intensity of the 7.73 ppm peak decreased by 75%,

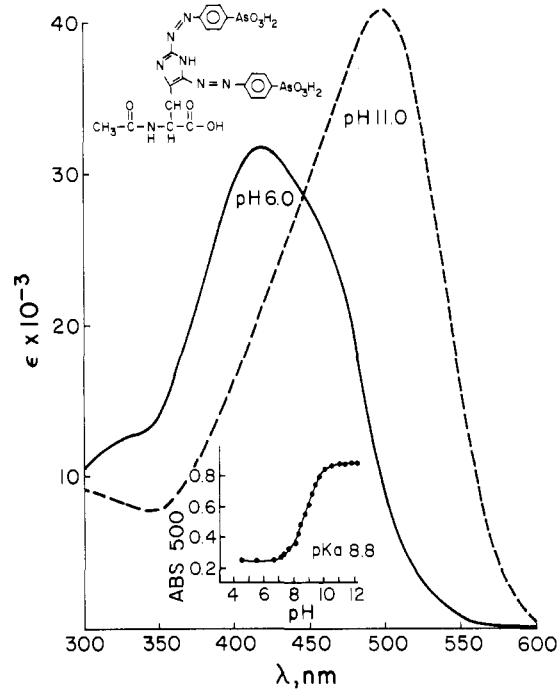

FIGURE 6: Visible spectral properties of NABAAH. The inset shows the change in absorbance at 500 nm with pH used to calculate the pK_a of the imidazole.

consistent with the assignment of this resonance as the C-2 methine. Heating a basic D_2O solution of NAMAAH(C-2) for 4 h did not decrease the intensity of the 7.18 ppm resonance (Table IV).

Three states of ionization are available to the imidazole ring of histidine: the positively charged imidazolium, the neutral imidazole, and the negatively charged imidazolate. These states are distinguished by spectrophotometric titration (Table V) and are characterized by isosbestic behavior and reversibility. The pK_a values are listed in Table III.

The 2- and 4-monoazohistidines differ significantly in their visible spectral properties (Table V, Figures 4 and 5). For both derivatives, the spectra are relatively independent of pH between 6 and 11, consistent with the fact that the imidazolium pK_a is lowered significantly upon azo coupling (Table III). The spectral behavior of NAMSAH(C-2) and NAMSAH(C-4) exhibits the same trend although there are quantitative differences (Tables III and V).

Table VI: Comparison of Molar Absorptivities Used in the Equations of Tabachnick and Sobotka[a] with Those Determined in This Study

compd	molar absorptivity ($\times 10^{-4}$ M^{-1} cm^{-1}) in 0.1 N NaOH	
	460 nm	500 nm
mono(arsanilazo)(chloroacetyl)tyrosine[b]	0.960	1.05
NAMAAT	1.00	1.14
NAMAAH[b]	1.65	0.256
NAMAAH(C-4)	0.629	0.0458
NAMAAH(C-2)	2.34	0.355

[a] From Tabachnik & Sobotka (1960). [b] From Tabachnick & Sobotka (1959).

In contrast to the mono(arsanilazo)histidines, the NABAAH spectrum is strongly pH dependent (Figure 6) since the imidazolate pK_a is lowered to 8.8 (Table III). The 419-nm maximum at pH 6 for imidazole gives rise to a 498-nm maximum at pH 11 for the imidazolate (see Tables III and V). It is the presence of NABAAH that results in the apparent pH dependence of the mono(arsanilazo)histidine spectrum as reported by Tabachnick & Sobotka (1959; see Figure 3).

Discussion

The major goal of this study was to isolate pure solids of all the arsanilazo and sulfanilazo derivatives of *N*-acetyl-L-tyrosine and N^{α}-acetyl-L-histidine in quantities sufficient for unambiguous structural assignments. This was imperative to the development of a spectral technique for the characterization of azo proteins. The isolation of pure solids proved exceedingly difficult since the combination of ionic and aromatic properties causes them to act like detergents. No effort was made to maximize yields.

Tabachnick & Sobotka (1959) reported elemental analyses of the presumed mono- and bis(arsanilazo)tyrosine derivatives that appear very good despite the fact that these are heterogeneous samples. This is due to the manner in which the compounds were isolated. After complete reaction of *stoichiometric* quantities of the diazotized arsanilic acid and N-blocked tyrosine, essentially everything in solution was precipitated in acid. Even though the precipitate contained the N-blocked tyrosine, monoazotyrosine, bisazotyrosine, and the self-coupling product, the elemental analyses were consistent with the monoazo product as a result of the ratio of starting materials. Similarly, acid precipitation of the reaction mixture containing a 2-fold excess of the diazonium salt and the N-blocked tyrosine also produced a solid with elemental composition consistent with the bis(arsanilazo)tyrosine derivative. Chemical analyses for the histidine derivatives were not reported, and the molar absorptivities were calculated from the concentration of the limiting reactant.

In the method of Tabachnick & Sobotka (1960), two equations employing molar absorptivities at two different wavelengths as constants are used to calculate the monoazotyrosine and monoazohistidine content in a modified protein from absorption spectral data at these wavelengths. Several assumptions must be shown valid for this method to yield correct results. The most important is that the molar absorptivities from the azo amino acid models are representative of the species presumed to be detected in the azo proteins. Since the azo amino acid models used by Tabachnick and Sobotka were not pure, this assumption is not valid. Tabachnick & Sobotka (1959) also did not consider the presence of the two structural isomers of monoazohistidine in their equations. Furthermore, as shown in the following paper (Pielak et al., 1984), the C-2 and C-4 isomers rarely occur in a one-to-one ratio in azo proteins, as implied by Tabachnick & Sobotka (1960). Therefore, three rather than two equations are necessary to correctly determine a system consisting of the three monoazo derivatives (monoazotyrosine, 2- and 4-monoazohistidine). This discussion is also limited to the mono(arsanilazo) derivative since there is no evidence for bisazo derivatives in proteins (see following paper).

A comparison of the molar absorptivities used in the equations of Tabachnick & Sobotka (1960) with those obtained for the pure compounds in this study, under identical conditions of pH and ionic strength, is given in Table VI. With these data it is possible to test the validity of the equations of Tabachnick & Sobotka (1960) by generating hypothetical arsanilazo proteins of known composition. Absorptivities of these hypothetical proteins at 460 and 500 nm in 0.1 N NaOH are simply linear combinations of the molar absorptivities of pure NAMAAT, NAMAAH(C-4), and NAMAAH(C-2) (Table VI). If the equations of Tabachnick & Sobotka (1960) are valid, then, by setting these equations equal to the absorptivities of each hypothetical arsanilazo protein, one would find that the solutions should equal the mono(arsanilazo)tyrosine and -histidine content originally assumed. Inspection of Table VII reveals not only severely over- and underestimated values but negative values as well. It is clear, then, that these equations cannot be used to characterize arsanilazo proteins. When mono(arsanilazo)-tyrosine is the only modification (Table VII, condition 1), its content is overestimated by only 10%. This may explain why there is reasonable agreement between the one mono(arsanilazo)tyrosine assigned spectrophotometrically and confirmed by sequence analysis (as tyrosine-248) in (arsanilazo)-carboxypeptidase A (Johansen & Vallee, 1971).

Since the absorption maxima of NAMSAH(C-2) and NAMAAH(C-2) differ at any given pH by about 40 nm from the maxima of NAMSAH(C-4) and NAMAAH(C-4) (Table V), it should be possible to differentiate between C-2 and C-4 azo coupling of histidine in sulfanilazo and arsanilazo proteins. Using structurally different diazonium salts to modify N^{α}-acetylhistidine methyl ester, Nagai et al. (1973) obtained spectra for the C-2 and C-4 azo derivatives that are very

Table VII: Arsanilazo Amino Acid Content of Hypothetical Proteins Using the Equations of Tabachnick and Sobotka[a]

condition	arsanilazo amino acid content of a hypothetical protein	mono(arsanilazo)-tyrosine/protein	mono(arsanilazo)-histidine/protein
1	1 mono(arsanilazo)tyrosine per protein	1.1	−0.04
2	1 4-mono(arsanilazo)histidine per protein	−0.1	0.4
3	1 2-mono(arsanilazo)histidine per protein	−0.4	1.6
	1 + 2	1.8	−0.4
	1 + 3	0.7	1.6
	2 + 3	−0.1	1.8
	1 + 2 + 3	1.0	1.8

[a] From Tabachnick & Sobotka (1960).

similar to those reported here. Consequently, it may prove possible to use these general spectral properties to characterize proteins azo coupled with a large variety of structurally different diazonium salts, providing pure and well-characterized model compounds are available. The utilization of the spectra to characterize azo proteins is presented in the following paper.

Registry No. NAMAAT, 39927-13-4; NABAAT, 88229-06-5; NAMAAH(C-2)·$^3/_2$Ba, 88211-97-6; NAMAAH(C-4)·$^3/_2$Ca, 88211-98-7; NABAAH·2Ca, 88229-07-6; NAMSAT·Ba, 88212-00-4; NABSAT·Ba, 88212-01-5; NAMSAH(C-2)·Ba, 88212-02-6; NAMSAH(C-4)·xBa, 88212-03-7; NABSAH, 88212-04-8; arsanilic acid, 98-50-0; arsanilic acid diazonium bromide, 88211-96-5; *N*-acetyl-L-tyrosine, 537-55-3; N^{α}-acetyl-L-histidine, 2497-02-1; sulfanilic acid, 121-57-3; sulfanilic acid diazonium bromide, 88211-99-8.

References

Johansen, J. T., & Vallee, B. L. (1971) *Proc. Natl. Acad. Sci. U.S.A. 68*, 2532.
Matsuo, H., Ohe, M., Sakiyama, F., & Narita, K. (1972) *J. Biochem. (Tokyo) 72*, 1057.
Nagai, W., Kirk, K. L., & Cohen, L. (1973) *J. Org. Chem. 38*, 1971.
Perrin, D. D. (1965) *Dissociation Constants of Organic Bases in Aqueous Solution*, Butterworths, London.
Pielak, G. J., Urdea, M. S., & Legg, J. I. (1984) *Biochemistry* (following paper in this issue).
Riordan, J. F., & Vallee, B. L. (1972) *Methods Enzymol. 25B*, 521.
Sachs, D. H., Schechter, A. N., & Cohen, J. S. (1971) *J. Biol. Chem. 246*, 6576.
Tabachnick, M., & Sobotka, H. (1959) *J. Biol. Chem. 234*, 1726.
Tabachnick, M., & Sobotka, H. (1960) *J. Biol. Chem. 235*, 1051.
Urdea, M. S., & Legg, J. I. (1979a) *Biochemistry 18*, 4984.
Urdea, M. S., & Legg, J. I. (1979b) *J. Biol. Chem. 254*, 11868.
Urdea, M. S., Pielak, G. J., Igi, K., & Legg, J. I. (1979) *Abstracts, Methods for Determining Metal Ion Environments in Proteins*, New Mexico State University, Las Cruses, NM, No. 56.
Warner, B. D., & Legg, J. I. (1979) *Inorg. Chem. 18*, 1839.
Warner, B. D., Boehme, G., Urdea, M. S., Pool, K. H., & Legg, J. I. (1980) *Anal. Biochem. 106*, 175.

Preparation and Characterization of Sulfanilazo and Arsanilazo Proteins†

Gary J. Pielak,‡ Mickey S. Urdea,§ and J. Ivan Legg*

ABSTRACT: A thorough study of the susceptibility of a series of proteins to azo coupling as a function of pH, ratio of diazonium salt to protein, and type of diazonium salt was undertaken. Spectral analysis of the azo proteins indicates that tyrosine is modified in preference to histidine, that while both possible structural isomers of azohistidine can be formed, azo coupling at the C-2 position is more common, and that little if any bis coupling occurs. Spectral analysis is based on data from chromatographically pure and chemically analyzed azo amino acids. Use of diazotized [^{35}S]sulfanilic acid indicates that the majority of residues modified in azo proteins is not detected by spectral analysis. These residues were identified as the products of the reaction of diazonium salt with free amines (i.e., N-terminal and/or the ϵ-amine of lysines) and with sulfhydryl groups. The former reactions are decreased or eliminated by acylation prior to azo coupling of the protein. The extent of azo coupling was found to be characteristic of the particular protein and not simply a function of the total number of potentially reactive residues in a protein.

Eighty years ago, Pauly (1904) reported the use of aromatic diazonium salts for the modification of proteins. Diazonium salts are known to react with various amino acid side chains of proteins to produce chromophores that absorb between 300 and 600 nm. Azo coupling of proteins[1] has found application in immunological investigations (Landsteiner & Lampl, 1917; Thorpe & Singer, 1969), the determination of structure–function relationships (Suh & Kaiser, 1976; Alter & Vallee, 1978; Fairclough & Vallee, 1970; Gorecki et al., 1978; Cueni & Riordan, 1978), the study of membrane proteins (Bell et al., 1979), a method for immobilizing proteins transferred to paper after electrophoresis (Renart et al., 1979), tumor imaging (Sundberg et al., 1974), and a novel approach to cancer therapy (Mizusawa et al., 1982). In a number of instances, it appears that azo coupling is specific for particular hyperreactive protein residues. We are interested in this modification as a particularly promising method for producing chelating agents on proteins. Once formed, substitution-inert Co(III) can be specifically incorporated into these chelating sites (Urdea & Legg, 1979a,b; Urdea et al., 1979).

Ascertaining conditions necessary for specific coupling of aryl diazonium salts to proteins requires a simple, accurate method for determining the type and number of residues modified. The method of choice is visible spectral analysis. However, attempts to utilize the spectral characteristics of simple azo model compounds to determine the extent of tyrosine and histidine modification in azo proteins (Tabachnick

† From the Department of Chemistry and the Program in Biochemistry and Biophysics, Washington State University, Pullman, Washington 99164. *Received March 28, 1983; revised manuscript received September 19, 1983.* This study was supported by the National Institutes of Health (GM 23081). M.S.U. was supported in part by a C. Glenn King Fellowship, Washington State University. Studies reported were submitted by G.J.P. in partial fulfillment of the requirements for the Ph.D. degree of Washington State University.

‡ Present address: Department of Biochemistry, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5.
§ Present address: Chiron Corp., Emeryville, CA 94608.

[1] The terms "diazotization" and "diazotized" have been used in previous work to indicate both the formation and coupling of diazonium salts. However, in this paper, in accordance with current usage, these terms only refer to diazonium salts while "azo coupling" and "coupled" refer to the reaction of diazonium salts.

0006-2960/84/0423-0596$01.50/0