Modification of Human Serum Albumin with Trifluoromethyl-Substituted Aryl Halides and Sulfonates

J. T. Gerig*[1a] and J. D. Reinheimer[1b]

Contribution from the Department of Chemistry, University of California, Santa Barbara, California 93106, and the Severance Chemical Laboratories, The College of Wooster, Wooster, Ohio, 44691. *Received May 17, 1974*

Abstract: Reactions of 4-fluoro-3-nitro-, 2-fluoro-4-nitro-, and 4-chloro-2,6-dinitrobenzotrifluoride and 2,6-dinitro-4-trifluoromethylbenzenesulfonate with various amino acid and peptide model compounds and with human serum albumin have been examined. Spectroscopic data for the derivatives of the model compounds are reported. Studies of the protein modification reactions suggest that these arylating reagents react by first forming an initial protein–small molecule complex within which an aromatic nucleophilic displacement takes place. It is probable that two lysine residues on the protein with p$K \sim 7.9$ are involved in reactions of these materials.

Nucleophilic displacement reactions with aryl substrates have found an important place in the melange of reactions available for the modification of proteins.[2] Of historical and continuing practical importance is the reaction of 2,4-dinitrofluorobenzene with the amino groups of proteins, first employed by Sanger.[3] 2,4-Dinitrochlorobenzene and picryl chloride have also been used to modify proteins as have 2,4-dinitrobenzenesulfonic acid and 2,4,6-trinitrobenzenesulfonic acid.[2a] The last-named compound has found utility as a colorimetric reagent for the determination of amino groups in amino acids and peptides.[4,5] In each of these cases, the reaction presumably proceeds by the mechanism established for aromatic nucleophilic substitution[6] with some group on the protein acting as the nucleophile, displacing the halide or sulfonate group on the activated aromatic ring of the reagent. The nucleophilic group could be an amino group from a lysine residue or the N-terminal of the macromolecule, a sulfhydryl group, the phenolic hydroxyl of tyrosine, or the imidazole ring of histidine. At the pH values typically used for such reactions, it is highly unlikely that hydroxyl groups of serine or threonine will be reactive.

The study of proteins which have been modified in some way to include a trifluoromethyl group in the structure has proved to be a useful means of probing protein behavior since fluorine magnetic resonance experiments with derivatives of this type afford chemical shift and nuclear relaxation time data which can reflect subtle changes in protein structure.[7] The present work was initiated to provide a series of reagents which will react with proteins in a manner similar to the aryl halides and sulfonates described above but which contain a trifluoromethyl group in their make-up. These materials offer additional reactions for introducing fluorine "reporter" groups into protein structures and, thus, expand the scope of this nmr technique. This paper describes the reactions of several arylating reagents of this type with some simple amino acids or peptides as well as the modification of reactive groups on human serum albumin by them.

Experimental Section

4-Fluoro-3-nitrobenzotrifluoride (Ia) was obtained from Pierce Chemical Co. and distilled through a 10-cm Vigreux column, boiling range 48–49° (0.5 mm).

2-Fluoro-4-nitrobenzotrifluoride (IIa) was obtained from Pierce Chemical Co. and distilled through a 10-cm Vigreux column, boiling range 64–65° (2 mm).

2,4-Dinitrofluorobenzene (Sigma) was similarly distilled, boiling range 136–137° (1 mm).

4-Chloro-3,5-dinitrobenzotrifluoride (IIIa, K&K Labs) was recrystallized twice from petroleum ether (30–60°) to give a material melting over the range 56.5–58.6° (lit.[8] 58°).

Sodium 2,6-dinitro-4-trifluoromethylbenzenesulfonate (IIIb) was prepared essentially according to the procedure of Adeniran, *et al.*[9] Ten grams of 4-chloro-3,5-dinitrobenzotrifluoride (37 mmol) was dissolved in 100 ml of 95% ethanol. This solution was added with stirring to a solution of 10 g of sodium sulfite (79 mmol) in 200 ml of water and 160 ml of ethanol at room temperature. After the solution was stirred for 20 min, 2 ml of concentrated hydrochloric acid was added and the ethanol removed on a rotary evaporator. Solid sodium chloride (~10 g) was added to the remaining hot aqueous solution. Upon cooling, the product precipitated; it was recrystallized several times from water by the same salting-out procedure, yield (2× recrystallized) 8.5 g (60%). The pmr spectrum of the product showed the expected aromatic proton singlet at 8.15 ppm from TMS in deuterium oxide solution.

The purity of IIIb was estimated by treating a known amount of material which had been dried several days in a vacuum desiccator over Drierite with an excess of glycylglycine at pH 7.4 or 8.0. Assuming quantitative formation of the glycylglycine derivative and the extinction coefficient given in Table I, it was estimated that this salt was at least 97% pure; two waters of crystallization were assumed in this calculation.

Table I. Spectral Properties of Trifluoromethyl-Substituted Aryl Derivatives

Aryl substituents	Amine	λ_{max}, nm[a]	$\epsilon \times 10^{-4}$, l. mol^{-1} cm^{-1} [a]
2-NO_2-4-CF_3	Glycylglycine	405.0	0.520
4-NO_2-2-CF_3	Glycylglycine	365.0	1.58
2-NO_2-4-NO_2	Glycylglycine	355.0	1.47
2,6-$(NO_2)_2$-4-CF_3	Glycylglycine	416.0	0.585
2,6-$(NO_2)_2$-4-CF_3	Glycinamide	416.0	0.568
2,6-$(NO_2)_2$-4-CF_3	Glycine	422.0[b]	0.585[b]
2,6-$(NO_2)_2$-4-CF_3	α-*N*-Acetyllysinamide	416.0	0.558

[a] At 25°, solvent was methanol. [b] Solvent was aqueous carbonate buffer at pH 10.

Sodium 2-nitro-4-trifluoromethylbenzenesulfonate (Ib) and sodium 4-nitro-2-trifluoromethylbenzenesulfonate (IIb) were prepared by similar reactions using the appropriate aryl fluoride. The reaction time in the case of compound IIb was 24 hr and with Ib, 60 hr. Isolation of the products required more careful control of the amount of added sodium chloride since these salts are more soluble in water than IIIb.

Glycine, glycylglycine, and glycinamide hydrochloride were obtained from Matheson Coleman and Bell, Mann Research Laboratories, and Mann, respectively. ***N*-α-Acetyl-L-lysinamide** acetate was a product of Cyclo Chemical Co.

Human serum albumin, Cohn fraction V, was purchased from Sigma Chemical Co. It was dialyzed against 10^{-3} *M* EDTA and doubly distilled water and then lyophilized. Acetylated albumin was prepared according to the procedure of Pinkard, *et al.*,[10] by

treating albumin with aspirin for 24 hr. The conditions used are reported to give an average of one acetyl group per molecule of albumin; we have not independently checked this result.

***N*-(2-Nitro-4-trifluoromethylphenyl)glycylglycine** was synthesized by dissolving 0.5 g (3.8 mmol) of glycylglycine and 0.44 g of sodium carbonate in a solvent composed of 18 ml of water and 20 ml of dimethyl sulfoxide. 4-Fluoro-3-nitrobenzotrifluoride (1.0 g, 4.8 mmol) was added and the reaction mixture stirred at room temperature overnight. The solution was then diluted with 100 ml of water, extracted with 50 ml of ether, and acidified with 6 *N* HCl. The yellow solid which precipitated was washed with water and recrystallized twice from methanol–water, crude yield 1.0 g, 82%. The pmr spectrum of this material in basic D_2O was consistent with the proposed structure showing aryl proton signals at 6.4, 8.0, and 8.1 ppm (total relative area 3) and two signals, a doublet and a broad singlet, at ~4.0 ppm (each relative area 2) corresponding to the two methylene groups of the dipeptide. The decomposition point was 237–238°.

***N*-(4-Nitro-2-trifluoromethylphenyl)glycylglycine** was prepared in 51% yield after two recrystallizations by treating 1.0 g of glycylglycine (7.5 mmol) in 7 ml of H_2O and 35 ml of dimethyl sulfoxide with 2.0 g (9.6 mmol) of 2-fluoro-4-nitrobenzotrifluoride. The pH was monitored continuously with a pH meter and was adjusted to pH 10.2 initially with 10% sodium hydroxide solution. As the reaction proceeded, NaOH solution was added dropwise to keep the pH in the range 10–11. After 3.5 hr, an equivalent amount of base had been added, and the reaction mixture was poured into 120 ml of water, extracted with ether, and acidified. The aqueous layer was extracted several times with ether, and the ether extracts were combined, dried, and evaporated to dryness to afford pale yellow crystals which, upon recrystallization twice from methanol–water, gave a decomposition point of 194–196°. The pmr spectrum, obtained using a dimethyl sulfoxide solution, was consistent with the expected structure, showing aryl protons (total relative area 3) at 6.7, 7.4, and 8.2 ppm and a singlet and a doublet for the methylene groups at 4.1 and 3.9 ppm, respectively.

***N*-(2,4-Dinitrophenyl)glycylglycine** was prepared by a procedure similar to the one described in the previous paragraph, giving a product with a decomposition point of 202–204°.

***N*-(2,6-Dinitro-4-trifluoromethylphenyl)glycine** was obtained by a procedure similar to that used for *N*-(4-nitro-2-trifluoromethylphenyl)glycylglycine. The pH during the reaction was controlled at 9–10, and the product was isolated as the sodium salt from the reaction mixture by saturating the solution with sodium chloride. The solid so formed was dissolved in water and acidified to produce the desired compound (88% crude yield) which was recrystallized from methanol–water, mp 189–192°.

The preparation of ***N*-(2,6-dinitro-4-trifluoromethylphenyl)glycylglycine** (80% crude yield) followed the same steps as the glycine derivative; this compound had a decomposition point of 216–218° after two recrystallizations from methanol–water.

***N*-(2,6-Dinitro-4-trifluoromethylphenyl)glycinamide** was prepared by the same procedure as that used for the glycine derivative except that acetone was used as the organic solvent. The pH was controlled in the range 9–10 during the reaction, and the product precipitated upon the addition of water at the end of the reaction (75% crude yield). After one recrystallization from methanol–water, the decomposition point of the product was 218–219°.

ε-*N*-(2,6-Dinitro-4-trifluoromethylphenyl)-α-*N*-acetyllysinamide was synthesized by the same procedure as used for the glycine derivative with pH control in the range 8.8–9.2. The product precipitated upon dilution of the basic reaction mixture with water and, after recrystallization from methanol–water four times, melted at 197–199° with a pre-melting transition at 194°.

Phosphate buffers at 0.1 *M* phosphate were used at pH values below 8; above pH 8, 0.1 *M* carbonate buffers were employed.

Ultraviolet–visible spectra were obtained on a Cary 15 or Beckman DU spectrophotometer. The Beckman, equipped with a Gilford 2000 recorder accessory and a thermostated cell compartment, was used in reaction rate determinations. Solution pH values were read with a Radiometer Model PHM63 or PHM52 pH meter equipped with Metrohm microelectrodes.

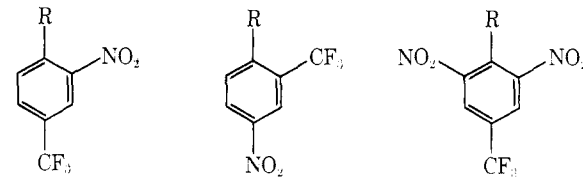
Kinetic Procedures. Reaction rates with the model compounds were observed under pseudo-first-order conditions with the concentration of amine about 0.01 *M* and the arylation reagent present at about 0.0001 *M*. Three milliliters of a stock amino acid or peptide solution in buffer was pipetted into a 1-cm quartz cell and allowed to thermally equilibrate in the spectrometer cell compartment. To start the reaction, 0.2 ml of a freshly prepared solution of the aryl halide in organic cosolvent was added and rapidly mixed. (For reactions with the sulfonate, an aqueous solution was used.) Optical density was recorded as a function of time, and the pseudo-first-order rate constant was computed from the slope of a plot of $\ln (OD_\infty - OD_t)$ *vs. t*, where OD_t is the absorbance measured at time *t*, and OD_∞ is the absorbance at completion of the reaction. Observed values of OD_∞ usually agreed well with those calculated using the extinction coefficients in Table I. Second-order rate constants were obtained by dividing the pseudo-first-order constants by total amine concentration.

The most reactive aryl halide (IIIa) hydrolyzes slowly at pH 10. However, this side reaction was shown to be negligible during the course of the reactions with the amino compounds. The 2,6-dinitro-4-trifluoromethylphenyl derivatives were not indefinitely stable at high pH and, after reaching a maximum, OD_t in these reactions slowly decreased, probably because of hydrolysis of the reaction products. The maximum OD observed with glycine corresponded to 100% reaction but was slightly less for the other nucleophiles. For these latter reactions, quantitative reaction was assumed, and a calculated OD_∞, obtained using the data in Table I, was used to extract the desired rate constants; the slope of the early part of the $\ln (OD_\infty - OD_t)$ *vs. t* plots was taken to be k_2'.

For reactions with human serum albumin, the aryl halide concentration was $\sim 10^{-3}$ *M* with the protein concentration at $\sim 10^{-5}$ *M*. Protein solutions were prepared gravimetrically assuming a molecular weight of 66,500. Fresh protein solutions were made up daily since unexplained spurious results with older solutions were sometimes noted. Methanol was used as the organic cosolvent for protein reactions. (It was found that the presence of dioxane essentially eliminated the "initial burst" such as shown in Figure 2.) Reactions were initiated by adding 0.3 ml of the aryl halide in methanol solution to 3.0 ml of protein solution in buffer contained in a 10-mm cuvette. For reactions with the benzenesulfonate, the arylating agent was water soluble and was simply dissolved in buffer. Reactions were followed at the same wavelength as that observed for the maximal absorption in the model compounds, and the extinction coefficients determined for these compounds were assumed to be good approximations for protein counterparts. The modified albumins were stable indefinitely at pH 7 or below.

Results

Since the trifluoromethyl group is highly electron withdrawing,[11] it seemed reasonable to consider compounds in which a trifluoromethyl group replaces a nitro group in the various nitro-activated aryl halides and sulfonates that have been used for protein modification. In this light, 2-nitro-4-trifluoromethylfluorobenzene (Ia) and 4-nitro-2-trifluoromethylfluorobenzene (IIa) can be regarded as analogs of Sanger's reagent (2,4-dinitrofluorobenzene). 2,6-Dinitro-4-trifluoromethylchlorobenzene (IIIa) is related to picryl chloride similarly, while the sulfonic acids (Ib, IIb, IIIb) can be thought of as stand-ins for dinitro- or trinitrobenzenesulfonic acids.



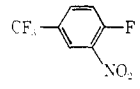
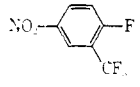
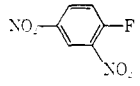
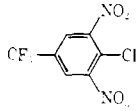
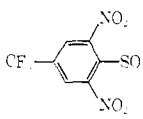
Ia, R = F
b, R = SO_3Na

IIa, R = F
b, R = SO_3Na

IIIa, R = Cl
b, R = SO_3Na

Model Reactions. In order to assess the relative reactivities of compounds I–III with amine nucleophiles, a series of exploratory reactions were run with glycine, glycinamide, glycylglycine, and α-*N*-acetyllysinamide. The products of these reactions were synthesized on a large scale and, after careful purification, had the ultraviolet spectral properties

Table II. Rate Constants for Arylation of Model Compounds by Trifluoromethyl Substituted Reagents[a]

Reagent	Amine	pH	$k_2' \times 10^3$,[b] l. mol^{-1} sec^{-1}	$k_2 \times 10^3$,[c] l. mol^{-1} sec^{-1}
	Glycylglycine	10.0	0.245	0.25
	Glycylglycine	10.0	0.045	0.045
	Glycylglycine	10.0	36.8	37
	Glycylglycine	10.0	7.0	7.1
	Glycinamide	10.0	41.0	41
	Glycine	10.0	42.0	61
	α-*N*-Acetyl-lysinamide	10.0	3.23	9.0
	Glycylglycine	7.97	23.6	59
	Glycinamide	7.85	23.3	45
	Glycine	7.93	4.47	239
	α-*N*-Acetyl-lysinamide	7.70	1.00	356

[a] At 30.2° with 6.3% dioxane as cosolvent for reactions of the first three aryl fluorides; 6.3% methanol was used for the aryl chloride. No organic solvent was used with the sulfonate. [b] Average of duplicate runs; agreement between runs was 5% or better. [c] Corrected for the amino group pK as indicated in the text.

indicated in Table I. Beer's law was obeyed by each derivative in methanol except for the 2,6-dinitro-4-trifluoromethylphenyl derivative of glycine. Beer's law was followed by this compound in aqueous solution at pH 10, however.

Reaction rates of the model compounds were observed under pseudo-first-order conditions with the amine component in large excess. The apparent second-order rate constants (k_2') obtained from these data are given in Table II. These second-order rate constants will be pH dependent since presumably it is only the free amino form of the nucleophile which is reactive. In order to get the rate data in Table II onto a common ground for comparison, we assumed that the pK_a's of the amino functions of glycine, glycinamide, glycylglycine, and α-*N*-acetyllysinamide were 9.65,[13] 7.81,[14] 8.15,[15] and 10.25,[16] respectively, at 30°, and that the pH dependence of k_2' is given by

$$k_2' = K_A k_2/(K_A + [H^+]) \quad (1)$$

where K_A is the dissociation constant for the ammonium ion under consideration and k_2 is the second-order rate constant for reaction of the free amino group. The values of k_2 computed on this basis are also given in Table II. Since several approximations are involved in obtaining values for k_2, these data cannot be regarded as highly precise. However, the relative reactivities of compounds Ia, IIa, IIIa, and IIIb with an amino function are clearly indicated. Qualitative observations made with the sulfonic acids Ib and IIb indicated that these compounds were very much less reactive than the corresponding aryl fluorides, and quantitative work with these compounds was not attempted.

A slow decrease in absorption at 416 nm due to product decomposition was observed in the reactions of IIIa with glycinamide and glycylglycine at pH 10 after the reaction had essentially run to completion. Bunnett and Herman have observed that 2,4-dinitrophenylglycine is unstable at high pH,[17] and it is known that trinitrophenyl derivatives of amino acids are unstable under similar conditions.[18] (Both types of compounds are also photolabile.)[1] Although base-catalyzed hydrolysis of the CF_3 group is a possible mode for decomposition of amino acid derivatives of I–III,[19] it is most likely that a nucleophilic displacement of the amino function by hydroxide ion accounts for the observed product instability at high pH. At pH 8 or lower in all cases, the reaction products were stable for at least several days.

Reactions with Human Serum Albumin. Plasma albumins have been the subject of a number of studies which involved modification of the proteins by 2,4-dinitrofluorobenzene,[20] picryl chloride,[21] and trinitrobenzenesulfonic acid.[21-23] With both bovine and human serum albumin, kinetic studies indicate several classes of nucleophilic reaction sites characterized by vastly different reactivities toward these reagents; for both albumins, a small number (one to two) of very reactive sites are observed,[16,18a] in addition to a larger number of sites which react more slowly. In exploring the utility of the reagents I–III described above, we examined the reaction of human serum albumin with reagents I–IIIa and IIIb, focusing our attention on the early stage of the reaction which is dominated by the behavior of the abnormally reactive nucleophilic groups on the protein.

Consider the reaction of a protein with n possible sites of reaction with one of the arylating reagents. Assuming a large concentration excess of aryl halide or sulfonate, if the reaction is monitored by the increase in absorbance at a wavelength appropriate for observation of the arylated site (*cf.* Table I) and each site on the protein reacts independently of the others, then the absorbance at time t, A_t, observed with the proper blank is given by eq 2, where ϵ_i is

$$A_t = P_0 \sum_{i=1}^{n} \epsilon_i (1 - e^{-k_i t}) \quad (2)$$

the extinction coefficient for the derivative of the ith site, P_0 is the molar protein concentration, k_i is the pseudo-first-order rate constant for reaction at ith site, and n is the number of reactive sites. To apply eq 2, we assume that the ϵ_i are identical for all reaction sites and equal to some average value, ϵ_{av}. Should it turn out that a number of the protein reaction sites have very similar values of k_i, eq 2 may be "clustered" to eq 3, where n_i corresponds to the number

$$A_t = \epsilon_{av} P_0 \sum_i n_i (1 - e^{-k_i t}) \quad (3)$$

of similarly reactive sites ($\Sigma n_i = n$). If one of these sets of sites is very much more reactive than the others, for a short enough time span the exponentials corresponding to the less reactive sites can be replaced by their series expansions. Truncating the series after the linear term changes eq 3 to eq 4. Here n_1 is the number of highly reactive sites. Equa-

$$A_t \cong \epsilon_{av} P_0 [n_1 (1 - e^{-k_1 t}) + \Sigma_{n_{i+1}} k_{i+1} t] \quad (4)$$

tion 4 can be rearranged to eq 5, where the constant a re-

$$A_t = at + \epsilon_{av} P_0 n_1 (1 - e^{-k_1 t}) \quad (5)$$

places the summation in eq 4. This constant as determined experimentally is not informative since it depends on a large number of inseparable constants. Equation 5 predicts an initially exponential increase in absorbance superimposed on a linear component corresponding to reaction of the "slow" groups of the protein; this behavior is essentially that observed in this work (Figure 1) and in previous studies of human and bovine albumin.[20,22]

Absorbance–time data for reaction of I–III with human albumin were obtained under several conditions and analyzed for the parameters n_1 and k_1 by fitting eq 5 to experimental data by the Marquardt technique for estimation of nonlinear parameters.[24] A program devised by Conway, *et*

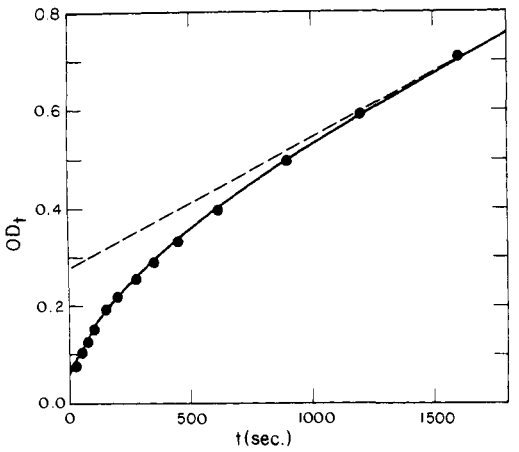


Figure 1. Observed (points) and calculated (solid line) absorbance-time curve for the reaction of human serum albumin with IIIa at pH 8.9. The dotted line represents the asymptote defined by the linear term in eq 5.

al., was modified for this purpose.[25] As Figure 1 illustrates, good fits to the experimental data could be obtained by this technique; Table III records the parameters so estimated

Table III. Arylation Reactions of Human Serum Albumin[a]

pH	[IIIa] × 10^3, M	[HSA] × 10^5, M	n_1	$k_1 \times 10^3$, sec^{-1} [b]	$k_c \times 10^3$, sec^{-1} [c]	Added reagents
With 2,6-Dinitro-4-trifluoromethylchlorobenzene (IIIa)						
7.52	0.82	1.99	1.4	1.6	12	
7.52	0.82	1.99	1.1	0.94		[Salicylic acid] = 0.01 M
7.54	0.82	2.27[d]	1.0	1.5		
7.74	0.91	2.22	1.4	3.8	19	
8.92	0.91	2.54	1.6	3.0	7	
9.00	0.91	1.69	1.8	6.6	15	
9.00	0.91	1.69	1.8	9.4		[Cu^{2+}] = 2×10^{-5} M
9.00	0.98	1.69	1.9	6.8		[Hg^{2+}] = 2×10^{-4} M
9.16	0.91	1.94	2.2	5.2	11	
With 2,6-Dinitro-4-trifluoromethylbenzenesulfonate (IIIb)						
	[IIIb] × 10^3, M					
7.36	1.53	2.33	1.0	0.25	1.4	
9.12	1.53	1.82	2.0	4.2	5.1	
9.18	1.53	1.82[d]	1.5	3.7		

[a] At 30.4° with 9.1% methanol present as cosolvent. [b] Computer-fitted apparent first-order rate constant. [c] Calculated from k_1 by assuming the equilibrium constants given in Table IV and an enzyme group pK = 7.9. [d] Aspirin-acetylated albumin used. See text.

for several reactions with human albumin. It was found that the number of "fast" reaction sites is about two in the pH range 7–9, in agreement with the findings of Goldfarb regarding the reaction of this protein with trinitrobenzenesulfonic acid.[22a]

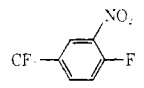
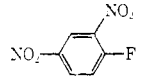
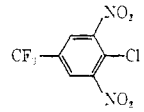
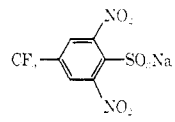
Interpretation of the pseudo-first-order constant k_1 depends upon the mechanism assumed for the arylation reaction. Most simply this reaction can take place by a simple bimolecular nucleophilic displacement on the aryl halide, or, alternatively, a protein–aryl halide complex can form within which the protein arylation takes place.[26] For the second possibility and assuming that the concentration of the arylating reagent is large enough to be constant over that phase of the modification reaction that is observed, the observed first-order rate constant k_1 is given by

$$k_1 = \frac{k_c'}{1 + (K/[\mathrm{ArR}])} \qquad (6)$$

where k_c' is the rate constant for reaction within the complex, K is the dissociation constant for the reagent–protein complex, and [ArR] is the concentration of the arylating reagent.[27,28] As indicated by eq 6, a hyperbolic dependence of k_1 on aryl halide or sulfonate concentration is expected if a discrete protein–reagent complex is formed. On the other hand, if the modification reaction is a simple bimolecular displacement, then the observed first-order rate constant should vary linearly with the concentration of the aryl halide or sulfonate.

A series of initial rate experiments for reaction of human albumin were performed at various initial concentrations of aryl halide. Treatment of the data according to eq 6 clearly showed that the mechanism involving complexation was operant, as has been found with bovine serum albumin.[20,23] Table IV gives the values of K/n and k_c' obtained for several of the trifluoromethyl-substituted reagents employed in this study.

Table IV. Initial Rate Studies of Reaction of Human Serum Albumin with Arylating Reagents[a]

Reagent	pH	$K/n_1 \times 10^3$ [b,c]	$k_c' \times 10^3$, sec^{-1} [c]	$k_c \times 10^3$, sec^{-1} [d]
	8.52	1.3	0.48	0.6
	7.66	2.4	13	35
	7.72	1.0	9.5	23
	7.55	0.5	4.0	13
	7.08	1.1	1.6	12
	6.80	1.0	0.8	11
	6.57	1.1	0.7	22
	5.66	0.8	0.1	17
	8.02	0.3	1.6	2.8

[a] At 30.3° with 9.1% methanol as cosolvent. [b] Substrate binding constant. [c] Obtained from double reciprocal plots of eq 6 in text. [d] Rate constant analogous to k_2 computed from k_c' by assuming an enzyme pK_a = 7.9.

The available data suggest that the binding interaction between human serum albumin and an activated aryl halide is not strongly pH dependent. However, k_c', the rate constant for reaction within the complex, varies greatly with pH. Assuming that this variation is a reflection of the pK_a of two (identical or nearly identical) nucleophilic groups on the enzyme which participate in this reaction, a p$K_a \approx 7.9$ can be estimated for these groups. The rate constants obtained in the initial rate studies (k_c') were corrected for this pH effect to give rate constants (k_c) for the reaction of the unprotonated nucleophiles within the protein–arylating reagent complexes. Similarly, the binding constants obtained in the initial rate experiments and the same pK_a were employed to obtain additional values of k_c from the computer-fitted rate constants (k_1) given in Table III. Although there is a fair amount of jitter in the data, the values for k_c for reactions of IIIa and IIIb obtained by initial rate studies and by computer fits of absorbance–time data from the "fast" phase of the modification reaction are in fair agree-

ment. The initial rate studies and the full-range kinetic studies reported in Table III were obtained over different pH ranges by different experimental techniques, and this agreement suggests that the simple model used for the protein reactions is reasonably valid.

Although several previously reported observations strongly implicate lysine residues as the two exceptionally reactive nucleophiles on the human albumin molecule,[23,29] the abnormally low pK_a for these groups suggested by our work generates some reluctance to accept this assignment. Several experiments were therefore carried out to delineate more clearly the nature of these reactive groups. Cupric ion is known to complex tightly with the N-terminal amino group of human serum albumin; the complexed amino group is unreactive toward dinitrofluorobenzene.[30,31] When a reaction with IIIa was run in the presence of 2×10^{-5} M Cu^{2+}, an amount equimolar to the protein concentration, no change in the number of reactive groups was observed (Table III), although a slight effect on k_1 was noted.[32] Similarly, the addition of mercuric ion at a concentration tenfold larger than the protein concentration had no effect on the number of fast-reacting groups modified; since mercury reacts avariciously with sulfhydryl groups, this experiment rules out the possibility that one or both of these groups is a sulfur-containing nucleophile.

Human serum albumin is acetylated by aspirin, and the number of acetyl groups transferred to the protein can be controlled by adjustment of pH and reaction time.[34] Hawkins, *et al.*, have shown that this acetylation reaction is inhibited by added salicylate,[29] and that acetylation takes place at a lysine residue.[35] We find that when human albumin which has been acetylated to the extent of about one acetyl group per protein molecule is used in the arylation reaction, the number of reactive groups detected drops substantially (Table III). Thus, it is probable that at least one of the exceptionally nucleophilic groups on the protein in reactions with I–III is identical with the lysine residue(s) acetylated by aspirin. In confirmation of this conclusion, we have observed (Table III) that added salicylate at 10^{-2} M also decreases the number of reactive groups detected, presumably because a salicylate binding site is identical with a binding site used by aspirin in the acetylation reaction.

Discussion

The results presented here confirm earlier indications that trifluoromethyl is not nearly as effective an activating group in aromatic nucleophilic substitution reactions as is the nitro function.[36] The aryl halides (Ia–IIIa) and the sulfonate substituted at both ortho positions with NO_2 (IIIb) are nevertheless still sufficiently reactive to be useful on a reasonable time scale in protein modification reactions. The sulfonates with a single nitro group reacted very slowly in model reactions and probably will have less utility in such applications. Of the two classes of materials, the sulfonates are the more stable in aqueous solution, and, in the absence of other considerations, these would appear to be the reagents of choice for protein modification.

Dinitrophenyl derivatives of amines typically absorb near 365 nm with an extinction coefficient of $\sim 1.6 \times 10^4$ M^{-1} cm^{-1}, while trinitrophenyl derivatives of glycine and *N*-acetyllysine absorb at 345 nm with an extinction coefficient of $\sim 1.5 \times 10^4$ M^{-1} cm^{-1}.[37] The replacement of nitro by trifluoromethyl at the para position in the corresponding arylamine reaction products thus results in a shift of the maximum absorption 40–60 nm to longer wavelength and a decrease in the extinction coefficient at the maximum to about one-third of the value found in the *p*-nitro-substituted derivatives. It is nevertheless easy to follow a protein

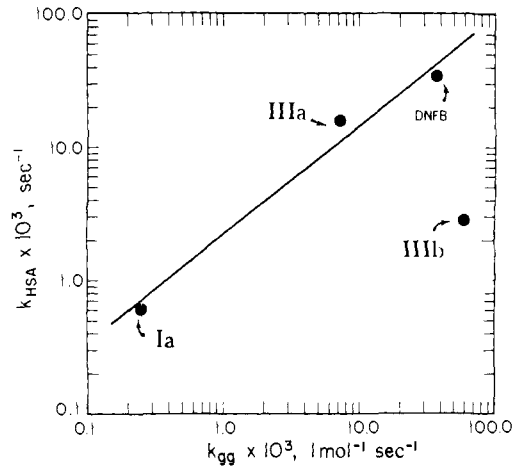


Figure 2. Correlation of the free energies of activation for reaction of IIa, IIIa, and 2,4-dinitrofluorobenzene with glycylglycine (k_{gg}) and human serum albumin (k_{HSA}). The sulfonate (IIIb) is less reactive with the protein than expected on the basis of this correlation.

modification by these reagents spectroscopically, and we have throughout this work assumed that the extinction coefficients found for the model compounds given in Table I are appropriate for use with modified proteins.[38]

One expects a linear correlation between the free energies of activation for the action of arylating agents such as I–III on simple amines and the free energies of activation for the corresponding reactions within a protein complex. Figure 2 shows that such a correlation obtains when the reactions of the halides Ia–IIIa with glycylglycine and human albumin are compared. With the sulfonate IIIb, a considerably slower rate of reaction with the protein is observed than is expected on the basis of the data obtained with the model reaction; the activated halides are uncharged molecules, and their reactions with the protein should not be greatly affected by charged groups near a protein binding center. However, the anionic sulfonate may well interact strongly enough with an ionized group on the protein near this site to impede its subsequent reaction. Thus, while IIIa and IIIb can be used to introduce the 2,6-dinitro-4-trifluoromethylbenzene moiety into a protein structure, a certain selectivity in reaction can be obtained by using one reagent rather than the other.

Both the aspirin-acetylation experiments[29] and this work show the presence of a limited number of unusually reactive nucleophilic groups on human serum albumin, at least one of which must be a lysine amino function. Most likely, both very reactive groups found on the protein in this work are lysine, although the pK value (~ 7.9) suggested by our experiments is rather low. However, an unusually reactive lysine (Lys-41) with a pK of 8.8 has been found in ribonuclease A,[39] and a lysine residue at the active site of acetoacetate decarboxylase has $pK = 5.9$.[40] It may be that the two super-reactive lysine residues of human albumin are close together in the protein structure so that one can electrostatically affect the acidity of the other in a manner similar to th effect of lysine-7 on lysine-41 of ribonuclease.[39]

In summary, we have examined a series of trifluoromethyl-labeled aryl halides and sulfonates which can be used in modification reactions of nucleophilic groups on proteins. A range of reactivity and selectivity is offered by these reagents so that some specificity in the introduction of fluorine-labeled nmr reporter groups into protein structures can be anticipated. Several proteins so modified are under study.

Acknowledgments. The authors gratefully acknowledge the support of the National Cancer Institute (Grant CA-

11220) for this work. J.T.G. is a PHS Career Development Awardee (Grant GM-70373 from the National Institute of General Medical Sciences).

References and Notes

(1) (a) University of California; (b) on leave from The College of Wooster, 1973–1974.
(2) (a) G. E. Means and R. E. Feeney, "Chemical Modification of Proteins," Holden-Day, San Francisco, Calif., 1971, p 118 ff; (b) S. J. Singer, *Advan. Protein Chem.*, **22**, 1 (1967).
(3) F. Sanger, *Biochem. J.*, **39**, 507 (1945).
(4) K. Satake, T. Okuyama, M. Ohashi, and T. Shinoda, *J. Biochem. (Tokyo)*, **47**, 654 (1960).
(5) A. F. Habeeb, *Anal. Biochem.*, **14**, 328 (1966).
(6) J. F. Bunnett and R. Zahler, *Chem. Rev.*, **49**, 273 (1951).
(7) W. H. Huestis and M. A. Raftery, *Biochem. Biophys. Res. Commun.*, **48**, 678 (1972); *Biochemistry*, **11**, 1648 (1972).
(8) L. M. Yagupol'skii and V. S. Mospan, *Ukr. Khim. Zh.*, **21**, 81 (1955); *Chem. Abstr.*, **49**, 8866 (1955).
(9) M. A. Adeniran, C. W. L. Bevan, and J. Hirst, *J. Chem. Soc.*, 5868 (1963).
(10) R. N. Pinkard, D. Hawkins, and R. S. Farr, *Arthritis Rheum.*, **13**, 361 (1970).
(11) The Hammett σ constant for p-CF_3 is +0.54 which may be compared with σ for p-NO_2 which is +0.78.[12]
(12) L. P. Hammett, "Physical Organic Chemistry," 2nd ed, McGraw-Hill, New York, N. Y., 1970, p 356.
(13) E. J. King, *J. Amer. Chem. Soc.*, **73**, 155 (1951).
(14) Estimated from the value at 24.3° [M. Zief and J. T. Edsall, *J. Amer. Chem. Soc.*, **59**, 2245 (1937)] by assuming same temperature dependence as glycylglycine.[15]
(15) Interpolated from the data of E. R. B. Smith and P. K. Smith, *J. Biol. Chem.*, **146**, 187 (1942).
(16) Estimated from the value for lysine at 25° [C. Tanford, *Advan. Protein Chem.*, **17**, 73 (1962)] by assuming same temperature dependence as ϵ-aminocaproic acid.[15]
(17) J. F. Bunnett and D. H. Herman, *Biochemistry*, **9**, 816 (1970).
(18) A. Kotaki and K. Sataki, *J. Biochem. (Tokyo)*, **56**, 299 (1964).
(19) T. T. Sakai and D. V. Santi, *J. Med. Chem.*, **16**, 1079 (1973).
(20) N. M. Green, *Biochim. Biophys. Acta*, **74**, 542 (1963).
(21) J. J. Arrotti and J. E. Garvin, *Biochim. Biophys. Acta*, **255**, 79 (1972).
(22) (a) A. R. Goldfarb, *Biochemistry*, **5**, 2574 (1966); (b) *Biochim. Biophys. Acta*, **200**, 1 (1970).
(23) L.-O. Andersson, J. Brandt, and S. Johansson, *Arch. Biochem. Biophys.*, **146**, 428 (1971).
(24) D. W. Marquardt, *J. Soc. Ind. Appl. Math.*, **11**, 431 (1963).
(25) G. R. Conway, N. R. Glass, and J. C. Wilcox, *Ecology*, **51**, 503 (1970).
(26) B. R. Baker, "Design of Active-Site-Directed Irreversible Enzyme Inhibitors," Wiley, New York, N. Y., 1967, p 122.
(27) E. W. Bittner and J. T. Gerig, *J. Amer. Chem. Soc.*, **92**, 2114 (1970).
(28) If n equivalent groups are reacting equivalently and simultaneously, K in eq 6 should be replaced by K/n.
(29) R. N. Pinckard, D. Hawkins, and R. S. Farr, *Arthritis Rheum.*, **13**, 361 (1970).
(30) R. A. Bradshaw and T. Peters, Jr., *J. Biol. Chem.*, **244**, 5582 (1969).
(31) D. W. Appleton and B. Sarker, *J. Biol. Chem.*, **246**, 5040 (1971).
(32) The dissociation constant for the Cu^{2+} protein complex should be about 10^{-16} so that an equimolar amount of Cu^{2+} should be sufficient to complex virtually all the terminal amino groups.[33]
(33) S.-J. Lau and B. Sarker, *J. Biol. Chem.*, **246**, 5938 (1971).
(34) D. Hawkins, R. N. Pinckard, and R. S. Farr, *Science*, **160**, 780 (1968).
(35) D. Hawkins, R. N. Pinckard, I. P. Crawford, and R. S. Farr, *J. Clin. Invest.*, **48**, 536 (1969).
(36) J. F. Bunnett, T. Kato, and N. S. Nudelman, *J. Org. Chem.*, **34**, 785 (1969).
(37) Reference 2a, pp 120 and 122.
(38) It has been found that the presence of sulfite ions can perturb the spectrum of trinitrophenylamino acid derivatives;[22a] we have not investigated the effect of sulfite on the trifluoromethylated compounds utilized in this work.
(39) A. L. Murdock, L. K. Grist, and C. H. W. Hirs, *Arch. Biochem. Biophys.*, **114**, 375 (1966).
(40) D. E. Schmidt, Jr., and F. H. Westheimer, *Biochemistry*, **10**, 1249 (1971).

Chemical Relaxation Studies of Metal Ion Activated Enzymes. I. Opposing Bicipital Relaxation Phenomena in the Bovine Carbonic Anhydrase System. Theory and Observation

Dennis E. Tallman,* George Graf, Timothy J. McNeese, and Mark M. Wilson

Department of Chemistry and Department of Biochemistry, North Dakota State University, Fargo, North Dakota 58102. *Received August 9, 1974*

Abstract: Temperature-jump investigations of bovine carbonic anhydrase have been carried out in solutions containing the enzyme, carbon dioxide, bicarbonate ion, and pH indicator in 0.15 *M* potassium fluoride over the pH range 6.0–8.0. A single relaxation has been observed, the amplitude of which is not only dependent on solution conditions but can actually be made to reverse its direction by appropriate adjustment of these conditions. The amplitude goes through a null point where, for a particular set of solution conditions, the amplitude of the relaxation is zero. This behavior, termed an opposing bicipital relaxation, is attributed to a change in the sign of the normal enthalpy for the overall hydration reaction which results from a coupling of the hydration reaction *via* proton to other proton ionization reactions in the system. An approach is presented for calculating the conditions necessary for a relaxation amplitude null in a general reaction system in which the observed reaction is coupled to several more rapid reactions. Agreement between the calculated and the experimental conditions required for a relaxation null may be used to further substantiate a reaction scheme proposed from relaxation time measurements.

The various roles that metal ions play in enzymatic reactions range from weak ionic interactions to specific associations with involvement of the metal ion in the catalytic mechanism.[1] Although numerous studies of metalloenzymes have been reported,[1,2] relatively little is known about the role of metal ions in the detailed catalytic mechanisms. We have initiated in our laboratories a series of investigations employing fast reaction methods for elucidating the kinetic and mechanistic role of metal ions in enzyme-catalyzed reactions. The application of fast reaction techniques permits kinetic experiments to be conducted under conditions which often result in direct observation of enzymatic intermediates.[3,4] Under such conditions, the role of the metal ion in the interconversion of the intermediates can be directly assessed. In this paper we describe the application of the temperature-jump relaxation method to the carbonic anhydrase system and the associated opposing bicipital relaxation phenomenon.

Carbonic anhydrase (carbonate hydro-lyase, E.C. 4.2.1.1) from bovine erythrocytes is a zinc-containing enzyme of the approximate mol wt 30,000. The enzyme catalyzes the reversible hydration of carbon dioxide.