Studies of the Location of Tyrosyl and Tryptophyl Residues in Protein. II. Applications of Model Data to Solvent Perturbation Studies of Proteins Rich in both Tyrosine and Tryptophan*

Theodore T. Herskovits and Sr. Margaret Sorensen†

ABSTRACT: The uses and applications of model compound data of the accompanying paper (Herskovits, T. T., and Sorensen, M. (1968), *Biochemistry 7*, 2533 (this issue; preceding paper) are described and as a means of illustration applied to the simultaneous study of the location of tyrosyl and tryptophyl residues in pepsin, rabbit muscle aldolase, and bovine serum albumin. From the study on pepsin with six perturbants, it is concluded that in the native enzyme 2–3 of the 5 tryptophyls and 10–12 of the 17 tyrosyls are exposed, with perhaps 1 or 2 of these residues being partly buried. In native aldolase, a much larger fraction of the tyrosyls and tryptophyls are found to be buried. Of the 47 tyrosyls 34–40% appear to be exposed, while of the 11–12 tryptophyls only about 15–20% are exposed; it is suggested that some of the observed perturbations may be due to a combination of exposed and partly exposed chromophores. In serum albumin, the two tryptophyl residues are found to be nearly fully exposed to perturbants with less than 4.4-Å diameter (deuterium oxide and ethylene glycol), while of the order of 50–70% exposure is observed with perturbants having greater than 5.2-Å diameters (glycerol and hexaethylene glycol). It is suggested that the two tryptophyls are probably located at the periphery of some of the conformational subunits of serum albumin postulated by Foster, or situated at other areas of polypeptide fold which must be largely accessible to solvent penetration. Previous studies on the location of the tyrosyls in bovine and human serum albumin (Herskovits, T. T., and Laskowski, M., Jr. (1962), *J. Biol. Chem. 237*, 2481) have suggested that about 70% of the 18–21 tyrosyls appear to be buried in the native folds of these proteins, with about 20% of 3–5 of these groups being located at areas of subunit contact. The latter key tyrosyls are not accessible to perturbants with molecular diameters greater than 5.2 Å.

In the accompanying paper (Herskovits and Sorensen, 1968) the solvent perturbation parameters of *N*-acetyl ethyl esters of tyrosine and tryptophan, obtained with nine of the commonly employed perturbants in the 350–240-mμ range, were reported, and the additivity of the tyrosine and tryptophan perturbation difference spectra necessary for the simultaneous assessment of the fraction of exposed tyrosyls and tryptophyl residues in proteins rich in both amino acids was tested. In the present paper, the range of applicability depending on chromophore composition of proteins (*i.e.*, the ratio of tyrosine to tryptophan in the protein) is explored, and the uses and applications of the model compound data are described and as a means of illustration applied to the study of the location of the tyrosyl and tryptophyl residues in rabbit muscle aldolase, pepsin, and bovine serum albumin.

Experimental Section

Materials. Most of the reagents, model compounds, and solvents employed were described in the accompanying paper (Herskovits and Sorensen, 1968). Two aldolase samples of essentially identical difference spectral behavior were employed, a twice-crystallized Mann-assayed sample and a second sample purchased from Worthington Biochemical Corp. Pepsin was also a Worthington product. Bovine serum albumin was a Pentex Inc. product.

Methods. The difference spectral methods were the same as those described previously (Herskovits and Laskowski, 1962a; Herskovits, 1967) and in the accompanying paper.

The mercaptoethanol reduction of pepsin was carried out essentially according to the procedure described by White (1961). Appropriate pepsin solutions in 10 M urea containing 2 μl of mercaptoethanol/mg of protein were adjusted to pH 8.5 with 1 M NaOH, purged with nitrogen, and after 4-hr standing in stoppered flasks were adjusted to the desired pH and ionic strength with two parts per ten of aqueous buffer and salt solutions. Rather than isolating the reduced protein, the adjusted reaction mixture in 8 M urea was then used as a stock solution from which dilutions were made with 8 M urea containing the appropriate 40% perturbant, as previously described (Herskovits and Laskowski, 1962a).

Concentrations were determined spectrophotometrically using the following molar extinction coefficients: 37,750 for lysozyme at 282 mμ (Sophianopoulos *et al.*,

* From the Department of Chemistry, Fordham University, New York, New York 10458. *Received March 15, 1968.* This investigation was supported by Grant 14468 from the National Institutes of Health, U. S. Public Health Service.

† Part of this work was taken from the thesis submitted in partial fulfillment of the requirements for the M.S. degree, Fordham University, 1968.

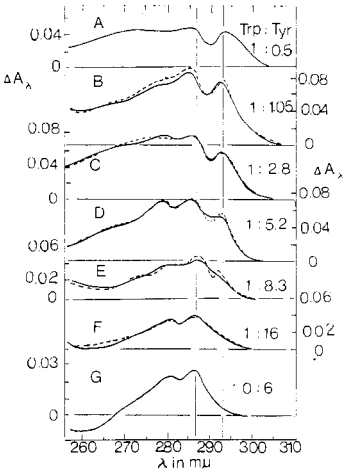

FIGURE 1: The additivity of solvent perturbation difference spectra. Solid lines represent the curves obtained on lysozyme–ribonuclease mixtures and pure ribonuclease with 20% ethylene glycol as perturbant, $\Gamma/2 = 0.1$, pH 6.8. Dotted lines represent calculated curves based on the pure lysozyme (Lys) and ribonuclease (RNase) data using the relation: $\Delta A_\lambda = c_1\Delta A_\lambda(\text{Lys}) + c_2\Delta A_\lambda(\text{RNase})$ where c_1 and c_2 represent the lysozyme and ribonuclease concentrations used. (A) Pure lysozyme, concentration = 2.83×10^{-5} M, chromophore ratio is 1:0.5. (B) 2.19×10^{-5} M RNase and 4.94×10^{-5} M Lys, giving a 1:1.05 tryptophan to tyrosine ratio, assuming 6 tyrosines per RNase and 6 tryptophans and 3 tyrosines per mole of Lys. (C) 6.55×10^{-5} M RNase to 2.83×10^{-5} M Lys mixture, giving a chromophore ratio of 1 tryptophan to 2.8 tyrosines. (D) 1.31×10^{-4} M RNase to 2.83×10^{-5} M Lys, giving a chromophore ratio of 1:5.2. (E) 8.77×10^{-5} M RNase to 1.12×10^{-5} M Lys, giving a chromophore ratio of 1:8.3. (F) 8.77×10^{-5} M RNase to 5.6×10^{-6} M Lys, giving a chromophore ratio of 1:16. (G) Pure RNase, 8.76×10^{-5} M, having a 0:6 chromophore ratio. $A_{276-278} = 0.64$–1.26; dynode voltage adjustment = 3.

1962), 9800 for ribonuclease at 277.5 mμ (Wetlaufer, 1962), 50,990 for pepsin at 280 mμ (Arnon and Perlmann, 1963), 148,200 for native aldolase at pH 6 and 280 mμ, and 131,450 for acid-denatured aldolase at pH 2.0 and 277 mμ. The aldolase values were based on a molecular weight of 158,000 (Kawahara and Tanford, 1966) and the $\epsilon_{280}^{1\%}$ of 9.38 and $\epsilon_{277}^{1\%}$ of 8.32, respectively (Donovan, 1964). The extinction of 46,000 for bovine serum albumin at 278 mμ was based on a $\epsilon_{278}^{1\%}$ of 6.67 (Sterman and Foster, 1956) and a molecular weight of 69,000.

Results and Discussion

Experimental Limitations. In the method of chromophore analysis to be described, it is necessary to be able to discern the first tryptophan difference maximum at 291–294 mμ and the first tyrosine difference maximum at 286–288 mμ in order to measure the exact absorbance difference due to tryptophan or tyrosine, respectively. To determine the ratio of chromophores at which this is not feasible, varying mixtures of proteins were analyzed. Lysozyme, a tryptophan-rich protein, and ribonuclease, which contains only tyrosine, were mixed to simulate a protein rich in both chromophores. Difference spectra were measured with 20% ethylene glycol as perturbant. The tryptophan maximum was eclipsed when the tyrosine to tryptophan ratio was about 10:1, whereas the tyrosine peak was eclipsed when the tryptophan to tyrosine ratio was about 1:1. Some of these spectra are illustrated in Figure 1. This apparent discrepancy is due to the facts that the perturbation difference spectra are about three times larger at 291–293 mμ than the tyrosyl difference spectra at 286–288 mμ and the fact that at the 286–288-mμ region tryptophan perturbations are significantly larger than tyrosine perturbations at 291–293 mμ (*cf.* Figures 1 and 6 and footnote 2 of the accompanying paper (Herskovits and Sorensen, 1968)).

In these experiments, the lysozyme content was considered to be six tryptophans and three tyrosines, while the ribonuclease chromophore content was taken to be six tyrosines with no tryptophan. It will be recalled, of course, that both ribonuclease and lysozyme have been studied thoroughly by the difference spectral method (Williams *et al.*, 1965; Herskovits and Laskowski, 1968), and it is now known that not all the chromophores in these proteins are accessible to solvent effects. However, one will not usually have such *a priori* information at hand when starting experimental work on a protein. It has been our intention to simulate approximate chromophore ratios of typical proteins for which the method can be expected to be useful.

Figure 1 also illustrates the fact that in actual protein measurements the shoulder or peak disappears at a point where calculations from the difference spectral values of the pure protein components (given as curves A and G) would still exhibit the shoulders, thus leading to a possible erroneous interpretation, *i.e.*, absence of exposed tryptophan residues. In this connection it is important to note that such shoulders may be quite discernible with one perturbant but not as readily apparent with the use of another perturbant. In the case of aldolase, for example, the difference spectral tracings of Figure 2 due to the perturbation of 20% ethylene glycol (curve C) exhibits less obvious shoulders than the similar perturbation spectra on aldolase obtained with 20% glycerol (curve D). When such shoulders are apparent, however (curves D and E of Figure 1, for example, having tyrosine to tryptophan ratios of less than 10:1), it is important to note that the discrepancies of peak heights between model and measured absorbance difference values are only of the order of 5–10%. Thus it is apparent that such measurements should lead to reasonable estimates of tryptophyl and tyrosyl exposure.

Estimation of Tyrosine and Tryptophan Exposure. A method of resolving the solvent perturbation difference spectra of proteins into their constituent tyrosyl and tryptophyl components by use of a set of simultaneous equations and appropriate tyrosyl and tryptophyl model compound data has been recommended (Herskovits, 1967). These equations are

$$\Delta\epsilon_{291-293}(\text{protein}) = a\Delta\epsilon_{291-293}(\text{Trp}) + b\Delta\epsilon_{291-293}(\text{Tyr}) \quad (1)$$

$$\Delta\epsilon_{286-288}(\text{protein}) = a\Delta\epsilon_{286-288}(\text{Trp}) + b\Delta\epsilon_{286-288}(\text{Tyr}) \quad (2)$$

where the coefficients a and b represent the apparent number of exposed tryptophyl (Trp) and tyrosyl (Tyr) residues in the protein to be studied. Here the

TABLE I: Parameters of Additives and Model Chromophores in Water and in 8 M Urea.[a,b]

Perturbant[c] (%)	Mean Diameter (Å)	Wavelength (λ) at First Trp Max (mμ)	$\Delta\epsilon_\lambda$ Ac-Trp-OEt	$\Delta\epsilon_\lambda$ Ac-Tyr-OEt	Wavelength (λ) at First Tyr Max (mμ)	$\Delta\epsilon_\lambda$ Ac-Trp-OEt	$\Delta\epsilon_\lambda$ Ac-Tyr-OEt
Molar Absorptivity Differences in Water							
Deuterium oxide (90)	2.0	292.0	−203.6	−12.2	285.5	−120.0	−67.1
Methanol (20)	2.8	291.5	235.4	16.8	285.5	135.9	75.5
Dimethyl sulfoxide (20)	4.0	292.5	489.5	35.5	286.0	168.4	213.7
Ethylene glycol (20)	4.4	292.0	305.1	16.1	285.5	172.2	92.1
Glycerol (20)	5.2	292.0	304.4	12.9	285.0	195.6	79.8
Erythritol (20)	5.8	293.0	200.0	4.2	286.0	66.3	60.3
Glucose (20)	7.2	292.5	192.2	6.2	286.0	100.0	41.5
Carbowax 300 (20)	9.2	291.0	518.7	57.1	286.0	94.3	187.0
Sucrose (20)	9.4	292.5	192.2	5.6	285.5	118.0	46.1
Molar Absorptivity Differences in 8 M Urea							
Methanol (20)	2.8	292.0	217.2[d]	21.8[d]	286.0	125.0[d]	98.0[d]
Dimethyl sulfoxide (20)	4.0	292.0	435.3	48.0	286.0	159.2	230.8
Ethylene glycol (20)	4.4	291.0	298.9	33.0	285.5	165.0	119.8
Glycerol (20)	5.2	291.0	289.1	14.2	285.5	173.9	79.1
Carbowax 300 (20)	9.2	292.0	387.4[d]	52.5[d]	286.0	70.7[d]	170.4[d]
Sucrose (20)	9.4	291.5	199.0	9.8	285.5	126.7	48.5

[a] In the presence of 0.1 M KCl and 0.01 M pH 6.8 phosphate. [b] Data taken from Tables I–III of the previous paper (Herskovits and Sorensen, 1968). The methanol and Carbowax 300 data are taken in part from Herskovits (1965) and Herskovits and Laskowski (1968). [c] With the exception of the sugars and D_2O, 20 volumes of liquid perturbant was used per 100 volumes of final solution. D_2O was 90 volumes/100 volumes of final solution. Sucrose, erythritol, and glucose solutions contained 21.6 g of sucrose, 21.36 g of erythritol, and 21.52 g of glucose per 100 ml of solution. Such a solution in water contains 20% sugar by weight. [d] For the major tyrosine and tryptophan peaks the molar absorptivity differences in 8 M urea were obtained by multiplying the $\Delta\epsilon_{max}/\epsilon_{max}$, taken from the literature, by the ϵ_{max} value reported in this paper, *i.e.*, 1420 for Ac-Tyr-OEt and 5870 for Ac-Trp-OEt. The tyrosine contribution at the 292-mμ peak and the tryptophan contribution at the 285.5- or 286-mμ peak were estimated by multiplying the aqueous $\Delta\epsilon$ values, given in columns 5 and 7, by the ratio of the $\Delta\epsilon_{max}$ values in 8 M urea and water, given in columns 4 and 8. These ratios are 0.92 and 0.75 for Ac-Trp-OEt and 1.3 and 0.91 for Ac-Tyr-OEt due to 20% methanol and Carbowax 300, respectively.

$\Delta\epsilon_\lambda$ values refer to the molar absorptivity differences of the protein and the free tryptophan and tyrosine model compound at the 291–293- and the 286–288-mμ difference spectral maxima, designated by the subscripts. The model compound data required, which serve to give the appropriate estimates of the apparent number or fraction of tryptophyls and tyrosyls exposed, are given in Table I. An estimate of chromophore exposure can be readily obtained by first neglecting the tyrosine contribution to the difference spectrum at the 291–293-mμ peak. Rearrangement of eq 1 gives us an approximate first value of a

$$a \cong \Delta\epsilon_{291-293}(\text{protein})/\Delta\epsilon_{291-293}(\text{Trp}) \qquad (3)$$

An approximate value for b is now obtained by use of eq 2 in the form

$$b \cong \frac{\Delta\epsilon_{286-288}(\text{protein}) - a\Delta\epsilon_{286-288}(\text{Trp})}{\Delta\epsilon_{286-288}(\text{Tyr})} \qquad (4)$$

Closer estimates of the a and b values can now be obtained by using eq 1 and 2. Successive readjustment of the a and b parameters is a straightforward procedure which, in a few steps of refinement, will give us a fairly close approximation of the experimentally obtained molar absorbance differences of the protein examined. By way of illustrating the use of model data for such calculations, the solvent perturbation data obtained on native pepsin, with 20% ethylene glycol as perturbant, have been analyzed and are compared in Table II. Changes in the calculated $\Delta\epsilon_{\lambda,max}$ values on successive refinement indicate that changes in the a and b parameters of the order of 5–10% are readily discernible.

It is important to emphasize the estimates of the apparent number of exposed chromophoric groups by means of eq 1 and 2 are really based on only two of the most readily discernible and accurately measurable of the solvent perturbation parameters and that the rest of the wavelength-dependent difference spectral data thus far have not been really utilized. Prudence would therefore dictate that the final choice of a and b values be

TABLE II: Calculated *a* and *b* Values for Native Pepsin in 0.1 M KCl and 0.01 M pH 5.2 Acetate Obtained with 20% Ethylene Glycol as Perturbant.

Trial	Equations Employed	Estimated Values of *a*	*b*	Calcd Molar Absorbance Differences, $\Delta\epsilon_\lambda$ At λ 292 mμ	At λ 285.5 mμ
I	3 and 4	3	6	1012	1069
II	1 and 2	2	10	771	1265
III	1 and 2	2.5	10	924	1352
IV	1 and 2	2	11	787	1357
V	1 and 2	2.3	11	879	1409
VI	1 and 2	2.4	11	909	1427
VII	1 and 2	2.5	11	940	1444
	Experimental value[a]			875 ± 60	1470 ± 85

[a] For native pepsin the actually observed maxima are at 291 and 285 mμ (see Figure 3).

made on the basis of a comparison of the complete protein difference spectrum, with the calculated curves based on a few of the closest *a* and *b* values obtained by use of the above equations. Figure 3 shows such a comparison of the pepsin data with the calculated model curves, employing some of the trial *a* and *b* values of Table II, together with tryptophyl and tyrosyl model data of Tables I and II of the preceding paper (Herskovits and Sorensen, 1968). The model curves as a function of wavelength, λ, have been calculated by use of the relation, $\Delta\epsilon_\lambda = a\Delta\epsilon_\lambda(\text{Trp}) + b\,\Delta\epsilon_\lambda(\text{Tyr})$.

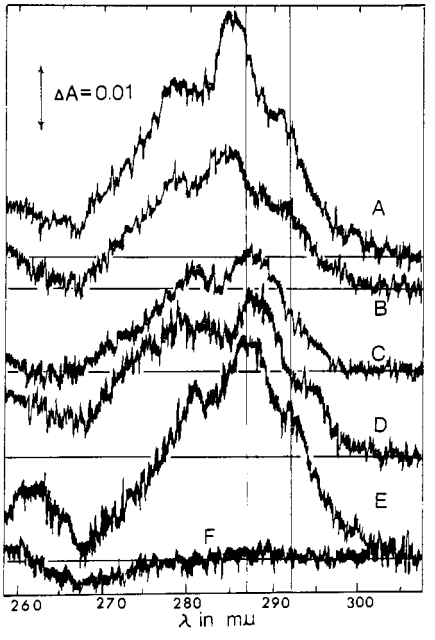

FIGURE 2: Typical Cary 14 tracings of proteins with significant tryptophyltyrosyl difference spectral contributions. The tracings were obtained with the use of the tenfold scale expander and dynode settings of 3. (A) 2.0×10^{-5} M pepsin at pH 5.6; difference spectrum due to 20% ethylene glycol. (B) 2.1×10^{-5} M pepsin, pH 5.6; difference spectrum due to 20% sucrose. (C) 1.1×10^{-5} M aldolase, pH 6.1; difference spectrum due to 20% ethylene glycol. (D) 1.2×10^{-5} M aldolase, pH 6.1; difference spectrum due to 20% glycerol. (E) 3.5×10^{-5} M bovine serum albumin, pH 5.6; difference spectrum due to 20% ethylene glycol. (F) Typical base line obtained with solvent and perturbant blanks, using the scale expander.

In view of the fact that the two protein maxima and minima centering at 291–293, 286–288, and 288–290 mμ tend to be obliterated and reduced in magnitude,[1] much is to be gained from such a comparison. Moreover, in this relation it is also important to appreciate that the experimental $\Delta\epsilon_{\lambda,\,max}$ values are often based on absorbance differences of the order of 0.008 to 0.01 optical density unit, as is the case, for example, for the 291–293-mμ tryptophyl shoulder of aldolase, shown in Figure 2.

We recommend the use of a number of perturbants of varying molecular diameters and range effects (Herskovits and Laskowski, 1962a, 1968) which in addition to giving further structural information, described later in the text, may also facilitate the location of often nearly obliterated peaks or shoulders in the main tyrosyl and tryptophyl difference spectral regions. For reasons not completely explained,[1] we have noted in a number of cases that fairly well resolved peaks are often obtained with certain perturbants but not with others, with the degree of resolution often varying from protein to pro-

[1] The obliteration of spectral and difference spectral detail and fine structure may be due to a number of causes other than lack of instrumental resolution or deviation from Beer's law. Variation of the dielectric constant and refractive index in the immediate environment of the protein chromophores due to varying neighboring amino acids in the primary structure and the side-chain environment caused by the particular folding of the native protein in the vicinity of the chromophores in question will largely determine the exact location and influence the broadness of spectral and difference spectral bands (Bayliss and McRae, 1954; Ito *et al.*, 1960; Yanari and Bovey, 1960; Wetlaufer, 1962). Short-range interactions and hydrogen-bond formation with solvent and perturbant molecules of varying hydrogen-bonding potential and possibly also with neighboring polar side chains is perhaps another source of spectral broadening to be considered (Wetlaufer *et al.*, 1958; Herskovits and Laskowski, 1962a). Finally, partly buried chromophores may be sterically blocked to the larger perturbants, but may be free to interact with the smaller perturbants, producing proportionately larger spectral shifts (Herskovits and Laskowski, 1960, 1962a, 1968).

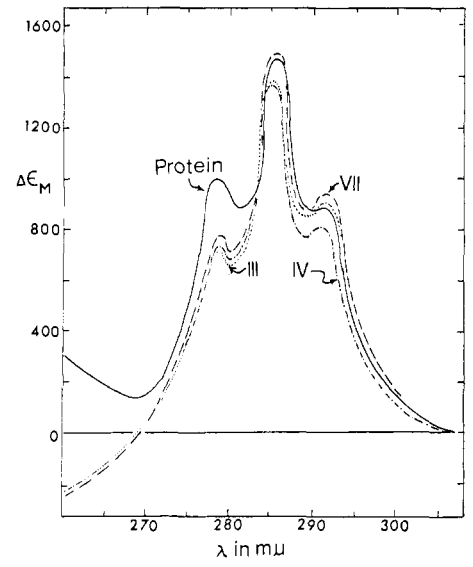

FIGURE 3: A comparison of experimental solvent perturbation difference spectra with calculated curves based on model compound data and *a* and *b* parameters obtained from the solutions of eq 1 and 2. The solid line represents the experimental data of native pepsin, pH 5.6, 0.1 M KCl, obtained with 20% ethylene glycol as perturbant. Curves III, IV, and VII are curves calculated with some of the trial *a* and *b* values of Table II and the wavelength-dependent model compound data of the accompanying paper (Herskovits and Sorensen, 1968). Curve III, $a = 2.5$, $b = 10$. Curve IV, $a = 2$, $b = 11$. Curve VII, $a = 2.5$, $b = 11$.

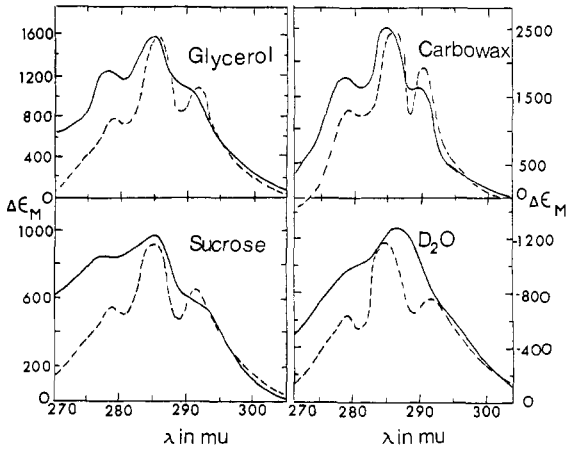

FIGURE 4: A comparison of the solvent perturbation data of native pepsin (pH 5.5–5.6, 0.1 M KCl) produced by 20% glycerol, Carbowax 300, sucrose, and 90% deuterium oxide, with calculated curves based on the *a* and *b* values of Table III. Protein concentrations: 2.0–3.5 × 10^{-5} M.

tein (see, for example, the 291–293-mμ tryptophyl shoulders of Figure 3 and Figure 4).

Finally, it should be emphasized, as has been done in previous studies on the subject (Herskovits and Laskowski, 1960, 1962a,b; Herskovits, 1967), that the solvent perturbation technique gives information only concerning the fraction or apparent number of exposed chromophoric groups. As a consequence, the interpretation of solvent perturbation data in terms of fully exposed and buried groups only is an oversimplification, or at best a shrewd guess concerning the actual spatial disposition of such groups found in the protein in question. The X-ray crystallographic studies on sperm whale myoglobin (Kendrew *et al.*, 1960), ribonuclease (Kartha *et al.*, 1967), and subtilisin-modified ribonuclease (Wyckoff *et al.*, 1967) indicate that some of the chromophoric groups in proteins will only be partly exposed, resulting clearly in fractional values of the *a* and *b* parameter, as is also found in the case of the *a* values of aldolase and serum albumin (Tables IV and V).

Protein Studies. The data obtained on pepsin best illustrate the uses and limitations of some of the perturbants commonly employed in solvent perturbation studies. The data obtained with six perturbants, ranging in diameters from 2.0 to 9.4 Å, are shown in Figures 3–5. With pepsin as well as other proteins, best fit of the data has been obtained with ethylene glycol, glycerol, dimethyl sulfoxide, and Carbowax 300 (hexaethylene glycol), making these four perturbants the most useful for such studies. The fit of experimental data on native pepsin, aldolase, and serum albumin is found to be less satisfactory with sucrose and deuterium oxide (Figure 4).

In the case of sucrose, the poor fit of the data around and below 280 mμ may be attributed to preferential interaction of the perturbant with the protein polypeptide chains, since such effects have also been noted in studies on ovomucoid (Herskovits and Laskowski, 1962b) and paramyosin (Riddiford, 1966). The relatively large spectral shifts induced by 90% deuterium oxide in the 286–288-mμ region and the loss of difference spectral detail in the 291–293-mμ region render the interpretation of the difference spectra due to this important perturbant difficult. Only preliminary estimates of tyrosyl and tryptophyl can be obtained at the present, due to largely unknown factors involving isotope substitution, which in addition to isotope shifts may also produce small conformation changes and ensuing spectral shifts (Herskovits and Laskowski, 1962a; Herskovits, 1967) in the neighborhood of the exposed chromophoric residues in question. The use of this important perturbant, due to its small diameter and other properties resembling water, will require further study.[2]

Table III gives a summary of the data obtained on native, urea-denatured, and disulfide-cleaved pepsin in 8 M urea. The calculated *a* and *b* values obtained from the best fit of the data indicate that about 40–60% of the tryptophyls and 60–75% of the tyrosyls are exposed to solvent access in the native enzyme. The data suggest no apparent gradation or decrease in "exposure" with increasing perturbant diameter, nor with the decrease in the range effect of the solvent, which has been used to probe the location of partly buried groups in ribonuclease (Herskovits and Laskowski, 1968), suggesting that only a small fraction of the chromophores are located in crevices or intrasurfaces formed by the protein fold, or are partly buried. As a result a discussion of the location of tyrosyls and tryptophyls in pepsin in terms of actually exposed chromophoric groups to specified perturbants is not unwarranted. It is fair to conclude that in the na-

[2] T. T. Herskovits and M. Sorensen, experiments in progress.

TABLE III: Difference Spectral Parameters of Native and Urea-Denatured Pepsin.

Perturbants (%)[a]	Mean Diameter (Å)	Molar Absorptivity Differences ($\Delta\epsilon_M$) Exptl λ (mμ) 290–292	Exptl λ (mμ) 285–288	Calcd λ (mμ) 291–292	Calcd λ (mμ) 285.5–286	App No. of Exposed Groups from Eq 1 and 2 *a*	App No. of Exposed Groups from Eq 1 and 2 *b*	Fraction of Residues Exposed[c] Trp	Fraction of Residues Exposed[c] Tyr
				Native[b]					
Deuterium oxide (90)	2.0	(−800)	−1280	(−757)	−1165	(3.0)	12	(0.60)	0.71
Dimethyl sulfoxide (20)	4.0	1370	2450	1334	2474	2.0	10	0.40	0.59
Ethylene glycol (20)	4.4	875	1470	940	1444	2.5	11	0.50	0.65
Glycerol (20)	5.2	1050	1560	1068	1545	3.0	12	0.60	0.71
Carbowax 300 (20)	9.2	1630	2510	1722	2433	2.0	12	0.40	0.71
Sucrose (20)	9.4	600	995	649	953	3.0	13	0.60	0.76
				8 M Urea Denatured (pH 5.8–6.6)					
Dimethyl sulfoxide (20)	4.0	3160	4790	2993	4719	5.0	17	1.00	1.00
Ethylene glycol (20)	4.4	1640	2450	1690	2458	4.0	15	0.80	0.88
Glycerol (20)	5.2	1580	2300	1557	2207	4.5	18	0.90	1.06
Sucrose (20)	9.4	1010	1610	1272	1570	5.5	18	1.10	1.06
					Av exposure	4.6 ± 0.5	17 ± 1		
				Disulfide Cleaved in 8 M Urea (pH 4.5–5.0)					
Dimethyl sulfoxide (20)	4.0	3450	5080	3427	4879	6.0	17	1.20	1.00
Ethylene glycol (20)	4.4	1720	2500	1724	2576	4.0	16	0.80	0.94
Glycerol (20)	5.2	1720	2560	1845	2380	5.5	18	1.10	1.06
Sucrose (20)	9.4	1110	1460	1251	1473	5.5	16	1.10	1.06
					Av exposure	5.3 ± 0.6	16.8 ± 1		

[a] Perturbant concentrations are defined in footnote *c* of Table I. [b] Native protein, pH 5.5–5.7, 0.1 M Cl^-; protein in 8 M urea, 0.1 M Cl^-; 2 μl of mercaptoethanol used/mg of protein. [c] Calculations were based on 5 tryptophans and 17 tyrosines per 35,000 g of pepsin.

TABLE IV: Difference Spectral Parameters of Native and Acid-Denatured Aldolase.

Perturbants (20%)[a]	Molar Absorptivity Differences ($\Delta\epsilon_M$) Exptl λ (mμ) 290–292	Exptl λ (mμ) 285–288	Calcd λ (mμ) 291–292	Calcd λ (mμ) 285.5–286	App No. of Exposed Groups from Eq 1 and 2 *a*	App No. of Exposed Groups from Eq 1 and 2 *b*	Fraction of Residues Exposed[c] Trp	Fraction of Residues Exposed[c] Tyr
			Native (pH 6.1)[b]					
Ethylene glycol	935	1915	1020	1904	2.5	16	0.22	0.34
Glycerol	910	2090	1006	2005	2.5	19	0.22	0.40
Carbowax 300	1625	3005	1692	3133	1.5	16	0.13	0.34
			Acid denatured (pH 2)[b]					
Ethylene glycol	4005	6685	4418	6405	12	47	1.04	1.00
Glycerol	4275	6485	4285	6267	12	49	1.04	1.04
Carbowax 300	6550	9725	7295	9451	9	46	0.78	0.98
				Av value	11 ± 1	47.3 ± 2		

[a] Perturbant concentrations are defined in footnote *c* of Table I. [b] Native protein, pH 6, 0.05 M Cl^-, 0.01 M pH 6 acetate; acid-denatured protein, pH 2, 0.05 M Cl^-. [c] Calculations are based on 11 tryptophyls and 47 tyrosyls per 158,000 g of aldolase.

tive enzyme 2–3 of the 5 tryptophyls and 10–12 of the 17 tyrosyls are exposed, with perhaps only 1 or 2 of these residues being partly buried. The average values of exposed tyrosyls correspond fairly well with the estimates of 8–10 groups obtained by Perlmann (1966) from chemical reactivity data toward acetylimidazole.

Urea (8 M) has nearly completely unfolded pepsin according to our criteria, or at least rendered the tyrosyl and tryptophyl residues largely accessible to perturbant penetration. The somewhat lower than theoretical *a* value obtained suggest that perhaps 1 or 2 of the chromophores (tryptophyls) adjoin disulfide bridge, as has been found in the case of lysozyme (Williams *et al.*, 1965), ribonuclease (Herskovits and Laskowski, 1968), and bovine serum albumin (Herskovits and Laskowski, 1962a). It should be noted here that this is not the case for α-chymotrypsinogen where the amino acid sequence shows no such chromophore adjoining disulfides and in acid-denatured aldolase which also contains no disulfide bridges (Table IV).

It is worth noting that in the case of the unfolded disulfide-cleaved protein, average values of exposure are close to 5 tryptophyls and 17 tyrosyls. Brand's (1946) amino acid data give 4 tryptophyls and 16 tyrosyls, while more recent studies (Blumenfeld and Perlmann, 1959; Perlmann, 1964, 1966) give 6 tryptophyls and 16 tyrosyls per mole, having a molecular weight of 35,000. Based on the molar extinction coefficients of tryptophan (ϵ_{278} 5550) tyrosine ($\epsilon_{274.5}$ 1340), and cystine (ϵ_{280} 150) (Wetlaufer, 1962) and the average value of 5 tryptophyls, 17 tyrosyls, and 3 cystinyl residues, we obtained a molar extinction value of 50,980 which is in close agreement with the experimental value of 50,990 reported by Arnon and Perlmann (1963).

The acid-denatured aldolase data summarized in Table IV are also in satisfactory accord with the literature, namely, 11–12 tryptophyls, based on the amino acid analysis of Velick and Ronzoni (1948), and 47 tyrosyls, based on the spectrophotometric data of Donovan (1964), lending additional confidence in our method of analysis. These values for tyrosine and tryptophan have been recalculated on the basis of the more recent higher value of 158,000 for the molecular weight of aldolase reported by Kawahara and Tanford (1966).

In sharp contrast with the findings on native pepsin, in native aldolase first studied by Donovan (1964) using

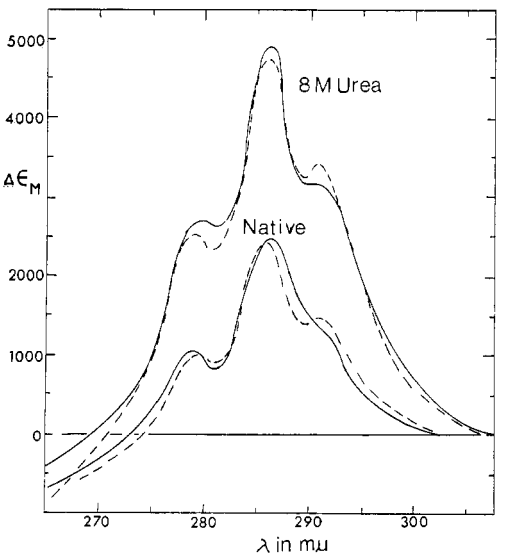

FIGURE 5: Solvent perturbation difference spectra of native (pH 5.2, 0.1 M KCl) and 8 M urea-denatured pepsin (pH 6.6, 0.1 M KCl, 0.01 M acetate) obtained with 20% dimethyl sulfoxide as perturbant. Solid lines represent the experimental data, while dashed lines represent theoretical curves calculated with *a* and *b* values of 2 and 10, and 5 and 17, respectively. Protein concentrations of 1.2 and 3.5 × 10^{-5} M.

TABLE V: Difference Spectral Parameters of Native and Acid Form of Bovine Serum Albumin.

Perturbants (%)[a]	Mean Diameter (Å)	Molar Absorptivity Differences ($\Delta\epsilon_M$) Exptl λ (mμ) 291–293	Exptl λ (mμ) 286–288	Calcd λ (mμ) 291–292	Calcd λ (mμ) 285.5–286	App No. of Exposed Groups from Eq 1 and 2 *a*	*b*	Fraction of Residues Exposed[d] Trp	Tyr
				Native[b]					
Deuterium oxide (90)	2.0	(−680)	−760	−505	−777	(2)	8	(1.0)	0.38
Ethylene glycol (20)	4.4	675	1020	678	1047	1.8	8	0.90	0.38
Glycerol (20)	5.2	520	800	547	852	1.5	7	0.75	0.33
Carbowax 300 (20)	9.2	715	1015	804	1029	1.0	5	0.50	0.24
				Acid Form[c] (pH 2.0–2.1)					
Ethylene glycol (20)	4.4	650	1070	678	1047	1.8	8	0.90	0.38
Glycerol (20)	5.2	830	1440	777	1428	2.0	13	1.00	0.62
Carbowax 300 (20)	9.2	1140	1740	1136	1796	1.2	9	0.60	0.43

[a] Perturbant concentrations are defined in footnote *c* of Table I. [b] pH 5.6–6.8, 0.1 M KCl. [c] 0.1 M Cl^-. [d] Based on 2 tryptophyls and 21 tyrosyls per mole (Brand, 1946), taking the molecular weight at 69,000.

20% ethylene glycol as perturbant, a much larger fraction of the tryptophyls and tyrosyls seem to be buried. The data of Table IV and Figure 6 indicate that only about 15–20% of the 11–12 tryptophyls and 35–40% of the 47 tyrosyls are exposed. These observations are in accord with the unusually low chemical reactivity of the tryptophyl with α-dimethylaminobenzaldehyde observed by Velick and Ronzoni (1946) and the estimate of 10–15 exposed tyrosyls which were found to react with acetylimidazole (Pugh and Horecker, 1967). Within the present uncertainties of the experimental method and the limited number of perturbants employed so far, little more can be said concerning these estimates. The fact that the major tyrosyltryptophyl peak, usually located at 286–288 mμ, and the tryptophyl shoulder at 291–293 mμ shown in the spectrophotometric tracings of Figure 2 are shifted to somewhat longer wavelength than in either the case of pepsin or serum albumin spectra would suggest that some of the perturbed chromophores may be partly buried. It is known, for example,

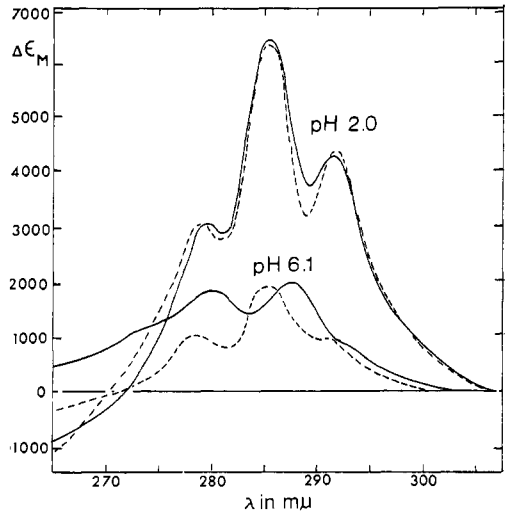

FIGURE 6: Solvent perturbation difference spectra of native aldolase (pH 6.1, 0.05 M KCl, 0.01 M acetate) and acid-denatured aldolase (pH 2.0, 0.05 M Cl^-) obtained with 20% glycerol as perturbant. Solid lines represent the experimental curves, while the dashed lines are calculated curves calculated with *a* and *b* parameters of 2.5 and 19, and 12 and 49, for the native and denatured aldolase, respectively. Protein concentrations: 0.68–1.2 × 10^{-5} M.

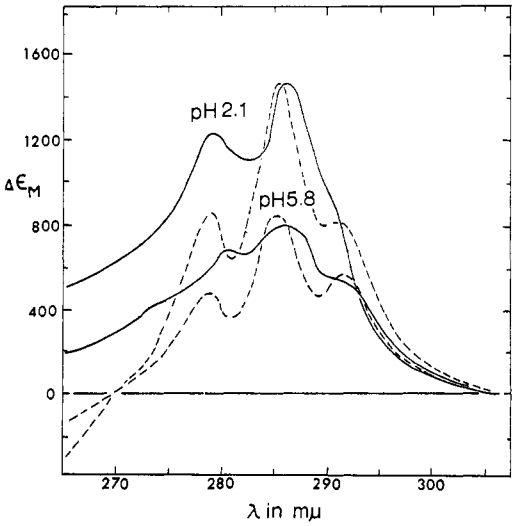

FIGURE 7: Solvent perturbation difference spectra of native (pH 5.8, 0.1 M KCl) and acid-expanded bovine serum albumin (pH 2.1, 0.1 M Cl^-) obtained with 20% glycerol. Solid curves represent the experimental data, while the dashed lines represent theoretical curves calculated with *a* and *b* values of 1.5 and 7, and 2 and 13, for the native and the acid-expanded protein, respectively. Protein concentrations: 3.5 × 10^{-5} M

that the incorporation of tyrosyls or tryptophyls into the polypeptide matrix results in "red shifts" or spectral shifts to longer wavelength (Beaven and Holiday, 1952; Wetlaufer, 1962; Donovan, 1964; Herskovits, 1965). The *a* value of 1.5–2.0 may therefore represent the possible difference spectral contribution of perhaps only one exposed and a few partly buried residues, or perhaps several partly buried tryptophyl residues.

Studies on the tyrosyls in human and bovine serum albumin (Herskovits and Laskowski, 1962a) and the tyrosyls and tryptophyls in L-glutamate dehydrogenase (Cross and Fisher, 1966) have suggested that the systematic use of perturbants of varying diameters may be used to explore the surface and geometry of subunit and multichained proteins. From a pH study on the two serum albumins it was concluded that perturbants with molecular diameters greater than 5.2 Å (glycerol, sucrose, and Carbowax 300) are excluded from the subunit intersurfaces of the native form of the albumins. In the case of the acid-expanded form below pH 4, it was found that about 20% or four to five of the tyrosyls became accessible to these perturbants. With perturbants of less than 4.4-Å diameter, however (ethylene glycol and dimethyl sulfoxide), essentially no pH dependence was noted, suggesting that these key tyrosyls must be located in crevices or areas of subunit contact, accessible to the latter smaller perturbants in both the native and acid-expanded forms. On the basis of a variety of physicochemical measurements Foster (1960) proposed a model for the serum albumins consisting of four flat conformational subunits stacked one on top of another. The albumin molecule, folded from a single polypeptide chain, is visualized as having three subunit interfaces, which on the basis of solvent perturbation studies (Herskovits and Laskowski, 1962a; Leonard and Foster, 1961) appear to be at least partly water permeable.

Table V and Figure 7 present a summary of our present data on bovine serum albumin analyzed in terms of both tyrosyl and tryptophyl contributions to the solvent perturbation spectra. Significantly, the two tryptophyl residues in serum albumin are found to be appreciably more exposed than in either pepsin or in aldolase. Moreover, in addition to being significantly more exposed than the tyrosyl residues,[3] there is also a more pronounced increase in partial chromophore exposure as the size or diameter of the perturbant is decreased. The fact that the tryptophyl residues seem to be 75–100% exposed with perturbants of less than 5.2 Å would suggest that the two tryptophyls are probably located at the periphery of some of the postulated conformational subunits (Foster, 1960) or in areas of polypeptide folding, which as far as the exposure of side chains are concerned are only marginally affected by changes in pH. While further more detailed studies will be required to substantiate fully these contentions,[2] our present findings amply illustrate the kind of information one may obtain concerning the location of tyrosyl and tryptophyl residues by the systematic use of perturbants of varying diameters.

References

Arnon, R., and Perlmann, G. E. (1963), *J. Biol. Chem. 238*, 963.

Bayliss, N. S., and McRae, E. G. (1954), *J. Phys. Chem. 58*, 1002, 1006.

Beaven, G. H., and Holiday, E. R. (1952), *Advan. Protein Chem. 7*, 319.

Blumenfeld, O. O., and Perlmann, G. E. (1959), *J. Gen. Physiol. 42*, 563.

Brand, E. (1946), *Ann. N. Y. Acad. Sci. 47*, 187.

Cross, D., and Fisher, H. (1966), *Biochemistry 5*, 880.

Donovan, J. W. (1964), *Biochemistry 3*, 67.

Foster, J. F. (1960), *in* The Plasma Proteins, Putnam, F. W., Ed., New York, N. Y., Academic, Chapter 6.

Herskovits, T. T. (1965), *J. Biol. Chem. 240*, 628.

Herskovits, T. T. (1967), *Methods Enzymol. 11*, 748.

Herskovits, T. T., and Laskowski, M., Jr. (1960), *J. Biol. Chem. 235*, PC56.

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

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

Herskovits, T. T., and Laskowski, M., Jr. (1968), *J. Biol. Chem. 243*, 2123.

Herskovits, T. T., and Sorensen, M. (1968), *Biochemistry 7*, 2523 (this issue; preceding paper).

Ito, M., Inuzuka, K., and Imanishi, S. (1960), *J. Am. Chem. Soc. 82*, 1317.

Kartha, G., Bello, J., and Harker, D. (1967), *Nature 213*, 862.

Kawahara, K., and Tanford, C. (1966), *Biochemistry 5*, 1578.

Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C., and Shore, V. C. (1960), *Nature 185*, 422.

Leonard, W. J., Jr., and Foster, J. R. (1961), *J. Biol. Chem. 236*, PC73.

[3] Previous studies on the location of the tyrosyls in human and bovine serum albumin (Herskovits and Laskowski, 1962a) have suggested that in the native form of these proteins only about 30% of the tyrosyls are exposed, with an additional 20% or three to five groups being located in crevices or areas of subunit contact, not accessible to bulkier perturbants having greater than 5.2-Å diameter. These estimates were based on the comparison of the difference spectral constants (at 286–288 mμ) with values obtained on the fully unfolded, disulfide-cleaved proteins in 8 M urea, or model compound mixture having the molar tyrosine to tryptophan composition of the proteins. Since the tryptophyl residues are more exposed than the tyrosyls (Table V), such a comparison should lead to slight overestimates of the exposed tyrosyls. Having realized at the time that the small but significant tryptophyl contribution to the 286–288-mμ absorbance differences of bovine serum albumin may lead to slight overestimates or underestimates of tyrosyl exposure, depending on whether the 2 tryptophyls are on an average more exposed or perhaps less exposed than the 21 tyrosyls (Herskovits and Laskowski, 1962a), we have also studied the location of tyrosyls in human serum albumin. Human serum albumin contains only 0.6 mole of tryptophan and 18 moles of tyrosine per average mole of protein (Brand, 1946). Since the tryptophan contribution to the 286–288-mμ difference spectrum is sufficiently small and can be neglected, the formal estimates of the tyrosyl exposure based on the findings on both of these proteins will be only slightly altered, with the general conclusions concerning the structure of the serum albumins remaining the same.

Perlmann, G. E. (1964), *J. Biol. Chem. 239*, 3762.
Perlmann, G. E. (1966), *J. Biol. Chem. 241*, 153.
Pugh, E. L., and Horecker, B. L. (1967), *Arch. Biochem. Biophys. 122*, 196.
Riddiford, L. M. (1966), *J. Biol. Chem. 241*, 2792.
Sophianopoulos, A. J., Rhodes, C. K., Holcomb, D. M., and Van Holde, K. E. (1962), *J. Biol. Chem. 237*, 1107.
Sterman, M. D., and Foster, J. F. (1956), *J. Am. Chem. Soc. 78*, 3652.
Velick, S. F., and Ronzoni, E. (1948), *J. Biol. Chem. 173*, 627.
Wetlaufer, D. B. (1962), *Advan. Protein Chem. 17*, 303.
Wetlaufer, D. B., Edsall, J. T., and Hollingworth, B. R. (1958), *J. Biol. Chem. 233*, 1421.
White, F. H., Jr. (1961), *J. Biol. Chem. 236*, 1353.
Williams, E. J., Herskovits, T. T., and Laskowski, M., Jr. (1965), *J. Biol. Chem. 240*, 3574.
Wyckoff, H. W., Hardman, K. D., Allewell, N. M., Inagami, T., Johnson, L. N., and Richard, F. M. (1967), *J. Biol. Chem. 242*, 3984.
Yanari, S., and Bovey, F. A. (1960), *J. Biol. Chem. 235*, 2818.

Acylated β-Caseins. Effect of Alkyl Group Size on Calcium Ion Sensitivity and on Aggregation*

Peter D. Hoagland

ABSTRACT: A series of acylated β-caseins ending with hexanoylated β-casein has been prepared. The nature of the alkyl group affects both the calcium ion sensitivity and aggregation of modified β-casein. The increase in net negative charge resulting from acylation apparently reduces calcium ion sensitivity and aggregation through electrostatic repulsion. A marked dependence of solubility in 0.025 M $CaCl_2$ of some acylated β-caseins upon pH near 6 was observed. Ionization of imidazole groups and/or phosphate groups is implicated in calcium ion binding. Increase of alkyl chain length of the acylating agent increased the sedimentation coefficient of the aggregate. Hydrophobic bonding is likely responsible for this enhanced aggregation.

β-Casein is one of the major components of bovine casein. In neutral solution this protein aggregates above 4° (von Hippel and Waugh, 1955; Sullivan *et al.*, 1955; Payens and van Markwijk, 1963) and above 18° exhibits calcium ion sensitivity, *i.e.*, it can be precipitated by calcium ions (Zittle and Walter, 1963). A high proline content, coupled with the absence of cysteine and cystine (Peterson *et al.*, 1966), confer a disordered conformation upon β-casein monomers (Herskovits, 1966; Noelken and Reibstein, 1967). Succinylation of the amino groups of β-casein reduces aggregation markedly and prevents precipitation by calcium ions (Hoagland, 1966). At pH 7, the net negative charge of β-casein A[1] of 11/molecule (mol wt 23,000–24,000; Peterson *et al.*, 1966; Sullivan *et al.*, 1955; McMeekin, 1954) is increased some threefold by near-complete succinylation of the 12 amino groups. Succinylated β-casein is presumably maintained in monomeric form by intermolecular electrostatic repulsion. Bound calcium ions cannot reduce the electrostatic repulsion sufficiently to permit precipitation. On the other hand, acetylation of the amino groups of β-casein roughly only doubles the net negative charge at pH 7. Aggregation of this derivative occurs, but is less than for β-casein; moreover, acetylated β-casein exhibits calcium ion sensitivity, although to a lesser extent than β-casein (Hoagland, 1966). Since alkyl groups probably participate in aggregation of β-casein, which is rich in hydrophobic amino acid residues (von Hippel and Waugh, 1955), the methyl groups of acetylated β-casein were replaced by alkyl groups of increasing chain length. The effect of larger alkyl groups upon calcium ion sensitivity and aggregation of acylated β-casein is the subject of this report.

Experimental Section

Materials. β-Casein A[1] (Peterson *et al.*, 1966) was prepared from milk of a typed cow by the urea fractionation method of Aschaffenburg (1963). A 4.5 × 25 cm DEAE-cellulose column was equilibrated at 20° with 0.01 M imidazole–3.3 M urea–0.1% mercaptoethanol (Thompson, 1966) buffer adjusted to pH 7.0 with 1 N HCl. A 100-ml solution of 5 g of β-casein in this buffer was added to the column. Contaminating κ-

* From the Agricultural Research Service, U. S. Department of Agriculture, Eastern Regional Research Laboratory, Philadelphia, Pennsylvania 19118. *Received March 19, 1968.*

[1] Mention of products or companies does not constitute an endorsement by the Department of Agriculture over others of a similar nature not mentioned.