values of the dihedral angles may be in error by 10° or more, since relatively small changes in the positions of the carbon atoms correspond to large changes in these angles. The S–S–S bond angle of 113° appears to be rather larger than normal, comparing with the values of 104 ± 5° found in dimethyltrisulfide[2] and 105 ± 2° in S_8 molecules in the vapor state.[11]

TABLE II

INTERATOMIC DISTANCES AND ANGLES IN 2,2′-DIIODODIETHYLTRISULFIDE

Bond distances, Å.		Bond angles	
S_1–S_2	2.05	S_1–S_2–S_1	113°
S_1–C_2	1.86	S_2–S_1–C_1	98°
C_2–C_1	1.55	S_1–C_2–C_1	114°
C_1–I	2.09	C_2–C_1–I	109°

The van der Waals Radius of Sulfur.—Each iodine atom makes two close contacts with other iodine atoms, and three close contacts with sulfur atoms. These contacts, which are indicated in Fig. 2, are I...I = 4.44 Å. and 4.50 Å., and I...S = 3.74 Å., 3.80 Å. and 3.85 Å. The shortest contacts between the layers in elementary iodine[12] are 4.35, 4.40 and 4.46 Å. As in diiododiethyltrisulfide, these distances are slightly greater than 4.30 Å., the distance predicted on the basis of the van der Waals radius of 2.15 Å.[13] If we assume a slightly increased van der Waals radius of 2.20 Å. for iodine, we obtain, by subtracting this value from the three I...S distances, values of 1.54, 1.60 and 1.65 Å. for the van der Waals radius of sulfur. These values compare with the radii derived from the closest S...S contacts between S_8 molecules in rhombic sulfur,[14] namely, 1.64 Å. and 1.66 Å. All of these values, however, are significantly smaller than the ionic radius of 1.85 Å., the value taken as the van der Waals radius.[13] It thus appears that a downward revision of about 0.25 Å. should be made in the van der Waals radius of bicovalent sulfur.

Acknowledgments.—I wish to thank Dr. E. W. Hughes of this Laboratory for pointing out the derivation giving the correct scale factor for use when atoms are omitted in the calculation of structure factors. Most of the calculations were performed by Mrs. June Jenkins and Miss Lillian Casler.

(11) C. S. Lu and J. Donohue, THIS JOURNAL, **66**, 818 (1944).

(12) "Strukturbericht," Vol. II, p. 5.

(13) L. Pauling, "The Nature of the Chemical Bond," Cornell University Press, Ithaca, N. Y., 1940, p. 164.

Summary

The data of Dawson and Robertson on the crystal structure of 2,2′-diiododiethyltrisulfide have been examined. Reassignment of the parameters of the carbon atoms leads to a structure which is in agreement with the Fourier projection of electron density on (010), and which gives much improved agreement of observed with calculated structure factors. In this structure the dihedral angles S–S–S–C and S–S–C–C are both close to 90°, while the group S–C–C–I is coplanar and *trans*. It is also suggested that the van der Waals radius of bicovalent sulfur be revised downward 0.25 to 1.60 Å.

(14) B. E. Warren and J. T. Burwell, *J. Chem. Phys.*, **3**, 6 (1935).

PASADENA 4, CALIF. RECEIVED NOVEMBER 4, 1949

[CONTRIBUTION FROM THE SLOAN-KETTERING INSTITUTE FOR CANCER RESEARCH]

Heterogeneity of the Binding Sites of Bovine Serum Albumin[1]

BY FRED KARUSH[1a]

In a previous communication[2] it was shown that the binding by bovine serum albumin of three alkyl sulfates could not be described by the simple mass action equation based on the assumption of a single intrinsic association constant, even if correction is made for electrostatic interaction. If, on the other hand, the free energies of binding of the various sites were assumed to obey a Gaussian distribution, then a theoretical expression deduced on this basis could, with proper parameters, yield an adequate description of the binding data.[2a]

The limited range of data available in that study, particularly with regard to the average number of anions bound per protein molecule, did not, however, permit a critical evaluation of this particular heterogeneity assumption. To achieve this, the binding of the anionic azo dye *p*-(2-hydroxy-5-methylphenylazo)-benzoic acid was investigated since it was anticipated that this dye would permit more extensive data to be secured, as indeed turned out to be the case. The use of a dye was especially advantageous in that it greatly simplified the analytical problem.

In addition to serving as a test of the heterogeneity theory the binding studies reported here

(1) Presented in part before the Division of Biological Chemistry at the Atlantic City meeting of the American Chemical Society, September, 1949.

(1a) Investigation conducted during tenure of a Fellowship in Cancer Research of the American Cancer Society, recommended by the Committee on Growth of the National Research Council; present address: The Neurological Institute, New York 32, N. Y.

(2) F. Karush and M. Sonenberg, THIS JOURNAL, **71**, 1369 (1949).

(2a) In the present paper the term "heterogeneity" is used in the general sense of non-homogeneity. Previously[2] it was employed in the more restricted sense of referring to a Gaussian distribution.

were directed toward the production of the maximum amount of thermodynamic data regarding the binding of the dye. This was achieved by working at two carefully controlled temperatures, and under conditions designed to ensure maximum accuracy.

Experimental

Materials.—The protein used in this study was crystallized bovine serum albumin obtained from Armour and Company. The moisture content of the material was determined by drying to constant weight at 105°. A value of 3.4% was found, which, together with a figure of 1.0% for the ash content, specified by the manufacturer, was employed to calculate the actual protein content.

The anionic dye, *p*-(2-hydroxy-5-methylphenylazo)-benzoic acid, was prepared by coupling diazotized *p*-aminobenzoic acid with five times the theoretical amount of *p*-cresol in alkaline solution. Purification of the dye was effected by precipitation of the colored material in the reaction mixture with acid, filtration, resolution of the precipitate with alkali at 70°, precipitation with acid, and filtration at the elevated temperature. The precipitated dye was then dissolved with alkali at room temperature, filtered, reprecipitated with acid, filtered and washed with copious quantities of distilled water. The dye was recrystallized twice from ethylene glycol monomethyl ether, the product washed with distilled water and thoroughly dried. The material melted at 250–252° with decomposition.

Dialysis Method.—The binding of the dye to the protein was determined by the method of equilibrium dialysis. Fifteen-ml. portions of a solution of protein in 0.05 *M* phosphate, *p*H 7.0, contained in dialysis bags prepared with Visking sausage casing, were equilibrated against equal volumes of solutions of dye in the same buffer. For this purpose glass vials with rubber-lined screw-cap covers were found convenient. These were placed in a horizontal position on a rack immersed in a constant temperature water bath and gently rocked at two cycles per minute. Experiments were conducted at 25.0 ± 0.1° and 5.0 ± 0.1°. For the latter case a constant temperature water bath was kept in a cold room whose temperature was maintained at 3–4°. Rocking times sufficient to ensure equilibrium were ascertained by determining the outside dye concentrations after various periods of rocking with the usual concentration of protein inside and with the dye initially inside and outside. It was established that eight hours and twenty-four hours were sufficient at 25° and 5°, respectively. For most of the experiments a protein concentration of 3.00×10^{-5} *M* was used, based on the amount of albumin weighed out and a molecular weight of 69,000.[3] This was checked by measurement of the light absorption in alkali at λ291 of a diluted portion. All binding determinations were done in duplicate.

Because the dye is rather strongly adsorbed on the casing, it is necessary to make a correction for this factor in the calculation of the amount of dye bound to the protein. For this purpose the adsorption of dye on casing was determined over a range of concentrations of free dye at the two temperatures. With these results curves were prepared showing the amount of dye adsorbed in per cent. of the free dye, as a function of the free dye concentration. At 25° the correction ranged from 5% at 8×10^{-5} to 10% at 1×10^{-5} *M* and at 5° from 10% at 6×10^{-4} to 16% at 1×10^{-5} *M*.

Analytical Method.—Dye concentrations were determined by measurement of the light absorption with the model DU Beckman spectrophotometer. The absorption spectra and molar extinction coefficients were obtained for the dye in 0.05 *M* phosphate buffer, *p*H 7.0 and in 0.1 *N* alkali. In the former solvent the $\lambda_{max.}$ is 330 mμ with the molar extinction coefficient, ε, equal to 2.10×10^4; in alkali $\lambda_{max.}$ is 490 mμ and ε is 1.10×10^4. Because of the presence in the outside solutions of impurity absorbing slightly at λ330, it was considered unsafe to measure the dye concentrations in these solutions at this wave length. Instead, an appropriate dilution in alkali (*p*H ~ 13) of the dye solution was employed. The concentration of free dye in equilibrium with the bound dye was taken to be equal to the outside dye concentration since the Donnan correction was negligible under the prevailing conditions.

(3) E. J. Cohn, W. L. Hughes, Jr., and J. H. Weare, THIS JOURNAL, **69**, 1753 (1947).

Results

In Table I are summarized the binding data of three separate experiments carried out at 25° with an initial albumin concentration of 3.00×10^{-5} *M*. The results are expressed in terms of *r*, the average

TABLE I

BINDING OF ANIONIC DYE BY BOVINE SERUM ALBUMIN IN 0.05 *M* PHOSPHATE BUFFER

Initial dye concn., m./l. × 10^5	Concn. of free dye (c), m./l. × 10^5	r	r/c × 10^{-4}
Initial protein concn. = 3.00×10^{-5} *M*, T = 25°			
158.2	55.0	13.71	2.495
78.9	22.93	9.64	4.20
49.8	12.45	7.50	6.03
33.7	7.25	5.92	8.16
25.7	4.91	5.00	10.19
18.9	3.06	4.04	13.21
12.88	1.76	3.00	17.07
7.97	0.90	1.99	22.1
5.96	0.625	1.52	24.4
3.98	0.386	1.04	27.0
11.76	1.57	2.77	17.6
7.85	0.91	1.95	21.4
5.90	0.63	1.503	23.9
3.95	0.384	1.035	27.0
17.73	2.76	3.89	14.10
11.80	1.57	2.78	17.70
7.86	0.916	1.95	21.3
5.90	0.644	1.49	23.2
3.96	0.393	1.03	26.25
Initial protein concn. = 3.00×10^{-5} *M*, T = 5°			
78.9	20.55	10.94	5.33
49.8	10.78	8.48	7.86
33.7	6.06	6.62	10.92
25.7	4.04	5.49	13.60
18.9	2.38	4.47	18.8
12.88	1.29	3.29	25.5
7.97	0.62	2.18	35.2
5.96	0.43	1.65	38.4
3.98	0.26	1.125	43.3
Initial protein concn. = 10.0×10^{-5} *M*, T = 25°			
195.5	37.2	11.54	3.11
157.0	25.9	10.08	3.90
116.9	15.52	8.30	5.35
78.3	7.62	6.15	8.07
Initial protein concn. = 3.00×10^{-5} *M*, T = 25°, Cl^- = 0.100 *M*			
17.73	3.33	3.47	10.43
11.80	2.11	2.39	11.33
7.86	1.29	1.675	13.0
5.90	0.94	1.278	13.6
3.96	0.608	0.875	14.4

number of moles of dye bound per mole of protein and r/c where c is the free equilibrium dye concentration. In Fig. 1 they are plotted in terms of r/c *vs.* r, the reason for which will be indicated shortly. It may be noted that though volume changes of the inside and outside solutions of as much as 1 ml. do occur during the dialysis, they do not significantly affect the values of r and c. The data for the low temperature binding are also given in Table I and plotted in Fig. 1.

It has been observed by Klotz and Urquhart[4] that an increase of albumin concentration from 0.2 to 1% results in a considerable decrease in r at constant c when the binding is determined at 0°. We have investigated this question by comparing the binding data obtained at 25° with an albumin concentration of 10.0×10^{-5} M (0.69%) with that for 3.00×10^{-5} M (0.207%). The results are shown in Table I and plotted in Fig. 1. It is clear from the plot that there is no detectable concentration effect with respect to the relation between r and c. It is not unlikely that the absence of the effect in our experiment is due to the higher temperature employed.

Discussion

On the basis of the law of mass action it readily follows that if there are n binding sites per protein molecule which have intrinsic association constants K_i and if there is no interaction among the bound ions, then the dependence of r on c is given by

$$r = \sum_i \frac{K_i c}{1 + K_i c} \quad (i = 1, 2, \ldots n) \qquad (1)$$

It has been suggested by Scatchard[5] for the special case that the K_i's are equal that if (1) is rearranged to give

$$r/c = nK - rK \qquad (2)$$

a plot of the experimental data in terms of r/c *vs.* r will permit the evaluation of n and K from the appropriate intercepts. He pointed out, furthermore, that if the experimental points do not fall on a straight line it may be inferred that the K_i's are not equal and/or there is interaction among the bound ions. The intercept on the r axis will still give the value of n and, as follows directly from (1)

$$\lim_{c \to 0} \frac{r}{c} = \sum_i K_i \qquad (3)$$

We have found it convenient, particularly because of the wide range of r, to plot our data as suggested by Scatchard. A casual examination of the points in Fig. 1 reveals a striking deviation from linearity and thereby poses the problem of an adequate quantitative representation and interpretation of our results.

State of Dispersion of Dye.—It is well known that many dyestuffs in aqueous solution undergo dimerization and even more extensive aggregation. Since such an effect could cause deviation from linearity, it was necessary to ascertain whether this phenomenon was involved in the dye solutions employed in our experiments. An examination was made therefore of the equilibrium distribution of the dye between the aqueous phase and an organic solvent over the concentration range of interest. To 10-ml. aliquots of various dye solutions in 0.05 M phosphate, pH 7.0, were added 10-ml. portions of a mixed organic solvent consisting of 50% $CHCl_3$ and 50% CCl_4 by volume. The two-phase systems were mixed and centrifuged. The equilibrium concentration of the dye in the aqueous phase was then determined spectrophotometrically at λ490 by appropriate dilution in alkali. The extractions were done at room temperature (22°) in duplicate. The results are shown in Table II in terms of per cent. of original dye extracted. The essential constancy of this value over a 50-fold range of concentration renders it highly probable that the dye is molecularly dispersed in the aqueous phase (as well as in the organic phase).

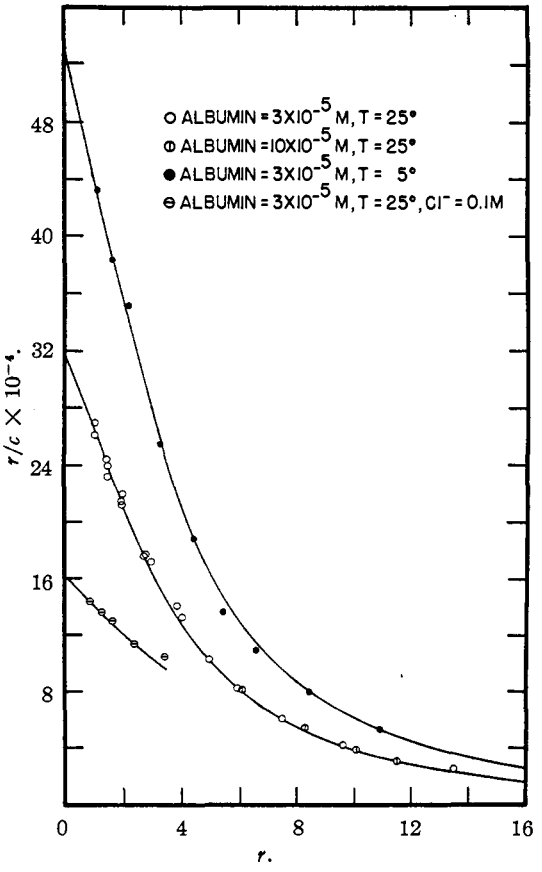

Fig. 1.—Binding of an anionic azo dye by bovine serum albumin in 0.05 M phosphate buffer, pH 7.0. The points are experimental and the upper two curves are theoretical: O, albumin 3×10^{-5} M, T 25°; ⊕, albumin 10×10^{-5} M, T 25°; ●, albumin 3×10^{-5} M, T 5°; ⊖, albumin 3×10^{-5} M, T 25°, Cl^- 0.1 M.

(4) I. M. Klotz and J. M. Urquhart, *J. Phys. Coll. Chem.*, **53**, 100 (1949).

(5) G. Scatchard, *Ann. N. Y. Acad. Sci.*, **51**, 660 (1949).

TABLE II

PARTITION OF DYE BETWEEN 0.05 M PO_4, pH 7.0 AND ORGANIC SOLVENT (50% $CHCl_3$, 50% CCl_4) AT 22°

Initial aqueous dye concn. × 10^5	Final aqueous dye concn. × 10^5	Extracted, %
98.0	40.8	58.4
39.5	16.6	58.0
19.7	8.48	57.1
5.96	2.53	57.6
2.01	0.85	57.8

Further evidence bearing on this question was adduced from a study of the absorption spectrum of the dye as a function of concentration at room temperature. Absorption curves were determined for the dye in 0.05 M phosphate pH 7.0 over a 100-fold concentration range, from 100 × 10^{-5} M to 1.00 × 10^{-5}. The results, given in Table III, demonstrate that within experimental error the molar extinction coefficient at 330 mμ remains constant. Furthermore, the shape of the absorption curve as judged by the ratio of the optical density at 330 mμ to that at 385 mμ also is unchanged. The measurement at 385 mμ is convenient for this purpose because the curve has a shoulder at this wave length. Thus the spectral results, in agreement with the partition experiments, substantiate the conclusion as to the monomolecular dispersion of the dye.

TABLE III

SPECTRAL ABSORPTION OF DYE IN 0.05 M PHOSPHATE, pH 7.0 AS A FUNCTION OF CONCENTRATION

Concn. m./l. × 10^5	Optical path length in mm.	ϵ at λ330 × 10^4	ϵ at λ385 × 10^4	$\epsilon_{\lambda 330}/\epsilon_{\lambda 385}$
100	0.17	2.19	0.965	2.27
20.0	1.00	2.12	.941	2.26
1.00	10.00	2.11	.93	2.27

Electrostatic Interaction.—Because the binding of the anionic dye may cause a net change in the average electrostatic charge of the protein, it is necessary to consider the effect on the free energy of binding of the ith anion of the $i - 1$ dye ions already bound. This may be done by the use of a modified form of (2) recently developed by Scatchard.[5]

$$(r/c)e^{2w'r} = nK - rK \qquad (4)$$

where $w' = (1 + 1/n)\,w$ and the value of w is calculated from the Debye-Hückel theory. As is evident from this theory and the relation

$$\Delta F_{\text{elec.}} = 2RTw \qquad (5)$$

w is a measure of the increase in free energy of binding, due to electrostatic interaction, of the ith dye ion compared to that for the $i - 1$ ion.

We have calculated the value of w at 25°, in the manner previously described for alkyl sulfates,[2] using a value of 0.1092 for the ionic strength of the medium. This gives w equal to 0.0313 and the electrostatic free energy per mole of dye equal to 37.0 cal. A convenient way to determine the significance of the electrostatic effect for the interpretation of our data is to calculate empirical values of w' from (4) for several values of r on the basis of this data. For this purpose we take (see below) $Kn = 32.0 \times 10^4$ and $n = 22$. The results, which are summarized in Table IV, show that the empirical values of w' average about two and one-half times the theoretical value. We are thus led to the conclusion that the electrostatic effect is inadequate to account for the deviation of the dye binding from the simple theory. Furthermore, it is apparent from the lack of constancy of the w' values that equation (4) cannot give a quantitative representation of the data even if w' is treated as a parameter. The physical implication of this fact is that the non-linearity of the dye binding curve cannot be accounted for on the assumption that, aside from the statistical factor, the consecutive binding of dye anions involves a constant differential increase per anion in the free energy of binding. The validity of this statement is independent of the nature of the change or interaction which is responsible for the constant increase.

It is likely that the binding of dye anions involves a displacement of buffer ions with a much smaller charge effect than assumed above.[5a] In view of this and because of the approximations involved in the calculation of w, we have neglected the electrostatic correction in our subsequent considerations.

TABLE IV

CALCULATION OF EMPIRICAL VALUES OF w' FOR BINDING OF DYE AT 25°, BASED ON EQUATION (4)

$Kn = 32.0 \times 10^4$, $n = 22$, theoretical value of $w' = 0.0327$

r	r/c	w'
3	16.9 × 10^4	0.0820
5	10.3 × 10^4	.0875
10	4.0 × 10^4	.0735

Gaussian Distribution Theory.—It has previously been shown[2] that the binding of alkyl sulfates to bovine serum albumin can be described in terms of a Gaussian distribution in the free energy of binding. The wider experimental range of r for the dye binding compared to that for the detergent binding makes possible a more severe test of the suitability of this function for the representation of binding data. The equation with which we are concerned[2] is

$$r/n = 1 - f(c) \qquad (6)$$

where

$$f(c) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{e^{-\alpha^2}}{1 + K_0 c e^{\alpha\sigma}}\, d\alpha \qquad (7)$$

and $\alpha = \ln (K/K_0)/\sigma$. K_0 is an average binding constant which is equal to $1/c$ for $r = n/2$ and σ is a measure of the variation of the binding sites with respect to their free energy of binding. In fitting equation (6) to the alkyl sulfate data σ was

(5a) L. C. Longsworth and C. F. Jacobsen, *J. Phys. Coll. Chem.*, **53**, 126 (1949).

determined by treating it as a parameter and selecting a value which gave the best fit. However, if binding data for small values of r are available which will permit the determination of an extrapolated value of r/c for zero concentration, then σ can be calculated from experimental quantities as shown by the following considerations. If

$$\lim_{c \to 0} r/c = A \tag{8}$$

then from (6)

$$\lim_{c \to 0} \frac{r}{nc} = \lim_{c \to 0} \frac{1 - f(c)}{c} = \frac{A}{n} \tag{9}$$

Since $[1 - f(c)]/c$ is indeterminate as $c \to 0$

$$\lim_{c \to 0} \frac{1 - f(c)}{c} = \frac{A}{n} = - \lim_{c \to 0} \frac{\partial f(c)}{\partial c} \tag{10}$$

Substituting (7) and differentiating, one finds

$$\frac{A}{n} = \frac{K_0}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\alpha^2 + \alpha\sigma} \, d\alpha \tag{11}$$

which can be integrated by writing

$$\frac{A}{n} = \frac{K_0}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(\alpha^2 - \alpha\sigma + \sigma^2/4) + \sigma^2/4} \, d\alpha = \frac{K_0 e^{\sigma^2/4}}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-(\alpha - \sigma/2)^2} \, d\alpha = K_0 e^{\sigma^2/4} \tag{12}$$

with the result that

$$\sigma = 2 \sqrt{\ln \frac{A}{K_0 n}} \tag{13}$$

For the dye binding at 25° we take $A = 32.0 \times 10^4$ and $n = 22$. Then from the experimental curve we find $K_0 = 3.18 \times 10^3$ which on substitution in (13) gives $\sigma = 2.47$. With these values of σ, K_0 and n, r/c was calculated for several values of c on the basis of equations (6) and (7). The integral in (7) was evaluated numerically by the Gauss quadrature formula as described by Greenwood and Miller.[6] When these theoretical values of r/c are compared to the experimental ones a substantial disagreement is observed. In particular, on the r/c *vs.* r plot the theoretical curve drops much too rapidly for small values of r. We conclude, therefore, that for the dye binding a Gaussian distribution of the free energies of binding does not result in an adequate representation of the data.

Two Constant Equation.—We have attempted to fit the data by an equation of the form of (1) on the assumption that the binding can be described by only two different values of K. This is a procedure originally employed by Scatchard and Scheinberg.[6a] As will be seen shortly the binding results can be accurately described on this basis. In this case (1) reduces to

$$\frac{r}{c} = \frac{n_1 K_1}{1 + K_1 c} + \frac{n_2 K_2}{1 + K_2 c} \tag{14}$$

where

$$n = n_1 + n_2 \tag{15}$$

and

$$\lim_{c \to 0} \frac{r}{c} = n_1 K_1 + n_2 K_2 = A \tag{16}$$

The values of A for the two temperatures were determined by linear extrapolation. Since data were not available for values of r less than one, there is some uncertainty in the A's though the error is probably rather small if the assumption of linearity is justifiable. In this way it is found that at 25° $A = 32.0 \times 10^4$ and at 5° $A = 52.8 \times 10^4$.

An examination of the binding curve for high values of r suggests that a reasonable value of n would be in the neighborhood of 20. An accurate value of n cannot, however, be obtained by extrapolation; rather, it is preferable to treat n as a parameter which is limited to a relatively narrow range. We have attempted to fit the data to (14) with $n = 20$ and 22, and have found that there is a small but significant difference between the two theoretical curves obtained, only the larger value resulting in an accurate fit over the whole range of r studied.

At this point we may interject a warning, which we have previously[2] emphasized, regarding the physical interpretation of n. It cannot be inferred that 22 is the average maximum number of binding sites per native albumin molecule. It need not even be so for the sites which can bind the dye. All that we can say is that we need only assume 22 sites to account for the data. However, if there are more sites in the native protein capable of binding dye, then it may be expected that their binding constants will be substantially less than the smaller of the two constants found for the 22 sites.

To determine the n's and K's equations (15) and (16) were employed together with two sets of values of r/c and c in (14), namely, those corresponding to $r = 3$ and 10. The results are given in Table V for 25° and 5°. Theoretical curves based on these constants are shown in Fig. 1 from which it is evident that there is excellent agreement between the curves and the experimental points over the entire range of r investigated. The fact that the values of n are not integral need cause no concern because, in the first place, they are average values and, in the second place, the protein may not be molecularly homogeneous with respect to its binding properties. We have attempted to fit the data to equation (14) using integral values of n, namely, $n_1 = 5$ and $n_2 = 17$, and best values of K_1 and K_2. This does not, however, give as

TABLE V

CONSTANTS FOR EQUATION (14) AND THERMODYNAMIC DATA FOR BINDING OF ONE MOLE OF DYE BY BOVINE SERUM ALBUMIN IN 0.05 *M* PHOSPHATE BUFFER, *p*H 7.0

	Group 1		Group 2	
	$T = 5°$	$T = 25°$	$T = 5°$	$T = 25°$
n	4.82	4.66	17.18	17.34
K	9.93×10^4	6.16×10^4	0.29×10^4	0.190×10^4
ΔF^0, kcal.	−6.36	−6.54	−4.41	−4.47
ΔH^0, kcal.	−3.93	−3.93	−3.50	−3.50
ΔS^0, E. U.	8.75	8.75	3.26	3.26

(6) R. E. Greenwood and J. J. Miller, *Bull. Am. Math. Soc.*, **54**, 765 (1948).

(6a) G. Scatchard and I. H. Scheinberg, personal communication; G. Scatchard, I. H. Scheinberg and S. Howard Armstrong, Jr., THIS JOURNAL, **72**, 540 (1950).

good agreement between theory and experiment as is obtained with the values of Table V.

These values suggest that the dye binding sites may be divided into two quite distinct groups. The first (Group 1) would consist of about 5 sites per protein molecule with a relatively large binding constant and the second (Group 2) would comprise about seventeen sites with a much smaller affinity for the dye. It would be carrying the physical interpretation beyond justifiable limits to infer that the sites in each group have identical constants. Rather, at the present stage, it is more judicious to regard K_1 and K_2 as average values around which the individual values in each group are distributed. Even the division into two groups may not reflect accurately the actual empirical situation. Because of the geometrical simplicity of the binding curve and the fact that there are about 22 K's, it is undoubtedly possible to fit the binding data with an equation of the form of (1) in which there is a gradual change in the K's from one extreme to the other. However, in the absence of any evidence to the contrary, we shall base our considerations on the existence of these two groups.

Once we grant the existence of two such groups of sites the question arises as to whether they occur on the same albumin molecule or whether they represent differences among the protein molecules. The latter possibility is suggested by the failure of crystallized serum albumins to fulfill the solubility criterion for molecular homogeneity. On the basis of our data we are, of course, unable to decide between these alternatives. For the purposes of our investigation the question is only of secondary importance. We would expect, however, that a considerable part of the heterogeneity is due to differences among the sites on the same albumin molecule, otherwise, it might be expected that there would be a substantial difference in solubility of two or more albumin species and a different behavior with respect to crystallization aids.

Thermodynamic Data.—By the usual thermodynamic methods the standard changes in free energy, enthalpy and entropy for the binding of one mole of dye have been calculated for the two groups of sites and listed in Table V. Because of the possible competitive role of the buffer ions in the binding,[4] the K values are to be regarded as being dependent on the buffer composition. With respect to the thermodynamic quantities this means that the standard state includes buffer which is 0.05 M phosphate, pH 7.0. The values of ΔH^0 and ΔS^0 are also subject to a possible error which would arise if the competitive buffer ion binding were significant and if it were considerably different at 25 and 5°.

A comparison of the data for the two groups reveals the rather striking fact that the divergence in the binding constants is almost entirely due to the difference in the entropy changes associated with the binding process (8.7 E. U. *vs.* 3.3 E. U.). What this difference means in terms of the structural details characteristic of each group of sites we cannot say with any assurance at the present time. It does suggest, however, that aside from the necessity of cationic loci[7] which are undoubtedly available in both groups, such details play a decisive role in binding phenomena. One possibility which merits serious consideration is that the positive sites of Group 2 are linked to nearby anionic carboxyl groups through electrostatic interaction. The binding of dye anions by these sites would require the breaking of these bonds and would be accompanied by the release of an equal number of carboxyl groups. If the binding by Group 1 sites, on the other hand, involves a net neutralization of charge, then this binding would result in a positive entropy change due to the liberation of water molecules from their association with the ions. Such an entropy increase would not characterize binding by the Group 2 sites. Thus the difference in the values of ΔS^0 for the two groups could be accounted for at least qualitatively.

The contribution of the entropy term to the free energy of binding of the dye is substantially less both absolutely and relative to ΔH^0 than in the case of the alkyl sulfates.[2] It is nevertheless significant that the entropy change is positive as has also been found with the anionic dyes methyl orange and azosulfathiazole.[8] It is also noteworthy that ΣK_i for our dye is 6.4 times as large as that for methyl orange (32.0×10^4 *vs.* 5.0×10^4) at 25° and paralleling this the average value of ΔH^0 for the two groups is almost twice that for methyl orange (−3.7 kcal. *vs.* −2.0 kcal). These facts suggest that the hydroxyl group ortho to the azo group may play an important role in the binding.

Inhibition Effects.—Evidence has recently been presented[4,9] that chloride ion is bound to bovine serum albumin. We have attempted to ascertain whether this binding would compete with the binding of the dye anion as would be revealed by an inhibition of the latter. This has indeed been found to be the case when the dye binding is determined at 25° in 0.05 M phosphate buffer, pH 7.0 in the presence of 0.100 M sodium chloride. The results are given in Table I and shown in Fig. 1 in the form of a plot of r/c *vs.* r. From this one can reasonably infer that there are several sites on the protein molecule which are capable of binding both kinds of anions. On the assumption that chloride ion is bound with equal affinity by the 22 sites which bind the dye, it can be simply calculated with the extrapolated value of r/c taken equal to 16.3×10^4 that its binding constant (K) is 9.6 at 25°. This value of K corresponds to $\Delta F^0 = -1.34$ kcal. per mole of chloride ion

(7) I. M. Klotz and J. M. Urquhart, THIS JOURNAL, **71**, 1597 (1949).

(8) I. M. Klotz and J. M. Urquhart, *ibid.*, **71**, 847 (1949).

(9) G. Scatchard, A. C. Batchelder and A. Brown, *ibid.*, **68**, 2320 (1946).

bound. In view of the assumptions involved in its determination, it is in reasonable agreement with the value of 20 found by Klotz and Urquhart at 0°.[4]

The interesting observation has been made that dissolved toluene can inhibit the binding of the dye. This observation arose in connection with the development of a new method for studying protein binding. This procedure, called partition analysis, will be described elsewhere. It involves the equilibration by gentle rocking of an aqueous solution of albumin and dye with an appropriate immiscible organic dye solvent. The solvent is selected so that the partition coefficient of the dye is of the order of magnitude of one. Such equilibration does not lead to any measurable loss of the protein by precipitation. The inhibitory effect of toluene was demonstrated in the following way. A 3×10^{-5} M albumin solution in 0.05 M phosphate, pH 7.0 and containing dye was rocked with some toluene at 25° to obtain a saturated solution of toluene in the aqueous phase. This phase was removed and its total dye concentration found to be 22.15×10^{-5} M. Dye binding at 25° was then determined by equilibrium dialysis using 15-ml. portions of this solution inside the bags and 15-ml. portions of 0.05 M phosphate buffer, which had been similarly saturated with toluene, outside. Under these conditions it was found that $r = 3.93$ and $r/c = 8.54 \times 10.^4$ It will be observed that when this pair of values is plotted in Fig. 1 the corresponding point falls substantially below the 25° curve.

The reversibility of the inhibitive effect of toluene was also demonstrated. Duplicate solutions of 3×10^{-5} M albumin in buffer, saturated with toluene, were dialysed in the cold against several changes of buffer for two days to remove the toluene. The concentrations of protein in these solutions were determined by the biuret method and the binding measured in the usual way. It was found that the binding of these albumin solutions was practically equivalent to that of control albumin solutions which had not been exposed to toluene. Thus, at a concentration of 5.3×10^{-3} M, its solubility in water at 25°,[10] toluene can substantially interfere with the albumin binding of the dye. It appears to us that this is most likely due to competitive binding of the toluene by uncharged groups of the protein which are also involved in the binding of the dye. This finding, that toluene associates with albumin, is undoubtedly involved in its efficacy as a crystallization aid in the preparation of crystalline serum albumins.[3]

An experiment was carried out to determine whether there might be interaction between bovine albumin and bovine γ-globulin (Armour's Fraction II) in 0.05 M phosphate, pH 7.0, which would be detectable by its effect on the dye binding. Buffered solutions containing 3.00×10^{-5} M albumin and 0.5% γ-globulin were equilibrated at 25° against buffered dye solutions with initial concentration 17.67×10^{-5} M. The results showed a very slight but insignificant increase of binding associated with the presence of the γ-globulin. It can be concluded that, under our conditions, there is no globulin–albumin interaction which affects the binding of the dye. Furthermore, the experiment also serves to establish the fact that the dye is not significantly bound by bovine γ-globulin.

The Configurational Adaptability of Serum Albumins.—From the considerable amount of work which has been done to date on protein binding[11] there emerges the fact that the serum albumins are outstanding, in fact almost unique, among soluble proteins in their capacity to form reversible complexes with a variety of ions and molecules.[11a] It is not the fact that these proteins form complexes which is distinctive since this is found, as well, with enzymes and antibodies in their interactions with substrates and homologous haptens, respectively. Rather, it is the ability to bind molecules with the most diverse configurations, most of them unphysiological, which sets the albumins apart. This striking contrast with respect to specificity is clearly shown, for example, by the ability of a specific antibody to discriminate between the sulfonate group, when it occurs in a hapten as phenylsulfonic acid, and the carboxyl group in a corresponding hapten.[12] Serum albumin, on the other hand, binds strongly dyes containing either anionic group, the dye used in this study and methyl orange, for example. Further emphasis of this point is found in the observation that the specific purified antibody homologous to the p-azophenylarsonate group binds the anionic dye p-(p-hydroxyphenylazo)-phenylarsonic acid with great tenacity (ΔF^0 at 29° is −7.7 kcal. per mole of dye).[13] On the other hand, when tested with methyl orange the antibody does not bind any significant amount of this dye.[14] We conclude from this that it is not the unavailability of cationic loci which prevents the association with methyl orange, since these are in all likelihood involved in the interaction with the homologous dye. This discrimination is to be attributed to the relatively rigid and therefore selective configuration associated with the antibody binding site.[15]

The serum albumins not only bind a great variety of organic anions but the same sites are involved in the association with many of these an-

(10) Landolt-Börnstein, "Physikalisch-Chemische Tabellen," 5th edition, Erg. III, p. 672.

(11) For a summary of the work with plasma proteins see A. Goldstein, *J. Pharmacol. Exptl. Therap.*, Part II, **95**, 102 (1949).

(11a) We omit from consideration here the binding properties of insoluble proteins and the changes attendant upon insolubilization. For studies of the anion binding of wool see J. Steinhardt, *Ann. N. Y. Acad. Sci.*, **41**, 287 (1941).

(12) K. Landsteiner, "The Specificity of Serological Reactions," Harvard University Press, Cambridge, Mass., 1945, p. 164.

(13) H. N. Eisen and F. Karush, THIS JOURNAL, **71**, 363 (1949).

(14) H. N. Eisen and F. Karush, unpublished experiments.

(15) L. Pauling, D. H. Campbell and D. Pressman, *Physiol. Rev.*, **23**, 203 (1943).

ions. This can be inferred, for example, from the displacement of methyl orange from the protein by various organic acids as shown by spectral observations.[16] In a study described in the next paper, we have found that there is competitive binding by bovine albumin of dodecyl sulfate and the anionic dye used here. Furthermore, the data show that those sites with the highest value of K for the dye are, for the most part, the sites which bind the alkyl sulfate most strongly. As noted above chloride ions are also bound to at least some of the sites which interact with the dye. These facts suffice to invalidate the view that the binding sites of the albumin have fixed configurations which are appropriate for the various complexing molecules and ions.

In addition to the electrostatic attraction between the protein site and the bound ion, the formation of complexes also involves interaction with uncharged groups of the albumin. This is seen from the fact that generally the strength of binding is enhanced by an increase in the size of the nonpolar portion of the bound molecule. Specifically, the increase in the association constants of the alkyl sulfates with increase in chain length[2] points to the significance of such interaction. The competitive effect of toluene, described above, contributes additional evidence in this direction. Finally, the difference between the binding constant of chloride ion on the one hand and that of the dye anions on the other leaves little doubt as to the role of non-electrostatic interaction.

Such interaction is significant in terms of the configurational relation it implies between the bound ion and the binding site. In order for the non-electrostatic interactions to contribute significantly to the free energy of binding, *e. g.*, in the case of our anionic dye, there must be considerable structural complementariness of the site with respect to a large portion of the dye. This is necessary so that the small net interactions of the various groups of the dye, *e. g.*, hydroxyl, azo, phenyl, with adjacent groups of the protein will yield a large total effect, either in terms of ΔH or ΔS or both, for the molecule as a whole. Although it is not possible to quantify this picture at the present time, it is suggestive that ΔF^0 at 25° for the binding of our dye to the group I sites is rather close to that for the binding to antibody of the arsonic acid dye mentioned above, −6.36 kcal. *vs.* −7.7 kcal., particularly in view of the structural similarity of the dyes. Furthermore, it is evident from hapten inhibition studies[17] that the binding site of the antibody homologous to the *p*-azophenylarsonate group is structurally complementary to the whole group, the fit being within less than 1 Å. It thus appears likely that at least the corresponding portion of our dye is fitted sufficiently well to the albumin site to permit electronic van der Waals attraction, hydrogen bond formation, etc., to become important, forces which are very sensitive to distance.

When the lines of evidence which have been summarized above are taken in their conjunctive significance, the view emerges that the distinctive feature of serum albumins is their configurational adaptability. By this is meant that there exist a number of sites on the protein, each associated probably with several side chains, which to a varying extent can assume a large number of configurations in equilibrium with each other and of approximately equal energy. In the presence of an organic anion, for example, that configuration is stabilized which, by virtue of its structural relation to the anion, permits the various portions of this anion to interact with appropriate groups of the protein. This is manifested by the formation of a complex. That the relative positions of the terminal groups of the side chains may change would seem to be a reasonable possibility in view of the fact that the side chains are single bonded systems. What structural factors in the protein, *e. g.*, the order of amino acids, the manner of folding of the polypeptide chain, permit configurational alternatives of the side chains, as we assume for the serum albumins, but make for configurational rigidity in other proteins (enzymes, antibodies, perhaps globulins in general) we do not know at the present time.

Aside from its usefulness in accounting in a general way for albumin binding, this hypothesis has another interesting implication. In physical terms, an aqueous solution of bovine albumin for example, would be viewed as containing not one structural species but a large number of approximately equal energy and in thermodynamic equilibrium with each other. Thus the dissolved protein would be characterized by a large configurational entropy. If, as is likely, the number of configurations permissible in the solid state is much less, then the process of solution would involve an entropy increase from this source with a consequent enhancement of the solubility of the protein. This argument fits in with the fact that the proteins showing non-specific binding are those of the highest solubility, *i. e.*, the albumins. Indeed, one of the important molecular factors underlying the solubility distinction between albumins and globulins may be this configurational factor.

To return to our multi-component protein solution, we can readily understand the role of crystallization aids in the preparation of crystalline albumins.[3] By virtue of their combination with the protein, these molecules stabilize one or several of the many configurations originally present thereby reducing the loss of configurational entropy involved in the transition to the crystalline state.

Albumin Binding in Hapten Inhibition Studies.—During the past several years, Pauling and his co-workers[18] have been conducting extensive quan-

(16) I. M. Klotz, THIS JOURNAL, **68**, 2299 (1946).

(17) L. Pauling and D. Pressman, *ibid.*, **67**, 1003 (1945).

(18) D. Pressman, J. H. Bryden and L. Pauling, *ibid.*, **70**, 1352 (1948), and earlier papers.

titative investigations of the inhibition by haptens of the precipitation of antibody homologous to antigenic groups structurally similar to these haptens. Practically all of the experiments reported to date were performed with whole antiserum and, consequently, in the presence of relatively high concentrations of albumin. The exact concentration of the albumin depended on the final dilutions employed, usually about three-fold, although they varied in different experiments. Because of the close structural similarity of the dye used in the present study to the haptens studied with anti-*p*-azobenzoate-serum,[19] the importance of the neglected factor of albumin binding of haptens can be assessed.

These authors have attempted to describe quantitatively the inhibitory power of the haptens by the assignment of relative inhibition constants. These were based on data secured by measuring the decrease in the amount of specific precipitate formed with an appropriate antigen as a function of the amount of hapten added. The largest inhibition constants were generally observed with the homologous azo haptens, one-half inhibition occurring, *e. g.*, at a total dye concentration of roughly 2×10^{-5} *M*. In view of the binding results reported in this paper and the binding by serum of dyes and other molecules,[11] there is little doubt that most of the azo hapten was bound to albumin. This would introduce considerable error in the estimation of the inhibition constant for the azo haptens. In fact these values may be off by a factor of several-fold. The decisive role of albumin in binding polyhaptenic azo dyes has recently been demonstrated by studies with purified antibody.[20] It may be noted that our remarks are applicable not only to the azo derivatives of benzoic acid but are equally relevant to other anionic azo haptens.

The presence of albumin in these experiments would also cause a considerable error in the inhibition constants of some of the colorless haptens though less than that for the dyes. The extent of this effect would depend on the albumin affinity for these haptens and the concentration range employed. The possibility that the binding by albumin in serum is much less than that observed with the purified protein due to competition with the hapten of naturally occurring substances is not very likely. The binding studies with sulfapyridine, by human plasma[21] and by crystallized bovine albumin,[22] suggest only a minor competitive role for such substances.

On the whole it is reasonable to conclude that the failure to take into account albumin binding has invalidated the quantitative interpretation of hapten inhibition in some cases. For many more the correctness of the inhibition constants stands in question until compensation for albumin binding is made, or its effect shown to be negligible. In spite of these shortcomings, it appears to us that the structural conclusions derived by these authors from their hapten inhibition studies are, for the most part, valid.

Acknowledgment.—I am indebted to Dr. C. P. Rhoads for making available to me during the conduct of this investigation the excellent laboratory facilities of the Sloan-Kettering Institute for Cancer Research. This work was supported by the Office of Naval Research.

Summary

The reversible binding of an anionic azo dye by bovine serum albumin has been investigated at two temperatures over a wide range of concentration. It is shown that the data can be accurately described on the assumption that there are 22 binding sites per protein molecule and that these are divided into two groups, each group characterized by a particular value of the association constant. The values of ΔF^0, ΔH^0 and ΔS^0 for the binding to the two groups were determined and the similarity of ΔH^0 and difference in ΔS^0 noted. A 3-fold increase in albumin concentration does not affect the binding at 25°. Toluene and chloride ion significantly inhibit binding whereas 0.5% bovine γ-globulin does not. An interpretation of the distinctive binding properties of serum albumins is presented in terms of the ability of these proteins in solution to exist in many molecular configurations of approximately equal energy. This permits various sites to take on configurations complementary to the structure of an anion and leads to complex formation. This picture also implies a large configurational entropy for the dissolved protein and, consequently, an enhancement of its solubility. Finally, the failure to take into account albumin binding of haptens has called into question the quantitative significance of recent hapten inhibition studies.

NEW YORK, N. Y. RECEIVED NOVEMBER 7, 1949

(19) D. Pressman, S. M. Swingle, M. L. Grossberg and L. Pauling, THIS JOURNAL, **66**, 1731 (1944).

(20) A. B. Pardee and L. Pauling, *ibid.*, **71**, 143 (1949).

(21) B. D. Davis, *J. Clin. Invest.*, **22**, 753 (1943).

(22) I. M. Klotz and F. M. Walker, THIS JOURNAL, **70**, 943 (1948).