Chloride- and Hydrogen-Ion Binding to Ribonuclease

G. I. Loeb* and H. A. Saroff

From the National Institute of Arthritis and Metabolic Diseases, National Institutes of Health, Public Health Service, U.S. Department of Health, Education, and Welfare, Bethesda, Md.

Received July 13, 1964

The binding of chloride ion and hydrogen ion to bovine ribonuclease was studied in the carboxyl-titration region of *p*H. We have found, in agreement with a previous study in this laboratory, that chloride ion is indeed bound in the acid *p*H region under the conditions normally encountered during measurement of proton binding; the chloride binding increases to 4.7 ions/molecule at *p*H 2.6 in 0.035 M total chloride concentration. Analysis indicates that there are six sites which bind chloride ion with sufficiently high binding constants to implicate cooperating positive charges. The *p*H dependence of the binding of both the hydrogen and chloride ions leads to the postulate that the carboxylate ions enter into competition with chloride ions for the binding site. The clustering of both positive and negative charges implied by this interpretation of the binding data imposes sufficient restrictions on the folding of the peptide backbone of ribonuclease to enable the use of these six sites as an aid toward the elucidation of the three-dimensional structure of ribonuclease.

Anions have been reported to exert large effects on the properties of ribonuclease. Studies on ribonuclease by electrophoresis (Crestfield and Allen, 1954, Rosemeyer and Shooter, 1963), optical rotation (Sela and Anfinsen, 1957), phosphorylation (Taborsky, 1959), alkylation (Neumann *et al.*, 1962; Crestfield *et al.*, 1963), and dialysis (Craig *et al.*, 1963) have demonstrated phosphate-, sulfate-, and chloride-ion effects. The binding of both chloride and sulfate ions to ribonuclease in the region of *p*H from isoionic to the point where the amino and imidazole groups were protonated has already been reported (Saroff and Carroll, 1962). This binding of anions indicated that the titration curves of ribonuclease in the *p*H region where the carboxylate ions were protonated could be properly interpreted only when anion-binding measurements were coupled with the hydrogen-ion-binding measurements. The protonation of the carboxylate ions of ribonuclease is already anomalous when anion binding is assumed to be zero (Tanford and Hauenstein, 1956) and becomes even more anomalous on consideration of anion binding. This paper reports on the results of a study on both hydrogen- and chloride-ion association in the region from *p*H 4.5 to 2.5. In view of a significant amount of chloride binding found in this region an explanation of the anomalous behavior of the carboxylate ions is offered based on structural considerations in the ribonuclease molecule.

Materials and Methods

Ribonuclease.—The ribonuclease was Sigma lots 33B-781-4A and 63B-8620. All of the measurements except the one with 12.2 protons bound per mole of ribonuclease were done on lot 33B-781-4A. The measurement with 12.2 protons bound per mole of ribonuclease was done on lot 63B-8620 (see Table I). Although we suspect conformational changes in the various preparations we have no explanation for the variations in the points with 12.2 and 12.9 protons bound per mole (see Table I and Fig. 1). The protein was dissolved in distilled and deionized water, and then passed three times through a deionizing column of the type described by Dintzis (1952). Prior to acidification the resulting material was stored in closed polypropylene vessels placed in a second closed bottle containing a layer of Ascarite. The *p*H of such preparations was 9.6; the conductivity was equivalent to $\sim 2 \times 10^{-4}$ M NaCl. Protein concentration was determined by diluting a weighed portion of the solution in a spectrophotometer curvet, and measuring optical density with a Cary Model 14 spectrophotometer. The value used for the molar extinction coefficient was ϵ 9800. The concentrations of the stock solutions after deionization were found to be approximately 2.4% ribonuclease. The *p*H and ionic strength of the solutions were adjusted by adding solutions of HCl and NaCl by weight to weighed aliquots of the stock solutions. All inorganic materials were reagent grade and all water was distilled and deionized before use.

*p*H *Measurements* were made with a Beckman Model G *p*H meter (calibrated with Beckman commercial buffer solutions) on the samples used for measurement of chloride-on activity after the chloride-activity measurements were completed.

The values of the binding constants were determined by dilution of the most concentrated sample of a series with water in order to measure the chloride-ion activity in a series of solutions over a range of concentration.

Chloride-ion activity was measured at room temperature (24° ± 1.5°) in a cell of the same type as that used previously in this laboratory (Saroff and Healy, 1959), and modified by the substitution of Teflon stopcock plugs for glass. This removed much of the inconvenience caused by grease at the liquid junctions; otherwise the method was identical with that used in the previous studies.

Briefly, the cell consists of two chambers machined in Lucite blocks and separated by a permselective membrane; for measurement of chloride-ion activity the membrane is positively charged (Nalfilm-2, furnished by Nalco Chemical Co., Chicago). The sample solution is placed in one compartment, and solutions of known sodium chloride concentration in the other. The potential difference between the two compartments is recorded for at least two known concentrations of sodium chloride solutions, so as to give small values (0–3 mv) both positive and negative with respect to the sample containing protein. The plot of emf versus log NaCl concentration was interpolated to determine that concentration of NaCl which would give zero potential with respect to the protein solution, and which was taken to have the same chloride-ion activity as the solution containing protein.

* Present address: U. S. Naval Research Laboratory, Washington, D. C.

TABLE I
CHLORIDE- AND HYDROGEN-ION-BINDING DATA FOR RIBONUCLEASE

Ribonuclease Concentration (M × 10^3)	pH	H+ Bound per Mole ribonuclease	Total Concentration Chloride (M × 10^3)	Free Concentration Chloride (M × 10^3)	$\bar{\nu}_{Cl^-}$	Net Charge of Ribonuclease, Z_p
1.474	4.43	9.3	34.87	30.9	2.6 ± 0.1	+6.7
0.579	4.42	9.3	13.7	12.3	2.3 ± 0.1	+7.0
0.193	4.43	9.2	4.56	4.35	1.0 ± 0.1	+8.2
0.0625	4.55	9.0	1.52	1.47	0.8 ± 0.15	+8.2
1.835	3.19	12.2	33.81	28.2	3.1 ± 0.1	+9.1
1.481	3.71	12.9	34.52	29.3	3.5 ± 0.05	+9.4
0.493	3.80	12.7	11.5	10.0	2.9 ± 0.2	+9.8
0.165	3.80	11.7	3.84	3.64	1.2 ± 0.1	+10.5
0.0552	4.05	11.5	1.286	1.27	0.2 ± 0.1	+11.3
1.473	2.59	15.6	34.42	27.4	4.7 ± 0.15	+10.9

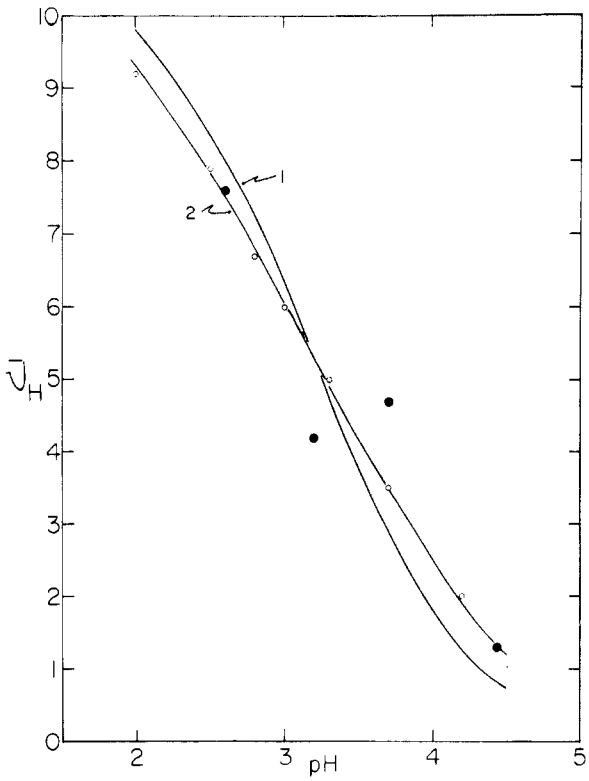


FIG. 1.—Experimental data and calculated curves at ionic strength = 0.03–0.035 in NaCl for the titration of the carboxyl groups of ribonuclease. $\bar{\nu}_H$ = number of protons bound to isoionic ribonuclease minus 8 (for 1 phenolic ion, 3 uncharged amino groups, and 4 uncharged imidazol groups, see text). Curve 1 calculated from the expression $\nu_H = k_{HC_H}e^{-2Z_p w Z}n/(1 + k_{HC_H}e^{-2Z_p w Z})$ for n = 11, w = 0.093 and $k_H = 10^{3.9}$. Curve 2 calculated from equation (8) and parameters of Table III. ○, data from Tanford and Hauenstein (1956); ●, data from Table I.

RESULTS

Addition of hydrogen ion to isoionic ribonuclease should protonate one phenolic ion, the three uncharged amino groups (one α and two ϵ), and the four uncharged imidazole groups before protonating the eleven carboxylate ions (Tanford and Hauenstein, 1956; Hirs *et al.*, 1960, Tanford *et al.*, 1955). The *p*H at which the carboxyl groups would begin to modulate the binding of chloride ion is that occurring after eight protons are bound per isoionic ribonuclease molecule (mw = 13,684). Since we assume that the sites binding anions are the positively charged nitrogen centers (ammonium, imidazolium, and guanidinium ions), the control of the anion binding exerted by the carboxyl group could be (a) one of direct competition at the binding sites or (b) one reflecting changes in the net charge of the ribonuclease molecule. The experiments for which data are summarized in Table I were designed to give proton as well as anion binding in order to evaluate the contributions of these two effects. At the conditions of *p*H, ionic strength, and temperature of these experiments the reversible conformational changes (Hermans and Scheraga, 1961; Ginsburg, 1964) are minor and were ignored.

Apparent association constants and apparent number of binding sites may be calculated from these data. The expression for the binding of anions (Saroff, 1957) to a set of n nitrogen sites containing an apparent number of available sites, n', is

$$\bar{\nu}_A = \frac{k_A'' c_A n'}{1 + k_A'' c_A + \frac{1}{k_H'' c_H}} \qquad (1)$$

where $\bar{\nu}_A$ is the average number of anions bound to such sites on 1 mole of protein and k_A'' and k_H'' are the apparent intrinsic association constants for the binding of anions and hydrogen ions, respectively, at the concentrations of c_A and c_H of the free anion and hydrogen ion, respectively. Activity coefficients of chloride ions were taken as unity. Since all the groups containing ionizable nitrogens may be considered to be protonated below *p*H 4.5, i.e., $(k_H'' c_H)^{-1} << 1$, this expression reduces to

$$\bar{\nu}_A = \frac{k_A'' c_A n'}{1 + k_A'' c_A} \qquad (2)$$

Equation (2) and the data of Table I yield the following values for k_A'' and n_A': at *p*H 4.43 k_A'' = 200, n_A' = 4.0; at *p*H 3.19–4.05 k_A'' = 250, n_A' = 2.9. The curves calculated from these values are illustrated in Figure 2. Measurements on diluted solutions at *p*H 2.59 were not made because the number of bound protons cannot be kept constant on dilution.

The carboxylate-ion behavior in native ribonuclease has already been cited as anomalous (Tanford and Hauenstein, 1956) toward hydrogen-ion binding. The anomalous behavior is evident in not only the *pK* of the ionizing protons and w values required to fit the data but also in the change with w as a function of ionic strength. Table II summarizes the apparent w values for ribonuclease derived from plots of *p*H − log $\alpha/(1 - \alpha)$ versus Z_p, where α is fraction of the carboxyl groups ionized and Z_p is the net charge on the protein molecule. The interpretation of Tanford and Hauenstein (1956), whereby the phenolic values of w are taken as normal, yields the data of Table II which indicate

TABLE II
APPARENT VALUES OF w FOR RIBONUCLEASE

	Ionic Strength		
	0.01	0.03	0.15
Phenolic values from titration followed spectrally[a]	0.112	0.093	0.061
Carboxylate-ion titration[a]	0.12	0.119	0.09
Carboxylate-ion titration with chloride-ion binding included	0.15 ± 0.01	0.17 ± 0.01	

[a] (Tanford and Hauenstein, 1956).

that the carboxyl groups of ribonuclease are subjected to very strong electrostatic effects (compared to those obtaining for the phenolic groups) and that these electrostatic effects are relatively insensitive to ionic strength.

Figure 1, curve 1, illustrates the fact that the ordinary titration equation is inadequate to fit the hydrogen-ion data for the carboxyl groups of native ribonuclease. The usual expression, $\bar{\nu}_H = k_H'/c_H n/(1 + k_H'c_H)$, was employed where $\bar{\nu}_H$ is the average number of hydrogen ions bound to 1 mole of ribonuclease at the concentration of hydrogen ions c_H, $k_H' = k_H e^{-2Z_p w Z}$ where k_H is the intrinsic association constant for the binding of hydrogen ions to the n carboxyl groups available for binding, Z_p is the net charge on the ribonuclease molecule, and Z represents the charge on the hydrogen ion. Curve 1 represents the best fit (*vide infra*) possible for this equation. The value of k_H to give this fit is $10^{3.9}$. The values of Z_p employed were calculated from the data of Table I and the value of w was that from the phenolic data (0.093). Comparison of curve 1 of Figure 1 with the titration data demonstrates that the usual titration equation must be modified in order to explain the titration of the carboxyl groups of ribonuclease.

INTERPRETATION OF THE DATA

The two experimental facts to be explained are (1) the presence of at least five sites to which chloride ions bind with an unusually high association constant and (2) the anomalous titration behavior of the carboxyl groups as reflected in the unusually high value for the parameter w, the lack of dependence of w on the ionic strength, and the inability to fit the hydrogen-ion titration data with the common titration expression and with a reasonable pK for the carboxyl groups.

With 4.7 chloride ions bound to 1 mole of ribonuclease at pH 2.59 the maximum number of chloride ions binding with the anomalously high association constant is at least five. At pH 2.59 there are 2.4 carboxylate ions still unprotonated. In the more acid region the data of Table I indicate that with each two protons added to ribonuclease one chloride ion is bound. On the basis of these facts and the curve-fitting accomplished below we propose that the maximum number of binding sites for chloride ions is six per mole of ribonuclease. From considerations previously outlined on the binding constants for chloride ions to protonated nitrogen groups we have already proposed that the two anion-binding sites of ribonuclease at pH 4.5, were each composed of triplets of protonated nitrogen groups (Saroff and Carroll, 1962). We now postulate that the six sites binding anions in ribonuclease are each composed of similar triplets of protonated nitrogen groups. There are a total of nineteen nitrogen centers in ribonuclease (four histidine; ten lysine; four arginine; and

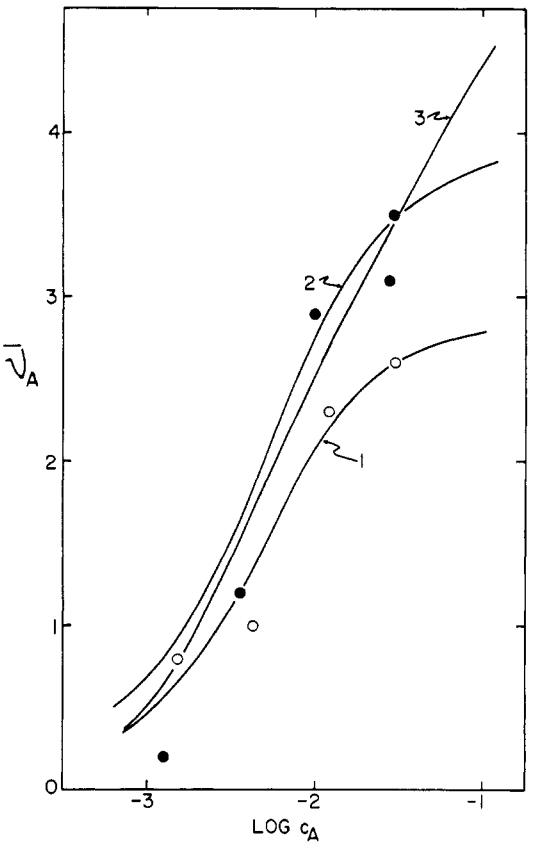


FIG. 2.—Experimental data and calculated curves for the binding of chloride ions to ribonuclease. Curve 1 calculated from equation (2) and parameters $k_A'' = 250$ and $n' = 29$. Curve 2 calculated from equation (2) and parameters $k_A'' = 200$ and $n' = 4.0$. Curve 3 calculated from equation (9) and parameters of Table III. ○, data from Table I at pH 4.43 representing 9.0–9.3 protons bound; ●, data from Table I at pH 3.19–4.05 representing 11.5–12.9 protons bound.

one α-amino of lysine-1) so that 18 of these nitrogens could participate in the positively charged clusters binding anions.

The low apparent pK value of the carboxyl groups of ribonuclease indicates that the strong electrostatic effects derive from the proximity of the carboxyl groups to positive rather than negative charges. Thus the carboxyl groups appear to be near the groupings of positive charges. If the carboxylate ions are very near the groupings of three positively charged nitrogen centers the predominating electrostatic effect will be that from the localized net charge of the resulting cluster of ions rather than that deriving from the net charge of the whole molecule. An indication of this proximity is given by the fact that the value of w is independent of the ionic strength.

The Debye-Hückel theory allows an approximate estimation of the variation of w with both size and ionic strength. For the smeared-charged model of Debye and Hückel, Linderstrøm-Lang (1924) derived the value of w given by the expression

$$w = \frac{Ne^2}{2DRT}\left[\frac{1}{b} - \frac{\kappa}{1 + \kappa a}\right]$$

where N is Avogadro's number, e is the charge on the electron, D is the dielectric constant of the medium, R is the gas constant, T is the absolute temperature, b and a are the radii of the assumed impenetrable sphere and distance of closest approach respectively, and κ is the reciprocal of the thickness of the Debye-Hückel

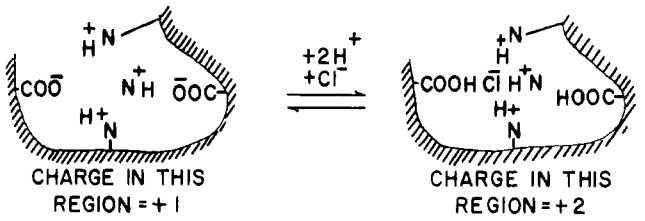


FIG. 3.—Sketch of the postulated cluster giving rise to the binding of anions. In the structure at the left, each carboxylate ion is hydrogen bonded to a protonated nitrogen. In the structure at the right, the chloride ion is hydrogen bonded to three protonated nitrogen atoms and one carboxyl group.

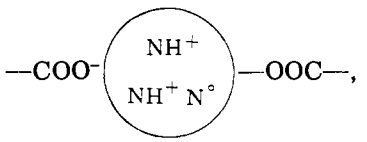
ionic atmosphere. This expression shows that as both b and the ionic strength decrease (higher κ) the value of w increases; however, the dependence of w upon the value of κ (or the ionic strength) becomes less as the value of b becomes smaller.

The experimental finding that the value of w for the carboxyl groups of ribonuclease is large and also relatively insensitive to ionic strength we interpret to mean that carboxyl groups *and* nitrogen centers are clustered together in the ribonuclease molecule with these clusters giving rise to both the anion-binding and proton-binding anomalies.

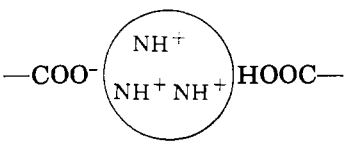
Structure of the Clusters.—The crystal structure of L-cystine hydrochloride (Steinrauf *et al.*, 1958) suggests a possibility for the structure of the clusters to which anions bind in ribonuclease. The chloride ion in crystalline L-cystine hydrochloride is hydrogen bonded to each of three protonated amino groups and one uncharged carboxyl group in an essentially tetrahedral array. The Cl^-—H—O_2 (HO_2 from a carboxyl group) distance is 2.98 A and the Cl^-—H—N distances are 3.08, 3.25, and 3.27 A.

Ribonuclease has eleven carboxyl groups per mole (mw = 13,684). All the carboxyl groups are assumed to be near centers of positive charge. The structure of the cluster of carboxyl and nitrogen groups then becomes a central core of three positive charges with two carboxylate ions in the immediate vicinity (close enough to form hydrogen bonds). The reaction involving the binding of chloride ion would then be a competitive one between chloride and the carboxylate ions for one group of the three positive nitrogen centers. A *speculative* representation of this cluster is drawn in Figure 3.

Calculation of the Titration Curves.—One test of this hypothesis involves the formulation of the expression for the binding of both hydrogen ions and anions based on such a model and the comparison of calculated titrations with the experimental data. For five clusters, each composed of three nitrogenous groups and two carboxyl groups, and one cluster composed of three nitrogenous groups and one carboxyl group, the expressions for the number of bound hydrogen ions ($\bar{\nu}_H$) and chloride ions ($\bar{\nu}_A$) are given in equations (8) and (9)[1] (see Appendix). In these equations a local electrostatic factor (Saroff and Lewis, 1963) is included in addition to that electrostatic factor which is a function of the net charge of the protein molecule. The local electrostatic effect obtains from the charge in the immediate domain of the cluster illustrated in Figure 3. Thus for the reaction, $—COO^- + H^+ = —COOH$, occurring at a cluster in the form,



the local electrostatic factor is the operational quantity ϵ_H and when the cluster has the form,



the local electrostatic favor is ϵ_H^2 (see Appendix). The values of the carboxyl group parameters capable of giving a good fit to the experimental data are given in Table III. The calculated curves are illustrated in Figures 1 and 2.

The first set of values in Table III are those which gave the best fit[2] for the calculated curve and the experimental data on allowing the listed values a free range on the computer computation. The second set of values are those giving an equivalent fit when the value of k_A was not allowed to rise above 1000. The third set of values are those for the conditions in which the pK of the carboxyl groups was not allowed to rise above 4.80. These sets of values illustrate that, for this model, the localized electrostatic effect and anion binding are sufficient to explain the anomalous hydrogen binding to ribonuclease and that the hydrogen bond appears to be of lesser importance.

Discussion

Some chemical evidence on the active site of ribonuclease already exists (Barnard and Stein, 1960; Stark *et al.*, 1960; Hirs, 1962; Anfinsen *et al.*, 1962; Richards and Vithayathil, 1960) to indicate a cluster of histidine 12, histidine 119, and lysine 41. When these amino acid residues are brought together in space the three carboxyl groups of aspartic acid 14, aspartic acid 121, and valine 124 are already in the immediate vicinity. The model developed in this communication postulates five additional groupings of basic amino acids and associates each of these six groupings with the available carboxyl groups of ribonuclease. The postulated model treats five of these six clusters of amino acids in an identical manner with the sixth cluster the same except for a missing carboxyl group. This is, of course, an approximation since it is shown from the anion-binding data that even the other five clusters cannot be identical in their behavior. Thus in the stage of anion binding between pH 2.6 and 3.2 (where only the carboxyl groups are presumed to be titrated), one chloride ion is bound for each of two protons bound. At the pH of 4.5 when eight protons are bound to isoionic ribonuclease, the first two chloride ions are bound (Saroff and Carroll, 1962) even though it is assumed that there are not enough bound hydrogen ions to protonate any of the carboxylate ions. It is obvious from these data that the binding sites behave differently toward the chloride ion in the different regions of pH. From the standpoint of hydrogen ion binding it appears that these differences are minor.

[1] Equations (8) and (9) were derived for the clusters with only one hydrogen-bonded form allowed for species 2, 3, 4, 5, and 9. These equations were found to give more consistent results than those with the multiple hydrogen-bond forms allowed for each of the above species (see Appendix).

[2] The best fit is defined as that which minimizes the sum of the squares of the differences between the observed and calculated values of $\bar{\nu}$. All of the calculations in this paper were done on either an IBM 1620 or Minneapolis-Honeywell 800 computer. We are indebted to Jonathan Schwartz, Computation and Data Processing Branch, N.I.H., for programing the equations and calculating the curves.

TABLE III

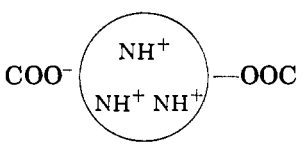
VALUES FOR THE PARAMETERS OF EQUATIONS (8) AND (9) WHICH GIVE THE CALCULATED CURVES OF FIGURE 1.

k_{H_1}	k_A	ϵ	k_{D_2}'[b]	k_{D_1}'[c]	Conditions from which computer selected listed parameters as best fit
$10^{5.0}$	1400	4.5	1.0	25	Free range on all values. Fit H^+-binding data
$10^{4.85}$	1000	3.8	0.35	23	Free range on k_{H_1}, ϵ, k_{D_2}' and k_{D_1}'; $k_A' \leq 1000$. Fit H^+-binding data
$10^{4.75}$	870	3.4	0.0	6	Free range on k_A', ϵ, k_{D_2}' and k_{D_1}'; $k_{H_1} \leq 10^{4.80}$. Fit H^+-binding data
$10^{4.70}$	820	4.5	1.0	25	Fit Cl^--binding data with values consistent with H^+-binding data

[a] $w = 0.093$; NaCl concentration = 0.03 M for H^+ binding [b] Value for five clusters with two carboxyl groups in each. [c] Values for single cluster with one carboxyl group.

The value of the localized electrostatic factor, ϵ, may be examined by comparing the values of 3.4–4.5 found in this work with that of an existing model compound with dimensions similar to those found for the amino groups in crystalline L-cystine hydrochloride. The distance between the protonated nitrogen atoms in crystalline L-cystine hydrochloride is about 5.3 A. A model constructed to represent citric acid shows an approximate maximum separation of the charged carboxylate ions to be about 4.9 A. If the three carboxylate ions of citric acid are assumed to be equivalent $k_1/k_2 = 3\epsilon_H$, $k_1/k_3 = 9\epsilon_H^2$, and $k_2/k_3 = 3\epsilon_H$. From the pK values (Bates and Pinching, 1949) of citric acid (3.13, 4.76, and 6.40) and these three ratios, the calculated values of ϵ_H are 12, 14, and 15. The values of ϵ_H calculated for citric acid vary in this manner because the three carboxyl groups are not equivalent. The values, 12–15, differ only by a factor of three from the values 3.4 to 4.5, found for the localized electrostatic effects required for application of equations (8), (9), and (10).

While it is apparent that this model fits both the hydrogen and chloride binding in the acid region it is necessary that the same test be applied to the alkaline region. For this we assume a single cluster of the form,



and allow the dissociation of the three identical nitrogen groups of the cluster while maintaining the charge on the carboxyl groups. Since ribonuclease does not bind chloride ions above pH 8 the value of k_A' was taken as zero and equation (10) was derived (see Appendix) for the ionization of the three nitrogen centers of the cluster. The titration curve predicted by this equation (with eighteen hydrogen-bonded species) was then compared with that predicted for the ionization of three independent sites where the unperturbed equation $\bar{\nu}_H = k_{H_2}c_H n/(1 + k_{H_2}c_H)$ applies with the value of $n = 3$. The calculated data and parameters given in Table IV illustrate that three nitrogen groups with the intrinsic pK of 8.5–8.8, when clustered with two carboxylate ions in a manner such that there exists very weak hydrogen bonding (or no hydrogen bond) and a localized electrostatic effect of 3.5, will yield a titration curve almost identical to that of three independent nitrogen groups of intrinsic $pK = 9.0$.

Clustering of this kind will increase rather than decrease the apparent value of the pK of the participating amino groups. In summary, for this cluster, the effect will be an apparently normal amino group, anomalously high anion-binding constants, and carboxylate ions with apparently lower and more heterogeneous pK values. While titration of the amino groups in the alkaline region will not reveal the structure, anion and hydrogen binding in the acid region will.

TABLE IV

IONIZATION OF THREE NITROGEN GROUPS IN A CLUSTER COMPARED WITH THAT OF THREE INDEPENDENT GROUPS WITH $K_{H_2} = 10^{9.6}$

Hydrogen-Ion-binding Equation	k_{H_2}	$(k_D')_1$	$(k_D')_2$	$(k_D')_3$	ϵ	Value of Sum of the Squares of the Differences in $\bar{\nu}_H$ from Each Equation[a] (14 points compared)
Equation (10)	$10^{8.8}$	0	0	0	3.5	0.25
	$10^{8.5}$	0	0.1	0.4	3.5	0.05

[a] Equation for independent sites: $\bar{\nu}_H = 3k_{H_2}c_H/(1 + k_{H_2}c_H)$.

The presence of six clusters involving 29 amino acids places considerable limitation upon the three-dimensional folding of the ribonuclease molecule. Previous studies (Saroff and Carroll, 1962) have implicated the four histidine residues as those taking part in the two anion-binding sites at pH 4.5. With the chemical evidence coupling the histidine residues, 12 and 119 with lysine, 41, the other two histidine residues are then paired in a second cluster. When histidine 48 and histidine 105 are brought together, the amino group of lysine 104 and the carboxyl group of glutamic acid 49 are already nearby and the carboxyl group of aspartic acid 53 is easily available. Held in place by the disulfide bond of the cystine residues 40–95 are the four basic amino acids lysine 37, arginine 39, lysine 41 and lysine 98. With lysine 41 assigned to a histidine cluster, the necessary three nitrogen centers remain clustered beside the cross-link of the cystine residues 40–95. The carboxyl group of aspartic acid 38 is also fixed into position by the disulfide bonds to furnish a carboxyl group for the third cluster, while the second carboxyl group for this cluster can come from the terminal amino acid, valine 124. The two carboxyl groups of aspartic acid 83 and glutamic acid 86 are astride arginine 85 and the disulfide bond of the cystine residues 26–84 holds lysine 31 and arginine 33 nearby. These constitute four easily identified clusters of the kind postulated and are summarized in Table V. Two other clusters

TABLE V

POSSIBLE CLUSTERS OF THREE BASIC NITROGEN CENTERS PLUS TWO CARBOXYL GROUPS

Site	Nitrogen Centers	Carboxyl Groups
1	Histidine 12	Aspartic acid 14
	Histidine 119	Aspartic acid 121
	Lysine 41	
2	Histidine 48	Glutamic acid 49
	Histidine 105	Aspartic acid 53
	Lysine 104	
3	Lysine 37	Aspartic acid 38
	Arginine 39	Valine 124
	Lysine 98	
4	Lysine 31	Aspartic acid 83
	Arginine 33	Glutamic acid 86
	Arginine 85	
5	Lysine 1–2 NH_2	Glutamic acid 2
	Lysine 61	Glutamic acid 111
6	Lysine 66	Glutamic acid 9
	Lysine 7	
	Arginine 10	

involving primarily the path for winding the *S*-peptide (Richards, 1958) (amino acids 1–20) are included in Table V. It is apparent that this fragment must take part in a considerable number of interactions to complete these sites. The S-peptide contains five basic nitrogen centers (out of a total of nineteen) and three carboxyl groups (out of a total of eleven).

APPENDIX

An approximate calculation of the titration curves for the binding of hydrogen ions and anions to ribonuclease in the region acid to *p*H 4.5 may be made by considering the ionization of the carboxylate ions of the structure of Figure 3. The unit of three positive charges is designated (3+) in the following reactions describing the species involved in hydrogen ion binding and hydrogen bonding.

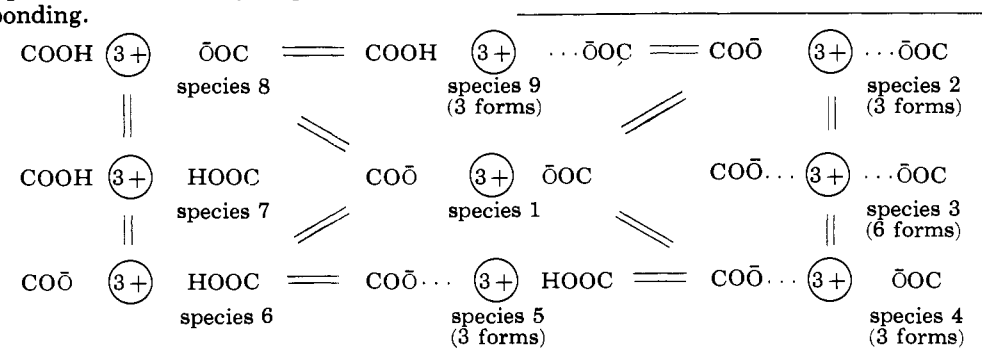
$$\mathrm{COO^-},\ \mathrm{H^+N} = \mathrm{COO^-}\cdots\mathrm{H^+}\cdots\mathrm{N} \qquad (6)$$

In the above reactions –COOH and $-\mathrm{COO^-}$ represent an isolated carboxyl group and carboxylate ion, respectively, and $-\mathrm{COO^-}$, $\mathrm{H^+N}$ and $\mathrm{COO^-}\ldots\mathrm{H^+}\ldots\mathrm{N}$ represent a carboxylate ion and charged nitrogen center in the separated and hydrogen-bonded form, respectively.

In the following equations the primed constants, k_A' and k_H', will be those including only the electrostatic effect of the net charge of a whole molecule such that $k_H' = k_H e^{-2p_W Z}$. For the binding of a proton to a cluster of equation (3) the apparent association constant will include a localized electrostatic effect resulting from the charge of the cluster as well as that deriving from the charge of the molecule. When a localized electrostatic factor, ϵ_H, is used to represent the correction factor for the localized charge of the cluster when the charge is unity then the ratio of the fraction of a single cluster in the form of species 6 to that of species 1 becomes $k_H' c_H/\epsilon_H^2$ since the charge on species 6 is +2. Similarly, with the definition of k_A of equation (4) the apparent associotion constant for a chloride ion is $k_A e^{-2Zp_W Z}/\epsilon_A$, where ϵ_A refers to the localized effect of the charge of one carboxylate ion (k_A refers to the reaction of the anion with the (3+) grouping; see equation 4). With these definitions

$$k_{H_1}' c_H = \frac{[\text{species 6}]}{[\text{species 1}]}\epsilon_H^2 = \frac{[\text{species 8}]}{[\text{species 1}]}\epsilon_H^2 = \frac{[\text{species 7}]}{[\text{species 6}]}\epsilon_H^3 = \frac{[\text{species 7}]}{[\text{species 8}]}\epsilon_H^3 = \frac{[\text{species 5}]}{[\text{species 4}]}\epsilon_H^2 = \frac{[\text{species 9}]}{[\text{species 2}]}\epsilon_H^2 \qquad (7)$$

$$k_A' c_A = \frac{[\text{species 7}']}{[\text{species 7}]} = \frac{[\text{species 6}']}{[\text{species 6}]}\epsilon_A = \frac{[\text{species 8}']}{[\text{species 8}]}\epsilon_A = \frac{[\text{species 1}']}{[\text{species 1}]}\epsilon_A^2 \text{ and } k_D' = \frac{[\text{species 2}]}{[\text{species 1}]} = \frac{[\text{species 4}]}{[\text{species 1}]} = \frac{[\text{species 3}]}{[\text{species 2}]}$$



(3)

Of these nine species only these numbered 1, 6, 7, and 8 may undergo a reaction with chloride ion as follows:

$$\underset{\text{species 1}}{\mathrm{COO^-}\ (3+)\ \mathrm{\bar{O}OC}} + \mathrm{Cl^-} = \underset{\text{species 1}'}{\mathrm{COO^-}\ (2+)\ \mathrm{\bar{O}OC}}$$

We designate the equilibrium constant, k_A, to describe the unperturbed reaction,

$$\underset{\text{species 7}}{\mathrm{COOH}\ (3+)\ \mathrm{HOOC}} + \mathrm{Cl^-} = \underset{\text{species 7}'}{\mathrm{COOH}\ (2+)\ \mathrm{HOOC}} \qquad (4)$$

The equilibrium constant, k_H, describes the unperturbed reaction,

$$\mathrm{COO^-} + \mathrm{H^+} = \mathrm{COOH} \qquad (5)$$

The equilibrium constant, k_D, refers to the reaction[3]

where the expression within the brackets designates the fraction of a single cluster in the form of that species.[4] In the foregoing equation defining the hydrogen-bonding reactions, the equilibrium constant k_D' is different from the k_D of equation (6). The k_D of equation (6) refers to the association constant for an isolated pair consisting of one carboxylate ion and one ammonium ion capable of hydrogen bonding. The k_D' refers to the apparent constant for the reaction

OH

[3] The $-\mathrm{NH^+}$-----O=C— hydrogen bond is ignored in this treatment since its constant is considered small compared to that of the $-\mathrm{NH^+}$---$\mathrm{\bar{O}OC}$— bond.

[4] Some of these ratios involve obvious approximations. For example, the ratio of species 5 to species 4, given as $k_H' c_H/e_{H_2}$, is actually different from that of species 6 to species 1 because of the presence of the hydrogen bond. This difference is assumed to be small and is ignored.

of one of the carboxylate ions with the ammonium ions of the cluster. The value of k_D' will be larger than that of k_D by the factor resulting from the charge of the cluster. Only the value of k_D' is evaluated in this approximate treatment. Since the ribonuclease molecule contains eleven carboxyl groups and nineteen basic nitrogenous groups, we assume five clusters of two carboxyl groups and three charged nitrogen centers and one cluster of one carboxyl group and three charged nitrogen centers. When all the hydrogen-bonded forms of each of the species 2, 3, 4, 5, and 9 are allowed, the cluster with the single carboxyl group will have 7 species compared to the 26 species of the cluster with two carboxyl groups. When only one form each of species 2, 3, 4, 5, and 9 are allowed the cluster with a single carboxyl group will have 5 species and the cluster with two carboxyl groups will have 13 species. For these clusters, with the assumption that $\epsilon_A = \epsilon_H$, the expression for the binding of hydrogen ions, $\bar{\nu}_H$, is:

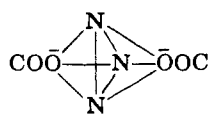
$$\bar{\nu}_H = \frac{\frac{2nk_{H_1}'c_H}{\epsilon^2}\left(1 + \frac{k_A'c_A}{\epsilon} + k_{D_2}'\right) + \frac{2nk_{H_1}'^2c_H^2}{\epsilon^5}(1 + k_A'c_A)}{\frac{2k_{H_1}'c_H}{\epsilon^2}\left(1 + \frac{k_A'c_A}{\epsilon} + k_{D_2}'\right) + \frac{k_{H_1}'^2c_H^2}{\epsilon^5}(1 + k_A'c_A) + \frac{k_A'c_A}{\epsilon^2} + (1 + k_{D_2}')^2} + \frac{\frac{k_{H_1}'c_H}{\epsilon^3}(1 + k_A'c_A)}{1 + \frac{k_{H_1}'c_H}{\epsilon^3}(1 + k_A'c_A) + \frac{k_A'c_A}{\epsilon} + k_{D_1}'} \quad (8)$$

The expression for chloride binding, ν_A, is

$$\bar{\nu}_A = \frac{\frac{nk_A'c_A}{\epsilon^2}\left(1 + \frac{2k_{H_1}'c_H}{\epsilon} + \frac{k_{H_1}'^2c_H^2}{\epsilon^3}\right)}{\frac{2k_{H_1}'c_H}{\epsilon^2}\left(1 + \frac{k_A'c_A}{\epsilon} + k_{D_2}'\right) + \frac{k_{H_1}'^2c_H^2}{\epsilon^5}(1 + k_A'c_A) + \frac{k_A'c_A}{\epsilon^2} + (1 + k_{D_2}')^2} + \frac{\frac{k_A'c_A}{\epsilon}\left(1 + \frac{k_{H_1}'c_H}{\epsilon^2}\right)}{1 + \frac{k_{H_1}'c_H}{\epsilon^3}\left(1 + k_A'c_A\right) + \frac{k_A'c_A}{\epsilon} + k_{D_1}'} \quad (9)$$

Equations (8) and (9) are written for the conditions for single forms of species 2, 3, 4, 5, and 9. If the multiple forms are allowed the term k_{D_1}' is replaced by $3k_{D_1}'$ and the term $(1 + k_{D_2}')^2$ is replaced by $(1 + 6k_{D_2}' + 6k_{D_2}'^2)$.

Equivalent equations may be written to examine the ionization of the nitrogen groups of a cluster in the alkaline region. Consider a cluster with three identical nitrogen centers and two neighboring carboxylate ions. The structure of the cluster is assumed to be as follows:



with either carboxylate ion capable of hydrogen bonding to any of the three protonated nitrogen atoms. Under these conditions there are 8 species with no hydrogen bonds and 36 species with hydrogen bonds describing the cluster during the ionization of the nitrogen atoms. If only one species of hydrogen bond is allowed between each combination of a carboxylate ion and the charged nitrogen groups then there are 8 species with no hydrogen bonds and 18 species with hydrogen bonds. The equation for the average number of hydrogen ions, $\bar{\nu}_H$, bound to one cluster is:

$$\bar{\nu}_H = \frac{3Mk_{H_2}c_H\epsilon^2 + 6Pk_{H_2}^2c_H^2\epsilon^3 + 3Qk_{H_2}^3c_H^3\epsilon^3}{1 + 3Mk_{H_2}c_H\epsilon^2 + 3Pk_{H_2}^2c_H^2\epsilon^3 + Qk_{H_2}^3c_H^3\epsilon^3} \quad (10)$$

where $M = 1 + 2(k_D')_1$ for conditions with both 36 and 18 H-bonded species, $P = 1 + 4(k_D')_2 + 2(k_D')_2^2$ for conditions for 36 H-bonded species and $[1 + (k_D')_2]^2$ for 18 H-bonded species, and $Q = 1 + 6(k_D')_3 + 6(k_D')_3^2$ for conditions with 36 hydrogen-bonded species snd $[1 + (k_D')_3]^2$ for conditions with 18 hydrogen-bonded species, and k_{H_2} is the unperturbed intrinsic association constant for hydrogen ions and an isolated nitrogen center, ϵ is the localized electrostatic effect of the cluster with a charge of unity, and $(k_D')_1$, $(k_D')_2$, and $(k_D')_3$ are the apparent constants for the hydrogen-bonding reaction between a single carboxylate ion and a protonated nitrogen center in a cluster with one, two, and three protonated nitrogen centers per cluster, respectively.

References

Anfinsen, C. B., Sela, M., and Cooke, J. P. (1962), *J. Biol. Chem. 237*, 1824.

Barnard, E. A., and Stein, W. H. (1960), *Biochim. Biophys. Acta 37*, 371.

Bates, R. G., and Pinching, G. D. (1949), *J. Am. Chem. Soc. 71*, 1274.

Craig, L. C., King, T. P., and Crestfield, A. M. (1963), *Biopolymers 1*, 231.

Crestfield, A. M., and Allen, F. W. (1954), *J. Biol. Chem. 211*, 363.

Crestfield, A. M., Stein, W. H., and Moore, S. (1963), *J. Biol. Chem. 238*, 2421.

Dintzis, H. (1952), Ph.D. dissertation, Harvard University.

Ginsburg, A. (1964), Ph.D. dissertation, George Washington University, Washington, D. C.

Hermans, J., Jr., and Scheraga, H. A. (1961), *J. Am. Chem. Soc. 83*, 3283.

Hirs, C. W. H. (1962), *Brookhaven Symp. Biol. 15*(BNL 738 (C-34)), 154.

Hirs, C. H. W., Moore, S., and Stein, W. H. (1960), *J. Biol. Chem. 235*, 633.

Linderstrøm-Lang, K. (1924), *Compt. Rend. Trav. Lab. Carlsberg, 15*, 7.

Neumann, N. P., Moore, S., and Stein, W. H. (1962), *Biochemistry 1*, 68.

Richards, F. M. (1958), *Proc. Natl. Acad. Sci. U. S. 44*, 162.

Richards, F. M., and Vithayathil, P. J. (1960), *Brookhaven Symp. Biol. 13*(BNL 608 (C 22)), 115.

Rosemeyer, M. A., and Shooter, E. M. (1963), *Proc. Intern. Congr. Biochem. 5th, Moscow, 1961, 9*, 105.

Saroff, H. A. (1957), *J. Phys. Chem. 51*, 1364.

Saroff, H. A., and Carroll, W. R. (1962), *J. Biol. Chem. 237*, 3384.

Saroff, H. A., and Healy, J. W. (1959), *J. Phys. Chem. 63*, 1178.

Saroff, H. A., and Lewis, M. S. (1963), *J. Phys. Chem. 67*, 1211.

Sela, M., and Anfinsen, C. B. (1957), *Biochim. Biophys. Acta 24*, 229.

Stark, G. R., Stein, W. H., and Moore, S. (1960), *J. Biol. Chem.*, *235*, 3177.

Steinrauf, L. K., Peterson, J., and Jensen, L. H. (1958), *J. Am. Chem. Soc. 80*, 3835.

Taborsky, G. (1959), *J. Biol. Chem. 234*, 2915.

Tanford, C., and Hauenstein, J. D. (1956), *J. Am. Chem. Soc. 78*, 5287.

Tanford, C., Hauenstein, J. D., and Rands, D. G. (1955), *J. Am. Chem. Soc. 77*, 6409.

Positions of Amino Acids in Mixed Peptides Produced from Collagen by the Action of Collagenase*

JACOB GREENBERG†, LOUIS FISHMAN, AND MILTON LEVY

From the Department of Biochemistry, New York University College of Dentistry, New York City

Received July 27, 1964

Pure collagen (ichthyocol) was nearly completely digested with purified collagenase. The enzymatic hydrolysate contained 220 peptides per 1000 amino acid residues. The enzymatic digest was degraded by stepwise application of the Edman procedure. It appears that glycine is most abundant in positions 1 and 4, proline in position 2, and hydroxyproline in position 3. There is no significant amount of proline in positions 1, 3, or 4 and no hydroxyproline in positions 1, 2, or 4 of the peptides. An equation correcting for lagging and masking effects during stepwise degradation is developed.

This paper describes attempts to find the positions of the various amino acids in the mixed peptides derived from collagen (ichthyocol) by the action of collagenase. The collagenase of *Clostridium histolyticum* was purified to the extent that it would not attack proteins other than collagen or gelatin derived from it. Its specificity, as determined by its action on peptides, involves a sequence of four amino acids given by Mandl (1961) as -X-proline-R_1-R_2-proline-Y-. Hydroxyproline as well as other amino acids may be at R_1. Hydrolysis is most rapid when R_2 is glycine, and it becomes slower as more complex amino acids occupy this position. The bond between R_1 and R_2 is hydrolyzed. R_1 becomes a carboxyl terminus and R_2 becomes an amino end of the resulting peptides. In conformity with this specificity, glycine is the main N-terminus of these peptides, with much smaller amounts of other amino acids (alanine, glutamic and aspartic acids) in this position (Michaels *et al.*, 1958). The preponderant second amino acid is proline (Grassmann *et al.*, 1961). Disagreement appears about the third position, particularly in the role of glycine and hydroxyproline in filling it. Schrohenloher *et al.* (1959) isolated large amounts of glycylprolylhydroxyproline whereas Grassmann *et al.* (1961) found only small amounts of hydroxyproline at the third position by the Edman method but reported that 26% of the total glycine was in this position. The cyclization step of the Edman method is more critical for glycine (Fraenkel-Conrat and Harris, 1954; Margoliash, 1962; Sjöquist, 1957) than for other amino acids and we have paid special attention to this step in our work. It would be expected from the current understanding of the X-ray diffraction pattern of collagen and the specificity of collagenase that the mixture of peptides should contain preponderantly three, six, nine, and so on, residue peptides.

Studies of the present type will not solve the specific sequence of collagen but will provide a kind of pattern to which known sequences should fit and also serve as a guide to the sequences to be looked for in collagen.

* This investigation was supported by research grants from the National Institutes of Health (GM 4902) and the National Science Foundation (GB 1460).

† On leave from Hebrew University, Jerusalem.

Experimental Procedures

Collagen.—Ichthyocol was prepared from carp swim bladders by extraction with citrate buffers and dialysis against dilute disodium phosphate (Gallop, 1955).

Collagenase.—Crude enzyme (Worthington Biochemical Corp.) was purified by chromatography on hydroxylapatite. The purified material contained 150 units (kinetic suspension assay of Gallop *et al.*, 1957) per mg of protein. It did not digest casein under conditions used for collagenase action.

Digestion.—The reaction mixture (806 mg ichthyocol and 3 mg collagenase in 110 ml of water) was brought from *p*H 5.70 to 6.95 by the addition of 0.13 ml of 0.5 N potassium hydroxide. Three drops of toluene were added and the entire solution was kept stirred at 40° and maintained at *p*H 6.95 by the operation of a *p*H-stat. After 18 hours calcium chloride was added to 0.005 M, and after 48 hours to 0.0075 M. Each addition resulted in an increase in the rate of hydrolysis at the time.

After 88 hours the suspension was centrifuged and the clear supernatant fluid was diluted to 120 ml with water. The solution is A in Table I. The precipitate was suspended in 20 ml of 0.005 M calcium chloride, 0.6 mg collagenase, and 5 drops of toluene, and digestion was continued another 21 hours. A very slight suspension remained which gave a tiny dark brown pellet on centrifuging. The clear supernatant was brought to a final volume of 20 ml (B, Table I).

Amino Acid Analysis.—Samples were hydrolyzed in evacuated sealed tubes with 6 N hydrochloric acid, (triple-distilled from glass) at 110° for 20 hours. The hydrolysates were taken to dryness in evacuated (to 1 or 2 mm Hg) desiccators over sodium hydroxide at room temperature. Amino acid analyses were carried out by the method of Piez and Morris (1960) using Aminex-50 × 12, 25 to 32-μ particles (Bio-Rad Co.), and a Technicon automatic analyzer. The precision in our hands is 5% or better and 0.04 μmole is the smallest amount determined with reasonable accuracy.