# Specificity of the Cleavage of Proteins by Dilute Acid. I. Release of Aspartic Acid from Insulin, Ribonuclease, and Glucagon*

JULIUS SCHULTZ, HOWARD ALLISON, AND MARJORIE GRICE

*From the Institute for Biochemical Studies in Cancer, Department of Biological Chemistry, Hahnemann Medical College, Philadelphia, Pennsylvania*

*Received November 27, 1961*

Aspartic acid was found to be released more rapidly than any other amino acid in the presence of 0.03 N HCl at 105° from three proteins of known structure (ribonuclease A, glucagon, and insulin). The high degree of specificity of the preferential cleavage of aspartyl residues was measured by examining the rate of appearance of free amino acids, which differed in the case of each protein and reflected the lability of particularly sensitive sequences, rather than the characteristics of the structure of the amino acids other than aspartic, to acid hydrolysis under these conditions.

Aspartic acid linkages in the peptide chain are unusually sensitive to acid hydrolysis, as indicated by the early appearance of the amino acid in the free state after treatment of proteins with acid (Bull *et al.*, 1949). Early observations of this reaction stimulated subsequent studies on its mechanism, kinetics, and sensitivity (Leach, 1955), but such data were understandably qualitative owing to the lack of automatic procedures for the quantitative determination of amino acids and the unavailability of proteins of known structure other than insulin. A better understanding of the factors involved in the selective cleavage of peptide bonds would increase the possible usefulness of partial acid hydrolysates of proteins for structural analysis (Schultz, 1961). In the present report, the degree of specificity of the liberation of aspartic acid from proteins in dilute acid was determined by following the time course of release of amino acids from insulin, glucagon, and ribonuclease A in 0.03 N HCl at 105° by use of the analytical procedure of Spackman *et al.* (1958). Aspartic acid was found to be released according to first-order kinetics. The presence of other amino acids in certain cases could be interpreted as a reflection of certain structural features of the proteins. Proteins of known structure used here were glucagon (Bromer *et al.*, 1957), insulin (Ryle *et al.* 1955), and ribonuclease A (Hirs *et al.*, 1960).

## Materials and Methods

A special preparation of ribonuclease A, obtained from Worthington Corporation, was found to contain the following molar ratios of amino acids, in agreement with the reports of Hirs *et al.* (1960) and Van Vunakis *et al.* (1960): Lysine (10.5), histidine (4), $NH_3$ (17), arginine (4), aspartic acid (15), threonine (10), serine (14), glutamic acid (12), glycine (4), methionine (3.4), valine (4), isoleucine (2), leucine (2), phenylalanine (3), and tyrosine (5). Glucagon was obtained from Eli Lilly and Company, as was the zinc insulin.

Each protein was dissolved in a concentration of approximately 1 to 2 mg per ml of 0.03 N HCl. Aliquots were sealed in 10-ml ampuls and placed in an oven at 105°. An additional aliquot was treated with an equal volume of 12 N HCl, sealed, and put in the same oven for 24 hours for complete hydrolysis. At various time intervals, as indicated in the results, ampuls were removed and cooled and their contents taken to dryness in a rotary evaporator. During the course of these experiments, duplicate samples were compared in the case of the partial hydrolysates in 0.03 N HCl to determine whether evaporation was necessary or whether these samples could be directly diluted with 2.25 M citrate buffer and analyzed. No difference could be observed, and hence the samples containing the 0.03 N HCl were cooled and diluted, and aliquots were placed at the 15-cm and 150-cm columns of the Phoenix Automatic Amino Acid Analyzer according to the procedure of Spackman *et al.* (1958).

## Results

Figures 1 to 5 contain the results obtained in these experiments. It can be observed that aspartic acid is rapidly released from each protein (Fig. 1). The initial rate of release is more than 100 times faster than that of other amino acids, the latter appearing only several hours after the release of one half of the expected yield of aspartic acid residues (Fig. 2). There are two exceptions to this statement, however: Threonine is rapidly released from glucagon (Fig. 3), which may be explained by the fact that the release of the aspartic next to the carboxyl-terminal threonine automatically releases free threonine. The other exception is the "isoleucine" peak in ribonuclease

* This work was supported in part by the National Institutes of Health, United States Public Health Service Grant No. C-3715 and the Damon Runyon Fund Grant No. DRG-473. The authors wish to acknowledge valuable discussion with Drs. S. Moore and W. H. Stein of the Rockefeller Institute concerning certain aspects of this problem.

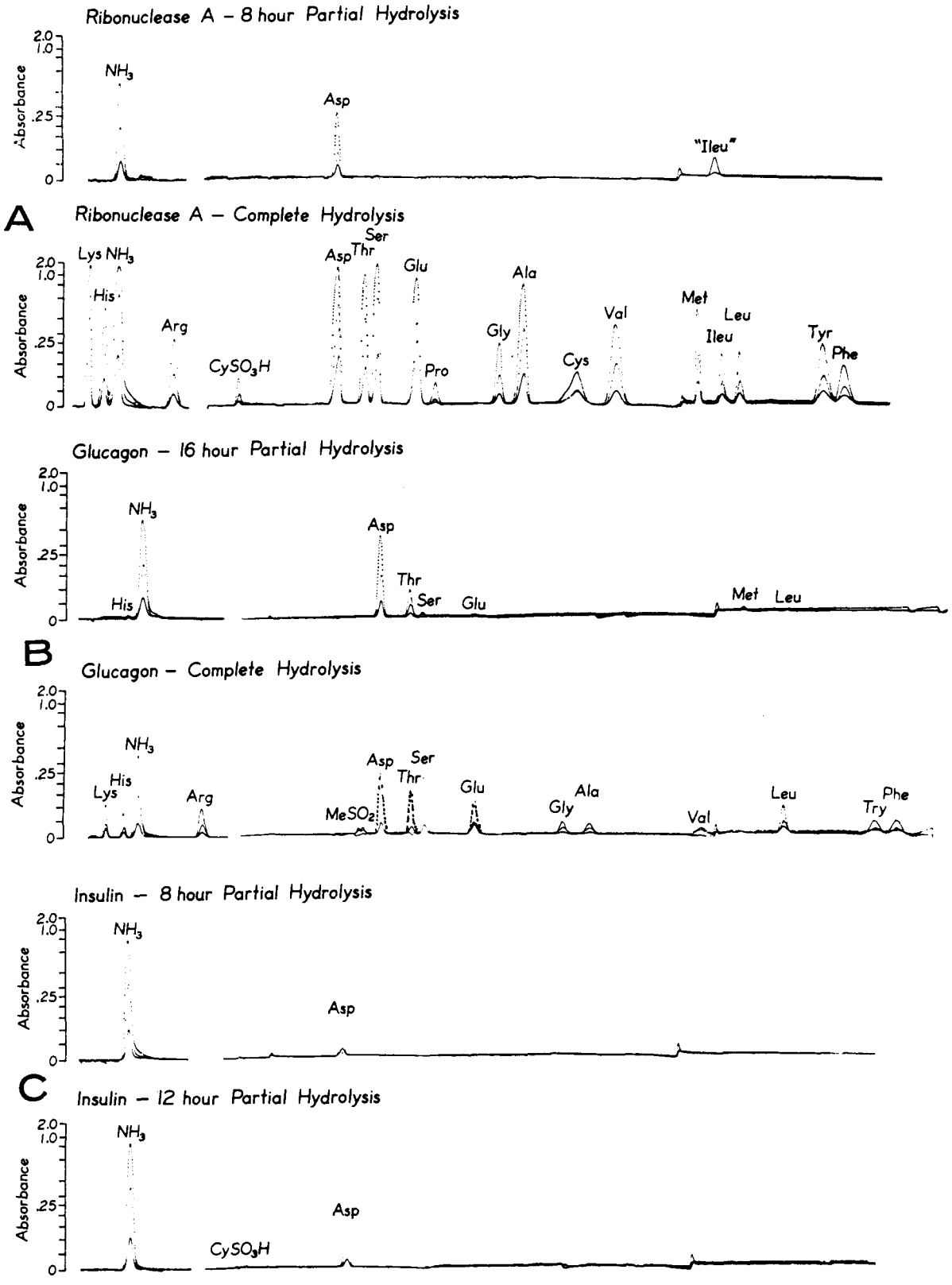


FIG. 1.—Amino acid elution patterns of partial acid hydrolysates of ribonuclease, glucagon, and insulin. *A*, comparison of elution patterns of 1.28 mg of ribonuclease A from an 8-hour hydrolysate in 0.03 N HCl at 105° with patterns from a complete hydrolysate of 1.96 mg of the same ribonuclease preparation. Chromatography was carried out in accordance with procedure of Spackman *et al.* (1958). The only amino acid peaks visible in the partial hydrolysate is aspartic acid; the "isoleucine" peak as explained in the text, is actually a peptide. *B*, glucagon. Each pattern was obtained on 0.293-mg aliquots per column of partial hydrolysates in 0.03 N HCl at 105° for time periods indicated. Other conditions same as in *A*. *C*, insulin. Conditions same as in *A* and *B*; 0.984 mg of insulin per column.

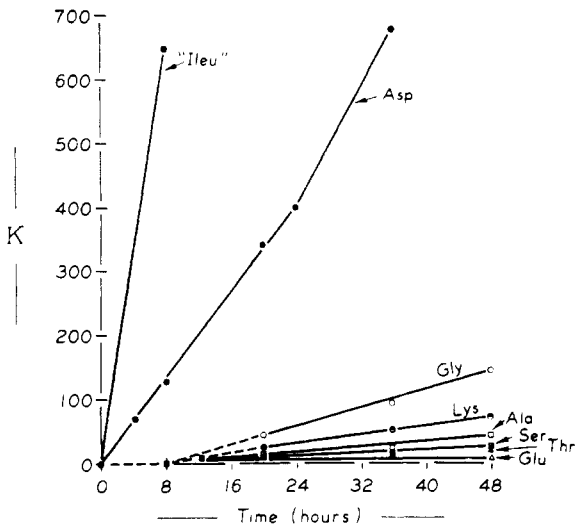


FIG. 2.—Rate of release of amino acids from 0.06 mM ribonuclease A (Worthington) in 0.30 N HCl at 105°. $K$ is equal to 100 [log $(a/a\text{-}x)$], where $x$ is the fraction of amino acid liberated of the total number of residues, $a$, found in the complete hydrolysate. The initial rate of liberation of aspartic acid is greater than 100 times that of any other amino acid. The "isoleucine" peak is not the free amino acid but rather the terminal peptide, alanyl-serylvaline.

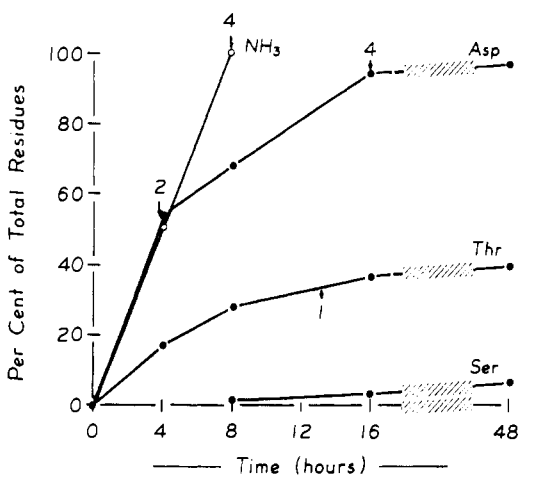


FIG. 3.—Liberation of amino acids from glucagon during partial hydrolysis in 0.03 N HCl at 105°. The arrows on each curve indicate the time interval at which the amount of amino acid released is equivalent to the number of residues per mole shown. The carboxyl-terminal aspartyl-threonine bond is not cleaved as rapidly as the other aspartyl bonds, since at least three of the four aspartic acid residues are released before all of the threonine appears (see text).

which appears as rapidly as the aspartic acid peak (Fig. 2). This, however, appears to be a peptide, as shown by the following experiment. A 4-hour partial hydrolysate of ribonuclease, prepared as described above, was subjected to electrophoresis-chromatography as described in a previous paper

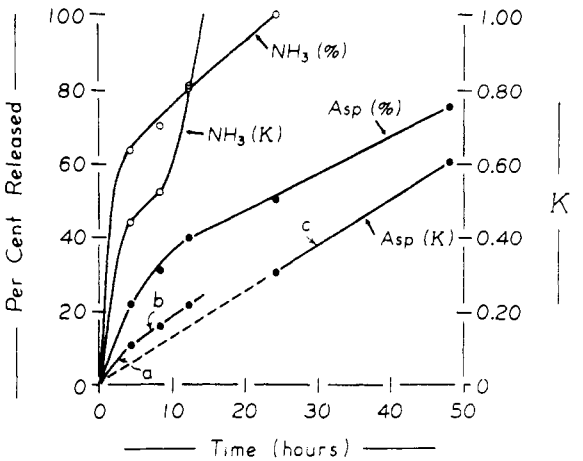


FIG. 4.—Rate of release of aspartic acid and ammonia from beef Zn-insulin during hydrolysis in 0.03 N HCl at 105°; concentration of insulin was 0.40 mg per ml. Amide nitrogen is completely released as ammonia in 24 hours, whereas only 2 of the 3 aspartic acid residues appear in 48 hours. Not indicated in the chart are other amino acids which do not appear in measurable quantities up to 24 hours; at 48 hours the following μmoles per mg insulin were found: arginine, 0.049; serine, 0.007; glutamic acid, 0.048; glycine, 0.032; alanine, 0.056; leucine, 0.025. $K$ is log $a/a - x$, where the $x$ is the fraction of aspartic acid, $a$, initially present, appearing at the time interval indicated. Percentage released is 100 $x$. See discussion for complete explanation of this chart.

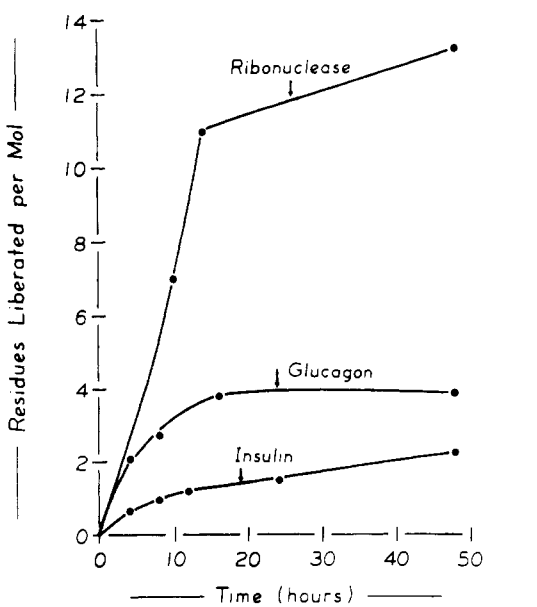


FIG. 5.—Comparative rates of release of aspartic acid by acid hydrolysis from insulin, glucagon, and ribonuclease. Conditions as described in legends of Figure 2, 3, and 4 (see text).

(Schultz, 1961). In this system of peptide mapping, the acidic and neutral peptides remain close to the origin on electrophoresis but slightly

separated. On chromatography in butanol–acetic acid–water (3:1:1), the acidic peptides rise about half way to the solvent front while the neutral peptides containing branched chain amino acids rise higher. In this region the 4-hour partial hydrolysate of ribonuclease yielded ninhydrin-positive material, which was cut and eluted in accordance with the staining procedure of Anfinsen *et al.* (1959). The eluate was divided into two parts. One was left intact and was placed directly on the amino acid analyzer; the other one was first hydrolyzed at 105° in 6 N HCl for 24 hours. The unhydrolyzed sample showed an "isoleucine" peak, whereas the hydrolyzed sample contained valine, alanine, and serine. This observation indicates that the "isoleucine" peak was most likely the carboxyl-terminal tripeptide adjacent to aspartic acid number 121.

## Discussion

A number of factors are involved in the rate of liberation of aspartic acid. Certain linkages of aspartic acid may be more resistant to cleavage than others, for one residue in insulin and one in ribonuclease appear to remain after 48 hours of hydrolysis, and both the proteins contain Asp.Cys sequences. Asparagine seems to be more slowly released than aspartic acid, as deduced from the more rapid appearance of free aspartic acid in glucagon than in insulin, where all of the aspartic acid is in the form of the amide. The data on insulin offer an opportunity for an interesting interpretation, in as much as the γ-carboxyls of all three aspartic acid residues are in the amide form, and the amide bonds must be split before the peptide bonds. In Figure 4, the percentage of aspartic acid released during the 48-hour period is plotted as a continuous curve and the first-order constants at certain time intervals are recorded as a discontinuous curve. Since only two equivalents are cleaved from the peptide chain in 48 hours and one equivalent in 8 hours, one may be justified in assuming that *two particular* residues are being liberated at independent rates. The rate for *one* would be indicated by *c*, as determined from values on the percentage curve during the 24 to 48 hour interval; this was extrapolated back to zero time. The rate at *a* was determined from the 4-hour value on the percentage curve. During this period, the two thirds of an equivalent of aspartic acid liberated was principally that of the other residue, for during this time only one-twelfth equivalent would be contributed by the residue's being split at the slower rate indicated at *c*. The slope at *b* is the same as *c*, suggesting that the predominant cleavage at that time was the residue responsible for rate *c*. The variation in the rate of liberation of the two individual residues may be due to the availability of the corresponding amide for hydrolysis; the peculiar break in ammonia curve corresponds to the marked change in rate from *a* to *b* wherein a two-fold reduction takes place in the first-order constant. Since there are also only three glutamine residues in insulin, the effect of a more resistant amide bond on a single asparagine residue could interfere with first-order kinetics: note that, unlike insulin, glucagon shows a remarkably linear rate curve for amide bond cleavage (Fig. 3). The peptides of 4-hour and 8-hour hydrolysates of insulin will be examined in future studies to clarify this interpretation. Finally, the abundance of aspartic acid in the molecule is an important factor, as indicated by the fact that more residues of aspartic acid are released from ribonuclease in 8 hours than the total number present in glucagon and insulin (see Fig. 5).

Although the absolute number of aspartic acid residues released is related to the abundance of aspartic acid in a protein, this is not true of the release of other amino acids. For, if one calculates from first-order constants the time required for the appearance of one equivalent of these amino acids in the case of ribonuclease the following values obtain: aspartic acid (less than 1 hour); lysine (30 hours); alanine (40 hours); threonine (44 hours); serine (48 hours); glutamic acid (54 hours); and glycine (56 hours). The order of abundance of these residues in ribonuclease is aspartyl > seryl > glutamyl > alanyl > threonyl > lysyl > glycyl. Since the order of cleavage is not related to the order of abundance of a given residue, the position of the residue in the peptide chain or its chemical nature, or both, must affect the sensitivity of the bonds to acid hydrolysis. If the sensitivity to acid cleavage were dependent primarily on the chemical nature of the amino acid *per se*, as in the case of aspartic acid, then the rate of appearance in the free state should be related to the abundance of a particular amino acid in each of the proteins. But since this is not so, as indicated above, it is possible that the chemical nature of the residue is not the determining factor, and that the position in the peptide chain or the nature of the adjoining residues must be considered as factors involved in the sensitivity of a given residue to cleavage from the peptide chain.

Application of dilute acid hydrolysis as a means of initial cleavage of large protein molecules into smaller units for microstructural studies was not encouraged by Sanger (1952) because the requirements for specificity were not met under the conditions used at the time. Leach (1955) examined amino-N and aspartic acid liberated during the course of partial acid hydrolysis of a number of proteins (at various hydrogen ion concentrations). Within the limits of the methods used, the aspartic acid liberated corresponded to the amino-N found at acid concentrations of *p*H of 2–3. Sanger's early statements were based on higher acid concentrations. Leach explained the effect of acid concentration by two kinds of cleavage, one through the free γ-carboxyl group of the aspartic acid residue, which is specific for aspartic acid, and the other a nonspecific cleavage by hydronium ions free in solution. Hence, at a *p*H where most of

the available protons would originate from the free carboxyl of aspartic acid a high degree of specificity would obtain. Detailed arguments for this view are reviewed by Leach (1955). The high yield of aspartic acid–free peptides obtained from ribonuclease has been reported (Schultz, 1961). Greater insight into the various degrees of nonspecific cleavage taking place at 105° in 0.03 N HCl is offered by the data presented here on three proteins of known amino acid sequence.

### References

Anfinsen, C. B., Aqvist, S. E. G., Cooke, J. P., and Jonsson, B. (1959), *J. Biol. Chem. 234*, 1118.

Bromer, W. W., Sinn, L. G., and Behrens, O. K. (1957), *J. Am. Chem. Soc. 79*, 2807.

Bull, H. B., Hahn, J. W., and Baptist, V. H. (1949), *J. Am. Chem. Soc. 71*, 550.

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

Leach, S. J. (1955), *in* Proceedings Intern. Wool Textile Res. Conf., Pt. I, Commonwealth Scientific and Industrial Research Organization, Melbourne, Australia, C-181.

Ryle, A. P., Sanger, F., Smith, L. F., and Kitai, R. (1955), *Biochem. J. 60*, 541.

Sanger, F. (1952), *Advances in Protein Chem. 7*, 1.

Schultz, J. (1961), *J. Polymer Sci. 49*, 25.

Spackman, D., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1190.

Van Vunakis, H., Leikim, E., Delaney, R., Levine, L., and Brown, R. K. (1960), *J. Biol. Chem. 235*, 3430.

---

# Studies of Soybean Trypsin Inhibitor. I. Physicochemical Properties*

YING VICTOR WU† AND HAROLD A. SCHERAGA

*From the Department of Chemistry, Cornell University, Ithaca, N. Y.*

*Received January 15, 1962*

Crystalline soybean trypsin inhibitor (STI) was found to be fairly homogeneous by ultracentrifugal measurements and by gradient chromatography on diethylaminoethyl cellulose (DEAE-cellulose). Its partial specific volume in 1 M KCl at 20.0° and its Archibald molecular weight are 0.698 ± 0.006 ml/g and 21,500 ± 800 g/mole respectively. A hydrogen ion titration curve of STI has been obtained in 1 M KCl at 25.0°. The ionizing groups and their intrinsic $pK$ values found are 35 carboxyl (4.21), 2 imidazole (6.45), 1 $\alpha$-amino (7.8), 11 $\epsilon$-amino (9.9), and 4 tyrosyl (9.5). The guanidyl groups affect the titration curve only through the net charge, $Z$, of STI, because of their high intrinsic $pK$, the large negative charge of the protein at high $p$H, and the large value of the electrostatic factor, $w$. The $pK$ values for the carboxyl groups are abnormally low; this suggests the presence of local interactions involving these groups. Empirical values of the electrostatic factor, $w$, are found to be 0.026 for carboxyl groups and 0.060 for tyrosyl groups in 1 M KCl. This increase in $w$ suggests that the conformation of the molecule may be dependent on $p$H, a conclusion which will be discussed in a forthcoming paper in connection with optical rotation and ultraviolet difference spectra measurements.

Theoretical and experimental studies of $p$H-dependent, reversible denaturation in dilute solution offer the possibility of providing us with information about the nature of side-chain interactions and conformation in protein molecules (Scheraga, 1960, 1961a,b). As a further experimental test of current ideas about denaturation, a study of crystalline soybean trypsin inhibitor (STI) was undertaken. This protein was selected since Kunitz (1947) had reported its denaturation to be reversible.

This paper is a report of various physicochemical experiments on STI, the results of which are necessary for the interpretation of the denaturation data to be reported in the forthcoming paper (Wu and Scheraga, 1962).

### Experimental

*Materials.*—The STI used was obtained from Worthington Biochemical Corp., Freehold, N. J. (five-times crystallized, Lot No. SI5433). Reagent grade diethylaminoethyl cellulose (DEAE-cellulose), with an exchange capacity of 0.85 meq per g, was purchased from Brown Co., Berlin, N. H. Two-times crystallized trypsin, containing

* This work was supported by research grant No. E-1473 from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health, Public Health Service, and by grant No. G-14793 from The National Science Foundation.

† Present address: Northern Regional Research Lab., Peoria, Illinois.