References

Buck, C. A., Glick, M. C., and Warren, L. (1970), *Biochemistry 9*, 4567.
Buck, C. A., Glick, M. C., and Warren, L. (1971), *Science 172*, 169.
Chanock, B. M., Hayflick, L., and Barile, M. C. (1962), *Proc. Nat. Acad. Sci. U. S. 48*, 41.
Gasic, G. J., and Gasic, T. (1962), *Proc. Nat. Acad. Sci. U. S. 48*, 1172.
Gesner, B. M., and Ginsberg, V. (1964), *Proc. Nat. Acad. Sci. U. S. 52*, 750.
Glick, M. C. (1971), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 30*, No. 3, 1944.
Glick, M. C., and Buck, C. A. (1973), *Biochemistry 12*, 85.
Hayflick, L. (1965), *Exp. Cell Res. 37*, 614.
Hitotsumachi, S., Rabinowitz, Z., and Sachs, L. (1972), *Int. J. Cancer 9*, 305.
Huberman, E., Salzberg, S., and Sachs, L. (1968), *Proc. Nat. Acad. Sci. U. S. 59*, 77.
Inbar, M., Ben-Bassat, H., and Sachs, L. (1972), *Nature (London), New Biol. 236*, 3.
Macpherson, I., and Montagnier, L. (1964), *Virology 23*, 292.
Meezan, E., Wu, H. C., Black, P. H., and Robbins, P. W. (1969), *Biochemistry 8*, 2518.
Pollack, R., Green, H., and Todaro, G. (1968), *Proc. Nat. Acad. Sci. U. S. 60*, 126.
Pricer, W. E., Jr., and Ashwell, G. (1971), *J. Biol. Chem. 246*, 4825.
Rabinowitz, Z., and Sachs, L. (1968), *Nature (London) 220*, 1203.
Rabinowitz, Z., and Sachs, L. (1970), *Int. J. Cancer 6*, 388.
Rabinowitz, Z., and Sachs, L. (1972), *Int. J. Cancer 9*, 334.
Sakiyama, H., and Burge, B. W. (1972), *Biochemistry 11*, 1366.
Shoham, J., Inbar, M., and Sachs, L. (1970), *Nature (London) 227*, 1244.
Sumner, A. T., Evans, H. J., and Buckland, R. A. (1971), *Nature (London), New Biol. 232*, 31.
Warren, L., Critchley, D., and Macpherson, I. (1972a), *Nature (London) 235*, 275.
Warren, L., Fuhrer, J. P., and Buck, C. A. (1972b), *Proc. Nat. Acad. Sci. U. S. 69*, 1838.
Wu, H. C., Meezan, E., Black, P. W., and Robbins, P. W. (1969), *Biochemistry 8*, 2509.
Yamamoto, T., Rabinowitz, Z., and Sachs, L. (1973), *Nature (London), New Biol. 243*, 247.

Purification and Partial Sequencing of Cyanogen Bromide Peptides from L-Asparaginase of *Escherichia coli* B†

Janet G. Gumprecht‡ and John C. Wriston, Jr.*

ABSTRACT: *S*-Aminoethylated asparaginase from *Escherichia coli* B has been fragmented with cyanogen bromide, and the seven expected peptides isolated in homogeneous condition. The total amino acid composition of the peptides agrees to within 4% of that reported for asparaginase itself. Four of the cyanogen bromide peptides have been sequenced by conventional methods; two of them are linked in the native enzyme by an intrachain disulfide bond. The N- and C-terminal peptides have been identified. The C-terminal peptide contains approximately 173 amino acids, accounting for over half of the amino acids per subunit. A Beckman Sequenator was used to establish the sequence of twelve amino acids at the N-terminal end of this large C-terminal cyanogen bromide peptide. The isolation and sequencing of an overlap tryptic peptide permitted the alignment of three of the interior cyanogen bromide peptides with respect to one another. The sequence of 78 residues, about 25% of the asparaginase molecule, has been established. The results presented here are supportive evidence for a four-subunit model in which the subunits are identical.

L-Asparaginase from *Escherichia coli* B is an enzyme of particular interest because of its ability to cause the regression of certain tumors in mice, and its antileukemic activity in some cases of human acute lymphatic leukemia (for reviews, see Adamson and Fabro, 1968, Capizzi *et al.*, 1970, Cooney and Handschumacher, 1970, and Wriston, 1971). The purification of this enzyme has been described (Whelan and Wriston, 1969) and certain properties of the enzyme, including molecular weight, amino acid composition, isoelectric point, and the existence of four identical or nearly identical subunits have been established (Wriston, 1971). There has been little work so far, however, on the primary structure of the enzyme. Arens *et al.* (1970) reported that asparaginase contained only leucine as the N-terminal amino acid, and established the sequence of ten amino acids at the N-terminal end. Greenquist and Wriston (1972) showed that asparaginase also contained only tyrosine as the C-terminal amino acid; established the sequence of four amino acids at the C-terminal end; and reported that the results obtained by preparing two-dimensional peptide maps of tryptic digests of reduced, aminoethylated asparaginase were in agreement with a four-subunit model in which the subunits are identical or nearly so. Here we report the initial results of sequence studies on the enzyme involving cleavage with cyanogen bromide, isolation in homogeneous condition of the expected number of peptides, and the se-

† From the Department of Chemistry, University of Delaware, Newark, Delaware 19711. *Received August 25, 1973.* This work was supported in part by Grant CA 06780 from the National Cancer Institute.

‡ Present address: Harvard Medical School, Hematology Unit, Massachusetts General Hospital, Boston, Mass. 02114.

quencing of several of them. A preliminary account of this work has been presented (Wriston *et al.*, 1972).

Experimental Section

Materials. E. coli B asparaginase ("Lyovac") was a generous gift from Merck, Inc. The material was dialyzed against water to remove added mannitol, and lyophilized. The enzyme was judged to be essentially homogeneous because it was indistinguishable in terms of specific activity, amino acid composition, and electrophoretic mobility on polyacrylamide gels from asparaginase previously isolated in this laboratory (Whelan and Wriston, 1969).

Trypsin (L-1-tosylamido-2-phenylethyl chloromethyl ketone treated), carboxypeptidase A and B (diisopropyl fluorophosphate treated), and chymotrypsin were obtained from the Worthington Biochemical Corp. Aminopeptidase M was obtained from Henley and Co.

Cyanogen bromide, thioglycolic acid, tryptophan, and Ultra Pure guanidine hydrochloride were obtained from Schwarz/Mann. The Sephadexes (fine) were purchased from Pharmacia Fine Chemicals, and analytical grade Dowex 50W-X2 and Bio-Rex 70 from Bio-Rad. Homoserine was purchased from Sigma Chemical Co., and *S*-aminoethylcysteine from British Drug Houses, Ltd. Other amino acid standards were obtained from Beckman. Various reagents for the amino acid analyzer were obtained either from Pierce Chemical Co. or Beckman. Polyamide sheets were purchased from Gallard-Schlesinger Chemical Mfg. Corp. Constant-boiling HCl and water were double glass distilled. Pyridine was redistilled over ninhydrin, and *N*-ethylmorpholine (Aldrich Chemical Co.) was also redistilled. All other reagents were either analytical or spectroscopic grade.

Amino Acid Analyses. Hydrolysis of protein and peptide samples in constant-boiling HCl was performed in sealed evacuated tubes at 110°, usually for 22 hr. Analyses were carried out on a Beckman 120C amino acid analyzer equipped for high-sensitivity analysis according to the general procedures of Spackman *et al.* (1958). Amino acid peaks were integrated with an Infotronics CRS-100A integrator. When tryptophan was to be determined, the HCl was also made 2% in thioglycolic acid (Matsubara and Sasaki, 1969). Tryptophan standards subjected to acid hydrolysis gave 90% recovery under these conditions. In some cases, enzymatic hydrolysis with aminopeptidase M was used to determine the presence of tryptophan (Elzinga *et al.*, 1968). In order to determine *S*-aminoethylcysteine and homoserine lactone, a 15-cm short column was used. The order of elution of the basic amino acids on this column is tryptophan, lysine, *S*-aminoethylcysteine, histidine, ammonia, homoserine lactone, and arginine. Homoserine and glutamic acid were resolved from each other on the long column according to Mills *et al.* (1970).

Cyanogen Bromide Cleavage. Asparaginase was first reduced and *S*-aminoethylated as previously described (Greenquist and Wriston, 1972). The RAE-asparaginase[1] was then dissolved in 70% formic acid (Steers *et al.*, 1965) at a concentration of 10 mg/ml and incubated with a 50-fold molar excess of cyanogen bromide to methionine at room temperature for 24 hr. The peptide mixture was diluted approximately 10-fold, shell frozen, and lyophilized.

Column Chromatography. Gel filtration was used in several stages of purification (see Figure 1). Samples were applied to 0.9 × 400 cm columns equilibrated with 20% acetic acid and eluted with the same solvent (10% acetic acid was used with the G-25 columns). Thiodiglycol (0.1% final concentration) was added to eluting solvents when separating peptides containing *S*-aminoethylcysteine, to minimize oxidation losses. Bio-Rex 70 chromatography utilized resin (minus 400 mesh) which was converted to the hydrogen form with 5% acetic acid. Fines were removed according to the procedure of Edmundson (1967), and columns were packed according to Hirs (1955) and then equilibrated with 5% acetic acid. Because of the relative insolubility of the samples, the gradient system of Humbel *et al.* (1968) was used. Dowex 50W-X2 (200–400 mesh) columns were prepared and developed with pyridine–acetate buffers according to Schroeder *et al.* (1962). Peptides were located by ninhydrin analysis after alkaline hydrolysis (Hirs, 1967). The ninhydrin solution was prepared with dimethyl sulfoxide as the solvent (Moore and Stein, 1954). Fractions comprising the various pools were combined and lyophilized.

Tryptic and Chymotryptic Digestion. Lyophilized peptides were dissolved or suspended in 0.5 ml of 0.2 M NH_4HCO_3 buffer (pH 8.1). Trypsin was dissolved in 0.001 M HCl (1 mg/ml) and cleavage of the peptide initiated by adding trypsin at a 1:100 (w/w) ratio of enzyme to substrate. The reaction was allowed to continue for 1 hr, when a second portion of trypsin was added, giving a final ratio of enzyme to substrate of 1:50. The reaction was terminated after 18 hr by evaporating to dryness at 40° under a stream of nitrogen. With chymotrypsin, 25 μl of a solution containing 1 mg/ml in water was added, and the reaction was allowed to proceed for 1 hr at 37°. More enzyme was then added, giving a final enzyme to substrate ratio (w/w) of 1:50, and digestion was allowed to continue for 2 more hr. The reaction mixture was acidified with acetic acid, and dried under nitrogen at 40°.

Carboxypeptidase Digestion. Carboxypeptidase A was solubilized according to the procedure of Ambler (1967) adapted from an earlier method (Fraenkel-Conrat *et al.*, 1955). An aliquot containing 10–20 nmol of peptide was dried under nitrogen at 40°, and then taken up in 0.25 ml of 0.2 M *N*-ethylmorpholine acetate buffer (pH 8.5), and carboxypeptidase A was added. Aliquots were removed at various times, and the reaction was stopped by acidification with acetic acid, followed by drying. The samples were taken up in 0.2 N sodium citrate buffer (pH 2.2), and applied directly to the amino acid analyzer for determination of free amino acids. Appropriate reaction conditions were determined for each peptide. Enzyme blanks were included in all runs. When carboxypeptidase A was to be used with tryptic peptides, lysine and arginine residues were first removed by hydrolysis with carboxypeptidase B (1 hr, 37°, 6.4 μg).

Cleavage at Aspartic Acid. The procedure of Schultz (1967) was used to bring about preferential cleavage at aspartic acid in one instance, using 0.03 M HCl for 16 hr at 110°.

Edman Degradation. The subtractive Edman method (Hirs *et al.*, 1960) as modified by Greene (personal communication) was used. Aliquots containing about 10 nmol of peptide were transferred to ignition tubes (18 × 150 mm) and dried at 40° under a stream of nitrogen; 0.1 ml of 50% aqueous pyridine (v/v) was added to each tube, followed by 0.1 ml of 10% ethyl isothiocyanate in pyridine (v/v), freshly prepared. The tubes

[1] Abbreviations used are: Tos-PheCH$_2$Cl, L-1-tosylamido-2-phenylethyl chloromethyl ketone; RAE-asparaginase, reduced aminoethylated asparaginase; HS, homoserine; AECySH, *S*-aminoethylcysteine; ETH, ethylthiohydantoin; PTH, phenylthiohydantoin; CPA, carboxypeptidase A; CPB, carboxypeptidase B; Quadrol, *N,N,N',N'*-tetrakis(*S*-hydroxypropyl)ethylenediaminetrifluoroacetic acid; dansyl, dimethylaminonaphthalene-5'-sulfonyl.

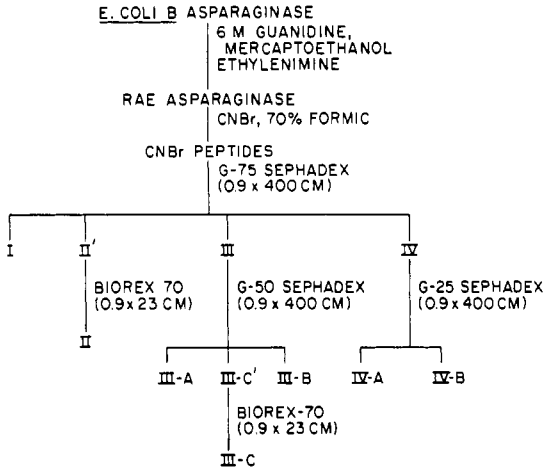

FIGURE 1: Purification scheme for cyanogen bromide peptides from *E. coli* B asparaginase.

were flushed briefly with nitrogen, stoppered with silicone stoppers, and the coupling reaction was allowed to proceed for 2 hr at 40°. After thorough removal of excess reagents under nitrogen, 0.2 ml of anhydrous trifluoroacetic acid was added, nitrogen was swept over the tubes before sealing, and cyclization was accomplished in 20 min at 40°. The samples were again dried, one tube was removed, and the same procedure was repeated until the desired number of cycles had been carried out. The thiazolinones were converted to ETH derivatives and by-products were removed by extracting twice with 1.5 ml of *n*-butyl acetate. The aqueous layer was evaporated to dryness, and the residue was subjected to amino acid analysis to determine the composition of the degraded peptide.

Dansylation. Peptides were dansylated according to the procedure of Gray (1972). After acid hydrolysis at 110° for 12 hr, the samples were dried under vacuum at 40°. They were then dissolved in 20 μl of 0.2 M NH_4OH, and 1-μl aliquots were spotted on polyamide layers for thin-layer chromatography. A standard mixture of dansylated amino acids was spotted on the back of the polyamide layer in the same position as the unknown spot (Hartley, 1970). Plates were developed with the solvent system of Woods and Wang (1967).

Automatic Sequenator. A 500-nmol sample of the large cyanogen bromide peptide CN-I was subjected to sequencing on the Beckman Spinco protein-peptide sequencer, Model 890. The initial solvent was water, and the "Quadrol Buffer" program was used. PTH derivatives were identified by gas chromatography or by thin-layer chromatography. A Beckman GC Model 45 was used for the gas chromatographic identification, with a glass column packed with SP-400. Eastman Chromagram silica gel with indicator plates were used for thin-layer chromatography, with chloroform as the solvent. When the PTH derivative could not be identified by these methods, recourse was made to base hydrolysis, using 0.1 N NaOH followed by amino acid analysis (Africa and Carpenter, 1966).

Nomenclature and Yield Calculations. CNBr peptides are designated by the prefix "CN," and tryptic and chymotryptic peptides by "T" and "C," respectively. Individual peptides are numbered based on their elution position during column chromatography. Peptide yields are calculated from amino acid analyses, and are not corrected for chromatographic losses. They are corrected for the amounts used for amino acid analysis and peptide detection if significant aliquots were withdrawn for these purposes.

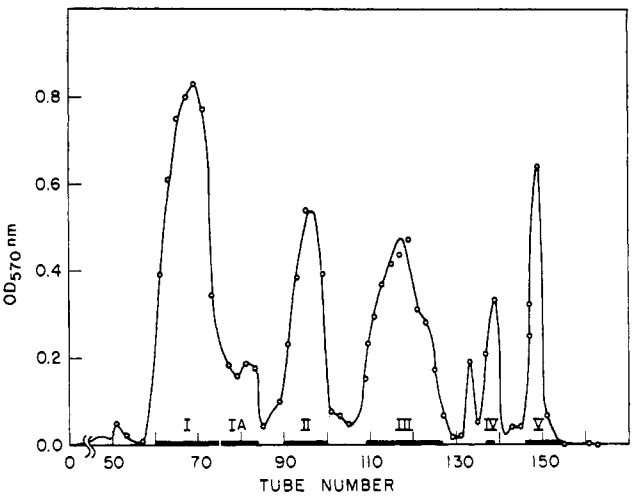

FIGURE 2: Sephadex G-75 separation of the cyanogen bromide peptides from RAE-asparaginase. Eluted with 20% acetic acid at a flow rate of 11 ml/hr. See Methods.

Results

Purification of Cyanogen Bromide Peptides. In a typical experiment approximately 70 mg (2.15 μmol) of RAE-asparaginase was cleaved with CNBr (see Methods). The purification scheme is outlined in Figure 1. Gel filtration on Sephadex G-75 gave five peptide peaks (Figure 2). The shoulder on the descending side of peak I (peak Ia) may represent a small amount of partially cleaved material, since it was absent in another cleavage experiment carried out under more stringent conditions. Neither peak IV nor peak V in Figure 2 was homogeneous, and in other experiments the two apparent peaks were not as well separated as they are here. They were recombined and fractionated on Sephadex G-25, giving two homogeneous peptides, CN-IV-A and CN-IV-B.

Purification of the three peptides contained in pool III was difficult because at least one of the peptides was insoluble in several solvents. Resolution of these peptides on Sephadex G-50 was not complete (Figure 3) but by combining fractions on the rising edge of the first peak into one pool, and on the

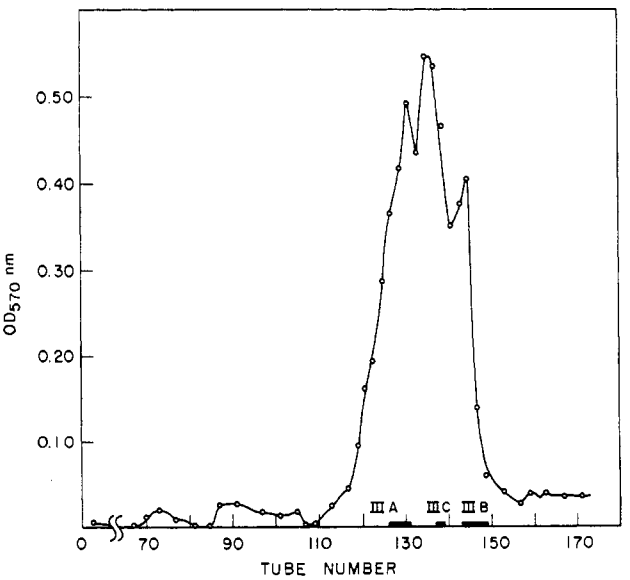

FIGURE 3: Gel filtration of pool III cyanogen bromide peptides on Sephadex G-50. Eluted with 20% acetic acid at a flow rate of 5 ml/hr. See Methods.

TABLE I: Amino Acid Composition of CNBr Peptides.[a]

Amino Acid	Whole Enzyme[c]	I	II	III-A	III-B	III-C	IV-A	IV-B	Total
Lys	21	11.5 (11)	3.65 (4)	2.99 (3)	2.00 (2)	1.00 (1)			21
His	3	1.77 (2)		1.00 (1)					3
Arg	9	5.82 (6)		0.91 (1)		0.84 (1)	0.97 (1)		9
Trp[b]	3			0.84 (1)					1
Asp	45	27.7 (28)	8.05 (8)	10.0 (10)	2.10 (2)	4.62 (5)			53
Thr	30	16.9 (17)	4.13 (4)	6.21 (6)	1.84 (2)	0.88 (1)	1.15 (1)		31
Ser	15	8.56 (9)	2.43 (2)	0.75 (1)		1.50 (1)	1.88 (2)		15
Glu	21	12.0 (12)	4.13 (4)		1.87 (2)				18
Pro	12	7.67 (8)	1.83 (2)		1.46 (1)	0.94 (1)	1.00 (1)		13
Gly	27	14.0 (14)	7.95 (8)	2.19 (2)		2.00 (2)		1.04 (1)	27
Ala	30	17.7 (18)	5.08 (5)	4.14 (4)	1.07 (1)	5.62 (6)		1.00 (1)	35
AECySH	2			0.97 (1)	0.70 (1)				2
Val	30	16.0 (16)	5.77 (6)	1.89 (2)	2.25 (2)	3.84 (4)		0.99 (1)	31
Met	6								0
Ile	12	4.92 (5)	4.03 (4)	1.61 (2)					11
Leu	21	13.7 (14)	3.45 (4)[d]	1.96 (2)	1.93 (2)	2.08 (2)			24
Tyr	12	7.69 (8)	0.82 (1)		0.76 (1)	0.85 (1)			11
Phe	9	4.85 (5)		1.00 (1)	0.64 (1)	1.29 (1)			8
Hs	6		0.30 (1)	0.27 (1)	0.40 (1)	0.38 (1)	0.45 (1)	0.45 (1)	6
Total	308	173	53	38	18	27	6	4	319
Yield (%)		50	23	37	37	11	24	41	

[a] In residues per subunit; assumed integral values are given in parentheses. [b] All cyanogen bromide peptides were analyzed for tryptophan. [c] Published values (Whelan and Wriston, 1969). [d] Integral value assumed to be 4 since removal of the N-terminal leucine gives 2.80 residues (3) in the remainder.

descending edge of the third peak into another, peptides CN-III-A and CN-III-B were obtained in homogeneous condition. Fractions in the central region of the G-50 separation were pooled and rechromatographed on Bio-Rex 70 (Figure 4), yielding homogeneous peptide CN-III-C. Both peptides CN-III-A and CN-III-B contain *S*-aminoethylcysteine, and the absence of this amino acid in CN-III-C is evidence of its homogeneity. This latter peptide appears to be responsible for the solubility problems experienced with pool III.

Amino acid analysis indicated that pool II from the initial G-75 column was not homogeneous. Application of this material to a Bio-Rex 70 column gave a single major peak. The contaminating peak appears to be CN-III-A, since it contains *S*-aminoethylcysteine. Pool I from the initial G-75 column appears to contain a single homogeneous peptide, since amino acid analysis showed no homoserine or homoserine lactone, no *S*-aminoethylcysteine, and no tryptophan.

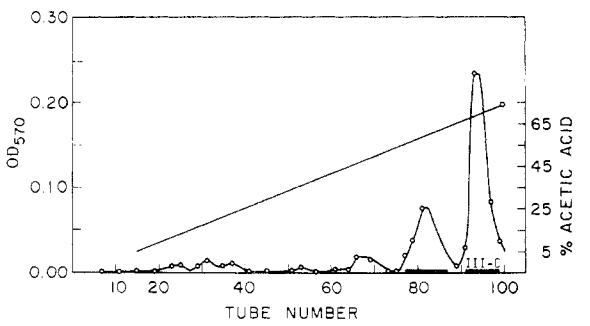

FIGURE 4: Purification of peptide CN-III-C on Bio-Rex 70. The middle peak from the gel filtration of pool III (Figure 3) was applied to a 0.9 × 23 cm column of Bio-Rex 70, equilibrated with 5% acetic acid and developed with a 5–90% acetic acid gradient.

The amino acid compositions of the seven cyanogen bromide peptides are shown in Table I.

Identification of the N- and C-Terminal Peptides. The absence of homoserine and homoserine lactone in peptide CN-I shows that it is C-terminal. Treatment with carboxypeptidase A verified this since amino acids were released in the same order as found earlier (Greenquist and Wriston, 1972) for the intact enzyme. The C-terminal sequence in Phe-Ile-Gln-Tyr.

Two cycles of Edman degradation were used to show that peptide CN-II is N-terminal. The first cycle removed leucine and the second proline, in agreement with the N-terminal sequence of asparaginase already established (Arens *et al.*, 1970). None of the other CNBr peptides have N-terminal leucine. When peptide CN-II was treated with carboxypeptidases A and B, no amino acids were released, not even homoserine, suggesting the possibility of another proline as the penultimate residue at the C-terminal end of peptide CN-II.

Sequence of Peptides CN-IV-A and CN-IV-B. These two peptides were sequenced in preliminary CNBr cleavage experiments (Greenquist, 1971) and the earlier unpublished results were confirmed in this work. The sequence of the tetrapeptide CN-IV-B was established by the subtractive Edman procedure as Val-Gly-Ala-Met. Aminopeptidase M digestion showed the absence of tryptophan.

Peptide CN-IV-A was also reexamined here to confirm the previous finding of N-terminal Arg-Pro, since this peptide is one of the few points of overlap with tryptic peptides yet established. The Arg-Pro sequence was confirmed by two separate Edman degradations, and again amino peptidase M showed no tryptophan to be present. The sequence of CN-IV-A is Arg-Pro-Ser-Thr-Ser-Met.

Sequence of Peptide CN-III-B. The N-terminal amino acid of CN-III-B was found by one cycle of Edman degradation to be

TABLE II: Tryptic Peptides from CN-III-B.

Amino Acid	CN-III-B	CN-III-B-T-A	CN-III-B-T-B	Total
Lys	2.00 (2)	1.00 (1)	1.00 (1)	2
AECySH	0.70 (1)		0.70 (1)	1
Asp	2.10 (2)	1.11 (1)	0.92 (1)	2
Thr	1.84 (2)	2.03 (2)		2
Glu	1.87 (2)	2.09 (2)		2
Pro	1.46 (1)		0.95 (1)	1
Ala	1.07 (1)	0.97 (1)		1
Val	2.25 (2)	1.00 (1)	1.39 (1)	2
Leu	1.93 (2)	2.19 (2)		2
Tyr	0.76 (1)	0.78 (1)		1
Phe	0.64 (1)	0.81 (1)		1
Hs	0.40 (1)		0.37 (1)	1
Total	18	12	6	18

Glx. CPA (enzyme to substrate molar ratio of 1:60) characterized the C terminus as -Val-Met. Since this peptide contains 18 amino acids it was further fragmented with trypsin, and the mixture gave two homogeneous peptides (CN-III-B-T-A and CN-III-B-T-B) on Dowex 50W-X2 chromatography (Figure 5). The amino acid composition of the tryptic peptides from CN-III-B is given in Table II.

PEPTIDE CN-III-B-T-B. This peptide is contained within the V-D-3 tryptic peptide found in a companion study in this laboratory (Bullis, 1971). Bullis found the N terminal to be CMCySH; in this work it is AECySH. This hexapeptide is the C terminus of CN-III-B since it contains homoserine. Assuming a Lys-Pro- unit, there are two possible sequences: (1) AECySH-Asp-Lys-Pro-Val-Met and (2) AECySH-Lys-Pro-Asp-Val-Met. The sequence was established by specific cleavage at aspartic acid (Schultz, 1967; see Methods). Free *S*-aminoethylcysteine was found in 42% yield on the 15-cm basics column, and an equal amount of aspartic acid on the long column. No other amino acids were released, a result which is only consistent with the first of the two possibilities mentioned above, and the sequence of peptide CN-III-B-T-B is AECySH-Asp-Lys-Pro-Val-Met.

PEPTIDE CN-III-B-T-A. This peptide contains 12 amino acids and must have C-terminal lysine. Three cycles of Edman degradation established the N-terminal sequence as: Glx-Glx-Thr-. CPB was used to release lysine from the C terminus of the peptide so that a time study with CPA could be carried out. This peptide was unusually susceptible to sequential release of amino acids when various reaction conditions were employed. Using an enzyme to substrate ratio of 1:165 and 1:65 in separate experiments at both 37° and room temperature and with times from 30 min to 22 hr, the definite assignment of four amino acids and the tentative assignment of five more was possible. The Asx residue is identified as aspartic acid by the enzymatic release. The tentative sequence of peptide CN-III-B-T-A is Glx-Glx-Thr-(Leu-Asp-Ala-Phe-Tyr)-Leu-Thr-Val-Lys.

Sequence of Peptide CN-III-A (See Table III). Approximately 600 nmol of this peptide was incubated with trypsin, and the tryptic peptides separated on a Dowex 50W-X2 column (Figure 6). Pools B, D, and E were judged to be homogeneous based on their amino acid compostions; pools A and C gave nonintegral values for certain amino acids, and were further purified by gel filtration on Sephadex G-25

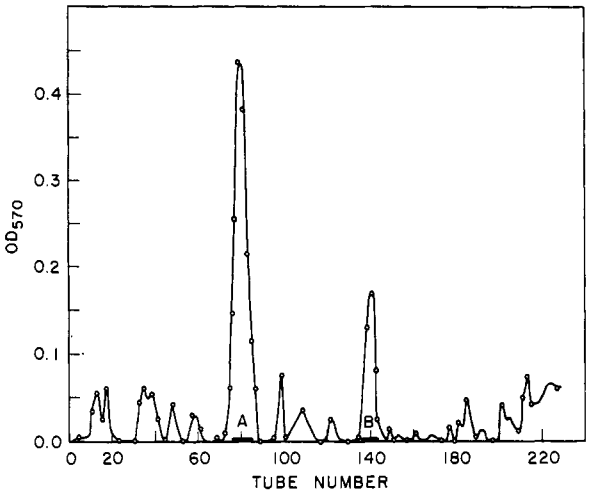

FIGURE 5: Separation of tryptic peptides from CN-III-B on Dowex 50W-X2. A 0.9 × 50 cm column was equilibrated with 0.2 M pyridine–acetate buffer (pH 3.1) at 40°. A 1-ml sample in 25% acetic acid was applied, and the column was developed with a linear gradient (250 ml of equilibration buffer and 250 ml of 2 M pyridine–acetate buffer, pH 4.9). The flow rate was 24 ml/hr, and 2-ml fractions were collected. Fifty milliliters of 8.5 M pyridine–acetate buffer (pH 5.4) was used to complete the elution.

columns. The amino acid compositions of the tryptic peptides from peptide CN-III-A are shown in Table IV.

PEPTIDES CN-III-A-T-C AND CN-III-A-T-D. The amino acid composition showed that these two peptides correspond to peptides isolated from tryptic digests of intact RAE-asparaginase and already sequenced (Greenquist, 1971). The sequence of CN-III-A-T-C is Ala-Ser-Ala-Asn-Arg and that of CN-III-A-T-D is Ile-Asn-Thr-Asp-AECySH-Asp-Lys.

PEPTIDE CN-III-A-T-E. This peptide, the most basic eluted from the Dowex 50W-X2 column, has the same composition as CN-III-A-T-D, except for an additional lysine residue. Data for the yields of the tryptic peptides (Table IV) indicate that one of the peptide bonds in CN-III-A potentially susceptible to trypsin cleavage has been only partially cleaved, since peptides CN-III-A-T-D and CN-III-A-T-E are present in approximately equal amounts, but in only about half the yield of the other tryptic peptides. End-group analysis was used to establish the

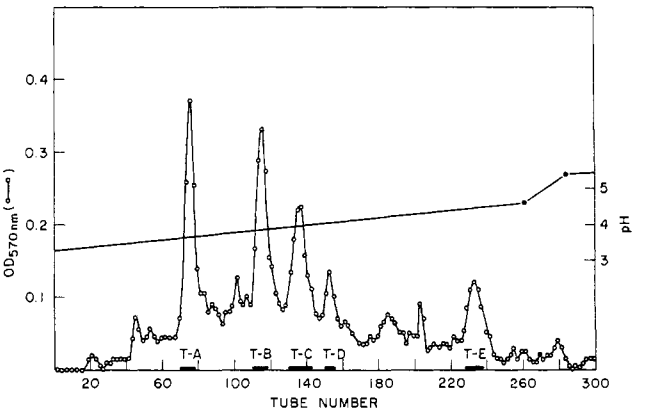

FIGURE 6: Separation of tryptic peptides from CN-III-A on Dowex 50W-X2. A 0.9 × 55 cm column was equilibrated with 0.2 M pyridine–acetate buffer (pH 3.1) at 40°. The sample was applied in 0.5 ml of 25% acetic acid, and eluted with a linear gradient using 300 ml each of equilibrating buffer and of 2 M pyridine–acetate buffer (pH 4.9). The flow rate was 17 ml/hr, and 1.6-ml fractions were collected. The elution was completed with 75 ml of 8.5 M pyridine–acetate buffer (pH 5.4).

TABLE III: Sequence of CN-III-A.

	1 5 10
	ASX-ASX-ASX-VAL-TRP-LEU-THR-LEU-ALA-LYS-ALA-SER-ALA-
CN-III-A-T-B	Asx- Asx- Asx- Val- Trp- Leu- Thr- Leu- Ala- Lys
CN-III-A-T-C	Ala- Ser- Ala-
CN-III-A-C-3B	Asx- Asx- Asx- Val- Trp
CN-III-A-C-2	Asx- Asx- Asx- Val- Trp- Leu
CN-III-A-C-1	Leu- Thr- Leu
	15 20 25
	ASN-ARG-LYS-ILE-ASN-THR-ASP-AECySH-ASP-LYS-THR-ASX-
CN-III-A-T-C	Asx- Arg
CN-III-A-T-D	Ile- Asn- Thr- Asp- AECySH-Asp- Lys
CN-III-A-T-E	Lys- Ile- Asn- Thr- Asp- AECySH-Asp- Lys
CN-III-A-T-A	Thr- Asx-
CN-III-A-C-7	Lys- Ile- Asn- Thr- Asp- AECySH-Asp- Lys- Thr- Asx-
	30 38
	GLY-PHE-VAL-ILE-THR-(ASX,ALA,GLY,THR_2,HIS)-ASP-MET
CN-III-A-T-A	Gly- Phe- Val- Ile- Thr- (Asx, Ala, Gly, Thr_2, His)- Asp- Met
CN-III-A-C-7	Gly- Phe
CN-III-A-C-3A	Val- Ile- Thr- (Asx, Ala, Gly, Thr_2, His)- Asp- Met

position of attachment of the lysine residue. CPB released only one lysine at the C-terminal end. One cycle of subtractive Edman degradation removed lysine and did not remove isoleucine, thus placing the second lysine at the N terminus. The sequence of CN-III-A-T-E is Lys-Ile-Asn-Thr-Asp-AECySH-Asp-Lys.

PEPTIDE CN-III-A-T-B. The N-terminal sequence of this decapeptide was tentatively established by subtractive Edman degradation as Asx-Asx-Asx-Val. Lysine was removed with CPB, and the use of CPA allowed verification of the N-terminal sequence of four amino acids. The first six amino acids at the C-terminal end are released very rapidly, but valine in contrast is released very slowly, presumably because of the three aspartic acid residues that follow it.

Subsequent experiments with chymotrypsin allowed the unambiguous sequencing of CN-III-A-T-B. Based on their amino acid compositions, three of the chymotryptic peptides obtained from peptide CN-III-A can only be part of peptide CN-III-A-T-B. Further, the sequence of two of these peptides (III-A-C-2 and III-A-C-3B) is established by their amino acid compositions and the known N-terminal sequence of CN-III-A-T-B. The sequence of the third of these small chymotryptic peptides (III-A-C-1) was established by Edman degradation as Leu-Thr-Leu. The sequence of peptide CN-III-A-T-B can then be established as shown in Table III.

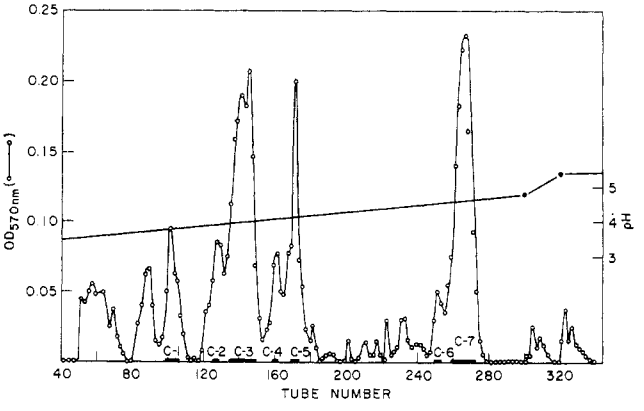

FIGURE 7: Chromatography of chymotryptic peptides from CN-III-A on Dowex 50W-X2. The sample was applied to a 0.9 × 60 cm column equilibrated with 0.2 M pyridine–acetate buffer and 0.1% thiodiglycol (pH 3.1). Eluted as in Figure 6.

PEPTIDE CN-III-A-T-A. This is the C-terminal fragment of CN-III-A since it is the only one of the tryptic peptides derived from CN-III-A that contains homoserine. Edman degradation was used to characterize the three amino acids at the N terminus as Thr-Asx-Gly. CPA (enzyme to substrate molar ratio of 1:165 at 37°) led to the removal of only homoserine and aspartic acid at the end of 5 hr, establishing the C terminus as -Asp-Met.

Based on their amino acid compositions, two more of the chymotryptic peptides obtained from CN-III-A (see next section), namely, peptides CN-III-A-C-7 and CN-III-A-C-3A, are contained in whole or in part in peptide CN-III-A-T-A, due to cleavage at the phenylalanine residue in this latter peptide. Digestion of the chymotryptic peptide that contains the phenylalanine (CN-III-A-C-7) with CPA (molar ratio of enzyme to substrate of 1:40 at 37°) for 30 min, gave 0.68 residue of phenylalanine and 0.08 residue of glycine, with no other amino acids being released. Since the chymotryptic peptide CN-III-A-C-3A contains the eleven amino acids making up the C-terminal part of CN-III-A-T-A, the N-terminal sequence is thus Thr-Asx-Gly-Phe. Three amino acids at the N-terminal end of CN-III-A-C-3A were established by Edman degradation as Val-Ile-Thr. Insufficient material was available to complete the sequencing of CN-III-A-T-A. The partial sequence established for this peptide is also shown in Table III.

CHYMOTRYPTIC DIGESTION OF PEPTIDE CN-III-A. Chymotrypsin was used to obtain overlap peptides to permit the ordering of the four tryptic peptides from CN-III-A and to provide scission at phenylalanine and leucine to aid in the sequencing of the tryptic peptides. The chymotryptic peptides were chromatographed on a Dowex 50W-X2 column (Figure 7). Pools 1 and 2 each contained a single essentially homogeneous peptide (CN-III-A-C-1 and CN-III-A-C-2, respectively) and pool 7 contained a peptide purified to homogeneity on a Sephadex G-25 column (peptide CN-III-A-C-7). Pool 3 was also rechromatographed on G-25 Sephadex,

TABLE IV: CN-III-A Tryptic Peptides.

Amino Acid	CN-III-A-T-A	CN-III-A-T-B	CN-III-A-T-C	CN-III-A-T-D[b]	CN-III-A-T-E	Total
Lys	0.30 (0)	0.99 (1)		1.00 (1)	2.00 (2)	3
His	1.00 (1)					1
Arg			0.91 (1)			1
Trp		0.86 (1)				1
AECySH				0.80 (1)	0.97 (1)	1
Asp	2.91 (3)	2.90 (3)	1.05 (1)	3.17 (3)	3.17 (3)	10
Thr	4.13 (3)	1.00 (1)		1.26 (1)	1.08 (1)	6
Ser			0.75[a] (1)			1
Glu						
Pro						
Gly	2.19 (2)					2
Ala	1.08 (1)	1.06 (1)	2.00 (2)			4
Val	0.94 (1)	0.95 (1)				2
Ile	0.65 (1)			0.87 (1)	0.96 (1)	2
Leu	0.21 (0)	1.96 (2)				2
Tyr						
Phe	1.00 (1)					1
Hs	0.27 (1)					1
Yield (%)	48	64	40	22	24	
Total	15	10	5	7	8	38

[a] 64-hr hydrolysis. [b] Not included in total; see text.

yielding two homogeneous peptides, CN-III-A-C-3A and 3B. The amino acid compositions of the chymotryptic peptides are shown in Table V. The remaining pools were not purified since overlap was established without them.

Peptide CN-III-A-C-3A is the C-terminal peptide of CN-III-A, contained in tryptic peptide CN-III-A-T-A, since it contains homoserine. Peptide CN-III-A-C-7 contains the amino acids found in tryptic peptide CN-III-A-T-E, together with four amino acids from the N-terminal end of CN-III-A-T-A. Peptide CN-III-A-T-C cannot be N-terminal in CN-III-A, and must therefore be between CN-III-A-T-B and CN-III-A-T-E, because it has been isolated independently as a tryptic peptide from intact RAE-asparaginase (Greenquist, 1971). The sequence of CN-III-A is shown in Table III.

The N-terminal amino acid was confirmed directly. Peptides CN-III-A and CN-III-A-T-B were both dansylated, hydrolyzed and the N-terminal dansylamino acids identified by thin-layer chromatography. In both cases, only *N*-dansyl-aspartic acid was found.

Peptide CN-III-C. Only a limited amount of information is available on this peptide, which contains 27 amino acids. Three cycles of the Edman degradation were carried out with no significant loss of amino acids. The results of incubation with carboxypeptidases A and B suggest that two valine residues and a leucine are in the immediate vicinity of the C-terminal methionine, but do not allow assignment of an unambiguous sequence.

Peptide CN-I. This peptide is C-terminal in asparaginase and contains over half of the amino acids in each subunit. The sequence of 12 amino acids at the N-terminal end was determined with a Beckman Sequenator (see Methods). The gas chromatographic technique used for identification cannot distinguish leucine from isoleucine, and base hydrolysis of the PTH derivatives, followed by amino acid analysis, was used to show that the sixth residue is leucine. Residue 11 was identified as threonine based on its position in a previously sequenced tryptic peptide (Greenquist, 1971). The N-terminal sequence of CN-I is Asn-Thr-Lys-Val-Asp-Leu-Gly-Arg-Asp-Val-Thr-Lys-.

Discussion

The isolation of the expected number of cyanogen bromide peptides in reasonably good yield, and the sequencing of four

TABLE V: Chymotryptic Peptides from CN-III-A.

Amino Acid	III-A-C-1	III-A-C-2	III-A-C-3A	III-A-C-3B	III-A-C-7
Lys					2.13
His			1.07		
Trp		0.90		+	
Arg					
AECySH					0.60
Asp		3.00	1.57	2.50	3.75
Thr	1.00		3.10		2.00
Ser					
Glu					
Pro					
Gly			1.06		0.88
Ala			1.00		0.65
Cys					
Val		0.94	0.53	1.00	
Met					
Ile			0.55		0.81
Leu	2.02	1.28			
Tyr					
Phe					0.73
Hs			0.29		
Total	3	6	11	5	12
Yield (%)	28	17	36	16	52

TABLE VI: Alignment of Peptides CN-III-B, CN-IV-B, and CN-IV-A.

	1 5 10
	GLX-GLX-THR-[LEU-ASP-ALA-PHE-TYR]-LEU-THR
CN-III-B	Glx- Glx- Thr- [Leu- Asp- Ala- Phe- Tyr]- Leu- Thr-
CN-III-B-T-A	Glx- Glx- Thr- [Leu- Asp- Ala- Phe- Tyr]- Leu- Thr-
	15 20
	VAL-LYS-AECySH- ASP-LYS-PRO-VAL-MET-VAL-GLY-
CN-III-B	Val- Lys- AECySH- Asp- Lys- Pro- Val- Met
CN-III-B-T-A	Val- Lys
CN-III-B-T-B	AECySH- Asp- Lys- Pro- Val- Met
CN-IV-B	Val- Gly-
V-D-3	CMCySH-Asp- Lys- Pro- Val- Met- Val- Gly-
	25
	ALA-MET-ARG-PRO-SER-THR-SER-MET
CN-IV-B	Ala- Met
V-D-3	Ala- Met- Arg
CN-IV-A	Arg- Pro- Ser- Thr- Ser- Met

of them without encountering ambiguities, supports the view that the four subunits in asparaginase are identical (Greenquist and Wriston, 1972). The total number of amino acids in the seven peptides, 319, is in reasonable agreement with that reported for asparaginase itself, 308. Eight amino acids have the same value and six more are within one residue of the values found for the whole enzyme. Only two amino acids (aspartic acid and alanine) vary by more than two residues, and since these two are present in high proportions (45 and 30 residues per subunit, respectively) it is possible that the values found for them in asparaginase are not correct.

In addition, the combined molecular weights of the seven cyanogen bromide peptides, based on their amino acid compositions, is about 34,000, in agreement with the generally accepted value of 33,000 per subunit (Frank *et al.*, 1970). The four peptides that have been sequenced (CN-IV-A, CN-IV-B, CN-III-A, and CN-III-B) together with the N-terminal part of CN-I, represent 75 residues, or about 25% of the residues in a subunit.

Peptides CN-II and CN-I are N and C terminal, respectively. The order of the remaining five cyanogen bromide peptides is not yet known, but it has been possible to align three of them with respect to each other by virtue of overlap with an undecapeptide isolated from a tryptic digest of *S*-carboxymethylated asparaginase (Bullis, 1971). This peptide, partially sequenced by Bullis, contains the C-terminal part of peptide CN-III-B, followed by the tetrapeptide CN-IV-B, permitting alignment of these two cyanogen bromide peptides (Table VI). There are also two methionine residues in the undecapeptide, the second appearing as a Met-Arg sequence at the C-terminal end. This creates another overlap, this time with the hexapeptide CN-IV-A, which is the only cyanogen bromide peptide with N-terminal arginine, and permits the alignment of three cyanogen bromide peptides with respect to one another.

An Arg-Pro linkage would normally be considered to be resistant to cleavage by trypsin. Partial cleavage of an Arg–Hyp bond in collagen has been reported, however (Butler and Ponds, 1971). Also, all nine arginines in asparaginase have been located in cyanogen bromide peptides, and all of the cyanogen bromide peptides have been characterized with respect to their N-terminal residues except for CN-III-C, which is refractory to Edman degradation. The arginine in CN-III-C is known to be near the C-terminal end, however, as a result of carboxypeptidase studies, and cannot be N terminal, and thus the only alignment consistent with the information at hand is that shown in Table VI.

In the original report on the amino acid composition of *E. coli* B asparaginase (Whelan and Wriston, 1968) three tryptophan residues per subunit were found using the spectrophotometric method of Beaven and Holiday (1952). Here, direct determinations were carried out on Merck aparaginase, as well as on each cyanogen bromide peptide, by acid hydrolysis in the presence of thioglycolic acid and only one tryptophan was found, in peptide CN-III-A.

The two *S*-aminoethylcysteines linked by an intrachain disulfide bond in native asparaginase have been located in peptides CN-III-B and CN-III-A. Neither of the *S*-aminoethylcysteine bonds was cleaved by trypsin, even in 21-hr digestions. This is not unexpected after analyzing the sequence around them. In both cases, *S*-aminoethylcysteine is followed by aspartic acid and such bonds have been found to be more resistant to trypsin cleavage. There is a small section of internal homology here. Both *S*-aminoethylcysteine residues are followed by an Asp-Lys- sequence. Since there is only one intrachain disulfide bond, and these particular adjacent amino acid residues easily form electrostatic bonds, it is quite possible that they are important in the conformation of the enzyme molecule.

Rauenbusch *et al.* (1971) treated Bayer asparaginase with cyanogen bromide and then subjected the entire reaction mixture to Edman degradation. Using thin-layer chromatography, they identified asparagine, aspartic acid, glutamic acid, valine, leucine, and arginine as N-terminal amino acids. The amino acid analyses reported for the Bayer material show only five methionines instead of six, as reported for the Merck and Squibb material, and Rauenbusch *et al.* conclude that the six cyanogen bromide peptides they found represent the appropriate number. The six N-terminal amino acids that they found are the same as those for six of the seven cyanogen bromide peptides we have isolated, however, and it seems more likely that the Bayer material also has six methionines, and that the apparent discrepancy is due to the cyanogen bromide peptide (CN-III-C) described here, which is refractory to Edman degradation.

Glossman and Bode (1971) have also studied cyanogen bromide cleavage of Bayer asparaginase, before and after per-

formic acid oxidation, by estimating the molecular weights of peptides by sodium dodecyl sulfate disc electrophoresis. They found six peptides in the performic acid treated material, but did not detect the two small peptides described here (CN-IV-A and CN-IV-B, with six and four amino acids, respectively). They also reported that the cysteic acid residues are present in two of the largest peptides, in disagreement with the sequence data reported here that locates the two half-cystines in peptides CN-III-B and CN-III-A with 18 and 38 amino acids, respectively. There appears to be no ready explanation for this discrepancy although it is possible that there was incomplete cleavage by cyanogen bromide.

The insolubility of several of the cyanogen bromide peptides has been an obstacle in their purification. Bio-Rex 70 has strong sorption properties, and was valuable in two purifications, those of peptides CN-II and CN-III-C. The pK_a of this weakly acidic resin is 5.9, and both purifications were carried out at a pH where the carboxylic acid groups on the resin would not be dissociated, so that nonspecific interactions were presumably important.

Acknowledgment

We are grateful to Dr. Robert E. Handschumacher and Mr. Robert Peterson of the Department of Pharmacology, Yale University Medical School, for their assistance with the dansylation procedure; and to Dr. Peter Mitchell and Mr. Francis Sweeney of the Franklin Institute, Philadelphia, Pa., for the sequence they established for us on an automatic sequenator in their laboratory.

Added in Proof

Using 5-diazo-4-oxo-[5-^{14}C]L-norvaline, an alternate substrate for L-asparaginase (Jackson and Handschumacher, (1970). Peterson *et al.* (1973) have recently isolated from an α-chymotryptic digest of the *E. coli* enzyme a peptide containing the majority of the radioactivity. The sequence they report for this catalytic site decapeptide matches that reported here for peptides CN-IV-B and CN-IV-A (see Table VI).

References

Adamson, R. H., and Fabro, S. (1968), *Cancer Chemother. Rep. 52*, 617.

Africa, B., and Carpenter, F. H. (1966), *Biochem. Biophys. Res. Commun. 24*, 113.

Ambler, R. P. (1967), *Methods Enzymol. 11*, 155.

Arens, A., Rauenbusch, E., Irion, E., Wagner, O., Bauer, K., and Kaufmann, W. (1970), *Hoppe-Seyler's Z. Physiol. Chem. 351*, 197.

Beaven, G. H., and Holiday, E. R. (1952), *Advan. Protein Chem. 7*, 375.

Bullis, W. F. (1971), Masters Thesis, University of Delaware.

Butler, W. T., and Ponds, S. L. (1971), *Biochemistry 10*, 2076.

Capizzi, R. L., Bertino, J. R., and Handschumacher, R. E. (1970), *Annu. Rev. Med. 21*, 433.

Cooney, D. A., and Handschumacher, R. E. (1970), *Annu. Rev. Pharmacol. 10*, 421.

Edmundson, A. B. (1967), *Methods Enzymol. 11*, 370.

Elzinga, M., Lai, C. Y., and Hirs, C. H. W. (1968), *Arch. Biochem. Biophys. 123*, 353.

Fraenkel-Conrat, H., Harris, J. L., and Levy, A. L. (1955), *Methods Biochem. Anal. 2*, 339.

Frank, B. H., Pekar, A. H., Veros, A. J., and Ho, P. P. K. (1970), *J. Biol. Chem. 245*, 3716.

Glossman, H., and Bode, W. (1971), *Hoppe-Seyler's Z. Physiol. Chem. 352*, 132.

Gray, W. R. (1972), *Methods Enzymol. 25*, 121.

Greenquist, A. C. (1971), Ph.D. Thesis, University of Delaware.

Greenquist, A. C., and Wriston, J. C. (1972), *Arch. Biochem. Biophys. 152*, 280.

Hartley, B. S. (1970), *Biochem. J. 119*, 805.

Hirs, C. H. W. (1955), *Methods Enzymol. 1*, 113.

Hirs, C. H. W. (1967), *Methods Enzymol. 11*, 325.

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

Humbel, R. E., Derron, R., and Neumann, P. (1968), *Biochemistry 7*, 621.

Jackson, R. C., and Handschumacher, R. E. (1970), *Biochemistry 9*, 3585.

Matsubara, H , and Sasaki, K. (1969), *Biochem. Biophys. Res. Commun. 35*, 177.

Mills, J. B., Howard, S. C., Scapa, S., and Wilhelm, A. E. (1970), *J. Biol. Chem. 245*, 3407.

Moore, S., and Stein, W. H. (1954), *J. Biol. Chem. 211*, 907.

Peterson, R. G., Richards, F. F., and Handschumacher, R. E. (1973), 166th National Meeting of the American Chemical Society, Chicago, Ill., Sept, Abstract BIOL, B-6.

Rauenbusch, E., Irion, E., and Arens, A. (1971), *Colloq. Int. Centre Nat. Rech. Sci. No. 197.*

Schroeder, W. A., Jones, R. T., Cormick, J., and McCall, K. (1962), *Anal. Chem. 34*, 1570.

Schultz, J. (1967), *Methods Enzymol. II*, 255.

Spackman, D. H., Stein, W. H., and Moore, S. (1958), *Anal. Chem. 30*, 1900.

Steers, E., Jr., Craven, G. R., and Anfinsen, C. B. (1965), *J. Biol. Chem. 240*, 2478.

Whelan, H. A., and Wriston, J. C. (1969), *Biochemistry 8*, 2386.

Woods, K. R., and Wang, K. T. (1967), *Biochim. Biophys. Acta 133*, 369.

Wriston, J. C. (1971), *Enzymes, 3rd Ed., 4*, 101.

Wriston, J. C., Greenquist, A. C., Wyatt, J., and Gumprecht, J. G. (1972), *Fed. Eur. Biochem. Soc. Meeting, 8th.*