Angeletti, R. H., and Bradshaw, R. A. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2417.
Angeletti, R. H., Bradshaw, R. A., and Wade, R. D. (1971), *Biochemistry 10*, 463.
Angeletti, R. H., Hermodsen, M., and Bradshaw, R. A. (1973), *Biochemistry 12*, 100.
Bocchini, V., and Angeletti, P. U. (1969), *Proc. Nat. Acad. Sci. U. S. 64*, 787.
Bradshaw, R. A., Babin, D. R., Nomoto, M., Srinivasin, N. G., Ericsson, L. H., Walsh, K. A., and Neurath, H. (1969a), *Biochemistry 8*, 3859.
Bradshaw, R. A., Garner, W. H., and Gurd, F. R. N. (1969b), *J. Biol. Chem. 244*, 2149.
Bray, G. A. (1969), *Anal. Biochem. 1*, 279.
Edman, P. (1956), *Acta. Chem. Scand. 10*, 761.
Gray, W. R., and Smith, J. F. (1970), *Anal. Biochem. 33*, 36.
Hartley, B. S. (1970), *Biochem. J. 119*, 805.
Hill, R. L. (1965), *Advan. Protein Chem. 20*, 37.
Hill, R. L., and Delaney, R. (1967), *Methods Enzymol. 11*, 339.
Konigsberg, W. J., and Hill, R. J. (1962), *J. Biol. Chem. 237*, 2547.
Levi-Montalcini, R., and Angeletti, P. U. (1968), *Physiol. Rev. 48*, 534.
Levi-Montalcini, R., Angeletti, R. H., and Angeletti, P. U., (1973), *in* The Structure and Function of Nervous Tissue, Bourne, G. H., Ed., Vol. 4, New York, N. Y., Academic Press (in press).
Moore, S., Spackman, D. H., and Stein, W. H. (1968), *Anal. Chem., 30*, 1185.
Raftery, M. A., and Cole, R. D. (1966), *J. Biol. Chem. 241*, 3457.
Waterfield, M., and Haber, E. (1970), *Biochemistry 9*, 832.
Woods, K. R., and Wange, K. T. (1967), *Biochim. Biophys. Acta 133*, 369.

Amino Acid Sequences of Mouse 2.5S Nerve Growth Factor. II. Isolation and Characterization of the Thermolytic and Peptic Peptides and the Complete Covalent Structure†

Ruth Hogue Angeletti,‡ Mark A. Hermodson, and Ralph A. Bradshaw*,§

ABSTRACT: The isolation and characterization of the thermolytic peptides of *S*-carboxymethyl nerve growth factor and the peptic peptides of native nerve growth factor are described. The peptides derived from the latter digestion were also used for the determination of the pairing of the half-cystinyl residues. In addition, the two fragments produced by cyanogen bromide cleavage were isolated and used, together with the analyses of the protein Sequencer, to establish the structure of the amino-terminal portion of the protein. It was determined that some preparations of nerve growth factor contained polypeptide chains that were eight residues shorter than the primary subunit. These chains have been designated A (long) and B (short) and apparently arise from limited proteolysis. Carboxypeptidase B digestion of *S*-carboxymethyl nerve growth factor defined the carboxyl-terminal portion of the molecule. These data are combined with those derived from the tryptic and chymotryptic peptides described in the preceding communication to construct an internally consistent primary and secondary structure of nerve growth factor. In addition, each of the side chains of Glx and Asx residues was identified as to its acidic or neutral character. Thus, mouse nerve growth factor consists, in the unaltered state, of a dimer of two polypeptide chains containing 118 amino acids, each with a molecular weight of 13,259. The disulfide pairs are formed by Cys-15 and Cys-80, Cys-58 and Cys-108, and Cys-68 and Cys-110.

The preceding communication (Angeletti *et al.*, 1973) describes the isolation and characterization of the soluble peptides from tryptic and chymotryptic digests of *S*-carboxymethyl nerve growth factor (NGF)[1] and the tryptic digest of *S*-aminoethyl-NGF. Because these data in their entirety did not provide sufficient information, especially with respect to the amino terminus, to construct a tentative sequence for NGF, additional peptides from digestion of NGF with thermolysin and pepsin were prepared and characterized. Furthermore, as it was known that the subunit of NGF is a single polypeptide chain containing three internal disulfide bridges and no free sulfhydryl groups (Bocchini, 1970; Angeletti *et al.*, 1971), the peptic digestion was performed on native NGF so as to allow the determination of the disulfide pairing. In addition, experiments were also conducted using the Beckman protein Sequencer, on the one hand, and carboxypeptidase digestion, on the other, to elucidate the nature of the amino and carboxy termini, respectively. The Sequencer

† From the Department of Biology, Washington University, St. Louis, Missouri 63130 (R. H. A.), from the Department of Biochemistry, University of Washington, Seattle, Washington 98105 (M. A. H.), and from the Department of Biological Chemistry, Washington University School of Medicine, St. Louis, Missouri 63110 (R. A. B.). *Received August 14, 1972.* This work was supported by research grants from the National Institutes of Health (AM-13362, NS-03777, and GM-15731).

* To whom to address correspondence at the Department of Biological Chemistry, Washington University School of Medicine.

‡ Present address: Laboratorio di Biologia Cellulare, CNR, Rome, Italy.

§ Research Career Development awardee of the National Institutes of Health (AM-23698).

[1] Abbreviations used are: NGF, nerve growth factor; CM-, carboxymethyl: Th-, thermolytic peptides; P-, peptic peptides; SE-, sulfoethyl; PTh-, peptides from the thermolytic digest of a peptic fragment; PC-, peptic acid insoluble peptides; CNBr-, cyanogen bromide fragments; Pth, phenylthiohydantoin.

analyses indicated some heterogeneity of the type expected from proteolytic cleavage at the amino terminus. To establish the correctness of the alignment of the two chains, cyanogen bromide cleavage was performed on the S-carboxymethylated derivative and the two fragments formed were isolated and characterized. The sum of these results, which are described below, allowed the elucidation of the complete covalent structure of mouse 2.5S NGF.

Experimental Procedure

Materials. Nerve growth factor and *S*-carboxymethyl nerve growth factor were prepared as described previously (Angeletti *et al.*, 1973). Pepsin and carboxypeptidases A and B were obtained from Worthington Biochemicals. Thermolysin was obtained from Daiwa Kasei K. K., Osaka, Japan. Cyanogen bromide was purchased from Eastman. SE-Sephadex was obtained from Pharmacia Fine Chemicals.

Methods. THERMOLYTIC DIGESTION. SCM-NGF (4.00 μmol), dissolved in 10 ml of H_2O, was digested for 22 hr at pH 8.0, 37° with thermolysin at a final concentration to substrate of 0.5%. The reaction was monitored in a Radiometer pH-Stat and terminated by acidification to pH 2 with 6 N HCl.

PEPTIC DIGESTION. Native nerve growth factor (4.40 μmol) was dissolved as a 1% solution in 5% formic acid to which 2% pepsin (w/w) had been added. Digestion was allowed to proceed for 18 hr, at which time the suspension was centrifuged to remove the insoluble material. The thermolytic digestion of peptide P-XVI was performed as described above, but with the pH of the reaction mixture maintained at 7, in order to avoid disulfide interchange.

CYANOGEN BROMIDE CLEAVAGE. NGF (0.75 μmol), dissolved in 70% formic acid (3.0 ml), was reacted with a 30-fold molar excess of cyanogen bromide for 24 hr at 20° in the dark. The reaction was stopped by diluting with 15 volumes of distilled water and followed by lyophilization.

CHARACTERIZATION OF THE AMINO TERMINUS OF *S*-CARBOXYMETHYL-NGF BY THE PROTEIN SEQUENCER. Approximately 300 nmol of SCM-NGF was suspended in 0.5 ml of 50% acetic acid and dried in the cup of a Beckman Sequencer. Twenty-four cycles of the Edman degradation were carried out following the general technique of Edman and Begg (1967) as adapted by Hermodson *et al.*[2] Dithioerythritol (0.2 mM) was included in the chlorobutane extractant as recommended by Hermodson *et al.* (1970) to improve the Pth-amino acid recoveries. The Pth-amino acids released were identified by gas–liquid chromatography by a method similar to that described by Pisano and Bronzert (1969).

CARBOXYPEPTIDASE DIGESTION OF SCM-NGF. SCM-NGF (2 mg) was suspended in 0.5 ml of Tris-HCl (0.01 M pH 8.5) and digested with carboxypeptidase B (1%) for 15 min. The sample was diluted with 0.2 M sodium citrate buffer (pH 2.2) and applied directly to the amino acid analyzer after centrifugation at 5000*g*.

PEPTIDE PURIFICATION. The procedures for isolating and characterizing peptides were essentially those described previously (Angeletti *et al.*, 1973), except for the fractionation of the peptic peptides, which was performed on SE-Sephadex at 55° as described by Walsh *et al.* (1970).

DETECTION OF DISULFIDE PEPTIDES AND IDENTIFICATION OF DISULFIDE PEPTIDE PAIRS. The column separation of the peptic digest on SE-Sephadex was monitored for cystine by testing aliquots removed from every other fraction with the cyanide–nitroprusside spray reaction (Toennies and Kolb, 1951). Semiquantitative values for cystine were obtained from acid hydrolysates of the corresponding peptides. Purified fractions known to contain unique disulfide peptides were subjected to mild performic acid oxidation as described by Hirs (1969), and the resultant mixture of oxidized peptides was separated and identified by amino acid composition.

[2] M. A. Hermodson, L. H., Ericsson, K. Titani, H. Neurath, and K. A. Walsh, manuscript in preparation.

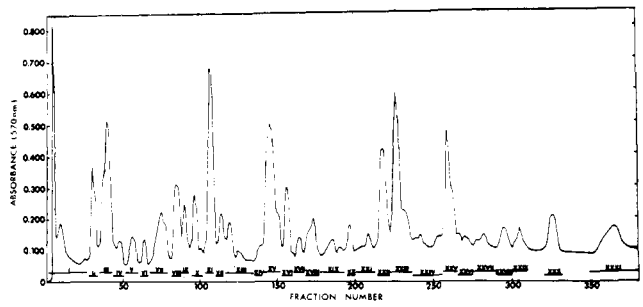

FIGURE 1: Elution profile of the thermolytic peptides of 4.00 μmol of *S*-carboxymethyl-NGF on a 0.9 × 20 cm column of Dowex 50-X8 at 55°. The column was developed at 30 ml/hr with a continuous gradient of pyridine acetate as described in the text, and monitored automatically by ninhydrin analysis after alkaline hydrolysis. Fractions of 3.0 ml were collected.

Results

Isolation of Soluble Thermolytic Peptides. Following digestion of SCM-NGF with thermolysin for 22 hr and acidification to pH 2, no insoluble material remained. Therefore, the entire digestion mixture was applied to a Dowex 50-X8 column and eluted with pyridine acetate buffers (Figure 1). Fractions were made as indicated by the bars and Roman numerals, and each was tested for heterogeneity by high-voltage paper electrophoresis at pH 3.75. Subsequent purification entailed fractionation on Dowex 1-X2 and Dowex 50-X2 columns and preparative high-voltage paper electrophoresis. Table I summarizes the amino acid composition yield, and method of purification of each of the 60 or more pure peptides obtained from this digest. The position of each of the purified peptides in the final sequence is listed in the last line. The characterization of those peptides which supplied further structural information is considered in detail subsequently.

FRACTION I. Subfractionation on Dowex 1-X2 yielded 2 peptides, in addition to the cysteic acid column marker. The first, Th-I-1, is a tripeptide which was not further characterized. The second, Th-I-2, a cysteine-containing peptide, was subjected to two rounds of Edman degradation (Table II).

FRACTION II. This fraction was separated by chromatography on a column of Dowex 1-X2 into two components, peptides Th-II-1 and Th-II-2. The former was entirely structured by three turns of Edman degradation. Electrophoretic analysis demonstrated that the aspartyl residue is present in the amide form. Peptide Th-II-2 was subjected to four turns of Edman degradation. The sequence data are summarized in Table II.

FRACTION III. From the purification of this fraction on Dowex 1-X2, 4 peptides were isolated (Table II). The first, Th-III-1, is a tetrapeptide which was fully characterized by 3 turns of subtractive Edman degradation. It is electrophoretically neutral, and therefore, both aspartic acid residues

TABLE I: Amino Acid Composition of the Thermolytic Peptides of *S*-Carboxymethyl Nerve Growth Factor.[a]

Amino Acid	Th-I-1	Th-I-2	Th-II-1	Th-II-2	Th-III-1	Th-III-2	Th-III-3	Th-III-4	Th-IV-1	Th-V-1
Lysine										
Histidine										
Arginine										
S-Carboxymethylcysteine		0.81 (1)		1.00 (1)				1.01 (1)		
Aspartic acid	1.17 (1)	1.09 (1)	1.00 (1)	1.27 (1)	2.16 (2)		1.00 (1)	1.18 (1)	1.00 (1)	
Threonine		1.06 (1)					2.18 (2)			1.10 (1)
Serine	0.80 (1)		1.02 (1)	2.30 (2)	1.00 (1)			0.98 (1)	0.94 (1)	
Glutamic acid						1.21 (1)				
Proline			1.27 (1)						0.98 (1)	
Glycine										
Alanine		1.00 (1)	1.00 (1)			1.18 (1)	1.28 (1)			
Valine				1.15 (1)		1.00 (1)		1.09 (1)		1.89 (2)
Methionine										
Isoleucine		0.91 (1)			1.00 (1)					
Leucine						0.95 (1)				
Tyrosine										
Phenylalanine				0.89 (1)						
Tryptophan	(1)[b]									
Total	3	5	4	6	4	4	4	4	3	3
Per cent yield	5	18	10	14	16	17	28	17	22	9
Purification procedure	DX1[c]	DX1	DX1	DX1	DX1	DX1, DX50	DX1	DX1	DX1	DX1
Residue no.	76–78	104–108	60–63	12–17	44–47	38–41	27–30	14–17	61–63	36–38

Amino Acid	Th-XII-2	Th-XIII-1	Th-XIV	Th-XV-1	Th-XV-2	Th-XV-3	Th-XV-4	Th-XV-5	Th-XVI	Th-XVII
Lysine			1.00 (1)						0.93 (1)	
Histidine		1.12 (1)		1.00 (1)	1.28 (1)	1.00 (1)				
Arginine							1.05 (1)	1.16 (1)		
S-Carboxymethylcysteine		0.80 (1)					0.95 (1)	0.93 (1)		
Aspartic acid			1.04 (1)					1.26 (1)	1.04 (1)	
Threonine		4.00 (4)	1.77 (2)	1.00 (1)	4.00 (4)			1.00 (1)	1.02 (1)	
Serine				1.79 (2)			0.87 (1)			
Glutamic acid						1.00 (1)	1.05 (1)			
Proline				(1)						
Glycine			1.09 (1)			1.12 (1)	2.02 (2)		0.96 (1)	
Alanine								1.23 (1)		
Valine	1.00 (1)		0.83 (1)			1.00 (1)	0.94 (1)		0.75 (1)	
Methionine						0.74 (1)				
Isoleucine								2.23 (2)		
Leucine										
Tyrosine		1.18 (1)								0.86 (1)
Phenylalanine	1.15 (1)					0.94 (1)				1.00 (1)
Tryptophan										
Total	2	7	6	5	5	6	7	7	5	2
Per cent yield	22	3	11	8	24	4	35	8	27	25
Purification procedure	DX1	DX1		DX1	DX1	DX1	DX1	DX1		
Residue no.	6–7 or 48–49	79–85	22–27	1–5	81–85	6–11	64–70	102–108	22–26	52–53

Th-VI	Th-VII-1a	Th-VII-1b	Th-VII-2	Th-VIII-1	Th-VIII-2	Th-IX	Th-X-1	Th-XI-1	Th-XI-2	TH-XI-3
									1.15 (1)	
							1.33 (1)			
			1.32 (1)				0.76 (1)			0.94 (1)
		1.01 (1)					1.11 (1)			
									0.99 (1)	
	0.97 (1)								1.92 (2)	
					1.00 (1)		1.78 (2)	0.96 (1)		
1.00 (1)							1.18 (1)		2.00 (2)	
							0.85 (1)			
							2.20 (2)			
1.13 (1)						(2)[e]				
	1.00 (1)	0.98 (1)	1.00 (1)[d]		1.37 (1)		0.96 (1)			
									1.02 (1)	
1.00 (1)										1.00 (1)
				1.00 (1)				1.06 (1)		
3	2	2	2	1	2	2	10	2	7	2
13	40	40	27	22	25	51	5	17	33	8
	DX1, PC	DX1, PC	DX1	DX1	DX1	DX1	DX1	DX1	DX1	DX1
39–41	36–37	42–43	109–110	NA	18–19	97–98	61–70	12–13	90–96	79–80

Th-XVIII-1	Th-XVIII-2	Th-XIX-1	Th-XX-1	Th-XX-2	Th-XXII-1	Th-XXII-2	Th-XXII-3	Th-XXIII-1	Th-XXIII-2	Th-XXIII-3
						0.97 (1)	1.44 (1)		1.51 (1)	1.12 (1)
1.00 (1)	0.81 (1)						1.37 (1)			
		1.25 (1)			1.00 (1)			1.00 (1)	1.24 (1)	1.11 (1)
									0.99 (1)	0.87 (1)
		1.00 (1)					0.89 (1)			
2.74 (3)		0.89 (1)	(2)					0.92 (1)	0.98 (1)	1.00 (1)
							1.14 (1)			
	1.00 (1)				1.10 (1)				1.04 (1)	1.10 (1)
	1.23 (1)									
						1.00 (1)		1.02 (1)	0.89 (1)	1.00 (1)
				1.00 (1)	0.93 (1)	0.80 (1)				
	0.57 (1)									
		1.82 (2)					1.36 (1)			
	0.63 (1)				1.00 (1)				1.10 (1)	2.00 (2)
				(1)[b]						
4	5	5	2	2	4	3	5	3	7	8
5	10	9	8	24	10	33	31	28	28	7
DX1	DX1	DX1	DX1	DX1	DX1	DX1	DX1	DX1	DX1	DX1
82–85	7–11	102–106	26–27, 81–82 or 82–83	20–21	48–51	87–89	71–75	116–118	54–60	53–60

TABLE I (*Continued*)

Amino Acid	Th-XXIV-1	Th-XXIV-2	Th-XXV-1	Th-XXV-2	Th-XXVII	Th-XXVIII	Th-XXIX-1	Th-XXX	Th-XXXI
Lysine				2.00 (2)		1.25 (1)		1.22 (1)	
Histidine									
Arginine	1.16 (1)	1.20 (1)	1.24 (1)		1.00 (1)	1.22 (1)	1.00 (1)	1.26 (1)	(1)[b]
S-Carboxymethylcysteine									
Aspartic acid									
Threonine			1.00 (1)						
Serine						0.95 (1)		0.92 (1)	
Glutamic acid	1.00 (1)	1.00 (1)		1.14 (1)					
Proline									
Glycine				0.96 (1)					
Alanine									
Valine						1.14 (1)			
Methionine									
Isoleucine				0.94 (1)	0.70 (1)				
Leucine						1.00 (1)		1.00 (1)	
Tyrosine		1.00 (1)							
Phenylalanine	0.97 (1)	1.00 (1)					0.85 (1)		
Tryptophan							(1)[b]		(1)[b]
Total	3	4	2	5	2	5	3	4	2
Per cent yield	4	4	12	39	32	17	8	25	47
Purification procedure	DX1	DX1	DX50	DX50			HVE		
Residue no.	49–51	49–52	117–118	31–35	102–103	111–115	99–101	112–115	99–100

[a] Values are given in residues/mol. The assumed integral values are given in parentheses. See Angeletti *et al.* (1973) for further details. [b] Assumed to be a residue (Angeletti *et al.*, 1973). [c] Abbreviations used are: DX1, Dowex-1-X2; DX50, Dowex 50-X2; PC, paper chromatography; HVE, preparative high-voltage electrophoresis; NA, no assignment. [d] 72-hr hydrolysis. [e] See text for details.

are present in the amide form. The second fraction was purified on a column of Dowex 50-X2 to remove minor contaminants, yielding a pure peptide, Th-III-2. The positions of the N-terminal two residues were determined by Edman degradation. Electrophoretic analysis established that the glutamyl residue is present in the acid form. The third fraction contained a pure tetrapeptide, Th-III-3, which was not further characterized. This peptide was electrically neutral, confirming the presence of asparagine. The last fraction also contained a pure tetrapeptide, Th-III-4, which was fully characterized by a combination of Edman degradation and digestion with carboxypeptidase A.

FRACTION IV. This fraction was slightly contaminated as judged by paper electrophoresis, and was further purified by Dowex 1-X2 chromatography. A single tripeptide, Th-IV-1, was recovered. Its N-terminal residue was confirmed by the dansyl-Edman technique (Table III).

FRACTION V. After passage on Dowex 1-X2, this fraction yielded a pure tripeptide, Th-V-1, whose sequence was determined by two rounds of Edman degradation.

FRACTION VI. This fraction contained a tripeptide, Th-VI, which required no further purification. Its sequence was determined by two rounds of Edman degradation, and electrophoretic analysis confirmed the presence of glutamic acid in the acid form.

FRACTION VII. This fraction was separated into 2 fractions on Dowex 1-X2 chromatography. The first fraction appeared to be homogeneous by both electrophoretic and chromatographic criteria. However, Edman degradation indicated it to be impure. Further separation on preparative paper chromatography yielded two dipeptides, Th-VII-1a and Th-VII-1b. Electrophoretic analysis showed the latter peptide to be neutral, confirming that the aspartyl residue is present as an amide.

The second fraction contained a pure dipeptide, Th-VII-2, whose sequence was confirmed by the determination of the amino-terminal residue.

FRACTION VIII. This fraction was a mixture of 2 peptides which were readily separated by chromatography on Dowex 1-X2. The first, Th-VIII-1, was free phenylalanine. Amino acid analysis of an unhydrolyzed sample confirmed that the structure was not Phe-Phe. One turn of Edman degradation was performed in order to characterize the dipeptide, Th-VIII-2.

TABLE II: Characterization of Selected Thermolytic Peptides from Fractions I, II, and III.

Fraction I:
Peptide Th-I-2:
Sequence: Ile-Asx-(Thr,Ala,Cys)
Fraction II:
Peptide Th-II-1:
Sequence: Ala-Ser-Asn-Pro
High-voltage electrophoresis, pH 6.5: Neutral
Peptide Th-II-2:
Sequence: Phe-Ser-Val-Cys-(Asx,Ser)
Fraction III:
Peptide Th-III-1:
Sequence: Ile-Asn-Asn-Ser
High-voltage electrophoresis, pH 6.5: Neutral
Peptide Th-III-2:
Sequence: Val-Leu-(Ala,Glu)
High-voltage electrophoresis, pH 6.5: Acidic
Peptide Th-III-3:
Sequence: (Thr,Ala,Thr,Asn)
High-voltage electrophoresis, pH 6.5: Neutral
Peptide Th-III-4:
Sequence: Val-Cys-Asp-Ser
Carboxypeptidase A: 4 hr, Ser, 1.00; Asp, 0.43

TABLE III: Characterization of Thermolytic Peptides from Fractions IV, V, VI, and VII.

Fraction IV:
Peptide Th-IV-1:
Sequence: Ser-(Asx,Pro)
Fraction V:
Peptide Th-V-1:
Sequence: Val-Thr-Val
Fraction VI:
Peptide Th-VI:
Sequence: Leu-Ala-Glu
High-voltage electrophoresis, pH 6.5: Acidic
Fraction VII:
Peptide Th-VII-1a:
Sequence: Val-Thr
Peptide Th-VII-1b:
Sequence: Val-Asn
High-voltage electrophoresis, pH 6.5: Neutral
Peptide Th-VII-2:
Sequence: Val-Cys

FRACTION IX. No further purification was required of this fraction which consisted of a single dipeptide, Th-IX. Amino acid analysis of both hydrolyzed and unhydrolyzed samples demonstrated that this fraction was Ala-Ala and not free alanine.

FRACTION X. After subfractionation on Dowex 1-X2, only one useful peptide was isolated from this fraction. Th-X-1 is a decapeptide containing both cysteine and proline. The positions of the first 7 residues were determined by Edman degradation (Table IV).

FRACTION XI. Three peptides were recovered from this fraction upon Dowex 1-X2 fractionation. The first, Th-XI-1, was a dipeptide which was not further characterized. The second pure peptide isolated contained 7 residues, 5 of which were placed by Edman degradation. This pepide was acidic when analyzed electrophoretically, confirming the amide assignments as shown in Table IV. The sequence of the third peptide Th-XI-3, a dipeptide, was determined by the dansyl-Edman procedure.

FRACTION XII. This pool was separated into two peptides using Dowex 1-X2 chromatography. Neither the 7-residue peptide, Th-XII-1, which was identical with Th-XI-2, nor the dipeptide, Th-XII-2, was characterized further.

FRACTION XIII. Purification of this fraction on Dowex 1-X2 column produced a single 7-residue peptide Th-XIII-1, which was not characterized further.

FRACTION XIV. Without further purification, this fraction yielded a pure 6-residue peptide, Th-XIV. One turn of Edman degradation demonstrated that the N-terminal residue was valine, and carboxypeptidase A digestion established the position of the two carboxyl-terminal residues (Table V).

FRACTION XV. Since electrophoretic analysis showed that this fraction contained a large mixture of peptides, they were subfractionated on Dowex 1-X2. Five pure peptides were isolated. The first, Th-XV-1, was subjected to Edman degradation, which yielded the complete structure of the peptide. The second, a pentapeptide Th-XV-2, contained only threonyl and histidyl residues, and was not further analyzed. Th-XV-3 was a 6-residue peptide containing methionine whose structure was unequivocally determined by 5 turns of Edman degradation. The fourth, Th-XV-4, was a 7-residue peptide containing both cysteine and arginine, and was fully characterized by Edman degradation. Electrophoresis at pH 6.5 showed it to be acidic, indicating the glutamyl residue to be

TABLE IV: Characterization of Selected Thermolytic Peptides from Fractions VIII, IX, X, and XI.

Fraction VIII:
Peptide Th-VIII-1:
Sequence: Phe
Peptide Th-VIII-2:
Sequence: Val-Ser
Fraction IX:
Peptide Th-IX:
Sequence: Ala-Ala
Fraction X:
Peptide Th-X-1:
Sequence: Ser-Asx-Pro-Val-Glx-Ser-Gly-(Cys,Arg,Gly)
Fraction XI:
Peptide Th-XI-2:
Sequence: Leu-Thr-Thr-Asx-Glx-(Lys,Glx)
High-voltage electrophoresis, pH 6.5: Acidic
Peptide Th-XI-3:
Sequence: Tyr-Cys

TABLE V: Characterization of Selected Thermolytic Peptides from Fractions XIV, XV, and XVI.

Fraction XIV:
Peptide Th-XIV:
Sequence: Val-(Gly,Asp,Lys)-Thr-Thr
Carboxypeptidase A: 1 hr, Thr, 1.67
Fraction XV:
Peptide Th-XV-1:
Sequence: Ser-Ser-Thr-His-Pro
Peptide Th-XV-3:
Sequence: Val-Phe-His-Met-Gly-Glx
Peptide Th-XV-4:
Sequence: Val-Glu-Ser-Gly-Cys-Arg-Gly
High-voltage electrophoresis, pH 6.5: Acidic
Fraction XVI:
Peptide Th-XVI:
Sequence: Val-(Gly,Asp,Lys,Thr)

in the acid form. Peptide Th-XV-5 contained redundant information and was therefore not further analyzed.

FRACTION XVI. This pool consisted of only one pure pentapeptide, Th-XVI. Since this peptide appeared to be one residue shorter than Th-XIV, only the amino-terminal residue was ascertained in order to position it in the final structure (Table V).

FRACTION XVII. This fraction contained a pure dipeptide, Th-XVII, whose sequence was determined by one round of subtractive Edman degradation (Table VI).

FRACTION XVIII. Purification of this fraction on Dowex 1-X2 yielded two pure peptides; one, a tetrapeptide, Th-XVIII-1, and the other, a pentapeptide, Th-XVIII-2. Since Th-XVIII-1 was only one residue shorter than Th-XV-2, and Th-XVIII-2 was one residue less than Th-XV-3, these peptides were not further characterized.

TABLE VI: Characterization of Selected Thermolytic Peptides from Fractions XVII, XX, and XXII.

Fraction XVII:
Peptide Th-XVII:
Sequence: Tyr-Phe
Fraction XX:
Peptide Th-XX-1:
Sequence: Thr-Thr
Peptide Th-XX-2:
Sequence: Val-Trp
Fraction XXII:
Peptide Th-XXII-1:
Sequence: Val-Phe-Arg-Gln
Carboxypeptidase A: 1 hr; Gln, 0.65
Peptide Th-XXII-2:
Sequence: Val-Lys-Ala
Peptide Th-XXII-3:
Sequence: Ile-Asp-Ser-Lys-His

TABLE VII: Characterization of Selected Thermolytic Peptides from Fractions XXIII, XXIV, XXV, XXVIII, and XXIX.

Fraction XXIII:
Peptide Th-XXIII-2:
Sequence: Phe-Glx-Thr-Lys-Cys-Arg-Ala
Peptide Th-XXIII-3:
Sequence: Phe-Phe-Glx-Thr-Lys-Cys-(Arg,Ala)
Fraction XXIV:
Peptide Th-XXIV-1:
Sequence: Phe-Arg-Gln
Carboxypeptidase A: 1 hr, Gln, 0.72
Peptide Th-XXIV-2:
Sequence: (Phe,Arg)-Gln-Tyr
Carboxypeptidase A: 1 hr, Tyr, 0.94; Gln, 0.68
Fraction XXV:
Peptide Th-XXV-2:
Sequence: Ile-Lys-Gly-Lys-Glx
Fraction XXVIII:
Peptide Th-XXVIII:
Sequence: Val-Leu-Ser-Arg-Lys
Fraction XXIX:
Peptide Th-XXIX-1:
Sequence: (Trp,Arg)-Phe
Carboxypeptidase A: 1 hr, Phe, 0.96

FRACTION XIX. One useful peptide Th-XIX-1, was isolated from this fraction after purification on Dowex 1-X2. This 5-residue peptide was not subjected to sequence analysis.

FRACTION XX. Subfractionation of this fraction on Dowex 1-X2 provided two usable peptides; Th-XX-1, as determined from amino acid analysis of both hydrolyzed and unhydrolyzed samples; the second, Th-XX-2, a tryptophan-containing dipeptide whose amino-terminal residue was positioned by the dansyl-Edman procedure (Table VI).

FRACTION XXI. This fraction appeared to be a mixture of several peptides in such low yield that no useful peptides were ever isolated from it.

FRACTION XXII. This fraction was primarily composed of 3 peptides which were separated from one another by chromatography on Dowex 1-X2. The first was a tetrapeptide, Th-XVII-1, whose complete sequence was unequivocally characterized by two turns of Edman degradation and carboxypeptidase A digestion. Th-XXII-2, a tripeptide, was fully sequenced by two turns of Edman degradation. The third peptide, Th-XXII-3, which contained one lysyl and one histidyl residue, was completely structured by 4 successful rounds of Edman degradation (Table VI).

FRACTION XXIII. Separation on a Dowex 1-X2 column effectively purified the 3 peptides present in this fraction. The first was a tripeptide, Th-XXIII-1, which was not subjected to sequence analysis. The other 2 peptides had identical compositions, except that peptide Th-XXIII-2 was shorter than Th-XXIII-3 by 1 phenylalanyl residue. Six rounds of Edman degradation indicated that the additional phenylalanyl residue of peptide Th-XXIII-3 was at the amino terminus (Table VII).

FRACTION XXIV. This fraction was composed of 2 related peptides, which were separated after chromatography on Dowex 1-X2. One turn of Edman degradation and carboxy-

peptidase A digestion confirmed the structure of the tripeptide, Th-XXIV-1. Th-XXIV-2, a tetrapeptide, contained 1 additional tyrosyl residue at the carboxy terminus (Table VII).

FRACTION XXV. Subfractionation of this fraction on Dowex 50-X2 yielded 2 peptides. The first, Th-XXV-1, was a dipeptide which was not further characterized. The second peptide, Th-XXV-2, contained 2 lysyl residues. The combined use of subtractive and dansyl-Edman procedures provided the complete structure of this peptide (Table VII).

FRACTION XXVI. Electrophoretic and amino acid analysis revealed that this fraction contained a mixture of many minor yield peptides. Further subfractionation was unsuccessful in isolating any useful peptides.

FRACTION XXVII. This fraction was comprised of a single dipeptide, Th-XXVII, which was not further analyzed.

FRACTION XXVIII. No purification was required of this fraction, which contained a pure pentapeptide, Th-XXVIII. Four successful turns of Edman degradation provided its complete sequence (Table VII).

FRACTION XXIX. Preparative high-voltage electrophoresis at pH 3.75 permitted the purification of the tripeptide, Th-XXIX-1, from contaminating material present in low yield. Carboxypeptidase A digestion revealed phenylalanine to be the carboxyl-terminal residue (Table VII).

FRACTION XXX. This fraction was found to contain a pure tetrapeptide, Th-XXX. No sequence analysis was performed.

FRACTION XXXI. The last fraction eluted from the column consisted of a pure dipeptide in high yield, Th-XXXI, which was not further characterized.

Isolation of the Soluble Peptic Peptides. The profile for the separation of the soluble peptic peptides of native NGF from a 2.0 × 35 cm SE-Sephadex column at 55° is shown in Figure 2. The 15 fractions are indicated by the bars and Roman numerals, and the amino acid compositions of the peptides purified from them are summarized in Table VIII which also lists the per cent yield, additional purification procedures and the residue numbers in the final sequence.

Characterization of the Soluble Peptic Peptides. FRACTION I. This fraction contained the cysteic acid marker, with no trace of peptidic material.

FRACTION II. Only free phenylalanine, P-II, was recovered from this pool. Unhydrolyzed samples confirmed that phenylalanine was present as the free amino acid.

FRACTION III. No peptide material was found in this fraction.

FRACTION IV. A pure pentapeptide, P-IV, was in this fraction, which required no further purification. Since this peptide contained redundant information, no additional characterization was carried out.

FRACTION V. This fraction was subfractionated on dowex 1-X2 to yield two peptides. The first peptide, P-V-1, was not further characterized. P-V-2, a tetrapeptide, was fully characterized by three rounds of Edman degradation (Table IX).

FRACTION VI. This fraction contained two peptides which were separated on a column of Dowex 1-X2. Because these peptides originate from a region of the molecule already well characterized, no further analyses were performed. It should be noted that P-VI-1 is related to P-VI-2, being only two residues longer.

FRACTION VII. No further purification was required of this fraction which contained a single 9-residue peptide. Four turns of Edman degradation were performed successfully (Table IX).

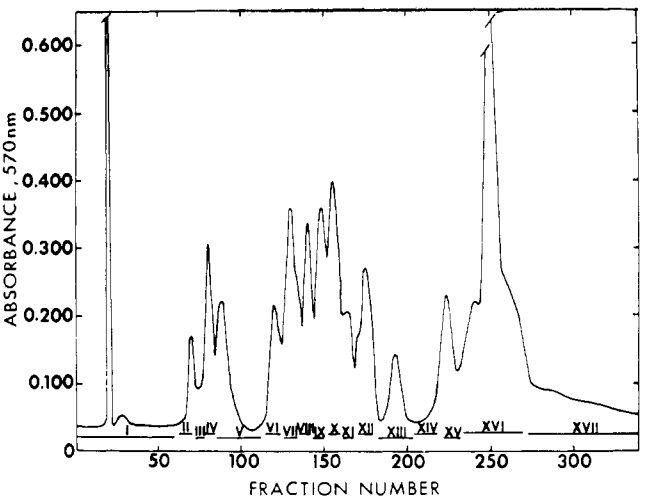

FIGURE 2: Elution profile of the soluble peptic peptides from native NGF on a 2.0 ± 35 cm column of SE-Sephadex C-25 at 55°. The column was developed at 30 ml/hr with a double linear gradient of pyridine acetate buffer. Fractions of 3.0 ml were collected. Bars and Roman numerals indicate the pools made.

FRACTION VIII. Subfractionation on Dowex 1-X2 yielded only one useful peptide, P-VIII-1. Four turns of Edman degradation permitted the alignment of this peptide (Table IX).

FRACTION IX. This pool was comprised of three peptides, which were separated on a column of Dowex 1-X2. The first, P-IX-1, was a tetrapeptide which was completely characterized by three turns of Edman degradation. P-IX-2 was not subjected to sequence analysis, but electrophoretic analysis showed it to be acidic thus confirming the glutamyl residue to be present in the acid form. P-IX-3 is one phenylalanine residue longer than P-IX-2, and four turns of Edman degradation unequivocally positioned all five residues (Table IX).

FRACTION X. The two peptides present in this fraction were separated on a column of Dowex 1-X2. The first, peptide P-X-1, was a dipeptide which was not further characterized. Peptide P-X-2 was found to contain six residues. The structure was derived from the Edman degradation procedure as shown in Table X.

FRACTION XI. This fraction contained two related peptides which were readily separated on Dowex 1-X2. P-XI-1 was a 14-residue peptide which was sequenced through the lysyl residue by 5 successful turns of Edman degradation. Peptide P-XI-2 corresponds to a portion of the carboxyl terminus of the previous peptide, and two turns of Edman degradation were performed. (Table X).

FRACTION XII. This fraction was comprised of a single tetrapeptide, P-XII, which required no further purification. Three turns of Edman degradation determined the sequence of this peptide. Electrophoretic analysis indicated that the glutamyl residue was present in the amide form (Table X).

FRACTION XIII. Upon subfractionation of this fraction on a column of Dowex 50-X2, a single pentapeptide, P-XIII-1, was isolated. Carboxypeptidase A digestion at pH 5, showed the carboxyl-terminal residue to be glutamic acid (Table X).

FRACTION XIV. This pool contained no peptide material.

FRACTION XV. Gel filtration on Sephadex G-25 in 0.1 M acetic acid permitted the removal of minor contaminants from this tryptophan containing peptide, P-XV-1. Six successful turns of Edman degradation were performed, completely characterizing the peptide (Table X).

TABLE VIII: Amino Acid Composition of the Soluble Peptic Peptides of Nerve Growth Factor.[a]

Amino Acid	P-II	P-IV	P-V-1	P-V-2	P-VI-2	P-VI-3	P-VII	P-VIII-1
Lysine					0.96 (1)	1.00 (1)	0.92 (1)	
Histidine								1.11 (1)
Arginine								
Aspartic acid		1.14 (1)			1.18 (1)	1.00 (1)	1.01 (1)	
Threonine		1.72 (2)	1.18 (1)	0.93 (1)	1.62 (2)	0.69 (1)	2.88 (3)	1.28 (1)
Serine								1.65 (2)
Glutamic acid		0.97 (1)			2.38 (2)	2.03 (2)		
Proline								1.00 (1)
Glycine							0.84 (1)	
Alanine					2.37 (2)	1.64 (2)	0.94 (1)	
Valine			0.78 (1)	1.91 (2)			1.34 (1)	1.00 (1)
Methionine								
Isoleucine								
Leucine		1.02 (1)	1.00 (1)	1.04 (1)	1.08 (1)			
Tyrosine								
Phenylalanine	1.00 (1)							0.91 (1)
Tryptophan							(1)[b]	
Total	1	5	3	4	9	7	9	7
Per cent yield	16	34	13	27	5	6	46	25
Purification procedure[c]			DX1[c]	DX1	DX1	DX1		DX1
Residue no.		90–94	37–39	36–39	90–98	92–98	21–29	1–7

[a] Values are given in residues/mol. The assumed integral values are given in parentheses. See Angeletti *et al.* (1973) for further

FRACTION XVII. No peptide material was present in this pool.

Thermolytic Digestion of Peptic Fraction XVI. Peptic pool XVI was the only one recovered among the soluble peptic peptides which contained disulfide bonds, as judged by the cyanide–nitroprusside spray reaction. The amino acid composition of this pool indicated that it was largely comprised of a single fragment covering about 50 residues containing all three disulfide bonds. Consequently, the fragment was subdigested with thermolysin at 40° for 20 hr at pH 7.0 in order to avoid interchange of the disulfide bonds. Acid-

TABLE IX: Characterization of Selected Peptic Peptides from Fractions V, VII, VIII, and IX.

Fraction V:
  Peptide P-V-2:
    Sequence: Val-Thr-Val-Leu
Fraction VII:
  Peptide P-VII:
    Sequence: Trp-Val-Gly-Asx-(Lys,Thr,Thr,Ala,Thr)
Fraction VIII:
  Peptide P-VIII-1:
    Sequence: Ser-Ser-Thr-His-(Pro,Val,Phe)
Fraction IX:
  Peptide P-IX-1:
    Sequence: Phe-Val-Lys-Ala
  Peptide P-IX-2:
    Sequence: (His,Met,Gly,Glu)
    High-voltage electrophoresis, pH 6.5: Acidic
  Peptide P-IX-3:
    Sequence: His-Met-Gly-Glx-Phe

TABLE X: Characterization of Selected Soluble Peptic Peptides from Fractions X, XI, XII, XIII, and XV.

Fraction X:
  Peptide P-X-2:
    Sequence: Asx-Ile-Lys-Gly-(Lys,Glx)
Fraction XI:
  Peptide P-XI-1:
    Sequence: Gly-Asx-Lys-Thr-Thr-(Ala,Thr,Asx,Ile,Lys,-Gly,Lys,Glx,Val)
  Peptide P-XI-2:
    Sequence: Thr-Asx-(Ile,Lys,Gly,Lys,Glx)
Fraction XII:
  Peptide P-XII:
    Sequence: Arg-Gln-Tyr-Phe
    High-voltage electrophoresis, pH 6.5: basic
Fraction XIII:
  Peptide P-XIII-1:
    Sequence: (Ile,Lys,Gly,Lys)-Glu
    Carboxypeptidase A: 1 hr, Glu, 0.96
Fraction XV:
  Peptide P-XV-1:
    Sequence: Lys-Glx-Ala-Ala-Trp-Arg-Phe

P-IX-1	P-IX-2	P-IX-3	P-X-1	P-X-2	P-XI-1	P-XI-2	P-XII	P-XIII-1	P-XV-1
0.98 (1)			1.00 (1)	2.35 (2)	3.20 (3)	2.18 (2)		2.00 (2)	1.08 (1)
	0.87 (1)	0.81 (1)							
							0.99 (1)		1.00 (1)
				1.10 (1)	1.79 (2)	0.94 (1)			
					2.60 (3)	0.80 (1)			
	1.01 (1)	1.23 (1)		1.12 (1)	0.95 (1)	0.96 (1)	1.15 (1)	1.10 (1)	0.97 (1)
	1.06 (1)	1.00 (1)		1.02 (1)	1.98 (2)	1.04 (1)		0.98 (1)	
					1.06 (1)				1.86 (2)
1.12 (1)			0.84 (1)		1.00 (1)				
0.90 (1)	1.00 (1)	0.37 (1)							
				0.92 (1)	0.76 (1)	0.88 (1)		0.64 (1)	
							0.98 (1)		
0.83 (1)		0.70 (1)					1.02 (1)		0.88 (1)
									(1)[b]
4	4	5	2	6	14	7	4	5	7
16	16	6	41	20	1	7	55	2	13
DX1	DX1	DX1	DX1	DX1	DX1	DX1		DX1	G-25
86–89	8–11	8–12	87–88	30–35	23–36	29–35		31–35	95–101

details. [b] Assumed to be one residue (Angeletti *et al.* (1973)). [c] Abbreviations used are: DX1, Dowex 1-X2; G-25, Sephadex G-25.

ification of the solution to pH 2.2 did not result in any precipitation so that the entire digest was applied to a 0.9 × 50 cm column of SE-Sephadex C-25, and eluted with a double linear gradient of pyridine acetate (Walsh *et al.*, 1970). The elution profile, shown in Figure 3, reveals the extent of cleavage of this fragment by thermolysin. Aside from the cystinyl peptides, the other peptides isolated had already been found in the soluble thermolytic peptides described above. Therefore, only the unique peptides containing disulfide bridges are described in detail below. Table XI summarizes the amino acid compositions and purification procedures of all peptides isolated.

Isolation and Characterization of Cystine-Containing Peptic–Thermolytic Peptides. FRACTION VII. No other peptides appeared to contaminate this fraction, whose composition is shown in Table XI. The entire fraction was subjected to gentle performic acid oxidation (Hirs, 1969) in order to cleave the disulfide bond and produce two cysteic acid containing peptides (P-XVI Th-VII-OX-1 and P-XVI Th-VII-OX-2). Preliminary attempts to purify these by column chromatography failed, so that preparative high-voltage electrophoresis at pH 3.75 was carried out. The amino acid compositions as well as amino-terminal analysis using the dansyl-Edman technique allowed unequivocal identification. Table XII summarizes the proposed structure of the cystinyl peptides, and their position in the linear sequence of the NGF polypeptide chain.

FRACTION XI. This fraction was found to contain the second cystinyl peptide, whose amino acid composition is shown in Table XI. It was oxidized as above and passed over a 0.9 × 50 cm column of Dowex 50-X2. The two cysteic acid containing peptides, P-XVI Th-XI-OX-1 and P-XVI Th-XI-OX-2 were obtained in pure form. The composition and the amino-terminal residues of these peptides were determined (Table XII). These data indicated that the residues of half-cystine present in these peptides correspond in the linear sequence to Cys-III and Cys-VI.

FRACTION XIII. This fraction contained the third disulfide peptide. After performic acid oxidation, the fraction was applied to a 0.9 × 50 cm column of Dowex 50-X2, from which the remaining two cysteic acid peptides, P-XVI Th-XIII-OX-1 and P-XVI Th-XIII-OX-2 were obtained. Their amino acid compositions are shown in Table XI. Table XII illustrates the proposed structures and positions, on the basis of amino acid composition and amino end-group data.

Isolation and Characterization of the Acid-Insoluble Peptic Peptides. As noted above, a small portion of the peptic

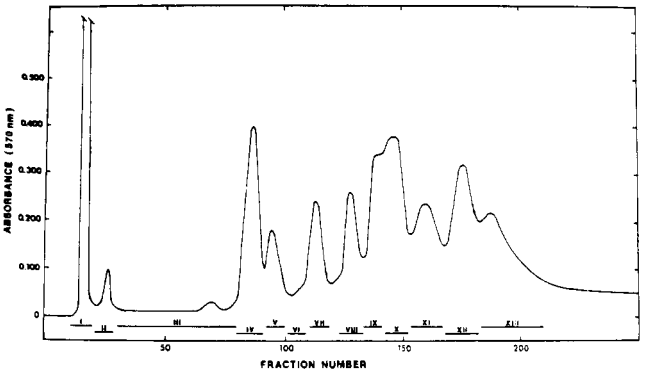

FIGURE 3: Elution profile of thermolytic peptides of fraction P-XVI on 2.0 × 35 cm column of SE-Sephadex C-25 at 55°. The column was developed at 30 ml/hr using the same conditions as in Figure 1.

TABLE XI: Amino Acid Composition of the Thermolytic Peptides of Peptic Peptide P-XVI.[a]

Amino Acid	Th-IV-1	Th-V-1	Th-V-2	Th-VII	Th-VIII	Th-IX-1	Th-X-1	Th-XI	Th-XII-1	Th-XIII
Lysine							1.04 (1)		1.09 (1)	0.97 (1)
Histidine					1.00 (1)		0.91 (1)			
Arginine						1.12 (1)		1.00 (1)	1.02 (1)	1.05 (1)
Aspartic acid			1.00 (1)	1.00 (1)			1.01 (1)			
Threonine				0.93 (1)	1.90 (2)	0.90 (1)				0.98 (1)
Serine	1.00 (1)		0.70 (1)	1.70 (2)			1.26 (1)	0.66 (1)	0.89 (1)	
Glutamic acid										
Proline										
Glycine								2.00 (2)		
Alanine					0.88 (1)					1.00 (1)
Cystine				(1)[d]				(1)[d]		(1)[d]
Valine				0.90 (1)				0.97 (1)		
Methionine										
Isoleucine							0.96 (1)			
Leucine									0.82 (1)	
Tyrosine				0.84 (1)						
Phenylalanine	0.98 (1)	(1)								
Tryptophan			(1)[b]							
Total	2	1	3	7	3	3	5	6	4	5
Per cent yield	13	9	9	17	9	18	16	23	19	25
Purification procedure	DX1[c]	DX1	DX1	HVE		G-50	G-50		G-25	

[a] Values are given in residues/mol. Assumed integral values are given in parentheses. See Angeletti *et al.* (1973) for further details. [b] Assumed to be 1 residue (Angeletti *et al.*, 1973). [c] Abbreviations used are: DX1, Dowex 1-X2; HVE, preparative high-voltage electrophoresis; G-25 and G-50, Sephadex gels G-25 and G-50. [d] Assumed to be 1 residue. See text for details.

digest was not soluble in 5% formic acid. This insoluble material became visibly evident after only 15-min digestion and was present throughout the remainder of the reaction. This precipitate, removed by centrifugation, was completely soluble in 3% pyridine, and was applied to a column of Dowex 1-X2. The elution pattern is shown in Figure 4. Significant amounts of peptide were found only in pools 3, 4, and 6. Their amino acid compositions are shown in Table XIII from which it is evident that the three peptides, PC-3, PC-4, and PC-6, are all related to the same region of the NGF molecule. Edman degradations were performed on each of these, with the results shown in Table XIV.

Isolation and Characterization of the Products of Cyanogen Bromide Cleavage. Since the NGF molecule is known to contain only one residue of methionine per polypeptide chain, two fragments should be obtained after cyanogen bromide cleavage (Gross and Witkop, 1961). Native NGF (0.74 μmol) was reacted as described above. The reaction was stopped by dilution with distilled water followed by lyophilization. The recovered material was applied to a column of Sephadex G-25 fine (0.9 × 125 cm) in 0.1 M acetic acid and eluted as shown in Figure 5. The fractions were

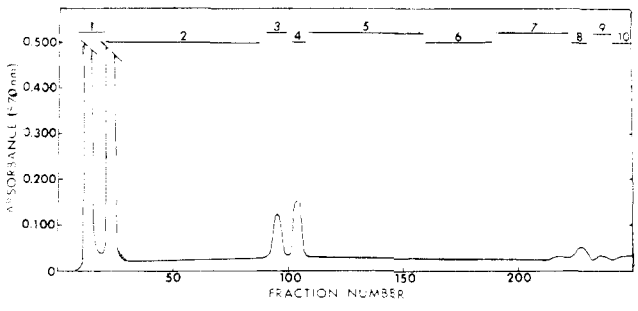

FIGURE 4: Elution profile of the acid-insoluble peptic peptides from native NGF on a 0.9 × 150 cm column of Dowex 1-X2 at 30°. The column was developed at 30 ml/hr with a four-chamber continuous gradient of pyridine acetate buffers. Fractions of 3 ml were collected Bars indicate the pools made.

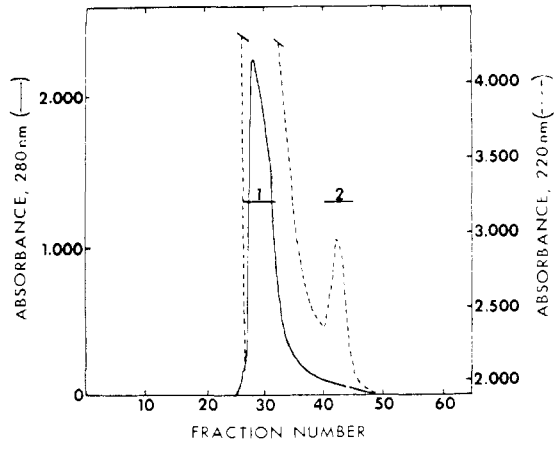

FIGURE 5: Elution profile of the cyanogen bromide fragments of native NGF from a gel filtration column 0.9 × 125 cm of Sephadex G-25 (fine) equilibrated with 0.1 M acetic acid. Flow rate was 9 ml/hr and 1.0-ml fractions were collected. The solid line indicates the absorbance at 280 nm and the dotted line, the absorbance at 220 nm.

TABLE XII: Characterization of Cysteine-Containing Peptic–Thermolytic Peptides.

	Position	Residue No. of Half-Cystinyl Residue
Fraction VII:		
Peptide P-XVI Th-VII-OX-1:		
Sequence: Ser-(Val,Cys,Asp,Ser)	I	15
Peptide P-XVI Th-VIII-OX-2:		
Sequence: Tyr-(Cys,Thr)	IV	80
Fraction XI:		
Peptide P-XVI Th-XI-OX-1:		
Sequence: Val-Cys	VI	110
Peptide P-XVI Th-XI-OX-2:		
Sequence: Ser-(Gly,Cys,Arg,Gly)	III	68
Fraction XIII:		
Peptide P-XVI Th-XIII-OX-1:		
Sequence: Ala-Cys	V	108
Peptide P-XVI Th-XIII-OX-2:		
Sequence: Thr-(Lys,Cys,Arg)	II	58

pooled as indicated. The major peak, pool 1, was reduced and S-carboxymethylated while pool 2 was hydrolyzed and analyzed directly. The amino acid composition of each of the two fragments is shown in Table XV, and the sequence of the smaller fragment, CB-II, in Table XIV. The sum of the two fragments is in excellent agreement with the amino acid composition of NGF (Angeletti *et al.*, 1971). The amino terminus of CB-II is a seryl residue, and the apparent carboxyl terminus is homoserine, indicating that CB-II is the amino-terminal fragment of the NGF.

Analysis of Nerve Growth Factor in the Protein Sequencer. When SCM-NGF was analyzed in the protein Sequencer, equal amounts of two nonidentical residues were released at every cycle. Since the original preparation had been thoroughly examined for homogeneity, and the subunit structure had been clearly ascertained as a dimer of two very similar or identical subunits (Angeletti *et al.*, 1971), it was initially difficult to rationalize these data. However, it was noted that there was a repetition in Pth-amino acids after 8 residues and after completion of 17 cycles it was apparent that the sample contained identical chains differing in length by 8 residues. Thus, twenty-four residues of the longer chain were identified in 24 degradation cycles, although residues 9–24 from the shorter chain also appeared in cycles 1–16. Cycles 17–24 yielded residues 17–24 from the longer chain and residues 25–32 from the shorter chain. The sequence information so obtained was entirely consistent with that obtained by conventional methods. These data are shown in Figure 6.

Carboxyl-Terminal Analysis with Carboxypeptidase B. As was noted previously (Angeletti *et al.*, 1973), no tryptic peptide lacking carboxyl-terminal lysine or arginine was

TABLE XIII: Amino Acid Composition of Acid-Insoluble Peptic Peptides.[a]

Amino Acid	PC-3	PC-4	PC-6
Aspartic acid	1.10 (1)	2.02 (2)	2.67 (3)
Serine	0.73 (1)		0.80 (1)
Glutamic acid		1.00 (1)	1.07 (1)
Alanine		1.00 (1)	0.82 (1)
Valine	1.11 (1)	1.21 (1)	1.77 (2)
Isoleucine		0.87 (1)	0.83 (1)
Phenylalanine	0.94 (1)		0.88 (1)
Total	4	6	10
Per cent yield	10	13	8
Purification procedure	DX1[b]	DX1	DX1
Residue no.	46–49	40–45	40–49

[a] Values given in residues/mol. Assumed integral values are given in parentheses. [b] Abbreviation used is: DX1, Dowex 1-X2.

TABLE XIV: Characterization of Acid-Insoluble Peptic Peptides and Cyanogen Bromide Fragment CB-II.

Peptide PC-3:
Sequence: Asx-Ser-Val-Phe

Peptide PC-4:
Sequence: Ala-Glx-Val-Asx-Ile-Asx

Peptide PC-6:
Sequence: Ala-Glx-Val-Asx-Ile-(Asx,Asx,Ser,Val,Phe)

Fragment CB-II:
Sequence: Ser-Ser-Thr-His-Pro-Val-(Phe,His,Hse)

TABLE XV: Amino Acid Composition of the Cyanogen Bromide Fragments of Native NGF.

Residue	CB-I	CB-II
Lysine	8.20 (8)	
Histidine	2.34 (2)	2.15 (2)
Arginine	6.65 (7)	
CM-Cysteine	5.80 (6)	
Aspartic acid	11.4 (11)	
Threonine	12.6 (13)	1.00 (1)
Serine	9.68 (10)	1.62 (2)
Glutamic acid	8.40 (8)	
Proline	1.27 (1)	1.00 (1)
Glycine	6.09 (7)	
Alanine	7.96 (8)	
Valine	11.60 (12)	0.84 (1)
Isoleucine	5.00 (5)	
Leucine	3.47 (3)	
Tyrosine	2.37 (2)	
Phenylalanine	6.00 (6)	1.00 (1)
Homoserine or HS-lactone	Trace	0.63 (1)
Tryptophan	(3)	
Total	111	9

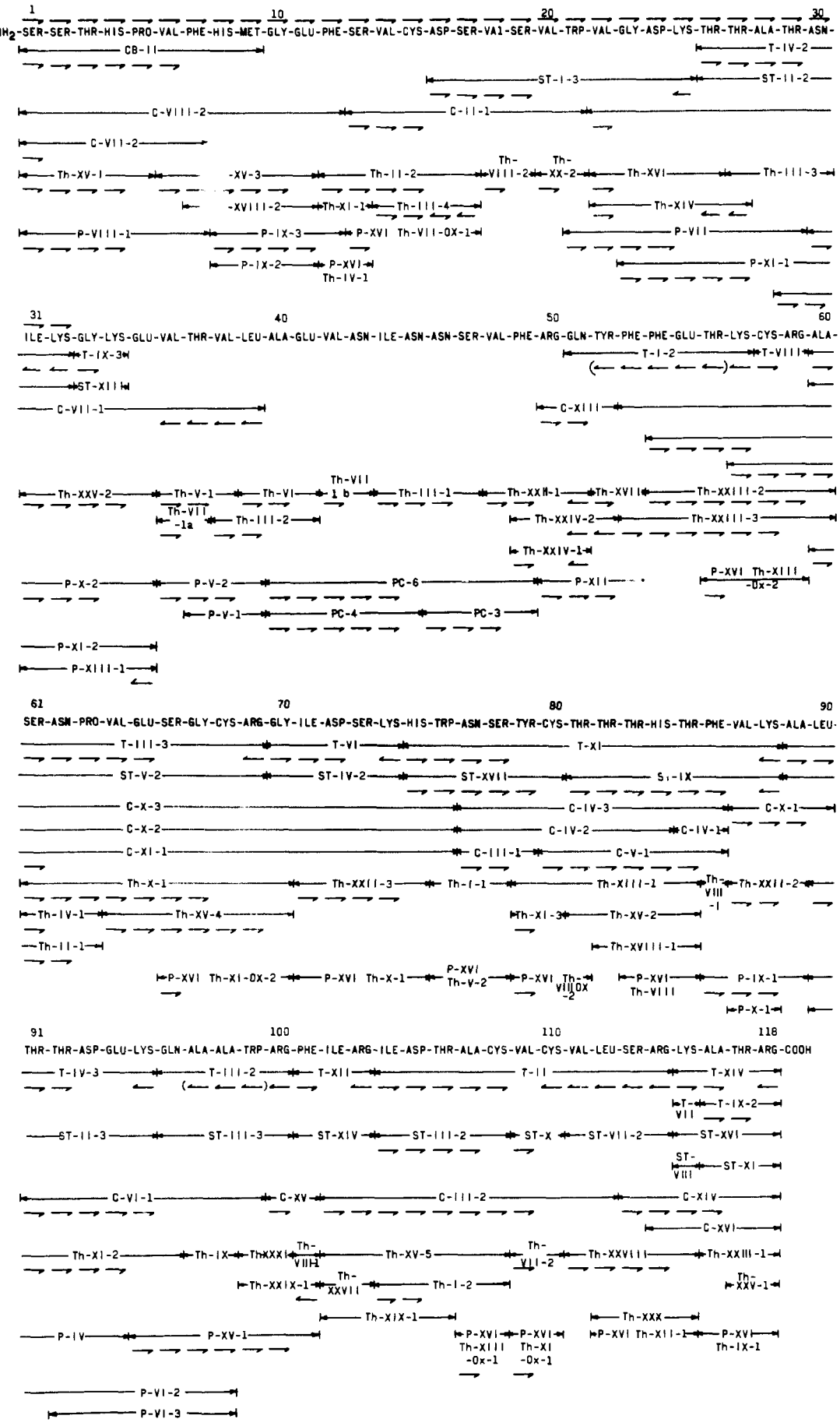

FIGURE 6: The amino acid sequence of 2.5S mouse submandibular gland nerve growth factor. The various peptides are indicated by double-headed arrows. Residues identified by Edman degradation (→), and carboxypeptidase A and/or B hydrolysis (←) are so indicated. Abbreviations used are: T-, tryptic; C-, chymotryptic; Th-, thermolytic; P-, peptic; PTh-, peptic-thermolytic; and CB-, cyanogen bromide.

recovered in the soluble fraction. This observation suggested that the carboxyl-terminal peptide was in the acid-insoluble fraction or that the polypeptide chain terminated in lysine or arginine. Digestion of SCM-NGF with carboxypeptidase A did not release any detectable amino acids. However, incubation with carboxypeptidase B at 40° for 15 min released 0.80 mol of arginine/mol of NGF (14,500 molecular weight). as can be seen in Figure 6, this observation is consistent with the deduced structure of the protein.

Discussion

Assembly of the Complete Sequence. Using the sum of the data described in this and the preceding paper (Angeletti *et al.*, 1973), it is possible to derive a complete and consistent amino acid sequence of mouse submandibular gland nerve growth factor. Figure 6 summarizes the requisite data. The peptides corresponding to each segment are listed below the designated residues in the order of hydrolysis by trypsin (SCM- and SAE-NGF), chymotrypsin, thermolysin, and pepsin. The details of the rationale for assembling this structure are discussed below.

Although quantitative amino-terminal analysis of NGF (Angeletti *et al.*, 1971) indicated that the amino-terminal residue was serine, no peptide containing amino-terminal serine was found among the soluble tryptic fractions. However, cyanogen bromide cleavage produced one large fragment plus a small nonapeptide, CB-II, which contained both amino-terminal serine and carboxyl-terminal homoserine. This peptide, along with the data of the automatic Sequencer, provide a firm basis for positioning the amino end of the polypeptide chain. The sequence of the residues in this region is confirmed in part by peptides P-VIII-1, C-VII-2, and Th-XV-1. By composition, C-VIII-2 extends this region by 3 residues to position 12. The sequence data for peptides Th-XV-3 and P-IX-3 confirm these assignments which are further substantiated by the composition data of Th-XVIII-2 and P-IX-2. In addition to the Sequencer data, further extension is achieved only by 1 residue overlap at positions 12-13 with the peptide Th-II-2. The sequence in the immediate vicinity of this half-cystine is reinforced by the composition of Th-III-4, and extended 4 residues by composition by C-II-1. Although the structure of this peptide was not fully delineated, sufficient information was obtained to indicate that ST-I-3 extended from this point to the first lysyl residue encountered, Lys-25. The internal sequence data for peptide ST-I-3 are confirmed in part by the dipeptides Th-VIII-2 and Th-XX-2.

The overlap with tryptic peptide T-IV-2 was first indicated by the two thermolysin peptides, Th-XVI and Th-XIV. By composition, both of these fit into the carboxyl-terminal portion of ST-I-3, and extend beyond the lysyl residue by 1 and 2 residues of threonine, respectively. The only soluble tryptic peptide containing threonine in the amino-terminal position is T-IV-2. By both sequence and composition, P-VII confirms the positioning of this tryptic peptide. P-XI-1 also supports this alignment, and extends, by composition, the sequence to residue 36, overlapping with the tryptic dipeptide, T-IX-3. The smaller peptic peptides P-XI-2 and P-XIII-1 provide additional proof from composition, and Th-XXV-2 and P-X-2 unequivocally establish the sequence in this region.

Further extension was rendered difficult because a soluble tryptic peptide with an amino-terminal sequence of Glu-Val was not isolated. An 18-residue chymotryptic peptide, C-VII-1, was isolated which by composition was comprised of T-IV-2, T-IX-3 the carboxyl-terminal portion of ST-I-3, and 5 additional residues. Although this peptide was obtained in insufficient quantities to subdigest, one turn of dansyl-Edman served to anchor its amino terminus at position 22. Digestion with carboxypeptidase A provided the carboxyl-terminal sequence, thus extending the chain to residue 39. A series of small peptides, Th-V-1, Th-VII-1a, P-V-1, and P-V-2 confirmed the sequence in this region. The thermolysin peptide Th-III-2 extends by a 2-residue overlap, the structure by 2 residues and with the related peptide Th-VI, provides the basis for the 2-residue overlap to PC-6 and PC-4. Along with PC-3, Th-VII-1b, and Th-III-1, the linear structure can be extended to residue 49.

Peptide Th-XXII-1 extends the chain to residue 51. A series of small peptides, Th-XXIV-1, Th-XXIV-2, ThXVII, C-XIII, and P-XII produce the overlapping structure through residue 53. This structure provides sufficient overlap to align tryptic peptide T-I-2. Unfortunately, the amino-terminal glutamine of this peptide was cyclized (pyrrolidonecarboxylic acid), and was not amenable to Edman degradation, but carboxypeptidase A and B digestion had yielded partial sequence data, and peptides from other digests confirmed these results.

The structural data needed were acquired from the 2 complementary peptides, Th-XXIII-2 and Th-XXIII-3, and from peptide P-XVI Th-XII-OX-2 that peptide T-VIII, the Cys-Arg dipeptide, followed T-I-2 and was, in turn, contiguous to a tryptic peptide beginning with alanine. This peptide was identified as T-III-3 from the isolation of the related peptides C-XI-1, C-X-2, and C-X-3 which are 20, 23, and 24 residues long, respectively. Both of the chymotryptic peptides, C-X-2 and C-X-3, contain histidine, tryptophan, and an additional radioactively labeled cysteinyl residue, indicating that at least a portion of another cysteinyl-containing tryptic peptide is included in the adjacent region. The internal structure of T-III-3 is provided by the thermolytic peptides Th-X-1, Th-IV-1, Th-XV-4, and Th-II-1. In addition, Th-XV-4 indicates that the amino-terminal residue of the next tryptic peptide begins with glycine. Only Tp-VI, which contains an amino-terminal glycyl residue, is consistent with the composition of C-X-3, C-X-2, and C-XI-1. This alignment provides an overlap consisting of a tryptophyl and histidyl residue. Only one tryptic peptide, T-XI, can be placed in this position as it contains the only two remaining histidyl residues in the structure. The sequence of T-XI is provided by ST-XVII and ST-IX and confirmed by a considerable number of chymotryptic and thermolytic peptides.

The series of short peptides P-IX-1, Th-XXII-2, and C-X-1 clearly provide the overlap to peptide T-IV-3 which begins at position 89. Although not structured completely, the difference in composition of C-VI-1 unequivocally identifies the adjacent tryptic peptide as T-III-2 which contains the only remaining tryptophyl residue. This overlap is also indicated by Th-XI-2 and compositionally by P-VI-2 and P-VI-3. The internal sequence of T-III-2 is provided by P-XV-1.

Peptides C-XV, Th-XXIX-1, and P-XV-2 place phenylalanine in position 101, but the specificities of chymotrypsin, thermolysin, and pepsin all permit cleavage around phenylalanine, and as a result, no peptide containing a more rigorous overlap was isolated. However, the only soluble tryptic peptide with amino-terminal phenylalanine not incorporated into the structure already is T-XII, which C-III-2 indicates to be contiguous with T-II. Again, only a 2-residue overlap is available, but no other soluble tryptic peptide possesses an Ile-Arg carboxyl-terminal sequence. From composition data, Th-I-2,

TABLE XVI: Assignment of Side-Chain Amides in Nerve Growth Factor.

Residue No.	Assignment	Method of Determination and Peptide[a]
11	Glu	HVE, P-IX-2; S
16	Asp	HVE, ST-I-3; S
24	Asp	HVE, ST-I-3; S
30	Asn	HVE, T-IV-2; S
35	Glu	CPA, P-XIII-1
41	Glu	HVE, Th-III-2
43	Asn	HVE, Th-VIII-1[b]
45	Asn	HVE, Th-III-1
46	Asn	HVE, Th-III-1
51	Gln	HVE, T-I-2, P-XII
55	Glu	HVE, T-I-2
62	Asn	HVE, Th-II-1
65	Glu	HVE, T-III-3, Th-XV-4
72	Asp	HVE, T-VI
77	Asn	HVE, C-III-1, ST-XVII
93	Asp	HVE, T-IV-3
94	Glu	HVE, T-IV-3
96	Gln	HVE, T-III-2
105	Asp	HVE, ST-III-2
5 Asp	6 Asn	6 Glu 2 Gln

[a] Abbreviations used are: HVE, high-voltage electrophoresis, pH 6.5; S, Sequencer; CPA, carboxypeptidase A.

Th-XIX-1, and Th-XV-5 support this hypothesis. The sequence of T-II was complete and confirmed by several peptides.

Peptide C-XIV extends the chain four residues beyond position 114. The structural information available concerning Th-XXVIII, Th-XXX, and C-XVI confirm the placement of T-XIV, and thus T-VII and T-IX-2, as the final tryptic peptides in the structure.

The positioning of T-XIV at the carboxyl terminus is consistent with the identification of arginine as the carboxyl-terminal residue by carboxypeptidase B digestion. This sequence (Figure 6) accounts for all peptides isolated, including the thermolysin digest in which all the peptides were soluble. In addition, the composition generated by this sequence is in reasonable agreement with that determined by amino acid analysis.

Assignment of the Amide Residues. High-voltage electrophoresis at pH 6.5, direct amino acid analysis of digestion mixtures of carboxypeptidase A and data from the protein Sequencer were used separately or in concert to distinguish between the acidic and amide forms of aspartyl and glutamyl residues. These results are summarized in Table XVI. Glu-11 is acidic as determined directly by the Sequencer and by the acidic mobility of P-IX-2. ST-I-3 was acidic at pH 6.5, indicating that both Asp-16 and Asp-24 are present in the acid form. Sequencer data corroborated these results. Residue 30 was identified as asparagine by electrophoresis of T-IV-2 in agreement with the Sequencer. However, one sample of T-IV-2 was neutral indicating the presence of an aspartic acid residue. The origin of this group is most likely deamidation and the native protein undoubtedly has asparagine in this position. Glu-35 was identified as acidic by direct amino acid analysis after carboxypeptidase A digestion of P-XIII-1. The acidic mobility of Th-III-2 demonstrated Glx-41 to be a glutamic acid residue. Th-VII-1b is neutral, confirming that Asx-43 is asparagine. Th-III-1 is also neutral, thus demonstrating that both Asx-45 and -46 are asparagine. The resistance of T-I-2 to Edman degradation implies that residue 51 is cyclized glutamine. This conclusion was confirmed by carboxypeptidase A digestion and by high-voltage electrophoresis of P-XII. The acidic mobility of T-I-2 is consistent with the identification of residue 55 as glutamic acid. Th-II-1 is neutral, demonstrating that Asx-62 is indeed asparagine. Given the assignment of Asn-62, the acidity of T-III-3 indicated that residue 65 is glutamic acid. The acidic nature of Th-XV-4 confirms these results. Electrophoretic analysis of T-VI reveals the peptide to be neutral, thus assigning aspartic acid to position 72. Peptide C-III-1 is electrophoretically neutral and ST-XVII basic, so that residue 77 can be identified as asparagine. Peptide T-IV-3 is acidic, so that both residues 93 and 94 must be in the acid form, aspartic acid and glutamic acid respectively. T-III-2 was not amenable to Edman degradation, indicating that residue 96 is cyclized glutamine. Its electrophoretic neutrality is consistent with this hypothesis. Residue 105 was initially reported to be asparagine (Angeletti and Bradshaw, 1971) but a confirmation of the electrophoretic mobility of ST-III-2 demonstrated it to be electrophoretically neutral, consistent with the presence of an aspartic acid residue instead. A similar conclusion was obtained with Th-XIX-1. A comparison of the electrophoretic mobilities of these samples to that when they were initially isolated does not indicate any change consistent with deamidation. It has been concluded therefore that the initial experiments were in error and the correct assignment of this residue is aspartic acid.

Of the 11 aspartic acid residues, 6 are present in the amide form, whereas 2 of the 8 glutamic acid residues are found as glutamine, or a total of 8 of the 19 carboxylate residues are present as the amide form in NGF.

Identification of the Disulfide Pairs. The NGF monomer contains three intrachain disulfide bridges, and no sulfhydryl residues (Bocchini, 1970; Angeletti *et al.*, 1971). All three of the disulfide bridges appeared to be located in one large peptide P-XVI. Thermolytic digestion of this fragment yielded a variety of peptides, among which were three separate disulfide-containing peptides.

Upon performic acid oxidation, P-XVI Th-VII yielded a tyrosyl-containing cysteinyl tripeptide, P-XVI Th-VII-OX-2 and a pentapeptide, P-XVI Th-VII-OX-1. Their compositions correspond to the regions surrounding Cys-15 and Cys-80, respectively, defining the first disulfide pair as I-IV. Cys-108 and Cys-110 are separated by a single valyl residue which was cleaved by thermolysin. Peptide P-XVI-Th-XIII-OX-1, Ala-Cys, can be derived only from Cys-108, and its pair, P-XVI Th-XIII-OX-2, Lys-Cys-Arg, can be unambiguously placed at Cys-58. In a like manner P-XVI Th-XI-OX-2 is clearly associated with Cys-68, whereas P-XVI Th-XI-OX-1 contains Cys-110. Thus, half-cystinyl residues comprising the disulfide pairs are paired as I-IV, II-V, and III-VI.

Comments on the Primary and Secondary Structure of Mouse 2.5S NGF. The assembly of the amino acid sequence of mouse NGF described above forms an internally consistent and coherent structure. However, the primary structure so described is 9 residues shorter than that expected from the amino acid composition. (Angeletti *et al.*, 1971). However, recalculation of the original composition data on the basis of a corrected molecular weight, minimizes the differences. In addition, the amino acid composition of the large cyanogen bromide fragment

CB-I (Table XV) is almost identical with the composition predicted from the sequence. This is to be expected, since the amino-terminal region of the molecule is quantitatively excised, yielding homogeneous polypeptide chains, whereas the native molecules appear to exist largely as a population of chains both with and without the amino-terminal fragment in the preparations used. These data plus the inherent consistency of the primary structure establish that native NGF monomer is a polypeptide chain of 118 amino acids, with a monomeric molecular weight of 13,259.

The data from the protein Sequencer revealed the presence of a significant proportion of polypeptide chains lacking the first 8 amino acid residues. This heterogeneity, which would probably have gone undetected in the conventional analysis, has been found in varying proportions in different preparations. In this regard, it is interesting to note that amino-terminal analyses of NGF by the cyanate method revealed serine as the amino-terminal residue (Angeletti *et al.*, 1971). The inaccuracy inherent in the determination of this residue was noted at that time. It has also been ascertained that the unidentified peak reported with those data was methionine sulfoxide. When properly calculated, there was 0.46 mol of methionine (uncorrected) in the end-group samples which is entirely consistent with the Sequencer results.

Experiments in progress in our laboratory have indicated that these two chains, which have been designated A for the long chain and B for the shorter one, are probably not the products of two distinct genes, but the result of proteolysis, either within the submandibular gland, or during the process of purification. The possible physiological significance of this phenomenon is now under study.

The primary structure data together with the results obtained from the determination of disulfide pairs impose some restrictions on the mobility and conformation of the polypeptide chain. The closing of disulfide bonds between half-cystines II-V and III-VI defines a 14-residue loop near the carboxyl end of the chain, which is more basic in character than the amino end.

The sequence of NGF has been compared to that of several other proteins (Angeletti and Bradshaw, 1971). Particular attention has centered around insulin because of its similarity of metabolic stimulatory effects on sympathetic nerve cells, (Angeletti *et al.*, 1965; Liuzzi *et al.*, 1968; Partlow and Larrabee, 1971) and structural similarity of the type associated with proteins related by evolutionary events has been proposed for these two proteins (Frazier *et al.*, 1972). Thus, the completion of the primary structure of mouse submaxillary gland NGF has stimulated the postulation of several experimentally accessible models for mechanism of action and structure–function relationships which are currently under investigation in our laboratory.

Acknowledgments

The authors express their thanks to Mr. Robert Sherman and Mr. Charles Yates and also to Mr. Delio Mercanti for his expert technical assistance.

References

Angeletti, P. U., Gandini-Attardi, D., Toschi, G., and Levi-Montalcini, R. (1965), *Biochim. Biophys. Acta 95*, 111.

Angeletti, R. H., and Bradshaw, R. A. (1971), *Proc. Nat. Acad. Sci. U. S. 68*, 2417.

Angeletti, R. H., Bradshaw, R. A., and Wade, R. D. (1971), *Biochemistry 10*, 463.

Angeletti, R. H., Mercanti, D., and Bradshaw, R. A. (1973), *Biochemistry 12*, 90.

Bocchini, V. (1970), *Eur. J. Biochem. 15*, 127.

Edman, P., and Begg, G. (1967), *Eur. J. Biochem. 1*, 80.

Frazier, W. A., Angeletti, R. H., and Bradshaw, R. A. (1972), *Science 176*, 482.

Gross, E., and Witkop, B. (1961), *J. Amer. Chem. Soc. 83*, 1510.

Hermodson, M. A., Ericsson, L. H., and Walsh, K. A. (1970), *Fed. Proc., Fed. Amer. Soc. Exp. Biol. 29*, 728.

Hirs C. H. W. (1969), *Methods Enzymol. 11*, 197.

Liuzzi, A., Pocchiari, F., and Angeletti, P. U. (1968), *Brain Res. 7*, 452.

Partlow, L. M., and Larrabee, M. G. (1971), *J. Neurochem. 18*, 2101.

Pisano, J. J., and Bronzert, T. J. (1969), *J. Biol. Chem. 20*, 5597.

Toennies, G., and Kolb, J. J. (1951), *Anal. Chem. 23*, 823.

Walsh, K. A., MacDonald, R. M., and Bradshaw, R. A. (1970), *Anal. Biochem. 35*, 193.