References
Akkerman, J. W. N., Gorter., & Kloprogge, E. (1982) *Thromb. Res. 27*, 59–64.
Bettex-Galland, M., & Lüscher, E. F. (1960) *Thromb. Diath. Haemorrh. 4*, 178–195.
Born, G. V. R. (1970) *J. Physiol. (London) 209*, 487–511.
Costa, J. L., & Murphy, D. L. (1975) *Nature (London) 255*, 407–408.
Detwiler, T. C., & Feinman, R. D. (1973a) *Biochemistry 12*, 282–289.
Detwiler, T. C., & Feinman, R. D. (1973b) *Biochemistry 12*, 2462–2468.
Feinman, R. D., & Detwiler, T. C. (1974) *Nature (London) 249*, 172–173.
Friedman, F., & Detwiler, T. C. (1975) *Biochemistry 14*, 1315–1320.
Holmsen, H., Kaplan, K., & Dangelmaier, C. A. (1982) *Biochem. J. 208*, 9–18.
Karnovsky, M. J. (1965) *J. Cell Biol. 27*, 137a.
Kolber, M. A., & Haynes, D. H. (1981) *Biophys. J. 36*, 361–391.
Latimer, P. J., Born, G. V. R., & Michal, F. (1976) *Arch. Biochem. Biophys. 180*, 151–159.
Lorez, H. P., Da Prada, M., & Pletscher, A. (1975) *Experientia 31*, 593–595.
Martin, B. M., Feinman, R. D., & Detwiler, T. C. (1975) *Biochemistry 14*, 1308–1314.
Massini, P., & Näf, U. (1979) *Biochim. Biophys. Acta 598*, 575–587.
Nishizuka, Y. (1983) *Trends Biochem. Sci. (Pers. Ed.) 8*, 13–16.
Patscheke, H. (1979) *Haemostasis 8*, 65–81.
Patscheke, H. (1981) *Haemostasis 10*, 14–27.
Patscheke, H., & Wörner, P. (1978) *Thromb. Res. 12*, 485–496.
Pletscher, A., Da Prada, M., Berneis, K. H., & Tranzer, J. P. (1971) *Experientia 27*, 993–1120.
Rink, R. J., Smith, S. W., & Tsien, R. Y. (1982a) *J. Physiol. (London) 237*, 53–54P.
Rink, R. J., Smith, S. W., & Tsien, R. Y. (1982b) *FEBS Lett. 148*, 21–26.
Rudnick, G., Fishkes, H., Nelson, P. J., & Schuldiner, S. (1980) *J. Biol. Chem. 255*, 3638.
Skaer, R. J. (1981) *Platelets in Biology and Pathology*, Vol. 2, pp 321–348, Elsevier/North-Holland Biomedical Press, Amsterdam.
Skaer, R. J., Flemans, R. J., & McQuilkan, S. (1981) *Br. J. Haematol. 49*, 435–438.
Traverso-Cori, A., Chaimovich, H., & Cori, O. (1965) *Arch. Biochem. Biophys. 109*, 173–184.

7S Nerve Growth Factor α and γ Subunits Are Closely Related Proteins†

Hans Ronne,* Helena Anundi, Lars Rask, and Per A. Peterson

ABSTRACT: The polypeptide composition and partial amino acid sequence of the 7S nerve growth factor (NGF) α subunit have been determined. Residues in 76 unique positions corresponding to 35% of the molecule were identified. The sequence shows that the NGF α subunit is closely related to the NGF γ subunit and thus a member of the same protein family as the serine proteases. This finding is unexpected since the NGF α subunit is devoid of detectable protease activity. However, the NGF α subunit differs in one important respect from the NGF γ subunit and related serine proteases. The highly conserved amino-terminal activation cleavage structure, common to most serine proteases, has been deleted, and an uncleaved activation peptide remains attached to the amino terminus of the mature NGF α subunit. It is suggested that this feature is causally related to the apparent lack of proteolytic activity.

The submaxillary gland of the adult male mouse is a rich source of nerve growth factor (NGF),[1] a polypeptide that promotes the differentiation and maintenance of certain nerve cells of neural crest origin (Server & Shooter, 1977). The growth factor is secreted into the saliva as a high molecular weight protein complex, 7S NGF (Varon et al., 1968; Burton et al., 1978). The core of the 7S NGF complex is a dimer of the biologically active polypeptide (the β subunit) and two molecules of a 28 000-dalton arginine esterase, the γ subunit (Stach et al., 1976). The 7S NGF complex also contains two molecules of the α subunit, which is a 27 000-dalton protein of unknown function (Stach et al., 1980). Like many other biologically active polypeptides, the NGF β subunit is synthesized as a precursor protein which is cleaved to yield the mature growth factor. The NGF γ subunit is responsible for this processing, and the enzyme remains bound to the growth factor after the cleavage has occurred (Berger & Shooter, 1977).

The NGF α and γ subunits have similar molecular weights and cross-react immunologically. When reduced and alkylated under denaturing conditions, the two proteins decompose into similar mixtures of 26 000-, 17 000-, 10 000-, and 6000-dalton chains (Anundi et al., 1978). These findings suggested that the two molecules may be related to one another. Sequence information on the NGF γ subunit (Thomas et al., 1981b) has demonstrated that it belongs to the trypsin branch of the serine proteases (DeHaën et al., 1975; Dayhoff, 1978). The aim of the present investigation was to obtain sequence information

† From the Department of Cell Research, The Wallenberg Laboratory, University of Uppsala, S-751 22 Uppsala, Sweden. *Received August 4, 1983.* This work was supported by grants from the Swedish Medical Research Council.

[1] Abbreviations: NGF, nerve growth factor; EGF-BP, epidermal growth factor binding protein; SDS, sodium dodecyl sulfate; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; PTH, phenylthiohydantoin.

0006-2960/84/0423-1229$01.50/0

on the NGF α subunit in order to determine the possible relatedness of this protein to the serine proteases.

Materials and Methods

7S NGF α Subunit. The 7S NGF α subunit was purified from the submaxillary glands of adult male NMRI mice (obtained from Anticimex, Stockholm, Sweden). The 7S NGF complex, isolated as described (Stach et al., 1977), was dissolved in 0.05 M sodium acetate, pH 5, containing 6 M guanidine hydrochloride and chromatographed on a column of Sepharose 4B (1.8 × 130 cm) equilibrated with the same buffer. This step separated the 27 000-dalton α and γ subunits from the 13 000-dalton β-NGF monomer. The 27 000-dalton material was dialyzed against water and lyophilized. Separation of the α and γ subunits was then achieved by preparative isoelectric focusing in the presence of 6 M urea. The procedure was that of Winter et al. (1975), using an LKB (Stockholm, Sweden) apparatus with a bed of Sephadex G-75 and pH 4–6 Pharmalytes (Pharmacia Fine Chemicals, Uppsala, Sweden). The purified α subunit was dialyzed against water and lyophilized.

Electrophoresis. Analytical SDS–PAGE was performed according to Laemmli (1970). Preparative SDS–PAGE was carried out as described by Maizel et al. (1970). Protein bands were cut out, and the gel pieces were minced and packed into small glass columns. The proteins were then eluted from the gel fragments by overnight electrophoresis into dialysis bags at the anodal end of the columns.

Amino Acid Analysis. Polypeptides were extensively reduced and alkylated (Wiman et al., 1979) and subjected to hydrolysis in vacuo at 110 °C in 6 M HCl containing 0.1% phenol. Hydrolysates were analyzed on a Beckman 121 M amino acid analyzer (Palo Alto, CA). Tryptophan and carboxymethylcysteine (Cm-cysteine) were not determined.

Amino Acid Sequence Determination. Amino-terminal sequencing was carried out in a Beckman 890 C liquid-phase sequencer (Trägårdh et al., 1979). PTH-amino acids were analyzed by reverse-phase high-pressure liquid chromatography (Fohlman et al., 1980). Yield calculations were corrected for background and carry-over.

Results

Polypeptide Composition of the NGF α Subunit. Like the NGF γ subunit (Thomas et al., 1981a), preparations of the NGF α subunit contain several polypeptide chains held together by disulfide bonds (Anundi et al., 1978). To isolate the individual polypeptides, the purified α subunit was reduced and alkylated under denaturing conditions and then applied to a column of Sepharose 6B, equilibrated with 6 M guanidine hydrochloride (Figure 1). Polypeptides in three peaks, corresponding to apparent molecular weights of about 17 000, 10 000, and 6000, respectively, were obtained in quantities sufficient for further analyses. Trace amounts of material with an apparent molecular weight of 26 000 also appeared in the chromatogram (Figure 1). This material is consistently found in preparations of the NGF α subunit (Anundi et al., 1978). The fractions shown in Figure 1 were pooled as indicated in the figure, dialyzed against distilled water, lyophilized, and analyzed by SDS–PAGE. The 6000-dalton material contained a single polypeptide chain, henceforth referred to as the 6K chain. The 17 000- and 10 000-dalton materials, in contrast, appeared heterogeneous on SDS–PAGE.

The 17 000-dalton material contained two polypeptides that migrated very closely together (Figure 2A). The two polypeptides will be referred to as the 17K-1 (slow migrating) and 17K-2 (fast migrating) chains. Part of the 17K material was

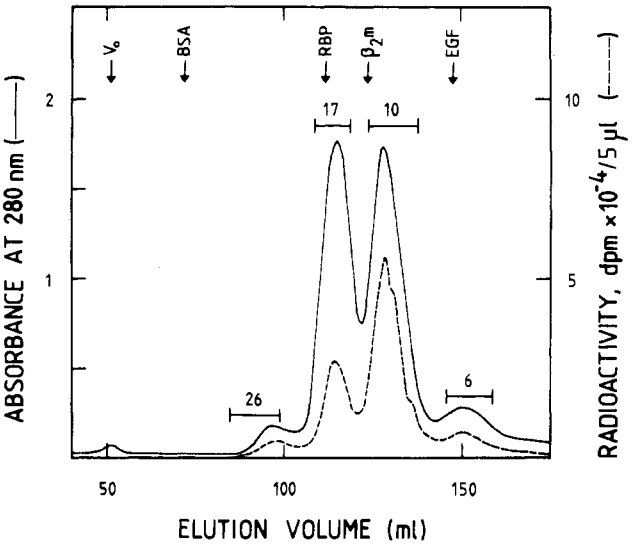


FIGURE 1: Separation of NGF α-subunit polypeptides. The protein, 30 mg, was reduced and alkylated under denaturing conditions. Alkylation was accomplished by using iodo[^{14}C]acetic acid. The polypeptides were separated on a column of Sepharose 6B (1.2 × 130 cm) equilibrated with 6 M guanidine hydrochloride in 0.05 M sodium acetate buffer, pH 5.0. The flow rate of the column was 3 mL/h. Protein in the eluate was monitored by the absorbance at 280 nm. Peptides containing Cm-cysteine were identified by measuring the ^{14}C radioactivity. Fractions were pooled as indicated by the bars. The markers and their molecular weights are the following: bovine serum albumin (BSA), 69 000; retinol-binding protein (RBP), 21 000; β_2-microglobulin (β_2m), 12 000; epidermal growth factor (EGF), 6000. The values above the peaks denote the apparent molecular masses in kilodaltons.

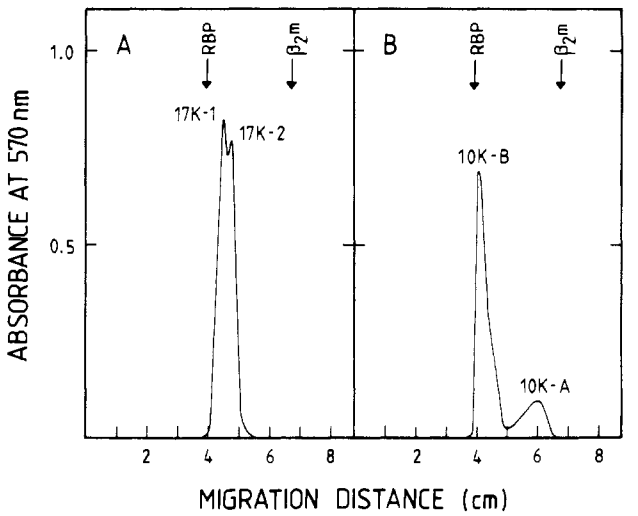


FIGURE 2: Analysis of NGF α-subunit polypeptides by SDS–PAGE. Aliquots of the pooled fractions in Figure 1 were analyzed on 12% polyacrylamide disc gels. The gels were fixed in 10% trichloroacetic acid, stained with Coomassie brilliant blue, and scanned at 570 nm. (A) 17K material; (B) 10K material. The markers and their molecular weights are the following: retinol-binding protein (RBP), 21 000; β_2-microglobulin (B_2m), 12 000.

subfractionated by preparative SDS–PAGE. Partial separation of the two chains was achieved. Thus, while the 17K-1 chain was obtained free from the 17K-2 species, only a fraction enriched for the 17K-2 form but still containing the 17K-1 chain was isolated.

The 10 000-dalton material (see Figure 1) resolved into two well-separated components on SDS–PAGE (Figure 2 B). A minor amount of the material migrated with an apparent molecular weight of about 13 000. This component will be referred to as the 10K-A chain. The major component in the 10K fraction (see Figure 1) migrated with an apparent mo-

Table I: Amino Acid Composition of the NGF α Subunit and Its Composite Polypeptide Chains

amino acid[b]	no. of residues[a]					
	α subunit[c]	17K-1	17K-2	10K-A	10K-B	6K
lysine	13.5 (14)	6.8 (7)	7.2 (7)	4.3 (4)	5.0 (5)	2.8 (3)
histidine	6.6 (7)	0.7 (1)	1.1 (1)	0.2 (0)	1.2 (1)	0.6 (1)
arginine	5.2 (5)	4.0 (4)	4.2 (4)	2.7 (3)	0.9 (1)	1.0 (1)
aspartic acid	25.3 (25)	18.6 (19)	18.5 (19)	14.5 (15)	12.1 (12)	4.6 (5)
threonine[d]	13.1 (13)	7.3 (7)	7.5 (8)	3.3 (3)	6.0 (6)	4.2 (4)
serine[d]	21.2 (21)	13.2 (13)	13.0 (13)	6.4 (6)	6.5 (7)	3.4 (3)
glutamic acid	28.4 (28)	15.9 (16)	15.2 (15)	11.2 (11)	10.3 (10)	4.3 (4)
proline	15.6 (16)	10.4 (10)	10.1 (10)	6.4 (6)	7.9 (8)	4.1 (4)
glycine	23.2 (23)	12.4 (12)	12.5 (13)	8.2 (8)	10.0 (10)	6.3 (6)
alanine	12.1 (12)	7.2 (7)	7.0 (7)	4.6 (5)	3.9 (4)	2.2 (2)
valine	11.0 (11)[e]	8.0 (8)	7.6 (8)	6.1 (6)	3.1 (3)	2.0 (2)
methionine	4.4 (4)	1.8 (2)	1.7 (2)	0.9 (1)	2.4 (2)	0.9 (1)
isoleucine	7.3 (7)[e]	3.5 (4)	3.8 (4)	1.7 (2)	4.6 (5)	2.5 (3)
leucine	18.8 (19)	14.1 (14)	14.3 (14)	8.2 (8)	8.1 (8)	5.4 (5)
tyrosine	6.5 (7)	4.3 (4)	4.4 (4)	3.6 (4)	2.8 (3)	0.9 (1)
phenylalanine	4.8 (5)	2.8 (3)	2.9 (3)	2.7 (3)	1.1 (1)	0.8 (1)
	217 (217)[f]	131 (131)[f]	131 (132)[f]	85 (85)[f]	86 (86)[f]	46 (46)[f]

[a] Values in parentheses are nearest integers. [b] Determined after a 24-h hydrolysis, except where noted. Tryptophans and Cm-cysteines were not determined. [c] Mean value of 24- and 72-h hydrolyses. [d] Corrected for hydrolysis. [e] 72-h value. [f] Total. For each polypeptide, the total number of residues was assumed to be equal to the number of residues found in the corresponding part of the NGF γ subunit (tryptophans and cysteines excluded).

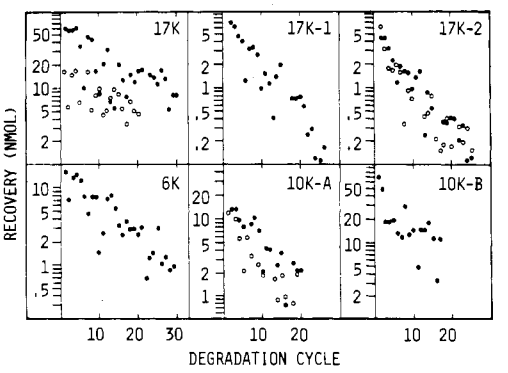
lecular weight of 20 000. This polypeptide will be referred to as the 10K-B chain. The 10K fraction was separated into highly purified 10K-A and 10K-B chains by using preparative SDS–PAGE.

To rule out that the 10K-B chain represents contaminating 17K material possibly present in the 10K fraction, aliquots of the 17K and 10K materials were labeled with ^{125}I and ^{131}I, respectively, and chromatographed together on a Sepharose 6B column equilibrated in 6 M guanidine hydrochloride. No cross-contamination was detected, and all of the 10K material was eluted in the expected position (data not shown). Therefore, it was tentatively concluded that the true molecular weight of the 10K-B chain is about 10 000 but that this polypeptide migrates anomalously on SDS–PAGE. This conclusion is supported by sequence and alignment data (see below).

Amino Acid Compositions. Amino acid composition data on the intact NGF α subunit and its purified constituent polypeptide chains are given in Table I. Obviously, the 17K-1 and 17K-2 polypeptides are almost identical. In contrast, the 10K-A chain does not appear to be related to the 10K-B chain. It has been suggested that limited endoproteolysis of the NGF α and γ subunits, respectively, may generate the various polypeptides from intact 27 000-dalton chains (Anundi et al., 1978; Thomas et al., 1981a). To examine this possibility, we compared the amino acid composition of the 17K-1 polypeptide with the combined compositions of the 6K chain and 10K-A and 10K-B chains, respectively (Table II). The comparison strongly suggests that the 17K-1 polypeptide may have generated the 10K-A and 6K chains. In contrast, the 10K-B and 6K chains cannot both be derived from the 17K-1 polypeptide (Table II).

Amino Acid Sequences. The amino-terminal amino acid sequences of the isolated NGF α-subunit polypeptides were determined. All polypeptides were degraded with excellent repetitive yields (Figure 3). Therefore, PTH-amino acids could be identified in a total of 147 positions.

The 17K material, comprising a mixture of the 17K-1 and 17K-2 chains, was sequenced for 30 steps. In most of the degradation cycles, two PTH-amino acids were obtained, but in different amounts. This difference in yield was sufficient to distinguish a major and a minor sequence from each other



17K
••• A P P V Q S Q V D C E N S Q P W H V A V Y R F N K Y L C G G
○○○ S Q V D C X N S Q P W H V A V Y R F N K Y

17K-1
••• X P P V Q S Q V D C E N S Q P W H V A V Y R F N K Y

17K-2
••• A P P V Q S Q V D C E N S Q P X H V A V Y R F N K
○○○ S Q V D C E N S Q P X H V A V Y R F N K Y L C G G

6K
••• S N D L M L L R L S K P A D I T D V V K P I T L P T E C P

10K-A
••• A P P V Q S Q V D C E N X Q P W H V A V
○○○ S Q V D C E N X Q P W H V A V Y R F N

10K-B
••• Y P D D L Q C V N L K L L P N E D

FIGURE 3: Amino-terminal amino acid sequences and PTH-amino acid yields of NGF α-subunit polypeptides. The amounts of material subjected to degradation were 180, 25, 31, 75, 159, and 73 nmol of the 17K, 17K-1, 17K-2, 10K-A, 10K-B, and 6K polypeptides, respectively. When two sequences are found in the same polypeptide fraction, the major one is denoted by solid dots and the minor one by open dots. Unidentified residues are denoted with an X.

(Figure 3). Thus, it could be concluded that the first five amino acid residues in the major sequence were absent in the minor sequence. Apart from this difference, the two sequences were identical.

The 17K-1 and 17K-2 fractions obtained by preparative SDS–PAGE were separately subjected to sequence determinations. The 17K-1 chain was sequenced for 26 steps. In each

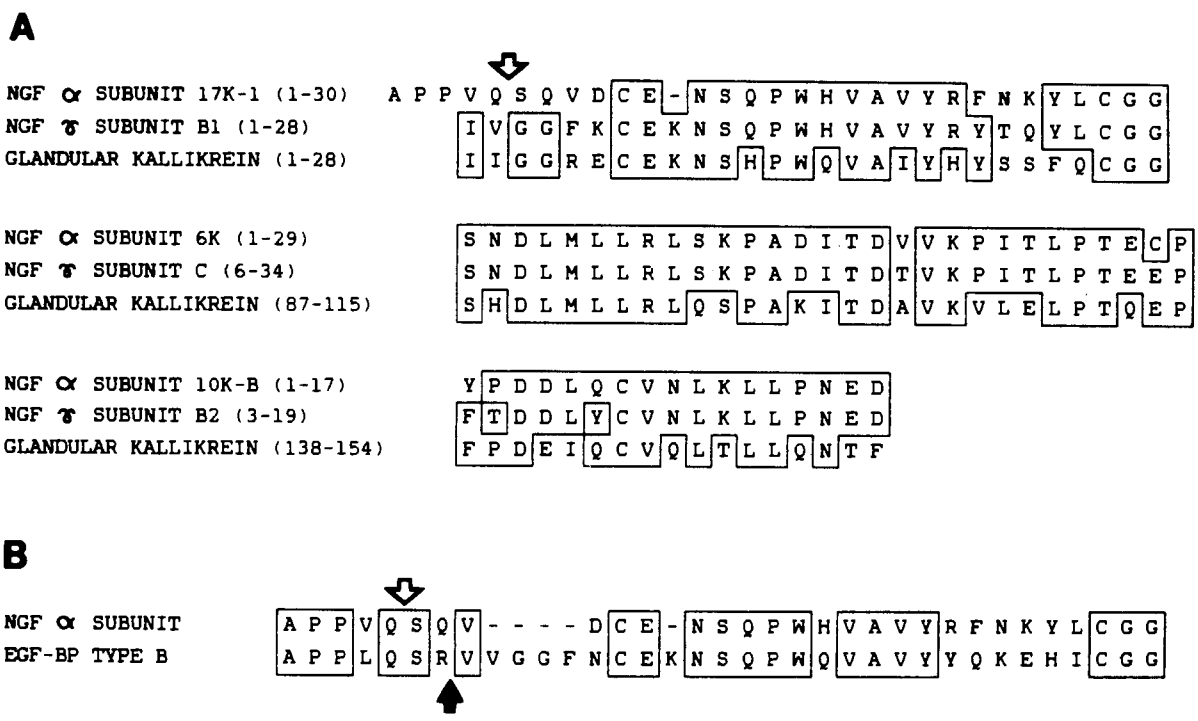


FIGURE 4: Alignment of the sequenced parts of the NGF α subunit to related serine proteases. Identical residues are enclosed within boxes. The site of partial endoproteolysis in the amino terminus of the NGF α subunit is indicated by an open arrow. Numbers within parentheses denote the residue numbers within sequenced fragments or, in the case of glandular kallikrein, within the total sequence. (A) Alignment of the NGF α-subunit sequences to homologous parts of the NGF γ subunit (Thomas et al., 1981b) and glandular kallikrein (Tschesche et al., 1979). (B) Alignment of the amino terminus of the NGF α subunit (the 17K-1 sequence) to the amino terminus and activation peptide of EGF-BP type B (Ronne et al., 1983). The solid arrow indicates the site of the activation cleavage that releases the first seven residues (the activation peptide) from the amino terminus of the EGF-BP zymogen.

Table II: Comparison of Amino Acid Compositions[a]

amino acid	17K-1	6K + 10K-A	6K + 10K-B
lysine	7	7	8
histidine	1	1	2
arginine	4	4	2
aspartic acid	19	20	17
threonine	7	7	10
serine	13	9	10
glutamic acid	16	15	14
proline	10	10	12
glycine	12	14	16
alanine	7	7	6
valine	8	8	5
methionine	2	2	3
isoleucine	4	5	8
leucine	14	13	13
tyrosine	4	5	4
phenylalanine	3	4	2

[a] Nearest integer values from Table I.

position, a single PTH-amino acid was found, and the sequence was identical with the major 17K sequence described above (Figure 3). The 17K-2 fraction was sequenced for 25 steps. The sequence obtained was the same composite sequence as for the unfractionated 17K material. However, the component giving rise to the minor sequence seemed to be substantially enriched, as compared to the original 17K material (Figure 3). In conclusion, the polypeptides in the 17K material appear to be derived from a single polypeptide, the 17K-1 chain. A minor fraction of this polypeptide seems to have been converted into the 17K-2 chain by endoproteolytic removal of an amino-terminal pentapeptide.

The 10K-A chain was sequenced for 20 steps. This sequence was identical with that of the 17K material (Figure 3). The 10K-B and 6K chains, finally, were sequenced for 17 and 29 steps, respectively. These sequences were unique and differed from each other as well as from the 17K and 10K-A sequences (Figure 3).

The amino acid composition data in Table II suggested that 10K-A and 6K chains are generated by limited endoproteolysis of the 17K material. This notion is reinforced by the finding that the amino-terminal sequences of the 10K-A and 17K chains are identical. The 6K chain must then be derived from the carboxy-terminal part of the 17K chain.

Alignment of the NGF α-Subunit Polypeptide Chains. The previously noted immunological cross-reactivity between the NGF α and γ subunits (Anundi et al., 1978) suggests that these proteins are structurally similar. Therefore, the amino acid sequences of the NGF α-subunit polypeptides were compared to the complete sequence of the NGF γ subunit (Thomas et al., 1981b). Figure 4A shows that all of the NGF α-subunit sequences are homologous to various regions of the NGF γ subunit. Therefore, the NGF α subunit, like the γ subunit, is a member of the large protein family to which the serine proteases belong (Dayhoff, 1978). This is also evident from a comparison of the NGF α-subunit sequences with the sequence of glandular kallikrein (Tschesche et al., 1979), another closely related serine protease (Figure 4A). The homologies in Figure 4A were tested for statistical significance by using the RELATE program (Dayhoff, 1976). The scores obtained (Table III) clearly demonstrate that the NGF α subunit is evolutionarily related to the serine proteases and that it stands particularly close to the NGF γ subunit within this protein family.

The alignments in Figure 4A show that the 17K-1 polypeptide represents the amino-terminal region of the NGF α subunit. The carboxy-terminal region corresponds to the 10K-B chain. Further endoproteolysis of the 17K-1 polypeptide obviously yields the 10K-A and 6K chains. The latter one corresponds to the carboxy-terminal region of the 17K-1 polypeptide. This is in agreement with the alignment deduced

Table III: Homology Scores for Alignments in Figure 4A

NGF α-subunit peptide	scores[a] for alignment to NGF γ subunit	scores[a] for alignment to glandular kallikrein
17K-1	11.3	9.3
6K	14.6	9.3
10K-B	8.5	8.0

[a] The scores given are the number of standard deviations of the mean of the real top scores above the mean of the top scores of 100 randomized alignments. The scores were calculated according to the RELATE program (Dayhoff, 1976) by using the mutation data matrix and a fragment length of 15 residues.

from the amino acid composition and sequence data (see above).

Discussion

The aim of the present investigation was to find out whether the NGF α subunit is structurally related to the NGF γ subunit. This is indeed the case, and the 76 positions identified in the α subunit can be aligned with the NGF γ subunit in a unique way due to the high degree of homology between the two proteins (Figure 4A).

Since the NGF α subunit is structurally homologous to the NGF γ subunit, it is a member of the same protein family as the serine proteases (Dayhoff, 1978). The extensive homology of the α subunit to the serine proteases, which are synthesized as single chains, lends strong support to the notion that the NGF α subunit is also synthesized as a single polypeptide chain. This is consistent with the observation that α-subunit preparations also contain an intact 26 000-dalton chain (Anundi et al., 1978).

There is one notable exception to the pronounced sequence homology between the NGF α and γ subunits. Thus, the homology disappears close to the amino terminus of the NGF γ subunit (Figure 4A). This is remarkable, since the amino terminus plays a crucial role in the activation of zymogen to enzyme (Bode, 1979) and therefore is an evolutionarily conserved part of the serine proteases (Dayhoff, 1978).

The amino-terminal part of the NGF α subunit does not bear any resemblance to the corresponding part of any known member of the serine protease family. However, a clue to the origin of this unusual structure was recently obtained from the sequence of a cDNA clone corresponding to EGF-BP (Ronne et al., 1983). EGF-BP is a submaxillary gland serine protease that is closely related to the NGF α and γ subunits. Sequence comparisons revealed that the seven amino-terminal residues of the NGF α subunit are highly homologous to the activation peptide of EGF-BP (Figure 4B).

The activation peptide precedes the amino terminus of the enzyme and is cleaved off during activation of the zymogen (Neurath & Walsh, 1976). The alignments in Figure 4 suggest that two important changes have occurred in the gene coding for the NGF α subunit:

(i) Five amino acid residues have been deleted in the NGF α subunit. This deletion (Figure 4B) covers precisely that part of the amino terminus which is important for the conformational transition from zymogen to enzyme (Bode, 1979). As a consequence of the deletion, the activation peptide has been shifted into the position previously occupied by the deleted segment. This accounts for the sudden lack of homology to the NGF γ subunit at the amino terminus (Figure 4A).

(ii) Among the serine proteases, the last residue in the activation peptide, at which the activation cleavage occurs, is always an arginine or a lysine (Neurath & Walsh, 1976; Dayhoff, 1978). The NGF α subunit, however, has a glutamine in the homologous position (Figure 4B). The presence of a glutamine instead of a basic residue would make cleavage of the activation peptide by a trypsin-like enzyme impossible. This, together with the deletion, may explain why the activation peptide remains attached to the amino terminus of the mature NGF α subunit.

The presence of the 17K-2 chain in preparations of the NGF α subunit suggests that some limited endoproteolysis does occur in the amino terminus of the protein. However, the new amino terminus thus generated is neither homologous to nor collinear with the amino terminus of an active serine protease (see Figure 4A). Therefore, it is doubtful whether this cleavage can induce the conformational change required for activation. Moreover, it is only a fraction of the molecules that are cleaved at this position. It should be pointed out that limited endoproteolysis is common in the submaxillary gland. For example, partial endoproteolysis removes an amino-terminal octapeptide from the NGF β subunit, which, of course, is not a serine protease (Wilson & Shooter, 1979). Accordingly, the limited proteolysis that does occur in the amino terminus of the NGF α subunit is not, in itself, evidence of zymogen activation.

It is of considerable interest to determine whether the NGF α subunit is an active protease. The present investigation has shown that it is a member of the same protein family as the serine proteases, yet the NGF α subunit has no esterolytic activity with standard substrates (Greene et al., 1968), and it does not interact with α_2-macroglobulin, which is a general protease inhibitor (unpublished results). This situation is not unprecedented, since the β chain of haptoglobin, another member of the serine protease family, may have lost its protease activity during its evolution into a serum transport protein (Kurosky et al., 1980). It is conceivable that the NGF α subunit, too, is permanently devoid of protease activity and that this loss is causally related to the structural alterations in the activation cleavage region.

However, alternative explanations cannot be excluded. The NGF α subunit may be an esterase displaying a very high degree of substrate specificity which so far has escaped detection. It is also possible that the α subunit is subject to a mechanism of activation which differs from that of most other serine proteases. To settle these questions, it seems necessary to obtain the complete sequence of the NGF α subunit.

Acknowledgments

The expert technical assistance of K. Andersson, J. Ericson, and I. Sjöquist is gratefully acknowledged.

Registry No. Nerve growth factor, 9061-61-4.

References

Anundi, H., Rask, L., & Peterson, P. A. (1978) *Zoon 6*, 213–216.

Berger, E. A., & Shooter, E. M. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 3647–3651.

Bode, W. (1979) *Naturwissenschaften 66*, 251–258.

Burton, L. E., Wilson, W. H., & Shooter, E. M. (1978) *J. Biol. Chem. 253*, 7807–7812.

Dayhoff, M. O. (1976) *Atlas of Protein Sequence and Structure*, Vol. 5, Georgetown University Press, Washington, DC.

Dayhoff, M. O. (1978) *Atlas of Protein Sequence and Structure*, Vol. 5, Suppl. 3, Georgetown University Press, Washington, DC.

DeHaën, C., Neurath, H., & Teller, D. C. (1975) *J. Mol. Biol. 92*, 225–259.

Fohlman, J., Rask, L., & Peterson, P. A. (1980) *Anal. Biochem. 106*, 22–26.

Greene, L. A., Shooter, E. M., & Varon, S. (1968) *Proc. Natl. Acad. Sci. U.S.A. 60*, 1383-1388.
Kurosky, A., Barnett, D. R., Lee, R.-H., Touchstone, B., Hay, R. E., Arnott, M. S., Bowman, B. H., & Fitch, W. M. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3388-3392.
Laemmli, U. K. (1970) *Nature (London) 227*, 680-685.
Maizel, J. V., Summers, D. F., & Scharft, M. S. (1970) *J. Cell. Physiol. 76*, 273-287.
Neurath, H., & Walsh, K. A. (1976) *Proc. Natl. Acad. Sci U.S.A. 73*, 3825-3832.
Ronne, H., Lundgren, S., Severinsson, L., Rask, L., & Peterson, P. A. (1983) *EMBO J. 2*, 1561-1564.
Server, A. C., & Shooter, E. M. (1977) *Adv. Protein Chem. 31*, 339-409.
Stach, R. W., Server, A. C., Pignatti, P-F., Plitch, A., & Shooter, E. M. (1976) *Biochemistry 15*, 1455-1461.
Stach, R. W., Wagner, B. J., & Stach, B. M. (1977) *Anal. Biochem. 83*, 26-32.
Stach, R. W., Pignatti, P.-F., Baker, M. E., & Shooter, E. M. (1980) *J. Neurochem. 34*, 859-865.
Thomas, K. A., Silverman, R. E., Jeng, I., Baglan, N. C., & Bradshaw, R. A. (1981a) *J. Biol. Chem. 256*, 9147-9155.
Thomas, K. A., Baglan, N. C., & Bradshaw, R. A. (1981b) *J. Biol. Chem. 256*, 9156-9166.
Trägårdh, L., Curman, B., Wiman, K., Rask, L., & Peterson, P. A. (1979) *Biochemistry 18*, 2218-2226.
Tschesche, H., Mair, G., Godec, G., Fiedler, F., Ehret, W., & Hirschauer, C. (1979) *Adv. Exp. Med. 120*, 245-260.
Varon, S., Nomura, I., & Shooter, E. M. (1968) *Biochemistry 7*, 1296-1303.
Wilson, W. H., & Shooter, E. M. (1979) *J. Biol. Chem. 254*, 6002-6009.
Wiman, K., Trägårdh, L., Rask, L., & Peterson, P. A. (1979) *Eur. J. Biochem. 95*, 265-273.
Winter, A., Perlmutter, H., & Davies, H. (1975) *LKB Application Note No. 198*, pp 1-9, LKB, Stockholm.

Conformational Studies of the Human Complex-Forming Glycoprotein, Heterogeneous in Charge: Protein HC†

José G. Gavilanes, Carlos López-Otín, Francisco Gavilanes, and Enrique Méndez*

ABSTRACT: Human complex-forming glycoprotein, heterogeneous in charge, is a single polypeptide chain widely distributed in physiological fluids. The conformation of the protein has been studied with attention to the secondary and tertiary structures. Circular dichroism and predictive methods from the amino acid sequence have been employed for the characterization of the secondary structure. This is composed of 20% α-helix, 21% β-structure, 29% β-turns, 30% aperiodic conformation, and an average number of residues per helical segment of nine. Titration of the protein indicated the existence of two groups for the tyrosine residues, each of them composed of three and five residues. The four tryptophan residues of the molecule are located in two different polarity microenvironments, according to the fluorescence studies. These observations are corroborated by studying the hydropathic profile of the protein. From this study, three different domains are observed in the protein, one of them being exposed and containing the main part of the unordered structure of the molecule. The chromophore naturally associated with the protein has been resolved in three fluorescent units not dependent on the protein conformation. These bands have been observed centered around 290, 360, and 410 nm, which do not correspond to any described chromophore.

Human complex-forming glycoprotein, heterogeneous in charge (protein HC), is a recently described low molecular weight glycoprotein originally isolated from normal human urine (Tejler & Grubb, 1976). It shows a considerable charge heterogeneity, carries an unidentified yellow-brown chromophore, and has been immunochemically demonstrated to occur in normal human plasma where a considerable part of the immunoreactivity is complexed with IgA (immunoglobulin A) (Tejler & Grubb, 1976). Protein HC displays appreciable charge heterogeneity on agarose gel electrophoresis and on isoelectric focusing, which does not diminish after desialylation (Tejler & Grubb, 1976). The complete amino acid sequence of protein HC isolated from the urine of an individual was reported in a preliminary paper (López et al., 1981). Further studies revealed the presence of an additional tryptophan residue at position 36 (C. López et al., unpublished results). No evidence for sequence variability of the single polypeptide chain of protein HC was found. However, protein HC isolated from a pool of urine from several individuals contains molecules with different COOH-terminal amino acid sequences (López et al., 1982). Computer analysis according to Dayfoff's program (Dayhoff, 1978) did not reveal any significant homology to any known protein. The protein is immunochemically and physicochemically related with two other recently described glycoproteins, α_1-microglobulin (Ekström et al., 1975; Ekström & Berggard, 1977; Svensson & Ravnskov, 1976; Takagi et al., 1979; Bernier et al., 1980) and α_1-microglycoprotein (Seon & Pressman, 1978), both isolated from the urine of patients with renal tubular dysfunction. Protein HC has been reported to be associated with lymphocyte cell surfaces (Tejler et al., 1976; Pearlstein et al., 1977) and to possess immunoregulatory properties (Logdberg & Akerstrom, 1981).

In this paper, we report the study of the secondary and tertiary structures of protein HC based on spectroscopic characterizations of the molecule and on the amino acid se-

† From the Department of Biochemistry, Faculty of Sciences, Complutense University, Madrid 3, Spain (J.G.G. and F.G.), and the Servicio de Endocrinología, Centro Ramón y Cajal, Carretera de Colmenar, Madrid 34, Spain (C.L.-O. and E.M.). *Received June 29, 1983.*

0006-2960/84/0423-1234$01.50/0