THE MOUSE NERVE GROWTH FACTOR GENE: ITS EVOLUTION, EXPRESSION AND REGULATION

by

MARK JOSEPH SELBY

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

MICROBIOLOGY

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

. . . . . .

Date University Librarian

JUN 14 1987

Degree Conferred: . . . . . .

DEDICATION

I dedicate my thesis to my family, specifically, to my wife, son, brother, parents and grandparents who helped see me through this arduous process. In particular, I would like to note special appreciation to my wife, Angelica, who provided constant encouragement. When the experiments were not working she gave me reason to continue. During my tenure at this university, we were most fortunate to bring into this world Nicolas, whose presence lights up our lives and gives us inner strength.

ACKNOWLEDGEMENTS

I would especially like to thank Dr. John Watson for his encouragement and for being a great friend. I also want to thank many former and current members of the Rutter lab/Hormone Research Institute for their friendship and guidance. I most appreciate that guidance and collaboration with Drs. Scott and Edwards. I must not forget to mention encouragement and discussions with my good friend, Dr. James Ou. I appreciate the thoughtfulness of Dr. Eric Shooter with regard to data and reagents. It was a pleasure to interact with him. I am indebted to Leslie Spector for typing not only this thesis but manuscripts as well. I will miss discussions with Leslie about topical polemic issues. I am most grateful to Dr. Bill Rutter for providing me with laboratory space for my thesis research and providing a stimulating atmosphere.

I respectively acknowledge that the work presented in Chapter I was a collaboration with Dr. James Scott; II, IV and VI with Dr. Robert Edwards; V with Pablo D. Garcia.

Portions of this dissertation (chapters I and II) are reprints of material as they appear in NATURE. Dr. Rutter (co-author) directed and supervised the research which forms the basis for these chapters.

COPYRIGHT PAGE

THE MOUSE NERVE GROWTH FACTOR GENE:

ITS EVOLUTION, EXPRESSION AND REGULATION

MARK JOSEPH SELBY

ABSTRACT

β-Nerve Growth Factor (NGF) is a neurotropic polypeptide that is required for the survival and differentiation of peripheral sympathetic and sensory neurons. It can be isolated from the mouse submaxillary gland where it is particularly abundant and is complexed with two other subunits, α and γ. I isolated a 1.3 kb mouse cDNA that predicted NGF is located at the carboxy-terminus of a higher molecular weight precursor. Subsequently, several additional cDNA's were described which have unique 5' ends, but the NGF coding region remains intact. I isolated and characterized the mouse NGF gene (≥45 kb) which consists of five small exons (33-167 bp) and a large 3' exon (894 bp) encoding NGF. The organization of the exons suggests that the transcript heterogeneity arises from independent promoter usage and alternatve splicing of exon II.

I isolated and characterized an NGF cDNA from the distantly related cobra. By comparing NGF precursors from cobra, mouse and several other species, I have identified potentially important functional residues. A greater sequence identity in the NGF moiety compared to the non-NGF region is suggestive of a structural role rather than a functional role for the latter. Multiple NGF transcripts were not observed in cobra venom gland RNA.

The coordinate expression of NGF transcripts during postnatal development and the inability of a neural network lesion to influence this expression suggest that NGF expression may be developmentally programmed and unresponsive to innervation per se.

The multiple NGF cDNA's predict two classes of precursors with molecular weights of ~34 kd and 27 kd. Interestingly, a signal sequence-like hydrophobic region is found either within the precursor (~34 kd) or at the N-terminus (27 kd). However, both precursors are completely translocated into the endoplasmic reticulum and are cleaved directly after the presumptive signal sequence. Further, in vivo-synthesized precursors give rise to NGF after incubation with the γ-subunit protease; α- subunit addition did not modify the cleavage pattern. The γ-subunit activity is not particularly specific since trypsin can effect the same cleavage. No similar processing pattern was observed using synthetic precursors, suggesting that they may not be folded correctly.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER VI

CHAPTER VIII

INTRODUCTION

The nervous system is a finely tuned biological system of enormous complexity. Specific nerves must innervate their appropriate targets in a timely and defined manner. Further, complex neural networks formed during development must be maintained throughout the life of the animal. To this end, one can envision the target providing a trophic influence on the nerves so as to ensure neuronal homeostasis. Moreover, the concentration of trophic factor in a given target would dictate the innervation density of that tissue. In the peripheral nervous system, one such trophic factor, nerve growth factor (NGF), has been identified which is, at least in part, target derived and supports the survival, differentiation, and maintenance of specific innervating neurons. Responsive sympathetic and sensory neurons take up NGF via specific receptors and transport it to the cell body where it can manifest its effects. In development, NGF determines the extent of sympathetic and sensory neuron survival and differentiation and in mature neural networks probably acts as a maintenance factor. In this thesis I report the following: the isolation and characterization of the NGF cDNAs; the corresponding gene structure; expression of the various transcripts in particular tissues and during postnatal development; and the proteolytic processing of the NGF precursors by an identified kallikrein protease. I present my experiments in the context of the discovery of NGF and its subsequent biochemical and physiological characterizations.

DISCOVERY OF NGF ACTIVITY

A tumor-derived factor was implicated in supporting a dense neurite outgrowth in chick embryos. Bueker (1948) grafted mouse sarcoma tissue

into the body wall of chick embryos and observed invasion of the tumor by nerve bundles from enlarged sensory ganglia. He concluded that increasing the peripheral field of innervation results in the enlargement of the involved sensory ganglia. Levi-Montalcini and Hamburger (1951) confirmed Beuker's results and extended them by demonstrating that adjacent sympathetic, as well as sensory, ganglia were enlarged compared to the normal contralateral ganglia. They modified the chick embryo experiments by implanting the tumors on the chorioallantoic membrane such that the circulation was the only contact between the tumor and the embryo. In this configuration, sympathetic ganlgia both proximal and distal to the tumor were enlarged. Moreover, neurites were extended into the adjacent viscera and through blood vessel walls where they formed agglomerations in the lumen (Levi-Montalcini, 1952). These observations were not compatible with a simple association of enlarged ganglia with an increased field of innervation. Rather, the data are consistent with the elaboration of a humoral factor by the tumor tissue which resulted in the extensive neurite outgrowth.

BIOASSAY/DISCOVERY OF SNAKE VENOM NGF

The characterization of this humoral factor, later termed nerve growth factor (NGF), was advanced by the development of an in vitro assay system. Levi-Montalcini et al. (1954) found that co-cultures of chick embryo ganglia and tumor fragments on semi-solid media resulted in a dense halo of ganglionic nerve fibers within 12 to 24 hours. Relying on this novel bioassay, Stanley Cohen fractionated tumor extracts and identified active fractions containing nucleoprotein. Crude snake venom (Agkistrodon piscivorus) containing phosphodiesterase was added to these fractions to determine if the activity remained following digestion of

the nucleic acids. Surprisingly, the venom control was more potent in the bioassay than the tumor fractions (Cohen & Levi-Montalcini, 1956)! The injection of NGF-enriched venom fractions into the yolk of 6-8 day old chick embryos produced qualitatively the same neurite outgrowth observed with mouse tumors (Levi-Montalcini & Cohen, 1956). Thus, a novel, enriched source of NGF was discovered. The presence of NGF was not restricted to the venom used; it has since been identified in the venoms of numerous other snakes (e.g., Hogue-Angeletti & Bradshaw, 1978).

Since the salivary gland is the phylogenetic homologue of the snake venom gland, Stanley Cohen serendipidously assayed extracts from mice submaxillary glands (SMGs) and identified an even more potent neurite growth-promoting activity (Cohen, 1960). While snake and mouse factors displayed similar sensitivities to heat, acid and base, their different behavior on ion-exchange chromatography suggested that the molecules are not identical. The levels of mouse SMG NGF are sexually dimorphic for more NGF activity was found in extracts of male glands compared to females. The high levels of SMG NGF are unique for other mouse tissues and the SMGs of other rodents do not contain similarly high NGF levels. Thus, a unique, highly enriched source of NGF has been identified based on similar activities and physical properties.

ANTIBODY INHIBITION IN VIVO; NGF ACTS ON SYMPATHETIC SYSTEM

Injection of NGF antiserum into newborn rodents resulted in the destruction of 97-99% of the neurons in the sympathetic superior cervical ganglia (SCG) (Levi-Montalcini & Booker, 1960). Overall, the antibodies were destructive for the adrenergic neurons of the paravertebral and rostroprevertebral sympathetic ganglia, but ineffective on

short adrenergic neurons. The ganglia of these refractile neurons are situated in close proximity to their targets (ex, urogenital tract, growth adipose tissue, central adrenergic neurons) and, predictably, are also unresponsive to systemically-provided NGF (Olson, 1967). The cytotoxic effect of this antiserum is also manifested at the biochemical level: the activities of tyrosine hydroxylase (TH) and dopamine-β-hydroxylase (DβH), catecholamine synthetic enzymes, are substantially reduced. The antiserum effects on neonatal sympathetic ganglia were irreversible: examination of mice months after the antibody therapy was terminated showed the same degree of destruction as observed earlier. In spite of this persistent effect, the animals were otherwise normal (Levi-Montalcini & Angeletti, 1966). Since sympathetic neurons develop in early postnatal life, the destructive antibody effects are consistent with a requirement for NGF in promoting the survival of most developing sympathetic neurons.

NGF antibodies became less cytotoxic for sympathetic ganglia with increasing age; irreversible damage to sympathetic ganglia no longer occurred beyond the third postnatal week (Goedert et al., 1978). In adult mice, a smaller proportion of the neurons in the ganglia were destroyed in contrast to an almost complete destruction in neonates (Angeletti et al., 1971a). Similarly, tyrosine hydroxylase and dopamine-β-hydroxylase showed more modest reductions in enzyme activities in the adult ganglia, although to a lesser extent than in the developing neonates.

AB INHIBITION IN VIVO: NGF ACTS ON SENSORY SYSTEM

NGF is required by sensory neurons during sensory embryonic development. Earlier studies using specific antiserum demonstrated that

sympathetic neurons require NGF for survival during early postnatal development; the same experiments showed that sensory ganglia were morphologically insensitive to neonatal antibody treatment (Levi-Montalcini & Angeletti, 1966). Since sensory neurons mature earlier than sympathetic neurons, it is not unexpected that major sensory effects were not observed in the newborns exposed to NGF antisera. Analysis of neonatal rats born to autoimmune mothers showed a nearly complete destruction of sympathetic ganglia (90%), while a 70% decline in neurons of the 8th cervical dorsal root sensory glangia (DRG) was observed (Gorin & Johnson, 1979; Johnson et al., 1980). Similar experiments performed in guinea pigs (because of a more efficient prenatal transfer of maternal antibody across the placenta) showed sensory defects (e.g. insensitivity to temperature extremes) in neonates from autoimmune mothers (Pearson et al., 1983). A significant reduction of 80-85% of cells in the eighth cervical DRG was observed in these offspring. Further, only neural crest-derived sensory neurons were affected by NGF antibodies during embryogenesis while neural crest-derived parasympathetic and placodally-derived sensory neurons were not (Johnson et al., 1980; Pearson et al., 1983). Thus, the survival of developing neural crest-derived sensory ganglia is dependent upon NGF.

Postnatally, sensory neurons are less responsive to NGF antibody treatment. While most studies find no morphological alteration associated with adult or postnatal antiserum exposure (ex, Schwartz et al., 1982), Yip et al. (1984) observed a modest reduction (18%) in the number and size distribution of lumbar DRG neurons. While the reduction was significant, it did not approach that observed following in utero antibody exposure. Biochemically, the levels of substance P (A marker

for many sensory neurons (Hokfelt et al., 1975; 1976)) are significantly decreased in DRG, spinal cord and skin (Otten et al., 1980; Goedert et al., 1981; Schwartz et al., 1982). If NGF were required for postnatal sensory survival, one would expect to observe a substantial neuronal death with antibody tratment. Since massive cell death did not occur, it would appear that mature sensory neurons require NGF for neuronal homeostasis and less so for survival. In total, the results of antibody experiments provide compelling evidence for the biological importance of NGF in both developing and matured sympathetic and sensory neurons.

ANTIBODY CYTOTOXICITY/COMPLEMENT CONTRIBUTION?

The destruction of sympathetic ganglia by NGF antiserum could be the result of complement activity or the consequence of insufficient NGF. Systemic administration of NGF up to 48 hr after antibody treatment can completely reverse the destructive antibody effects (Goedert et al., 1980). This tends to rule out a complement-mediated effect. In a more convincing experiment, sympathetic ganglia were analyzed after injection of anti-NGF serum into complement (C5)-deficient mice (DBA/2J): the antiserum caused a similar destruction of sympathetic ganglia (Ennis et al., 1979). Thus, it appears that the cytotoxic effect on sympathetic ganglia results from NGF neutralization and supports the conclusion that NGF is rquired for the survival of developing sympathetic and sensory neurons.

IN VIVO EFFECTS OF NGF

Systemically-administered NGF in neonates induces hypertrophy and hyperplasia in sympathetic ganglia. This accompanied increases in tyrosine hydroxylase (TH) and dopamine-β-hydroxylase (DβH) (Levi-Montalcini & Booker, 1960; Angeletti et al., 1971b; Thoenen et al.,

1971). The effects on the adult sympathetic ganglia are more modest: a less dramatic hypertrophy, no hyperplasia, and nominal increases in TH and DBH are observed (ibid; Olson 1967; Bjerre et al., 1975). Sympathetic neurons become less sensitive to NGF with age; the decline in NGF responsiveness in postnatal development coincides with the maturation of many sympathetic neurons (de Champlain et al., 1970). Thus, it would appear that NGF sensitivity decreases as a function of neuronal maturation. In early chicken embryos, NGF induces hyperplasia and hypertrophy of sympathetic ganglia which is also manifested by hyper-innervation (Levi-Montalcini & Hamburger, 1953). With regard to the sensory nervous system, NGF administration to newborn and adult rats causes a modest increase in substance P levels in the DRG (Otten et al., 1980; Kessler & Black, 1980; Goedert et al., 1981). Additionally, the decrease in substance P levels associated with administration of capsciacin (destroys peptidergic sensory neurons) to newborn rodents is muted by the concomitent injection of NGF (Otten et al., 1983). NGF injection into chick embryos reduced normal embryonic sensory (spinal ganglia) cell death (Hamberger et al., 1981). NGF has potent biochemical and morphological effects on developing sympathetic and sensory neurons and less so on mature neurons. Again, these observations are consistent with NGF promoting the survival and maintenance of developing and developed neurons, respectively.

SURVIVAL FACTOR IN VITRO

Levi-Montalcini and Angeletti (1963) showed that both developing chick sensory (7-10 days) and sympathetic (9-14 days) neurons cultured in Eagle's medium required exogenous NGF to survive. For sympathetic neurons, the longer the culture period with NGF, the better the survival

of neurons following NGF withdrawal. This is consistent with the observation that anti-NGF serum injected into neonatal rodents becomes less destructive for sympathetic ganglia with age and that sympathetic neurons become less sensitive to NGF in vivo. Cultured DRG derived from E8 chick embryos require NGF to survive; from this point to E18, the sensory ganglia became less dependent upon NGF for survival and show a proportional increase in the number of cells with neurite processes whose survival is independent of NGF (Greene, 1977b). The survival of neurons from E17 and E18 ganglia was not dependent upon NGF. The dispersed sensory neurons lose their NGF requirement for survival with increasing age. Similarly, the development of embryonic sympathetic ganglia (SCG) explants is dependent upon NGF from day 17 of gestation and not before (Coughlin et al., 1972). These in vitro results coupled with the in vivo experiments on systemic administration of NGF (yielding hypertrophy) and the in vivo antibody experiments support the general thesis that NGF promotes the survival of developing sympathetic and neural crest-derived sensory neurons.

NGF CHEMOTAXIS

The direction of axonal growth can be influenced by high concentrations of NGF. Gunderson and Barrett (1980) generated an NGF gradient in cultures of sensory neurons and observed neurons rapidly turned towards high concentrations of NGF. The higher concentrations of NGF in the gradient did not augment neurite growth or survival, suggesting that the response to the NGF gradient is chemotactic rather than trophic. An in vivo correlate of this experiment was performed by intracranial injection of NGF into newborn rats: sympathetic post-ganglionic fibers were found to be routed into the spinal cord and up the dorsal cord

toward the injection site (Menesini-Chen et al., 1978). Thus, neurons orient their axons towards high concentrations of NGF in vivo as well as in vitro. However, the pharmacologic doses used in these experiments and the fact that the levels of NGF are generally extremely low in vivo, suggest that these results may not represent a natural in vivo role of NGF. Further, it has been suggested that the chemotactic response may actually be a consequence of the adhesive interaction of NGF with the substratum (Lumsden & Davies, 1984).

NGF RECEPTORS

Similar to many hormones and other polypeptide effectors, NGF must first bind to extracellular receptors to be effective. Saturable NGF receptors have been characterized from responsive nerve cells and their membranes, as well as from human melanomas and rat pheochromocytoma (PC-12) cells (Herrup & Shooter, 1973; Banerjee et al., 1973; Herrup & Thoenen, 1979; Grob et al., 1985; Buxser et al., 1985). Scatchard analyses show two types of receptor binding, low ($Kd=1.7x10^{-9}M$ and high ($Kd=2.3x10^{-11}M$) affinity (Sutter et al., 1979). Immunoprecipitation of the NGF receptor shows bands of 70-85 kd and 180-200 kd. The NGF receptor is a glycoprotein, is phosphorylated on serine, and can be found as disulfide-linked dimers (Grob et al., 1985). cDNA and genomic clones were recently isolated corresponding to the low affinity ("fast") receptor (Chao et al., 1986; Johnson et al., 1986; Radeke et al., 1987).The high and low molecular weight receptor molecules correspond to the high and low affinity species, respectively. It is possible that there are two NGF receptor genes or that a 'converter' molecule associates with the low affinity subunit to transform it into the high affinity class (Landreth & Shooter, 1980; Schechter & Bothwell, 1981;

Green & Greene, 1986). The presence of a single copy gene and a single RNA species in cell types containing exclusively low or high affinity receptors argues against the former. Moreover, the receptors are related because they not only bind the same ligand but they share an epitope that is recognized by a monoclonal anti-receptor antibody (Ross et al., 1984; Chao et al., 1986). The possibility of a 'converter' subunit as a component of thei high affinity receptor may have a precedent: the high affinity IL-2 receptor appears to be converted from a low affinity form (Kondo et al., 1987). Interestingly, in both systems, the low affinity forms do not internalize the ligand (Hosang & Shooter, 1987). The high affinity NGF receptor is sensitive to mildly acidic conditions, suggesting that the receptor-ligand complex will dissociate in the acidic endosome. Early events associated with receptor binding include cell surface ruffling (Connolly et al., 1979) and a transient increase in c-fos RNA (Curran & Moragan, 1985; Milbrandt, 1986). NGF also induces TH, DβH and ornithine decarboxylase (Greene & McGuire, 1978; Feinstein et al., 1985). NGF receptors are found not only at the nerve terminals, but also along the nerve and on the perikaryon (Kim et al., 1979; Levi et al., 1980; Richardson & Riopelle, 1984; Carbonetto & Stach, 1982). The nerve terminal receptors presumably take up target-derived NGF; receptors located elsewhere on the nerve may take up NGF synthesized in accessory cells (Schwann cells) during nerve regeneration.

RETROGRADE TRANSPORT OF NGF

Once NGF is bound to its receptor on sympathetic and sensory neurons, it is transported in a retrograde fashion to the cell body. This was demonstrated in both sensory and sympathetic nerves by the

differential accumulation of iodinated NGF in the ganglia innervating the injection site (Johnson et al., 1978; Stockel et al., 1975a,b; Yip & Johnson, 1984). Specificity of transport was shown by the lack of retrograde transport of cytochrome C, which is similar in size and chemical charge to NGF. The retrograde transport of tetanus toxin with similar kinetics was the positive control. Transection of the innervating axons eliminated the preferential accumulation of label in the involved ganglia, thereby eliminating diffusion of label as an explanation for the accumulation. Further, autoradiography of labeled sensory ganglion showed radioactivity within the DRG neurons following injection of iodinated NGF (Yip & Johnson, 1984; Brunso-Bechtold & Hamburger, 1979). Biochemically, both substance P and TH levels were elevated in the appropriate ganglia by NGF injection (Goedart et al., 1981; Hendry, 1977). Thus, by a variety of criteria, NGF is taken up at its receptors and transported to the cell body in a specific and select manner.

Although NGF is transported to the cell body intact, it does not appear to be the intracellular effector because the direct introduction of NGF into (microinjection/fusion with NGF-laden erythrocyte ghosts) PC-12 cells did not induce nerve fiber outgrowth (Seeley et al., 1983; Heumann et al., 1981). Further, injection of NGF antibodies into PC-12 cells did not interfere with NGF-induced neurite outgrowth (Seeley et al., 1983). Thus, if these observation in PC-12 cells apply to nerves in vivo, then NGF exerts its effects via or in concert with its receptor. Since the NGF-receptor complex is apparently shuttled to the perikaryon, it may exert its activity here; alternatively, it may induce a secondary messenger upon receptor binding and is then transported to the cell body for processing and/or degradation. The high affinity receptor-associ-

ated 'converter' subunit might play a role in signal transduction (Hosang & Shooter, 1987).

SYMPATHETIC AXOTOMY RESCUE WITH NGF

The importance of NGF as a retrograde trophic factor is illustrated by its ability to prevent destruction of the perikaryia following axotomy. Surgical transection of postganglionic nerves, for example, causes degeneration of the sympathetic ganglia, as measured by loss of TH activity and decrease in neuronal cell number (Hendry, 1975c). NGF provided systemically restores to nearly normal the TH activity and the ganglia morphology. NGF also restores to normal choline acetyltransferase (ChAT) levels from preganglionic neurons that decrease in parallel with axotomized postganglionic nerves (Hendry, 1975a). NGF also rescues sympathetic ganglia that would otherwise degenerate following nerve terminal destruction by 6-hydroxydopamine (6-OHDA) (Levi-Montalcini et al., 1975). The destructive effects of axotomy are most marked with developing postnatal neurons than mature neurons. Postnatal axotomy leads to degeneration of the corresponding neurons in the ganglia. The effects on adult sympathetic ganglia are less marked and often do not result in massive degeneration so long as reinnervation of the target is unimpeded (Hendry, 1975b). These observations suggest that NGF can circumvent the deleterious effects of surgical and chemical axotomy on developing sympathetic ganglia by substituting for the lack of target-derived trophic support.

SENSORY AXOTOMY RESCUE

Sensory ganglia, which are pseudo-bipolar neurons with processes directed towards the spinal cord and periphery, also depend on NGF for trophic support. Both central (L5 dorsal root transection) and/or

peripheral (sciatic nerve crush) lesions led to dramatic decreases of L5 DRG neurons (Yip & Johnson, 1984). Further, substance P decreases dramatically in both sciatic nerve section and dorsal rhizotomy (Jessel et al., 1979). NGF treatment of central or central/peripheral lesions resulted in no loss of L5 DRG neurons (Yip & Johnson, 1984). Thus, as in axotomized sympathetic ganglia, NGF prevents the degeneration of sensory ganglia.

Since NGF can replace post-synaptically derived trophic substances NGF itself may be the principle agent responsible for the development and maintenance of neurons. Further, that neurons retrogradely transport NGF in both developing and mature neurons also suggests an important role for NGF. However, the effects of axotomy on developing neurons are more marked than in adults, demonstrating that NGF plays a more critical role in neuronal development. Similar results were observed in NGF antibody experiments: developing sympathetic and sensory ganglia are virtually destroyed by antibody exposure compared to very modest effects in mature ganglia (Angeletti et al., 1971). These observations are consistent with NGF as an indispensible survival factor for developing sympathetic and sensory neurons and as a maintenance factor for mature neurons.

NGF can also ensure survival of ganglia that would otherwise degenerate as a result of nerve terminal destruction by 6-hydroxydopamine (6-OHDA) (Levi-Montalcini et al., 1975). Since receptors are required for NGF uptake and intact nerve terminals are destroyed in both chemical (6-OHDA) and surgical axotomy, the NGF which supports ganglia survival under these conditions must enter via axonal or perikarya receptors. These receptors may, in fact, function principally during

regeneration to take up NGF presented by Schwann cells (Rush, 1984; Taniuchi et al., 1986).

NEURONAL CELL DEATH AND GUIDANCE

The responsiveness of embryonic neurons to NGF is consistent with target innervation and neuronal cell death but not with long-range neuronal guidance. Since neurite outgrowth from cultured mouse E9 trigeminal ganglia (sensory) towards its non-innervated target (epithelial cells of the maxillary process tissue) was unaffected by both anti-NGF antibodies and NGF protein, initial neurite outgrowth is independent of NGF (Lumsden & Davies, 1983,1986; Davies & Lumsden, 1984). Ganglia from later developmental stages (E10-E13) showed that neurite outgrowth was augmented by NGF. If NGF had a role in neuronal guidance, NGF responsiveness should appear before the nerves encounter the target. Since the appearance of NGF responsiveness correlates with the arrival of trigeminal ganglia nerves to their target and not before, a chemotactic role for NGF seems unlikely. Further supporting this contention is the observation that NGF protein and RNA can only be detected in the maxillary target beginning at E10.5, which coincides with initial nerve target contacts (Davies et al., 1987). The decrease in NGF responsiveness from E14 through postnatal day 4 coincides with normal neuronal cell death characterized by a loss of neurons and peripheral nerve fibers. Thus, the neurite outgrowth promoting activity of NGF is exerted coincident with the arrival of neurites to the target and continues for a time during which adjustments are made in the number of surviving neurons.

Cell death occurs in sensory and sympathetic neurons at characteristic times as part of their normal program of development. Cell

death is thought to be a consequence of competition for a limited amount of a target-derived trophic factor(s), such as NGF. For example, the injection of NGF into chick embryos significantly reduced normal embryonic sensory (spinal ganglia) cell death (Hamburger et al., 1981) and also reduced the increased cell loss associated with target removal (wing bud extirpation) (Hamburger & Yip, 1984). Similar observations apply to the sympathetic system where postnatal normal sympathetic neuronal cell death can be reduced substantially by systemically-provided NGF (Hendry & Campbell, 1976). NGF arrives at the cell bodies via retrograde transport and interference with this flow by axotomy results in the expected increased cell death in the involved ganglia; NGF administration can reverse this increased cell death. The effects of NGF in modulating cell death suggests that NGF itself could be the limiting survival factor for which the nerves compete.

CORRELATION: NGF PROTEIN AND SYMPATHETIC INNERVATION

The extent of sympathetic innervation of targets may be a reflection of the amount of NGF synthesized in the targets. Using a two-site enzyme immunoassay, Korshing and Thoenen (1983; 1985) observed a correlation between the density of sympathetic innervation and NGF levels in various target organs. NGF levels in sympathetic effector organs increase when their sympathetic nerve terminals are destroyed by 6-hydroxydopamine because NGF is no longer removed from the targets by retrograde transport. This supports the notion that innervated targets are providing the NGF trophic support for innervating neurons (Korshing & Thoenen, 1985).

BRAIN NGF

So far I have only examined the role of NGF in the peripheral

nervous system; NGF may also have a role in the brain. This is based on the following observations: (1) NGF has been detected in particular brain regions: the highest levels found in the regions innervated by the magnocellular cholinergic neurons (hippocampus, cortex and olfactory bulb) and the cell bodies of these neurons (septum, nucleus of the diagonal band of Broca, nucleus basalis) (Korshing et al., 1985). (2) Biochemical consequence of added NGF: Intracerebroventricular administration of NGF leads to increased ChAT levels in the cholinergic neurons of the basal forebrain (Gnahn et al., 1983; Mobley et al., 1986). Additionally, NGF addition to cultures of fetal rat striatum caused increased levels of ChAT (Martinez et al., 1985). This effect was abolished by anti-NGF antibodies and was not mimicked by physicochemically similar molecules (insulin, ferritin or cytochrome C). (3) Presence of a retrograde transport system: Identification of labeled basal forebrain cholinergic cell bodies following injection of iodinated-NGF or iodinated anti-NGF receptor monoclonal antibody into the cortex and hippocampus demonstrated retrograde transport (Schwab et al., 1979; Seiler & Schab, 1984; Taniuchi et al., 1986). (4) Identification of brain NGF receptors: Immunoprecipitation of rat brain NGF receptors showed a pattern identical to that observed using peripheral tissues (Taniuchi et al., 1986). While receptors were found throughout the brain, levels were higher (2-fold) in the medial septum, cerebellum and brainstem. The higher density of receptors in the aforementioned regions correlates fairly well with cholinergic forebrain neurons and the central processes of neural crest-derived sensory neurons. (5) Similar to peripheral axotomy rescue, NGF administration modestly attenuated degeneration of cholinergic septohippocampal neurons that

resulted from fornix-fimbria transections (Hefti, 1986; Kramer, 1987). The combined experimental results persuasively argue for an important function for NGF in some, but not all, central nervous system neurons.

THE NGF COMPLEX: BIOCHEMICAL CHARACTERIZATION

Mouse SMG NGF was characterized after purification of a high molecular weight NGF-containing complex. Neurite growth-promoting activity can be purified from the SMG at neural pH as a high molecular weight complex (~135 kd) called the 7S complex because of its sedimentation coefficient (Varon et al., 1968). This complex is comprised of three subunits with the following stoichiometry: $\alpha 2,\beta,\gamma 2$. All the NGF biological activity of the complex resides in the β subunit which is actually a homodimer of 26 kd (Server & Shooter, 1977). The 13 kd monomers are not held together by disulfides since only SDS is necessary to disrupt the components (Green et al., 1971). The α and γ subunits have been biochemically characterized and are members of the large (25-30) mouse kallikrein gene family (Shine et al., 1983). One to two gram atoms of zinc per complex also stabilizes the association of γ subunits in the complex (Server & Shooter, 1977).

NGF: MOLECULAR CHARACTERIZATION

A comparison of the amino acid sequences of mouse and cobra NGFs shows a modest homology to one another. Angeletti et al. (1973) characterized and sequenced proteolytic fragments of SMG NGF. It was determined to be a 118 amino acid protein with three intra-chain disulfide bridges and a molecular weight of 13,310. Cobra venom gland NGF was purified and characterized; it is also found as a homodimer (Server et al., 1976). The most favorable alignment of the 116 amino acid cobra NGF (Hogue-Angeletti et al., 1976) with the mouse sequence showed 64%

homology. Although cobra NGF elicits a qualitatively similar outgrowth of neurites in the chick embryo bioassay, it cannot displace about 20% of bound mouse NGF molecules. The 20% not displaced may represent mouse NGF bound to one of the two NGF receptors which is not competible by cobra NGF because of sequence divergence or may be due to some unknown receptor heterogeneity.

THE NGF COMPLEX: THE OTHER SUBUNITS

The 7S α subunit is an inactive serine protease. The amino acid sequence of α is consistent with the absence of catalytic activity: substitutions at the site of activation peptide removal and in the binding pocket may not allow maturation of the proenzyme and preclude correct substrate binding, respectively (Isackson & Bradshaw, 1984; Isackson et al., 1984; Evans & Richards, 1985). Since the α subunit has no catalytic activity, it has been suggested that its role may be to protect the β subunit from destruction by SMG proteases.

In contrast to α, the γ subunit is an active serine esteropeptidase. Purified γ subunit is catalytically active on the same synthetic arginine substrates as trypsin and is homologous to other characterized kallikreins (Server & Shooter, 1977). The activity of the γ subunit can be inhibited by α/β complexes, but not by α alone. NGF, at high concentrations, becomes an inhibitor of the γ subunit (Server & Shooter, 1977). However, when the terminal arginine is removed from NGF, β can no longer inhibit the activity of the γ subunit. Thus, the terminal arginine of β is required for the interaction with the γ subunit. The contribution of α to this reaction, if any, is unknown. To explain these observations in the context of complex formation, Server and Shooter have advanced the following hypothesis: NGF, which is

synthesized as a precursor with a carboxy-terminal extension, is cleaved by the γ subunit between the penultimate arginine and an unknown adjacent residue. Following catalysis, the γ enzyme remains bound to the processed NGF at its carboxy-terminal arginine. It is interesting that cobra NGF does not have a terminal arginine (Hogue-Angeletti et al., 1976) and as expected, does not complex with α and γ subunits to form a 7S equivalent (Server et al., 1976). Although NGF is found as a dimer in the venom gland, no higher molecular weight species have been detected.

NGF: SEXUAL DIMORPHISM AND CELLULAR LOCALIZATION

The levels of NGF are sexually dimorphic in the SMG. Using the chick embryo ganglia bioassay, Cohen (1960) quantitated NGF activity in the SMGs of developing male and female mice. Only trace levels of activity were detectable until about 17 days after birth. Subsequently, there were large increases until adult levels were achieved. The increase at 17 days coincides with the onset of puberty in mice. Thereafter, the levels were dimorphic, for higher specific activities were found in males compared to females. Thus, the levels of SMG NGF are either induced by androgens or repressed by estrogens. Experimental observations discussed below suggest the former is valid.

NGF has been localized in the glandular tubular cells (GTC) which along with acini represent the two morphological components of the SMG. The morphologic differentiation of the GTC is controlled by androgens (Chretien, 1977). Testosterone administration to females increases both tubule size and protease content, while castration of males decreases both parameters. NGF has been associated with the convoluted tubules based on the following observations: (1) Proteolytic enzymes (markers

for convoluted tubules) and NGF both appear in the SMG at puberty and are found at higher concentrations in males; tubules are undifferentiated in prepubescent mice (Caramia et al., 1962). (2) Both proteolytic enzymes and NGF increase and decrease similarly following testosterone administration to females and castration of males, respectively (Caramia et al., 1962). (3) Injection of anti-NGF antibodies into adolescents results in marked decreases in the tubules with regard to number, size and density of secretory granules. The epithelial cells of the tubules showed cytopathic effects following anti-NGF antibody administration, while acinar regions were virtually unmodified (Caramia et al., 1962). Goldstein and Burdman (1965) demonstrated via immunohistochemistry that NGF was localized to SMG tubule cells. Additionally, increased NGF levels in testosterone-treated females were manifested in increased staining of tissue sections with NGF antiserum and in the cytology of the tubule cells. These observations are consistent with the synthesis of NGF in the convoluted tubules of the SMG.

The cellular architecture of snake and mouse glands is similar. The principle secretory cells in the snake venom gland are epithelial cells that line the tubules, as in the mouse SMG (Hogue-Angeletti & Bradshaw, 1978). They contain secretory granules of a size and density comparable to those found in the mouse SMG. Therefore, it is likely that a similar cell type in these two evolutionarily distinct species synthesize NGF.

NGF SYNTHESIS IN THE SMG

The testosterone-mediated increase in SMG NGF is due to an increased rate of NGF synthesis. Ishii and Shooter (1975) calculated that testosterone injection into females increased the NGF content

20-fold (by RIA) and that conversely, castration produced a 96% decrease. By following the kinetics of immune-reactivity after castration and subsequent testosterone administration, they concluded that the increased NGF concentration is due to a 10-fold increase in the rate of synthesis and not to a decrease in the rate of degradation.

NGF is synthesized in the SMG and is androgen-responsive. Intravenous administration of tracer levels of iodinated NGF failed to accumulate in the SMG, suggesting that NGF is not taken up from the circulation and stored there (Ishii & Shooter, 1975). Burger and Shooter (1978) addressed the issue more directly by immunoprecipitation of de novo synthesized protein with anti-NGF serum. Specifically, immunoprecipitation of labeled extracts from cultured SMG slices revealed a labeled molecule of 13 kd that co-migrated with authentic NGF on SDS gels. Moreover, these two species had the identical tryptic peptide map. When using SMGs from castrated males, fewer labeled molecules of NGF were recovered; testosterone treatment increased the recovery of NGF to a level comparable with that of normal males. Thus, NGF is synthesized de novo in the SMG and the modulation of its synthesis by castration and subsequent testerosterone administration is suggestive of androgen regulation.

NGF PRECURSOR

Immunoprecipitation of de novo synthesized SMG proteins with affinity purified anti-NGF antibodies revealed two NGF species (13 kd and 22 kd), demonstrating that NGF is derived from a higher molecular weight precursor (Burger & Shooter, 1977). Pulse and pulse-chase labeling experiments showed that the 13 kd NGF is derived from the 22 kd precursor. Moreover, shared tryptic fragments between the 22 kd and 13

kd molecules supports the conclusion that SMG NGF is derived from a 22 kd precursor molecule.

PRECURSOR PROCESSING

The γ subunit appears to liberate mature NGF from the precursor. Burger and Shooter (1977) immunoprecipitated a labeled 22 kd precursor from anti-NGF antibodies and incubated it with purified γ subunit. After incubation, all the label was recovered at a molecular weight of 13 kd. Although the substrate remained bound to the antibody during the γ subunit digestion, it was apparently processed correctly into mature NGF. However, the γ subunit may not be uniquely required for precursor processing because purified EGF binding protein also catalyzed the same conversion.

The distribution of SMG NGF may reflect intracellular NGF precursor processing. Immunoreactive NGF was localized exclusively to secretory granules in the pyramidal secretory tubule cells using affinity purified NGF antibody coupled to horseradish peroxidase and formalin fixation (Schwab et al., 1976). Moreover, greater staining was observed in the apical part and in the duct lumen relative to the lateral part of the cells. Acinar cells were devoid of immunoreactivity. These observations are consistent with processing of the precursor while it is confined to the secretory granules. Mowry et al. (1984) localized, by immunocytochemistry, the α and γ subunits to the same secretory vesicles as β NGF. Thus, precursor processing and complex formation may occur in the same granules in the SMG.

STIMULATION OF NGF SECRETION

Secretagogues can discharge NGF from the convoluted tubules of the SMG. α-adrenergic (phenylephrine) and β-adrenergic (isoproterenol)

agents as well as a cholinergic stimulator (pilocarpine) caused, respectively, a marked, modest and no loss of immunoreactive granules from the convoluted tubules (Simson et al., 1978). This suggests that adrenergic nerves, particularly α-adrenergic, control the discharge of NGF from the tubule cells. Kallikreins and EGF are also discharged from the convoluted tubules by α-adrenergic agents. Burton et al. (1978) identified 7S NGF as the major form of NGF recovered from saliva following α-adrenergic stimulation. The 7S complex was very stable, most likely owing to its relatively high concentration (3.4 mg/ml) and the high levels of Zn^{++}, which stabilize the complex (Server & Shooter, 1977). These observations are consistent with previous data showing that α, β and γ subunits are found in the same secretory granules (Mowry et al., 1984).

PHYSIOLOGICAL ROLE OF THE SMG NGF

The high levels of NGF in the mouse SMG have no obvious physiological value. The following observations argue against a role for salivary gland NGF in the nervous system of mice: (1) Male and female mice have sympathetic nerve cells of identical size (Levi-Montalcini & Angeletti, 1966) and additionally, have the same specific activity of tyrosine hydroxylase (Hendry & Iversen, 1973). (2) NGF from the SMG appears to be secreted only into the saliva, not into the bloodstream. (3) Extirpation of the SMG has no effect on blood NGF (Levi-Montalcini & Booker, 1960) and on sympathetic neurons other than those which innervate the SMG itself (Levi-Montalcini & Angeletti, 1968). (4) Other rodents do not have similarly high levels of SMG NGF, with the exception of the African rodent Praomys (Mastomys) (Aloe et al., 1981). The guinea pig prostate also produces high levels of NGF (Harper et al.,

1979); its function is also unknown. The SMG also produces high levels of epidermal growth factor which may have a wound healing role; NGF might somehow participte in this process.

RESEARCH GOALS

The present level of understanding is largely a reflection of biologically and biochemically descriptive experiments using NGF antibodies as a probe. Molecular biological approaches should broaden the scope of available reagents and allow for more refined analyses. Specifically, the isolation of NGF cDNAs, elucidating the structure of the precursor(s), and determination of the corresponding gene structure should provide insight into the expression, processing and potential regulation of the gene.

The research described below addresses the following questions:

1. What is the size and nature of the precursor protein predicted from an NGF cDNA clone? From the predicted precursor structure, can one infer how mature NGF is derived? Further, using the clone as probe, it can be determined if the sexual dimorphism exists in NGF mRNA.

2. What is the structure of the mouse NGF gene? Can the organization of the gene account for all the NGF transcripts, i.e., are there other NGF-like genes? Are consensus splice sites and standard promoter elements present? How does the organization of this gene compare to the human gene?

3. How complex is the expression of the gene? Specifically, how many NGF-specific transcripts are present? Do they vary during development or in different tissues?

4. Can full-length NGF precursors be processed into mature NGF by the γ subunit? What is the specificity of this reaction?

5. A comparison of the evolutionarily distant cobra NGF cDNA sequence with that of mouse should allow for the identification of important residues or domains and thus to infer the biological importance of particular regions, especially the non-NGF domains. Further, primer extension analysis should determine if multiple transcripts are present in the cobra venom gland; from this data, inferences can be made with regard to a potential role for the transcript(s).

CHAPTER I

ISOLATION AND NUCLEOTIDE SEQUENCE OF A cDNA ENCODING THE PRECURSOR OF MOUSE NERVE GROWTH FACTOR

Reprinted from Nature, Vol 302, No 5908, pp 538-540, 7 April 1983

Isolation and nucleotide sequence of a cDNA encoding the precursor of mouse nerve growth factor

James Scott*†, Mark Selby*, Mickey Urdea‡, Margarita Quiroga‡, Graeme I. Bell‡ & William J. Rutter*

* Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143, USA
‡ Chiron Corporation, 4560 Horton Street, Emeryville, California 94608, USA

Nerve growth factor (NGF) is a polypeptide that enhances survival, nerve fibre outgrowth and neurotransmitter biosynthesis in sympathetic and sensory neurones[1-3]. Administration of antibodies against NGF to developing animals leads to atrophy of the sympathetic system[4]. NGF is not normally detectable in innervated tissues but ablation of the innervating neurones leads to the production of measurable NGF in the target tissue[5]. After transplantation of the denervated tissue, reinnervation occurs, then NGF decreases to undetectable levels. Thus NGF seems to act as a neurotrophic messenger and its level is regulated by innervating neurones. Because of the minute levels present it is very difficult to study NGF biosynthesis in innervated tissue. However, NGF can be isolated from male mouse submaxillary glands, where it exists in inexplicably high levels[6]. Its amino acid sequence has been determined[7], and the synthesis of NGF and its larger precursors has been demonstrated in cultured submaxillary glands[8]. We report here the nucleotide sequence of a submaxillary cDNA encoding the mouse NGF precursor (preproNGF). In contrast to previous suppositions[9] the NGF moiety is situated near the carboxy-terminus of the polyprotein precursor. It is flanked at the amino-terminus by 187 amino acids which may be cleaved at dibasic residues to generate three peptides; there are only two additional amino acids at the carboxy-terminus.

NGF-specific clones were isolated from a male submaxillary gland cDNA library by screening with synthetic oligodeoxynucleotide probes synthesized according to the nucleotide sequence predicted from the NGF amino acid sequence. To construct the library, poly(A)-containing RNA was isolated from the submaxillary glands of 60-day-old male Swiss-Webster mice[9,10]. Double-stranded cDNAs were prepared and inserted into the *Pst*I site of a pBR322 derivative (given by Dr R. A. Hallewell, Chiron Corporation) using the G-C tailing technique[11]. Tetracycline-resistant transformants of *Escherichia coli* HB101 were stored at −70 °C in microtitre dishes; 5,000 of the transformants were grown on Whatman 541 filter paper, the plasmids were amplified *in situ* with chloramphenicol and the DNA immobilized on the filters[11,12]. The oligonucleotide probe sequences used were constructed from the hexapeptide having the least degenerate codons within the NGF sequence: Gln-Tyr-Phe-Phe-Glu-Thr (excluding the third base of Thr). A mixture of all 32 possible 17-base oligonucleotides(3′-GT$^{T}_{C}$AT$^{A}_{G}$AA$^{A}_{G}$AA$^{A}_{G}$CT$^{T}_{C}$TG-5′) was synthesized by using solid-phase phosphoramidite methodology followed by isolation from 20% acrylamide gels (ref. 13 and M. Urdea *et al.*, in preparation). ^{32}P-end-labelled probes were used to screen the cDNA library[14]. Six colonies yielded strong signals with this probe; one contained a plasmid with an insert of ~1,100 base pairs (bp). This insert was used to rescreen the library and revealed one additional hybridizing colony. The relative abundance of NGF clones in the library (7 in 5,000) suggests that NGF mRNA comprises ~0.1‰ of the mRNA from this source.

Most of the amino acids encoded by NGF mRNA were deduced from the nucleotide sequence of the longest cDNA clone, which includes 1,068 bases complementary to NGF mRNA (Fig. 1, nucleotides 109–1,176), a 50-base poly(A) tract, in addition to the poly(dC) and poly(dG) tails[15]. The cDNA was used as a probe to establish the size of the NGF mRNA in both male and female submaxillary glands. Figure 2 shows that it is ~1,350 bases and the relative level of NGF mRNA is about 10-fold more abundant in the male. Thus, the well

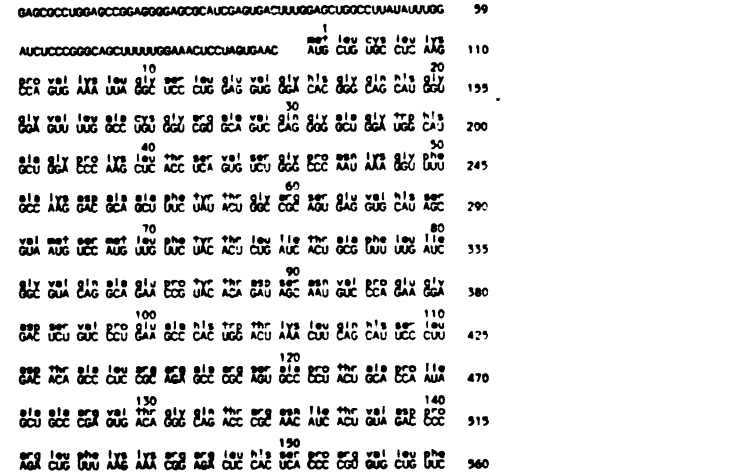


Fig. 1 Nucleotide sequence of the mouse submaxillary gland NGF mRNA and the derived amino acid sequence of the NGF precursor. The NGF mRNA sequence was determined by sequencing both strands of a 1,068-bp cDNA insert and from the sequence of cDNA generated by a primer extension reaction[16]. The number of nucleotides in the cDNA is indicated at the end of each line. The predicted amino acid sequence of mouse preproNGF is numbered by designating the first methionine as 1. The sequence corresponding to the mature NGF is labelled and underlined.

† Present address: Molecular Medicine Group, MRC Clinical Research Centre, Harrow, Middlesex HA1 3UJ, UK.

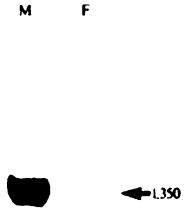


Fig. 2 Size determination of NGF mRNA from male and female mice. 10 μg of glyoxylated RNA obtained from male and female submaxillary glands were separated on a 2% agarose gel, transferred to a nitrocellulose filter and hybridized with nick-translated ^{32}P-labelled NGF cDNA insert[21]. After washing, the RNA blot was autoradiographed at −70 °C using an intensifying screen. Lane M, male RNA; F, female RNA. Glyoxylated *Hae*III-digested ΦX 174 fragments were used as size markers.

known difference in the levels of NGF protein between males and females is due to regulation at the level of mRNA. The size of the NGF mRNA suggests that the cDNA is missing 100–150 bp at the 5′ end. The remainder of the sequence was determined from a cDNA made by primed synthesis on the male mRNA using a ^{32}P-labelled 15-base oligonucleotide, 3′-CCACCTCAAAACCGG-5′, which is complementary to the mRNA segment encoding nucleotides 153–167 (Fig. 1)[18]. The sequence of the extended primer confirmed the region in common with the cDNA and provided an additional 108 bases. This RNA has a 5′ untranslated region of 100 bases, a single open reading frame which starts at the first methionine codon (nucleotides 96–98), and ends with the termination codon TGA (nucleotides 1,117–1,119). This nucleotide sequence predicts a 307-amino acid polypeptide which contains NGF (amino acids 188–305). The terminal 91 bases of the 160-base 3′ untranslated region is exceptionally A+T-rich (80% A+T); the variant polyadenylation signal, AUUAAA, is found 22 nucleotides from the poly(A) tail[17].

In common with other secreted proteins, NGF probably has an amino-terminal sequence which signals the translocation of the nascent protein into the cisternae of the endoplasmic reticulum (Fig. 3)[18]. Such signal peptides are usually 15–30 amino acids long, contain a central hydrophobic region and terminate in an amino acid possessing a small neutral side chain. The amino-terminal region of the NGF precursor contains two weakly hydrophobic regions (residues 7–14 and 22–23). The boundary of the signal peptide cannot be predicted from the amino acid sequence, but it will be evident when the amino-terminus of proNGF is determined. The predicted molecular weights of preproNGF (M_r = 33,800) and proNGF (M_r = 30,000–32,000) are in reasonable agreement with the values of ~36,000 and 29,000, respectively obtained for putative higher molecular weight precursors by T. Darling and E. Shooter (personal communication). The amino acid sequence predicted from the cDNA agrees with the sequence of Angeletti *et al.*[7] except at residue 30, in which position our sequence predicts an aspartate. Angeletti *et al.*[7] obtained equivocal evidence supporting the assignment of Asp or Asn at this residue, they chose Asn because of the possibility of deamidation. Therefore, it seems likely that this difference is due to an amino acid sequencing error, but the possibility of a DNA

Formation of NGF from its precursor requires proteolytic processing at both ends (Fig. 3). Cleavage sites are: Arg-Arg, 115–116; Lys-Lys-Arg-Arg, 144–147; Lys-Arg, 186–187 at the amino-terminus of NGF; and Arg-Arg, 305–306 at its carboxy-terminus. The residues Arg-Lys, 301–302, exist within the NGF sequence itself, but these residues are not apparently cleaved; they may be stabilized by the diacidic residues Asp-Glu at positions 280–281. Cleavage at Arg-Arg, 115–116 and Lys-Lys-Arg-Arg, 144–147 would result in NGF-containing moieties of M_r 22,000 and 18,000, respectively.

The extent and ordering of these possible processing reactions is unknown, but Berger and Shooter[8] have found a relatively stable 22,000-M_r NGF precursor, which must represent active cleavage at residues 115–116. Cleavage at all processing sites in the amino-terminal region would yield three peptides of ~90 amino acids (residues 25–114), 27 amino acids (residues 117–143), and 38 amino acids (residues 148–186). A computer search of a compendium of protein sequences revealed no homology between these peptides and other known sequences[19]. The functional role of these cryptic peptides is unknown, but these sequences obviously cannot represent the 26,000-M_r α- and γ-subunits of the 7S complex[1]. Perhaps proNGF is a member of the growing class of polyproteins that are processed into functionally distinct entities. The availability of the sequence of these peptides will allow direct examination of their biological properties.

There are several arginine-specific esteropeptidases in the submaxillary gland which could perform the proNGF processing reactions; one of these is the γ-subunit of the 7S complex. This esteropeptidase is bound to the carboxy-terminal arginine of the NGF monomer, and is probably the enzyme involved in processing the two carboxy-terminal amino acids. It seems likely that this esteropeptidase cleaves the amino-terminal sites also.

As reported previously, the amino acid sequence of mature NGF is approximately 25% homologous with proinsulin[20]. Alignment of the corresponding mRNA segments indicated that the nucleotide sequences possess 40% homology. However, the overall structure of the cDNAs and the positioning of the processing sites in the precursors are quite different. Thus, convergent evolution is as attractive as divergent evolution in explaining the structural resemblance between these molecules. The location of intervening sequences in the gene structure may help to decide this issue.

The availability of cDNA probes will facilitate the unambiguous determination of the sites of synthesis of NGF and the study of NGF gene expression during development, including the puzzling role of androgens in regulating the expression of this gene in the submaxillary gland. In addition, now it should be possible to isolate related genetic sequences from the same or different species, and to examine the role of NGF in diseases of the human nervous system.

We thank Drs Tom Darling and Eric Shooter for helpful discussions and communication of unpublished results, Drs Charles Craik and Richard Najarian for assisting with computer searches, Noel Fong for technical assistance, and Maureen Appling and Leslie Spector for typing the manuscript. J.S. was supported in part by the EMBO. This research was supported by NIH grant 21344 (to W.J.R.).

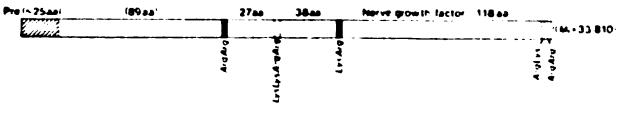


Fig. 3 Schematic representation of the mouse submaxillary preproNGF polyprotein. The basic di- and tetrapeptide possible cleavage sites are indicated; the sites at which cleavage is known to occur are indicated by solid boxes. The numbers in parentheses are tentative size designations; the others give accurate sizes.

Received 21 December 1982; accepted 27 January 1983

1. Bradshaw, R. A. *A. Rev. Biochem.* **47**, 191–216 (1978).
2. Thoenen, H. & Barde, Y. A. *Physiol. Rev.* **60**, 1284–1335 (1980).
3. Yankner, B. A. & Shooter, E. M. *A. Rev. Biochem.* **51**, 845–868 (1982).
4. Levi-Montalcini, R. & Booker, B. *Proc. natn. Acad. Sci. U.S.A.* **46**, 384–391 (1960).
5. Ebendal, T., Olson, L., Seiger, A. & Hedlund, K. O. *Nature* **286**, 25–28 (1980).
6. Cohen, S. *Proc. natn. Sci. U.S.A.* **46**, 302–311 (1960).
7. Angeletti, R. H., Hermodson, M. A. & Bradshaw, R. A. *Biochemistry* **12**, 100–115 (1973).
8. Berger, E. A. & Shooter, E. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3647–3651 (1977).
9. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
10. Goodman, H. M. & MacDonald, R. J. *Meth. Enzym.* **68**, 75–90 (1980).
11. Gergen, J. P., Stern, R. H. & Wensink, P. C. *Nucleic Acids Res.* **7**, 2115–2136 (1979).
12. Ish-Horowicz, D. & Burke, J. F. *Nucleic Acids Res.* **9**, 2989–2998 (1981).
13. Beaucage, S. L. & Caruthers, M. H. *Tetrahedron Lett.* **22**, 1859–1862 (1981).
14. Wallace, R. B. *et al.* *Nucleic Acids Res.* **9**, 879–894 (1981).
15. Maxam, A. M. & Gilbert, W. *Meth. Enzym.* **65**, 499–560 (1980).
16. Argos, K. L., Brunstedt, J. & Noyes, B. E. *J. biol. Chem.* **256**, 1023–1029 (1981).
17. Hagenbuchle, O., Bovey, R. & Young, R. A. *Cell* **21**, 179–187 (1980).
18. Lingappa, V. R. & Blobel, G. *Recent Prog. Horm. Res.* **36**, 451–475 (1980).
19. Dayhoff, M. O. *Atlas of Protein Sequence and Structure* (National Biomedical Research Foundation, Silver Spring, Maryland, 1978).
20. Frazier, W. A., Angeletti, R. & Bradshaw, R. A. *Science* **176**, 482–488 (1972).
21. Thomas, P. S. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).

APPENDIX TO CHAPTER I

These figures represent data that were either referred to in the Nature paper or are pertinent. Specifically, Figures 1-3 show the oligonucleotide primed reverse transcription product, its specificity, length and sequence. Figures 4 and 5 are genomic Southern blots using probes which represent virtually the entire mouse NGF cDNA. The cDNA probe used in Figure 4 was not full-length, thus prompting hybridization with the missing 5' end of the cDNA (Fig. 5). These data demonstrate that the NGF gene is most likely single copy.

FIGURE 1

Submaxillary gland primer extensions. 5 μg of polyA$^+$ RNA from male (M) and female (F) submaxillary glands were annealed to a ^{32}P-end labeled 16-base oligonucleotide primer (Scott et al., 1983) and extended with reverse transcriptase in the presence of deoxyribonucleotides at 42°C. The reactions were electrophoresed on an 8% sequencing gel. Labeled standards (HpaII-pBR) were electrophoresed in the adjacent lane. Arrows denote male-specific extension products. The extension product at ~170 bases represents full-length product while that at ~98 bases is most likely a prematurely terminated species. This primer extension experiment was undertaken to determine the number of nucleotides that were not represented in the original cDNA clone. The sexual dimorphism in extension products defined the specificity of the reaction since more (more than 10 fold) NGF RNA is found in male than female glands.

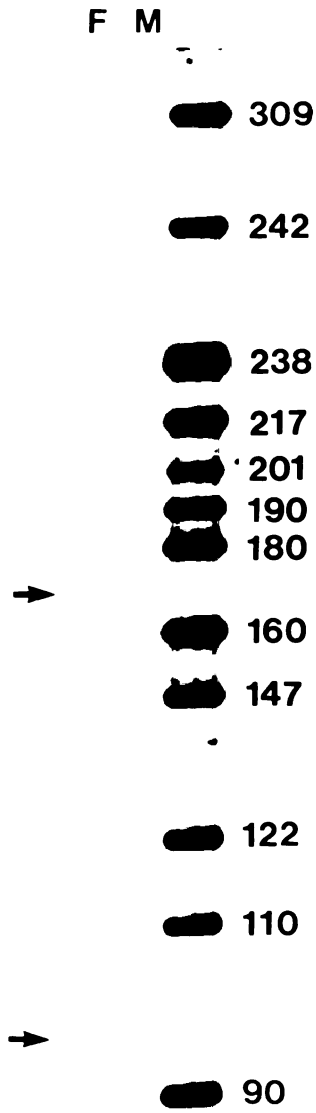

F M
309
242
238
217
201
190
180
160
147
122
110
90

FIGURE 2

Sizing of primer extension products. An end-labeled oligonucleotide (Scott et al., 1983) was annealed to 34 µg of polyA$^+$ RNA from male submaxillary glands. Two different amounts of the reaction products were loaded on 8% sequencing gels next to Maxam-Gilbert sequencing reactions. The full-length extension product of 170 bases is signified by an arrow as is a presumed premature stop at 98 bases.

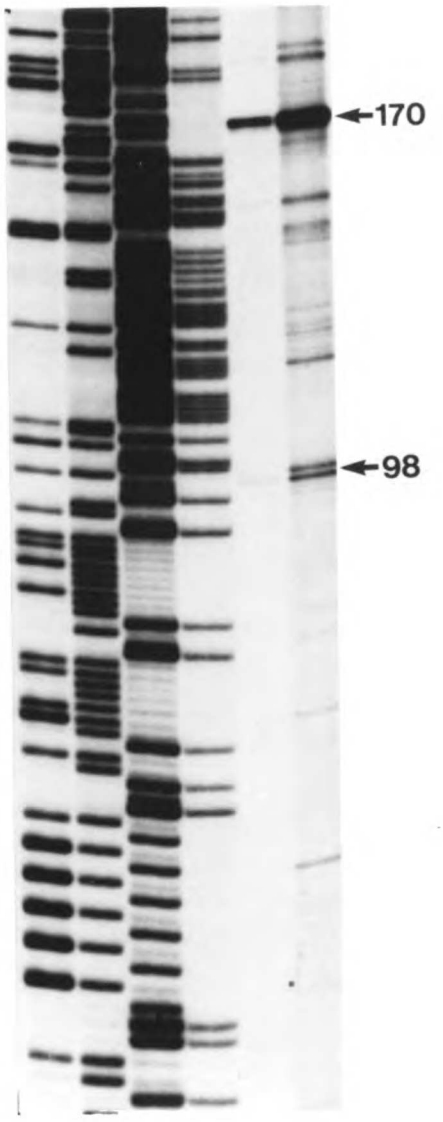

170
98

FIGURE 3

Primer extension sequencing. A preparative primer extension reaction using an end-labeled oligonucleotide (Scott et al., 1983) was electrophoresed on an 8% sequencing gel. The 170 base extension product was cut out, eluted and ethanol precipitated in the presence of salmon sperm DNA as carrier. The DNA was dissolved in water, aliquoted and sequenced (Maxam & Gilbert, 1980). The reactions were electrophoresed on 5% sequencing gels and then exposed to film. The corresponding nucleotide sequence is presented along side the sequencing reactions. This sequence accounts for the 5' end of the mRNA that was not present in the cDNA clone.

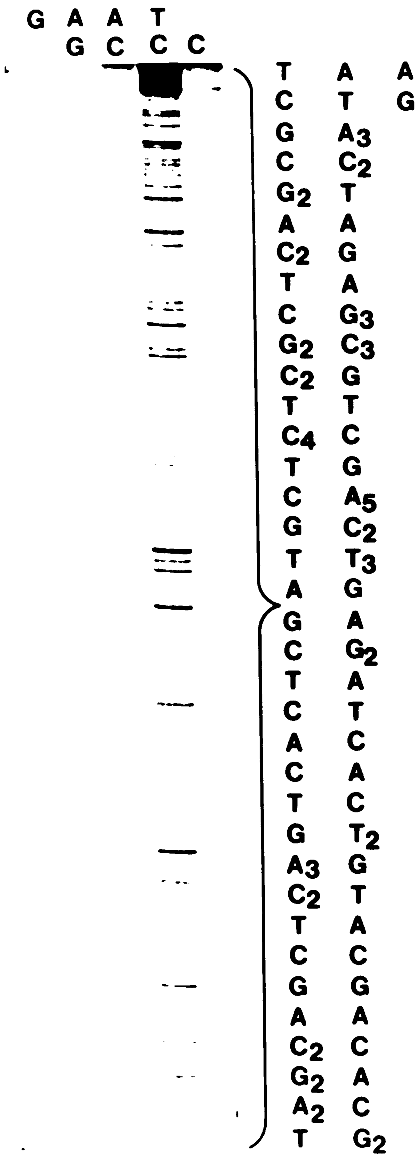

G A A T
G C C C
T C G C G_2 A C_2 T C G_2 C_2 T C_4 T C G T A G C T C A C T G A_3 C_2 T C G A C_2 G_2 A_2 T
A T A_3 C_2 T A G A G_3 C_3 G T C G A_5 C_2 T_3 G A G_2 A T C A C T_2 G T A C G A C A C G_2
A G

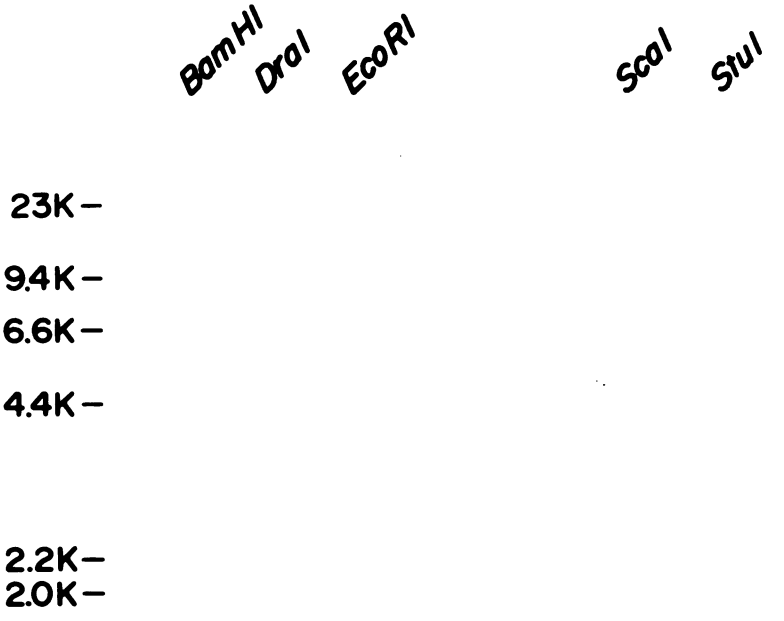
FIGURE 4

Genomic hybridizations. 10 µg of genomic DNA from mouse, human, monkey and rat were digested with EcoRI, fractionated on a 0.9% agarose gel, and transferred to nitrocellulose paper. The filter was hybridized with ^{32}P-NGF cDNA using 30% formamide, 5XSSC, 2X Denhardt's and 100 µg/ml denatured salmon sperm DNA at 42°C. The blot was washed at 45°C in 2XSSC and exposed at -70°C with an intensifying screen. The pattern of hybridization is not complex and suggests that the NGF gene is single copy. The isolation and characterization of NGF genomic phages supports this conclusion (Chapter IV).

Mouse
Human
Monkey
Rat

FIGURE 5

Oligonucleotide genomic hybridization. 10 µg of mouse genomic DNAs were digested with restriction enzymes, electrophoresed on a 0.9% agarose gel and transferred to nitrocellulose. The blot was probed with ^{32}P-labeled 32mer oligonucleotide representing the first exon that is not found in the NGF cDNA. The hybridization conditions were as prescribed by Wood et al. (1985) and the blot was washed in trimethylammoniumchloride at 70°C. The blot was exposed to film with an intensifying screen at -70°C. The restriction enzymes listed on the top of the figure correspond to the enzyme used in that particular lane. The numbers on the left represent the sizes of HindIII-digested λ DNA. The hybridizing bands correspond to restriction fragments containing DNA complementary to the 32mer oligonucleotide. This 32mer sequence is not found in the original NGF cDNA because it is not full-length; the simple pattern confirms that the gene is single copy. Note that the EcoRI fragment in this film does not appear on the EcoRI digest/hybridization using the cDNA as probe. This suggests, as was later confirmed (Chapter IV), that the corresponding exons are situated far away from one.


BamHI
DraI
EcoRI
ScaI
StuI
23K
9.4K
6.6K
4.4K
2.2K
2.0K

CHAPTER II

DIFFERENTIAL RNA SPLICING PREDICTS TWO DISTINCT NERVE GROWTH FACTOR PRECURSORS

Reprinted from Nature Vol. 319, No. 6056, pp. 784-787, 27 February 1986

Differential RNA splicing predicts two distinct nerve growth factor precursors

R. H. Edwards*†, M. J. Selby* & W. J. Rutter*‡

* Hormone Research Institute and Department of Biochemistry & Biophysics, and † Department of Neurology, University of California, San Francisco, California 94143, USA

Nerve growth factor (NGF) has a crucial role in the development of sensory and sympathetic neurones[1-6]. However, although it can affect other neural cell types under certain experimental conditions[7], no biological role has been convincingly demonstrated elsewhere in the nervous system. The 5′ end of the mouse NGF gene contains several relatively short exons[8]. The NGF messenger RNA contains two in-frame initiator methionine codons; the second precedes the signal peptide sequence. Studies of the translation of other eukaryotic mRNAs indicate that the first AUG is preferred, suggesting that the signal for secretion might be ambiguous. We have analysed the NGF mRNA species from various cell types, some of which (clonal myoblast and fibroblast cell lines) are known to secrete NGF[9,10], to search for different NGF transcripts. One pathway of RNA splicing generates the transcript already described from a submaxillary gland complementary DNA clone[11]. We demonstrate here that there is another splicing pathway, leading to a shorter transcript that lacks the second exon. This short transcript is the major form in most other mouse tissues and in the tissues of several other species, but both transcripts are usually present. In the short transcript, the initiator methionine is immediately upstream from a signal peptide-like sequence whereas in the long transcript the first methionine is 62 amino acids upstream from the signal peptide-like sequences. This may result in a different cellular localization of the NGF or alter the biological activity of the NGF precursor.

In the early experiments we found that the mouse G8 embryonic myoblast cell line contains NGF mRNA by Northern blotting. Using the cloned mouse submaxillary gland NGF cDNA[11] as a probe, we observed after high stringency washing a band representing about 0.001% of the cell mRNA and migrating at about 1,300 bases, approximately the size of NGF transcripts[12]. However, the 5′ end of the NGF gene contains several small (32-127-nucleotide) exons (M.J.S., in preparation), thus alternative splicing might yield mRNAs of slightly differing size that are not resolved on these gels. We therefore tested, using S_1 nuclease digestion, whether NGF mRNAs contained divergent 5′ ends. A 5′ end-labelled fragment extending from the *Nco*I site at +616 in the cDNA into the plasmid sequences flanking the insert was hybridized to poly(A)$^+$ RNA from G8 myoblast cells. After digestion with S_1 nuclease, several protected fragments were detected (Fig. 1). The longest protected fragment corresponds to the original submaxillary gland cDNA clone[11]; the shortest fragment, greatest in intensity, corresponds to a fragment cleaved very close to the splice junction at position +158. We observe a similar set of protected fragments in experiments using total RNA from female mouse submaxillary gland, but the longest fragment predominates.

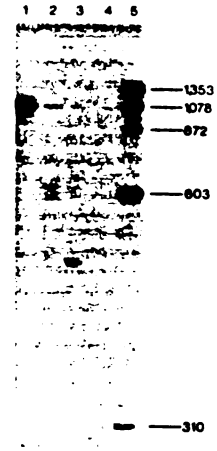


Fig. 1 S_1 nuclease analysis of the 5′ end of NGF mRNA. Lane 1, undigested probe; lane 2, 20 μg female mouse submaxillary gland total RNA; lane 3, 20 μg poly(A)$^+$ G8 cell RNA; lane 4, 20 μg yeast transfer RNA; lane 5, *M*, standards
Methods. RNA was prepared in guanidinium isothiocyanate[29], precipitated in lithium chloride[29], extracted and reprecipitated in ethanol. Poly(A)$^+$ RNA was selected on oligo(dT)-cellulose[30]. A NGF cDNA clone in pBR328 was digested with *Nco*I (+616), phosphatased and 5′ end-labelled with [γ-^{32}P]ATP and polynucleotide kinase. The DNA was recut with *Hin*dIII and the 5′ NGF fragment isolated on an 8% acrylamide gel. The probe (5,000 c.p.m.) was heated with the RNA in 80% formamide at 85 °C for 15 min, then hybridized at 45 °C for 3 h. S_1 nuclease digestion with 5 U of enzyme (PL Biochemicals) was performed at 37 °C for 1 h. The samples were extracted, precipitated twice and loaded onto a 5% denaturing acrylamide gel[31].

‡ To whom correspondence should be addressed

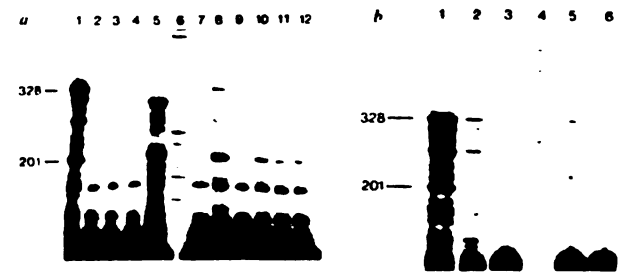


Fig. 2 **a**, Primer extension assays of 20 μg female mouse submaxillary gland total RNA (lane 1), 20 μg female rat submaxillary gland total RNA (lane 2), 20 μg mouse placental total RNA (lane 3), 20 μg rat placental total RNA (lane 4), 20 μg guinea pig prostate total RNA (lane 5), *Hpa*II digest of pBR322 (lane 6), 20 μg fused G8 cell total RNA (lane 7), 1 μg L-cell poly(A)$^+$ RNA with 19 μg tRNA (lane 8), 1 μg poly(A)$^+$ RNA from NIE-115 cells (which lack NGF transcripts by Northern blotting) with 19 μg tRNA (lane 9), and 20 μg total RNA from mouse hind limb at 17 days of gestation, birth and 2 days later (lanes 10–12). **b**, Primer extension assays using 10 μg total RNA from male (lane 1) and female (lane 2) mouse submaxillary glands, 20 μg tRNA (lane 3), 5 μg poly(A)$^+$ RNA from human placenta (lane 5) and 15-day rat brain (lane 6), standards in lane 4 are a *Hae*III digestion of φX174 DNA
Methods. RNA was prepared as for Fig. 1. The primer was prepared by annealing a 22-nucleotide primer (complementary to +328 to +306 in the original cDNA) to an M13 template containing NGF, extending with the Klenow fragment of DNA polymerase I in the presence of all four α-labelled deoxyribonucleotides, digesting with *Sph*I (which cleaves at position +199) and separating the strands on a 5% acrylamide-urea gel[31]. The 129-base, single-stranded primer (1.5×10^5 c.p.m. having a specific activity $\sim 10^9$ c.p.m. per μg) was then annealed to the RNA and extended with reverse transcriptase according to the procedure of McKnight and Kingsbury.

Primer extension experiments verified the presence of the short and long transcripts and several minor extension products in several tissues. In G8 myoblasts, the strongest band corresponds to a transcript which initiates about 35 bases 5' to the splice junction at +158 (Fig. 2*a*, lane 8), thus being shorter than and different from the original cDNA. In submaxillary gland RNA the most intense band is the longest and corresponds to the size expected for the original cDNA clone (Fig. 2*a*, lane 1). Similar experiments demonstrate the presence of short and long transcripts in various tissues from mouse, rat, guinea pig and human. The short NGF transcript is the predominant one in mouse L cells, placenta and in hindlimb at several stages of early development (Fig. 2*a*, lanes 8, 3, 10–12), in rat submaxillary gland, placenta (Fig. 2*a*, lanes 2, 4) and developing brain (Fig. 2*b*, lane 6). However, the short and long transcripts occur in about equal concentrations in human placenta (Fig. 2*b*, lane 5). The additional bands present especially in the submaxillary RNA probably represent stops to reverse transcription. The finding of relatively high levels of NGF mRNA in placenta is unexpected. Shelton and Reichardt[12] have shown a correlation between the levels of NGF mRNA and the extent of sympathetic innervation. The placenta is an exception to this rule; it has relatively high levels of NGF but is not innervated. Guinea pig prostate, another major source of NGF[13], contains the short transcript, but also displays an intense band somewhat shorter than the long transcript observed in the mouse salivary gland (Fig. 2*a*, lane 5). This finding is of interest because mature NGF purified from guinea pig prostate also displays different antigenic properties from mouse NGF[13].

To establish whether the shortest extension product is a true NGF transcript or the result of premature termination, we sequenced the G8 myoblast RNA primer extension product. The sequence was identical with the known NGF cDNA up to the +158 splice junction; it was followed by an interfering, unreadable ladder of ~35 bases (data not shown), then a faint signal corresponding to the known NGF sequence resumed. This pattern could have resulted from two mRNAs differing in sequence 5' to the +158 splice site.

We sought alternative NGF transcripts by screening a mouse submaxillary gland cDNA library specifically constructed to conserve the 5'-terminal sequences (see Fig. 3 legend). (Although this tissue has a lower proportion of the short transcript, it contains more of both forms than any other tissue known.) The unamplified library (10^6 plaque-forming units in *λgt10*) was screened with the NGF cDNA under high stringency, and with oligonucleotide probes from the second and third exons, respectively (Fig. 3*a*). Of 1,000 clones that hybridized with the cDNA, 100 hybridized with the oligonucleotide probe from the third exon but not with the probe from the second. Three such clones were sequenced and showed the precise juxtaposition of the first and third exons; the second exon contained in the previously reported cDNA was absent (Fig. 3*c*). Thus, the short and long transcripts originated by differential splicing from a single NGF gene.

The expression of several other genes of neurobiological interest also involves differential RNA splicing; the splicing can be extremely complex, developmentally regulated and tissue specific[14–18]. We infer that the differential splicing demonstrated here in the case of the mouse salivary gland also accounts for the primer extension patterns observed in other tissues of the mouse and in other species. Both splicing pathways described here adhere to the splice acceptor consensus sequence at the 5' ends of the second and third exons[19] (Fig. 3*b*). The predominance of the long mRNA in the male mouse submaxillary gland (from which the first NGF cDNA clones were derived) obviously decreased the probability of isolating the short NGF cDNA; in addition, the primer used by Scott *et al.*[11] to sequence the 5' end of the trancript adventitiously straddled the +158 splice junction and so precluded the detection of the alternative message. The mechanism involved in this differential splicing may result from different relative concentrations of *trans*-acting elements present in ribonucleoprotein particles that interact with these two splice acceptors[20,21]. Biological significance for the two pathways therefore demands a demonstration of different function for the two transcripts.

The absence of the second NGF exon may affect the biological

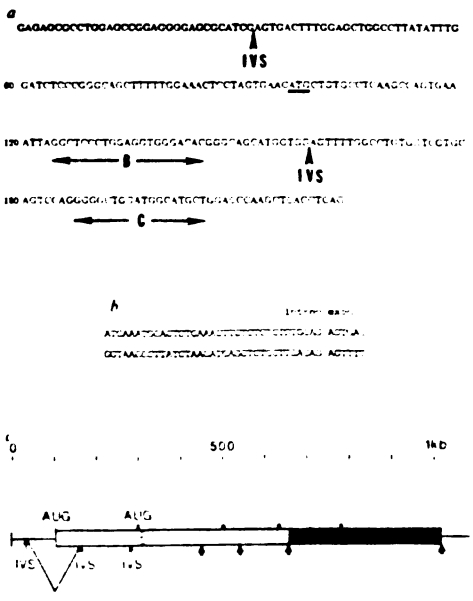


Fig. 3 *a*, Nucleotide sequence of the 5' end of the larger NGF transcript with arrows showing the position of intervening sequences (IVS). The shorter transcript described in the text lacks all the nucleotides between these IVSs. B, an oligonucleotide complementary to the second exon (+145 to +124). C, an oligonucleotide complementary to the third exon (+205 to +187)[34], used to screen the library. *b*, Nucleotide sequence of the introns just 5' to the second and third exons. *c*, Structure of the two NGF transcripts. The thin line indicates untranslated regions and the box translated regions. Thin arrows mark IVS and thick arrows potential cleavage sites; asterisks denote potential glycosylation sites. Length is measured in nucleotides (above), with zero the cap site of the larger RNA transcript.

Methods. Poly(A)$^+$ RNA was prepared as for Fig. 1 from the male mouse submaxillary gland. Reverse transcriptase synthesized first-strand cDNA from the RNA using random calf thymus DNA primers. The RNA-cDNA hybrid was tailed with dG and terminal transferase. Oligo(dC) in a molar ratio of 2:1 relative to the first strand was used to prime second-strand synthesis following the method of Okayama and Berg[33]. The ends of the double-stranded cDNA were made blunt first with Klenow enzyme, then with T4 DNA polymerase, an *Eco*RI linker adapter was added and the sequence inserted into λgt10 (ref. 36). *In vitro* packaging yielded a library of ~10^6 recombinants. The library was screened with an oligonucleotide (B) complementary to the second exon and another (C) complementary to the third exon[34], as well as the NGF cDNA. Of almost 1,000 clones that hybridized with the cDNA under high stringency and with oligonucleotides B and C, only about 100 hybridized with C but not B. Six of these were subcloned into M13 and the nucleotide sequence was determined by the dideoxy method in both directions[37]. Three showed a precise deletion of the second exon, with otherwise complete agreement to the original sequences down to the 5' end of the mRNA. Other clones showed premature termination of the larger transcript between oligonucleotides C and B.

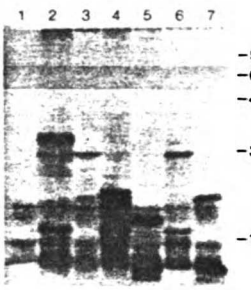


Fig. 4 *In vitro* translation of NGF RNA. For the longer construction containing both potential initiator AUGs (at +98 and +296 in the nucleotide sequence of the NGF mRNA), the cDNA was digested with *Sma*I (+66) and *Pst*I (+1027) and inserted into the SP65 vector. For the shorter construction containing only the second potential initiation codon, the cDNA was partially digested with *Bst*XI (+184), then *Pst*I and inserted into the same vector. Following linearization with the enzymes designated below, the two constructions (named LmRNA and SmRNA for the longer and shorter, respectively) were then used as templates for RNA synthesis with the SP6 polymerase[38]. This RNA was then translated *in vitro* by the wheat germ system in the presence of ^{35}S-methionine[39]. The protein products were precipitated in 10% trichloracetic acid and subjected to electrophoresis on 12% SDS-polyacrylamide. The gel was enhanced with 1 M sodium salicylate and exposed overnight at −70 °C with an enhancing screen. Lane 1, no RNA; lane 2, LmRNA linearized with *Hin*dIII (3′ to the inserted cDNA); lane 3, SmRNA linearized with *Hin*dIII; lane 4, LmRNA linearized with *Nco*I (at +616 in the cDNA); lane 5, SmRNA linearized with *Nco*I, lane 6, LmRNA linearized with *Sca*I (at +810 in the cDNA); lane 7, SmRNA linearized with *Sca*I; numbers on the right indicate the M_rs of the protein standards ($\times 10^3$).

properties of the NGF precursor. It produces a 5′-untranslated region of 172 bases and removes the first AUG. The initiation of translation would presumably occur at the next in-frame AUG. Both of these potential initiation codons have reasonable consensus sequences according to the Kozak rules[22]. Interestingly, in the short transcript, an out-of-frame AUG with good consensus sequence appears first in the mRNA, but a stop codon in the same frame immediately precedes the in-frame AUG. Translation of the short and long transcripts may explain more fully the observed complex pattern of the NGF precursor protein. Initiation at the first AUG would result in a protein of relative molecular mass (M_r) 34,000 (34K) and at the second, one of 27K. In salivary gland slices, Darling and Shooter have observed 35K and 29K NGF precursors after immunoprecipitation.

We have demonstrated use of the two potential initiation AUG codons to produce two distinct NGF proteins. A restriction fragment from the NGF cDNA containing both AUGs and the remainder of the protein coding sequence, and another fragment truncated between the two AUGs and thus containing only the second, were inserted into the SP65 vector system. Transcription of RNA from these templates and comparison of the protein products translated *in vitro* shows that the longer RNA yields proteins of roughly 34K and 27K M_r (corresponding to initiation of transcription at the first and second AUGs, respectively) (Fig. 4, lanes 2, 3) whereas the truncated form gives only the 27K protein. Truncation of both NGF RNAs at their 3′ end (by linearization of the template DNA with *Nco* and *Sca* at +616 and +810, respectively, in the cDNA) shortens the translated proteins by the expected amount, further establishing the origins of translation (Fig. 4, lanes 4–7). Although these *in vitro* results do not prove alternative use of the two AUGs *in vivo*, they do show that the translational apparatus can recognize both AUGs.

An understanding of the biological significance of differential splicing in the expression of NGF requires further information about the functions of the various precursor molecules. The NGF precursor contains several di- and tetrabasic residues that are potential peptidase cleavage sites (Fig. 3*c*). The small NGF precursor protein derived from the shorter transcript retains these sites, but the potential NH_2-terminal cleavage product would be truncated. If the full precursor or one of its derived peptides has biological activity, the two RNA splicing pathways could result in differing biological effects. It is intriguing that the two transcripts result in proteins with the hydrophobic sequence in different positions; in the long precursor, it resides in the middle of the molecule and might serve as a membrane anchor (similar to the epidermal growth factor precursor in the kidney[24]); in the short precursor, the hydrophobic region is at the N-terminus, and may act as a signal peptide to ensure secretion[25]. Differential splicing could therefore affect cellular localization, as with IgM and yeast invertase[26,27]. Although the short transcript presumably results in a secreted NGF that can act at a distance, the long transcript might produce a form of NGF that is membrane-bound and could interact with high efficiency at a receptor on a contiguous cell; these mechanisms could provide the basis for direct cell–cell communication.

We thank Dr David Standring for advice on the RNA analysis, Drs Jang Han and Jeff Edman for help with cDNA synthesis, Drs James Ou, Michael Walker and Eric Fodor for their thoughtful discussion, Drs Pablo Garcia and Peter Walter for assistance with *in vitro* translation, Dr Philip Barr for the oligonucleotides, and Ms Leslie Spector for preparing the manuscript. R.H.E. was supported by a postdoctoral fellowship from the NINCDS (NS07529-02). This research was supported by NIH grant AM 21334 (to W.J.R.).

Received 20 August 1985; accepted 2 January 1986.

1. Levi-Montalcini, R. & Angeletti, P. U. *Pharmac. Rev.* **18**, 619–628 (1966).
2. Gorin, P. D. & Johnson, E. M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5382–5386 (1979).
3. Stockel, K., Schwab, M. E. & Thoenen, H. *Brain Res.* **99**, 1–16 (1975).
4. Banks, B. E. C. *et al. J. Physiol., Lond.* **247**, 289–298 (1975).
5. Chun, L. L. Y. & Patterson, P. H. *J. Cell Biol.* **75**, 694–704 (1975).
6. Thoenen, H. & Barde, Y.-A. *Physiol. Rev.* **60**, 1284–1335 (1980).
7. Mobley, W. C. *et al. Science* **229**, 284–287 (1985).
8. Ullrich, A. *et al. Nature* **303**, 821–825 (1983).
9. Murphy, R. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 4496–4500 (1977).
10. Oger, J. O. *et al. Proc. natn. Acad. Sci. U.S.A.* **71**, 1554–1558 (1974).
11. Scott, J. *et al. Nature* **302**, 538–540 (1983).
12. Shelton, D. L. & Reichardt, L. F. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7951–7955 (1984).
13. Harper, G. P. *et al. Nature* **279**, 160–162 (1979).
14. Herbert, E. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 375–384 (1983).
15. Amara, S. G. *et al. Nature* **298**, 240–244 (1982).
16. Nawa, H., Kotani, H. & Nakanishi, S. *Nature* **312**, 729–734 (1984).
17. Falkenthal, S., Parker, V. P. & Davidson, N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 449–453 (1985).
18. Medford, R. M. *et al. Cell* **38**, 409–421 (1984).
19. Mount, S. *Nucleic Acids Res.* **10**, 459–472 (1982).
20. Steitz, J. A. *et al. Cold Spring Harb. Symp. quant. Biol.* **47**, 893–900 (1983).
21. Mount, S. M. *et al. Cell* **33**, 509–518 (1983).
22. Kozak, M. *Nucleic Acids Res.* **9**, 5233–5252 (1981).
23. Darling, T. L. *et al. Cold Spring Harb. Symp. quant. Biol.* **48**, 427–434 (1983).
24. Rall, L. *et al. Nature* **313**, 228–233 (1985).
25. Lingappa, V. R. *et al. Proc. natn. Acad. Sci. U.S.A.* **75**, 2338–2342 (1978).
26. Early, P. *et al. Cell* **20**, 313–319 (1980).
27. Carlson, M. & Botstein, D. *Cell* **28**, 145–154 (1982).
28. Chirgwin, J. M. *et al. Biochemistry* **18**, 5294–5299 (1979).
29. Cathala, G. *et al. DNA* **2**, 329–335 (1983).
30. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408–1412 (1972).
31. Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
32. Hu, N. T. & Messing, J. *Gene* **17**, 271–277 (1982).
33. McKnight, S. L. & Kingsbury, R. *Science* **217**, 316–324 (1982).
34. Wallace, R. B. *et al. Nucleic Acids Res.* **9**, 2647–2658 (1981).
35. Okayama, H. & Berg, P. *Molec. cell. Biol.* **2**, 161–170 (1982).
36. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning Techniques: A Practical Approach* (ed. Glover, D.) pp. 49–78 (IRL Press, Oxford, 1984).
37. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
38. Krieg, P. A. & Melton, D. A. *Nucleic Acids Res.* **12**, 7057–7070 (1984).
39. Erickson, A. H. & Blobel, G. *Meth. Enzym.* **96**, 38–50 (1983).

CHAPTER III

COBRA NERVE GROWTH FACTOR: STRUCTURE AND EVOLUTIONARY COMPARISON

ABSTRACT

We have cloned and sequenced an NGF cDNA from the cobra venom gland. The structure of the predicted cobra precursor resembles that of the mouse prohormone, with a highly conserved mature NGF protein at the C-terminus. In the non-NGF moiety, a putative signal peptide and two sets of basic residues presumably reflecting cleavage sites that are more conserved, while large interstitial regions are poorly conserved. This moiety may be involved in defining the folding of the precursor for subsequent proteolytic maturation and/or encode a bioactive polypeptide(s). Primer extension analysis indicates a single NGF transcript in the cobra venom gland.

INTRODUCTION

Nerve growth factor (NGF) regulates the development of specific neuronal populations (Thoenen & Barde, 1980). NGF was first identified as a diffusible factor from mouse sarcoma cells which elicited dramatic nerve fiber outgrowth in chicken embryos. Using an in vitro assay of neurite outgrowth from embryonic chick dorsal root ganglia (Levi-Montalcini, 1954), Cohen & Levi-Montalcini partially purified NGF. To determine whether the activity resided in nucleic acid, they treated the preparation with crude moccasin (Agkistron piscivorus) venom phosphodiesterase (Cohen & Levi-Montalcini, 1956). Surprisingly, the snake venom exhibited a higher activity than the tumor extracts. This led to the isolation of a protein with potent NGF activity from venom, and eventually to the discovery of the richest source of NGF, the mouse submaxillary gland (SMG), a phylogenetically homologous organ (Cohen, 1960), which has subsequently served as a source for the purification of NGF and for the production of antibodies to this protein. These in turn

have been used to elucidate the structure and biological properties of NGF.

We have previously studied the mouse NGF gene and its mRNA transcripts. Multiple mouse NGF transcripts encoding mature NGF and the previously undefined amino-terminal precursor regions (Scott et al., 1983; Edwards et al., 1986; Selby et al., submitted) can be explained by alternative splicing and independent promoter usage from a single gene. We now report the cloning of a virtually full-length NGF cDNA from the cobra (Naja naja siamensis) venom gland. Comparison of the sequence with that of the mouse as well as with other species, aids in the determination of residues critical for secretion and for correct prohormone processing.

EXPERIMENTAL PROCEDURES

Northern Blot Analysis

RNA was prepared from the cobra (Naja naja siamensis) venom gland (purchased from Miami Serpentarum Labs) following the guanidinium-LiCl method (Cathala et al., 1983). PolyA+ RNA was isolated by oligodT cellulose chromatography. Five micrograms of polyA+ RNA was fractionated on a 1.5% agarose-formaldehyde gel and transferred to nitrocellulose (Maniatis et al., 1983).

Southern Genomic Blot

Cobra genomic DNA was prepared from cobra liver as previously described (Maniatis et al., 1983). Ten micrograms of DNA was digested with several restriction enzymes according to the manufacturers specifications. The digested DNA was precipitated in 0.3M NaOAc and 2.5 volumes of ethanol, dried and resuspended in loading dye (Maniatis et al., 1983). The sample was heated to 65°C for 15 minutes and then fraction-

ated on a 0.9% agarose gel. The gel was then transferred to nitrocellulose.

cDNA Library Preparation

An oligodT primed cDNA library (Gubler & Hoffman, 1983) in λgt10 was generated from cobra venom gland polyA+ RNA (Huynh et al., 1985). Recombinant phage were plated at a density of ~25,000 plaque forming units per plate. Nitrocellulose lifts were made and processed for hybridization (Maniatis et al., 1983).

Hybridization

The mouse NGF cDNA was labeled (Maniatis et al., 1983) and used as a probe for Northern, Southern and phage screening. The hybridization was done at 42°C in the following solution: 35% formamide, 5xSSPE, 0.1% SDS, 5X Denhardt's, 0.1% pyrophosphate and 20 µg/ml denatured E. coli DNA. Following 18-24 hr hybridiziation, the filters were washed in 2x SSPE and 0.1% SDS at 43°C and then exposed to XAR-5 film at -70°C.

Primer Extension

Two 19mer oligonucleotides were end-labeled using ^{32}P-γ-ATP and polynucleotide kinase (Maniatis et al., 1983). The labeled oligonucleotides (one million cpm each) were then mixed with 1 µg of polyA+ cobra venom gland RNA, heated at 65°C for 15 minutes. Reverse transcriptase buffer was added (Maniatis et al., 1983) and then the reaction was slowly cooled to 46°C. All 4 deoxynucleotides were added to a final concentration of 1mM along with 20 units of reverse transcriptase. The reaction proceeded at 46°C for 60 minutes and was then terminated by phenol- chloroform extraction. The aqueous phase was lyophilized and then resuspended in formamide buffer (Maniatis et al., 1983). This was heated to 90°C for 3 minutes and then fractionated on a 5% acrylamide-

urea sequencing gel. A dideoxy sequencing reaction (Sanger et al., 1977) was electrophoresed in an adjacent lane.

Sequencing of Positive Clones

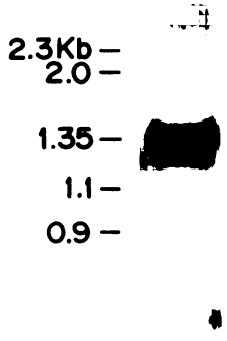
Clones positive by hybridization were purified, their DNA's subcloned into M13 and then sequenced (Sanger et al., 1977; Maniatis et al., 1983). One clone was sequenced in its entirety on both strands.

RESULTS AND DISCUSSION

Northern and Southern analyses under conditions of relaxed stringency demonstrated the feasibility of hybridization of a mouse probe to cobra NGF RNA and DNA. Northern blots demonstrated a strongly hybridizing band similar in size (about 1.3 kb) to that observed in the mouse submaxillary gland (Fig. 1A). A Southern blot of cobra genomic liver DNA also showed specific hybridization signals (Fig. 1B). Using the same hybridization conditions, we screened a λgt10 cDNA library derived from cobra venom gland mRNA and detected positive clones at a frequency of about 0.3%. The clone with the largest insert contained two in-frame amino-terminal initiation codons, but only the second AUG conforms well to the Kozak consensus sequence for the initiation of translation (Kozak, 1981). If this second AUG codon were used to start translation the cDNA would predict a 27 kd precursor. Additional protein coding sequence proximal to the two methionine residues is precluded because of an upstream in-frame termination codon (TGA). A signal peptide cleavage site is predicted (von Hjeine, 1986) between amino acids 18 and 19, both alanine, if the second methionine is the initiator codon. The location of the presumptive cleavage site is identical to that found in the mouse 27 kd NGF precursor (Chapter VI).

The predicted organization of the cobra NGF prohormone resembles

FIGURE 1. Northern and Southern blot analyses. (A) Five micrograms of cobra venom gland polyA+ RNA was fractionated on a 1.5% agarose-formaldehyde gel and transferred to nitrocellulose. B) Ten micrograms of restriction enzyme digested genomic cobra and mouse liver DNAs were fractionated on 0.9% agarose-TEA and transferred to nitrocellulose (Maniatis et al., 1983). Both blots were hybridized at low stringency (35% formamide) at 42°C with a ^{32}P-labeled mouse NGF cDNA. The filters were washed at 43°C in 2X SSPE and 0.1% SDS and autoradiographed at -70°C. A single hybridizing DNA species is observed; it is roughly the same size as mouse NGF RNA (~1300 bp). Hybridizing cobra and mouse genomic DNA bands appear in the blot using the mouse cDNA as probe. The typical bands are observed in mouse genomic DNA; no spurious bands appear, suggesting that the hybridizing bands in the cobra genomic DNA represent real sequence complements.


2.3Kb
2.0
1.35
1.1
0.9

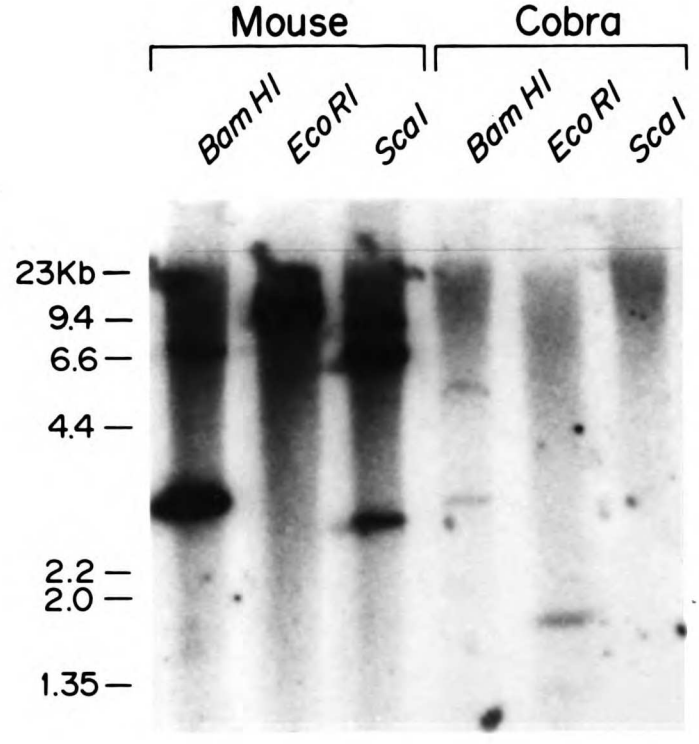
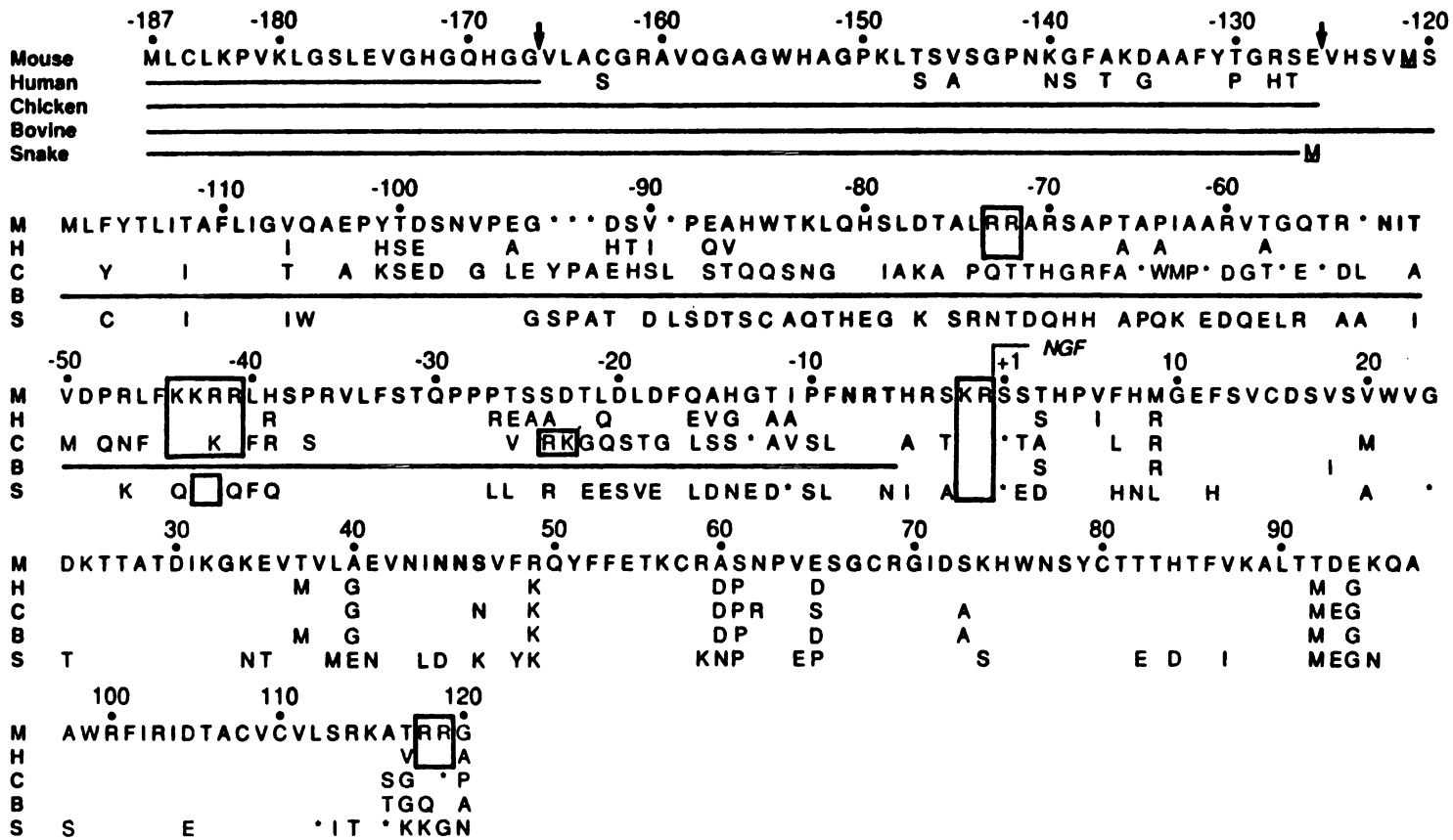

Mouse
Cobra
BamHI
EcoRI
ScaI
BamHI
EcoRI
ScaI
23Kb
9.4
6.6
4.4
2.2
2.0
1.35

FIGURE 2. Sequence of a cobra NGF cDNA predicts preproNGF. The sequence of a nearly full-length cDNA, obtained by the dideoxy method (Sanger et al., 1977) on both strands, is presented. Dibasic amino acid residues are boxed. Mature NGF begins at amino acid #1. The first methionine and the distal amino acids present before the second methionine are presented in lower case letters. The underlined nucleotides designated "A" and "B" represent the sequence from which complementary oligonucleotides were generated for primer extension. The amino-terminus is fairly hydrophobic--consistent with the presence of a signal sequence. A potential cleavage site is predicted between Ala -108 and Ala -107. Mature NGF is flanked at its amino-terminus by dibasic amino acid residues; cleavage here would liberate NGF from the precursor. No dibasic residues are found at the carboxy-terminus as in mouse NGF precursor.

```
                                                                  met val his ser
ACCTGAGGATATCCTCAACTGCTTCAGAGGTATCAATTGAGATCA ATG GTG CAT AGC

   -125                  -120
val Met Ser Met Leu Cys Tyr Thr Leu Ile Ile Ala Phe Leu Ile
GTA ATG TCC ATG CTG TGC TAC ACT CTG ATT ATA GCA TTT CTG ATT
    └──────────A──────────┘

   -110                                              -100
Gly Ile Trp Ala Ala Pro Lys Ser Glu Asp Asn Val Pro Leu Gly
GGC ATA TGG GCA GCA CCA AAG TCT GAA GAT AAT GTC CCT CTA GGC
                    └──────────B──────────┘

                        -90
Ser Pro Ala Thr Ser Asp Leu Ser Asp Thr Ser Cys Ala Gln Thr
TCC CCT GCA ACA TCT GAC CTT TCT GAC ACC AGC TGT GCT CAA ACT

    -80                                               -70
His Glu Gly Leu Lys Thr Ser Arg Asn Thr Asp Gln His His Pro
CAC GAG GGT CTG AAA ACA TCT CGA AAC ACA GAT CAG CAC CAT CCA

                        -60
Ala Pro Gln Lys Ala Glu Asp Gln Glu Leu Arg Thr Ala Ala Asn
GCT CCT CAG AAG GCA GAG GAT CAA GAA CTT AGA ACA GCA GCA AAC

    -50                                               -40
Ile Ile Val Asp Pro Lys Leu Phe Gln Lys Arg Gln Phe Gln Ser
ATC ATT GTG GAT CCC AAG CTT TTT CAG AAG AGG CAG TTC CAG TCG

                        -30
Pro Arg Val Leu Phe Ser Thr Gln Pro Pro Leu Leu Ser Arg Asp
CCT CGT GTT CTG TTC AGC ACT CAG CCC CCA CTG TTG TCA AGA GAT

    -20                                               -10
Glu Glu Ser Val Glu Phe Leu Asp Asn Glu Asp Ser Leu Asn Arg
GAG GAA AGT GTG GAG TTC CTG GAC AAT GAA GAC TCT CTT AAC AGG

                         1
Asn Ile Arg Ala Lys Arg Glu Asp His Pro Val His Asn Leu Gly
AAT ATC CGG GCC AAA CGT GAA GAT CAT CCT GTG CAT AAC CTA GGG

 10                                                20
Glu His Ser Val Cys Asp Ser Val Ser Ala Trp Val Thr Lys Thr
GAA CAT TCT GTG TGT GAC AGT GTC AGT GCC TGG GTT ACT AAA ACT

                        30
Thr Ala Thr Asp Ile Lys Gly Asn Thr Val Thr Val Met Glu Asn
ACA GCA ACA GAC ATC AAA GGC AAT ACG GTG ACT GTG ATG GAA AAT

 40                                                50
Val Asn Leu Asp Asn Lys Val Tyr Lys Gln Tyr Phe Phe Glu Thr
GTA AAT CTT GAT AAC AAA GTC TAC AAG CAG TAC TTT TTT GAA ACC

                        60
Lys Cys Lys Asn Pro Asn Pro Glu Pro Ser Gly Cys Arg Gly Ile
AAG TGC AAA AAT CCA AAT CCA GAA CCG AGT GGG TGC AGG GGC ATT

 70                                                80
Asp Ser Ser His Trp Asn Ser Tyr Cys Thr Glu Thr Asp Thr Phe
GAT TCC AGC CAT TGG AAT TCT TAT TGC ACC GAA ACA GAC ACA TTT

                        90
Ile Lys Ala Leu Thr Met Glu Gly Asn Gln Ala Ser Trp Arg Phe
ATC AAG GCA TTA ACC ATG GAA GGC AAT CAG GCA TCC TGG CGC TTC

100                                               110
Ile Arg Ile Glu Thr Ala Cys Val Cys Val Ile Thr Lys Lys Lys
ATT CGG ATT GAA ACT GCC TGT GTG TGT GTA ATC ACT AAA AAA AAA

    116
Gly Asn OP
GGA AAC TGA TTGCTCCAATACTCACTTTCTCCCTACCCCCTACCACAGCCTGTAAAT

TATTTTAAATTATGAGGGCTCTATTATATATTTATAGTTTATACAATGGAATAACTGAAT

CATCATAATTTATTAAATCCTTTTGAAACCTGGATACATGATGTGAGCACTTGGTTTCAA

AAAAGTACTGCTTGTAGACTTTGTTATCAAGATTTACATTTTCCATCACAGTCACCGTAT

TGCCTTTGATGTCTGTTGCTGTAGTTTTAGTAACCCAGGCACTGACACTGTCACACACAG

AATGTTCCCCTAGGTTATGCACAGGATGATCTTCACGTTTGGCCCGGATATTCCTGTTAA

CAGAGTCTTCATTGTCCAGGAACTCCACACTTTCCTCATCTCTTGACAACAGTGGGGGGC

CTGAGTGCTGAACAGGGCTCACAATTGCCACACAACATACGAGCCCGAAG
```

that of the 27 kd mouse precursor in size, in the relative position and sequence of mature NGF, and in selected regions of the non-NGF moiety (Fig. 2) (Edwards et al., 1986). The mature NGF portion is located at the C-terminus of the precursor and is highly conserved between cobra and mouse (66%) (Fig. 3). The sequence predicted by this cobra cDNA extends and corrects that obtained from direct sequencing of peptides derived from the mature protein (Hogue-Angeletti et al., 1976). Nineteen residues in total have been corrected on the basis of the cDNA sequence. The two major differences are: a glycine residue originally seen at position 23 is in fact missing, while our sequence reveals an additional lysine residue near the carboxy-terminus. Some differences could be due to polymorphisms. Cysteine and tryptophan residues thought to be important for activity are conserved among all the species examined (Frazier et al., 1973; Cohen et al., 1980). The non-NGF moiety is less well conserved (45%) (Fig. 3). Three regions of the non-NGF moiety, however display greater similarity between species: the presumptive signal sequence (77% similarity between cobra and mouse) (von Heijne, 1986) and two sets of dibasic amino acids, one in the central region of the precursor (-41/-42) and the other adjacent to the mature protein, presumably the site of cleavage to liberate NGF. It appears that both of these sites, not just that flanking NGF, are required for correct prohormone processing. Sequential cleavage steps may be required to generate a correctly folded mature protein. The interstitial regions separating these basic residues in the non-NGF moiety of the precursor have low homology with the mouse, considering even conservative amino acid substitutions. These rapidly evolving domains may be involved in the folding of the precursor, so that it undergoes correct processing in

FIGURE 3. Evolutionary comparison of NGF's. The amino acid sequences of human (Ullrich et al., 1983), chicken and bovine (Ebendal et al., 1986; Meier et al., 1986; Wion et al., 1986) NGF's are aligned with that of the mouse (Scott et al., 1983). Solid lines denote unknown sequence. Stars represent gaps in the sequences to obtain the best homologies. Dibasic amino acid residues are boxed and glycosylation consensus sequences are in bold type. Mature NGF begins at amino acid residue #1. Arrows denote the location of introns in the mouse (Chapter IV) and human (Ullrich et al., 1983) genes. The initiator methionine for the mouse 27K precursor is underlined; the first methionine codon in the cobra is underlined.


-187 -180 -170 -160 -150 -140 -130 -120
Mouse MLCLKPVKLGSLEVGHGQHGGVLACGRAVQGAGWHAGPKLTSVSGPNKGFAKDAAFYTGRSEVHSVMS
Human S S A NS T G P HT
Chicken
Bovine
Snake M
-110 -100 -90 -80 -70 -60
M MLFYTLITAFLIGVQAEPYTDSNVPEG***DSV*PEAHWTKLQHSLDTALRRARSAPTAPIAARVTGQTR*NIT
H I HSE A HTI QV A A A
C Y I T A KSED G LE YPAEHSL STQQSNG IAKA PQTTHGRFA*WMP*DGT*E*DL A
B
S C I IW GSPAT D LSDTSC AQTHEG K SRNTDQHH APQK EDQELR AA I
-50 -40 -30 -20 -10 +1 NGF 10 20
M VDPRLFKKRRLHSPRVLFSTQPPPTSSDTLDLDFQAHGTIPFNRTHRSKRSSTHPVFHMGEFSVCDSVSVWVG
H R REAA Q EVG AA S I R
C M QNF K FR S V RKGQSTG LSS*AVSL A T *TA L R M
B S R I
S K Q QFQ LL R EESVE LDNED*SL NI A *ED HNL H A *
30 40 50 60 70 80 90
M DKTTATDIKGKEVTVLAEVNINNSVFRQYFFETKCRASNPVESGCRGIDSKHWNSYCTTTHTFVKALTTDEKQA
H M G K DP D M G
C G N K DPR S A MEG
B M G K DP D A M G
S T NT MEN LD K YK KNP EP S E D I MEGN
100 110 120
M AWRFIRIDTACVCVLSRKATRRG
H V A
C SG *P
B TGQ A
S S E *IT *KKGN

the cell. Alternatively, these regions could, upon cleavage at the conserved sites, give rise to stable peptides with biological activity. Mammals do show better conservation in these segments of the non-NGF moiety, which may only reflect the considerably smaller evolutionary distance from mouse to human, or may indicate that these potential peptides have independent biological activity in mammals. Of other sequences conserved in different species, the Arg-Arg sequence found in mouse and human precursors at residues -73/-72 is not present in the cobra or in chicken. There are several N-linked glycosylation consensus sequences (Asn-X-Ser/Thr) in mouse and human precursors (two in the non-NGF moiety, one in NGF), while none are present in the cobra precursor. (The glycosylation site on mature mouse NGF is apparently not utilized (Angeletti et al., 1973)). As expected, the cobra cDNA hybridizes to mouse and human genomic DNA's. The hybridization pattern to cobra DNA is fairly complicated and thus one cannot be sure if the gene is single copy or not (Fig. 4).

Primer extension experiments show a single cobra NGF transcript (Fig. 5). Extension with either of two oligonucleotides (see Fig. 2) as primers yielded a single product that corresponds to a transcript initiating 12 bases upstream from the 5' end of the cDNA clone. This transcript is analogous to the shorter of the two principle transcripts found in mouse and humans (Edwards et al., 1986). This suggests that the longer mouse and human NGF precursor, containing 60 more N-terminal amino acids and an internal signal sequence, may have a function limited to mammals.

The presence of biologically active NGF in snakes suggests that it may also support the survival and differentiation of reptilian neurons,

although this has not been demonstrated. It is possible that NGF has a more restricted or even quite different role in lower vertebrates. A better understanding of its role in these animals might indicate the potential function of residues in the precursor that are conserved only in mammals. In addition, the explanation for such high concentrations in the snake as in the mouse salivary gland, remains a mystery.

FIGURE 4. Cross-species genomic Southern. 10 µg of DNA was digested to completion with the restriction enzymes designated in the figure. Genomic DNA's were isolated from cobra, mouse, and Xenopus livers. The probes were the two purified EcoRI fragments from the cobra venom gland NGF cDNA. The entire blot was washed at 2X SSPE (Maniatis et al., 1983) at 42°C. The blot was then exposed to autoradiography with an intensifying screen. The hybridization pattern of mouse and human DNA's is similar to that found when using the mouse cDNA as probe. The cobra pattern is complex and it cannot be stated whether or not the NGF gene is single copy.

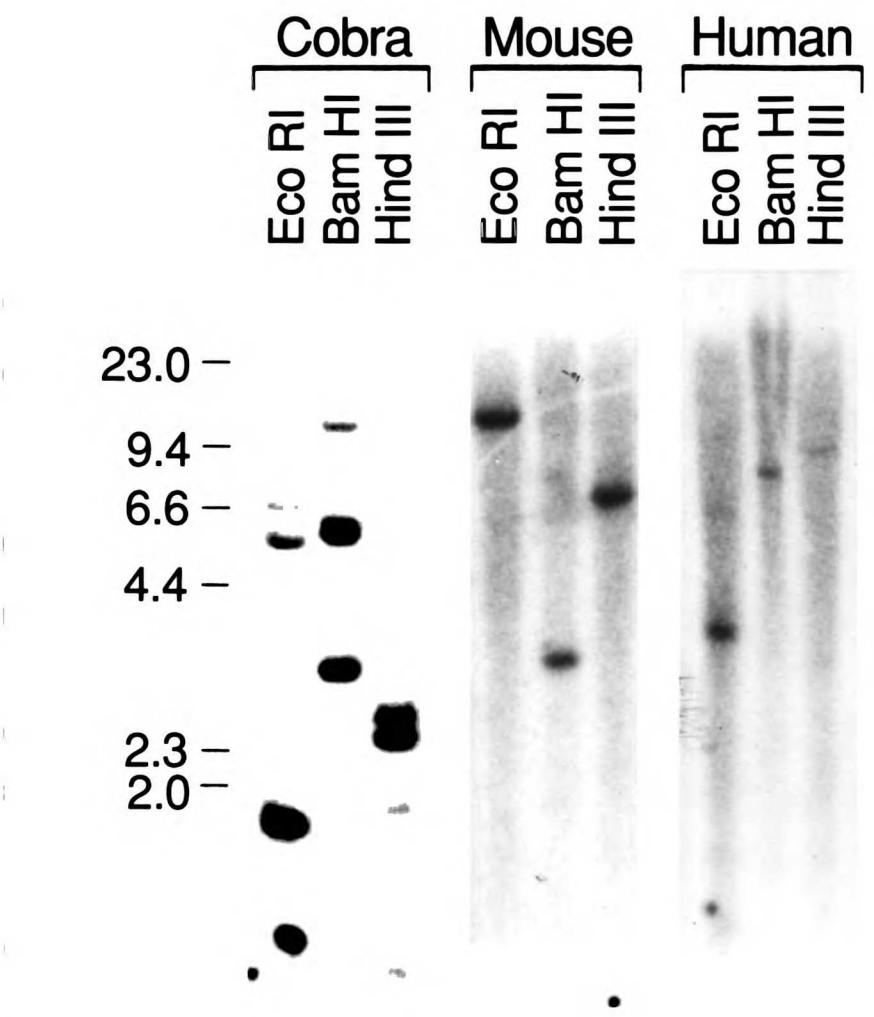

Cobra
Mouse
Human
Eco RI
Bam HI
Hind III
Eco RI
Bam HI
Hind III
Eco RI
Bam HI
Hind III
23.0
9.4
6.6
4.4
2.3
2.0

FIGURE 5. Primer extension of cobra NGF RNA. End-labeled oligonucleotides A and B (see Fig. 2) were annealed with 1 μg polyA$^+$ cobra RNA and extended with reverse transcriptase. At the end of a 60 minute incubation at 46°C, the reactions were then extracted, dried and resuspended in formamide dyes (Maniatis et al., 1983), heated to 95°C and loaded onto a 5% acrylamide-urea gel (lanes 1,2). A dideoxy sequencing reaction was loaded in an adjacent lane as standard. The gel was visualized by autoradiography. The extension product for each oligonucleotide is denoted by arrows and in number corresponding to the length of the product. Longer exposures fail to show any other specific extension products.

1 2 3

149 –

93 – [redacted]

CHAPTER IV

THE MOUSE NERVE GROWTH FACTOR GENE: STRUCTURE AND EXPRESSION

SUMMARY

The organization and biologically significant sequences of the entire mouse nerve growth factor gene have been determined. The gene spans 45 kb and contains several small 5' exons. Transcription of the gene results in four different mRNA species, which can be accounted for by alternative splicing and independent initiation from two promoters. These transcripts encode proteins which have divergent N-termini and the NGF moiety at their C-termini. The levels of the various NGF transcripts have been determined in different tissues and through postnatal development. We have also examined the expression of these transcripts in the brain in response to specific early sensory deprivation. The results suggest that the expression of NGF mRNA during postnatal development may be regulated independently of the formation of complex neural networks.

INTRODUCTION

Developing sympathetic and sensory neurons require nerve growth factor (NGF) (Levi-Montalcini, 1966; Thoenen & Barde, 1980) for survival. A significant proportion of these cells dies in the first few weeks after birth (approx 30% for the sympathetic neurons in the rat superior cervical ganglion). Injection of antibodies to NGF during the development of sympathetic and sensory neurons results in the death of these cells (Gorin & Johnson, 1979; Levi-Montalcini & Booker, 1960). Administration of purified NGF permits the survival of all these cells, including those that would normally die (Angeletti et al., 1971). NGF thus appears to be a modulator of normal, programmed cell death. During development, neurons which obtain NGF survive while those that do not, die.

Sympathetic and sensory neurons acquire NGF via a specific, retro-

grade axonal transport system (Stockel et al., 1975). Transecting the axons of developing sympathetic neurons in the superior cervical ganglion causes the lesioned cells to die. Similarly, death occurs after administration of antibodies to NGF. Systemic injection of NGF permits survival of the axotomized cells, presumably by providing NGF directly to the cell body (Hendry, 1975). Cells responding to NGF have NGF receptors on their processes and on their cell bodies (Herrup & Thoenen, 1979; Johnson et al., 1978). The presence of a retrograde axonal transport system for NGF suggests that neurons obtain NGF from the cells that they innervate. A correlation of NGF protein and mRNA levels with the density of innervation in peripheral organs supports this idea (Korshing & Thoenen, 1983; Shelton & Reichardt, 1984). The development of normal adult neural networks thus appears to depend on NGF gene expression in cells interacting with the dependent neurons.

An NGF cDNA derived from the mouse submaxillary gland has been isolated and characterized (Scott et al., 1983). It predicts a precursor protein with NGF moiety at the C-terminus. We have also described a cDNA representing a shorter NGF transcript that predominates in all tissues other than the mouse submaxillary gland and the placenta from several species (Edwards et al., 1986). The virtual identity of the nucleotide sequence of the two cDNAs suggests that they were derived from the same gene. A portion of the human gene corresponding to the mouse cDNA but lacking the 5' sequence has been reported (Darby et al., 1985; Ullrich et al., 1983).

We present here the structure of the entire mouse NGF gene as well as the sequences of two additional cDNAs, and we show that the four cDNAs result from alternative splicing and the use of independent

promoter elements. We also report the distribution of the NGF mRNAs in different tissues and through postnatal development. In sensory deprivation experiments, the level of steady-state NGF transcripts appears to be independent of the formation of certain neural connections.

MATERIALS AND METHODS

Isolation and Characterization of cDNA Clones

To conserve the full 5' ends of the mRNA, we constructed a λgt10 submaxillary gland cDNA library using random primers for first strand synthesis and a modification of the Gubler-Hoffman method for the second strand (Edwards et al., 1986; Gubler & Hoffman, 1983). We screened the library using oligonucleotides as previously described (Edwards et al., 1986) and sequenced the unique clones of interest (identified by comparing dideoxy T-tracks) on both strands (Maxam & Gilbert, 1980; Sanger et al., 1977).

RNA Analysis

Freshly dissected tissue from several animals in different litters was pooled and homogenized with a polytron in guanidinium isothiocyanate (Chirgwin et al., 1979). Total RNA was prepared by LiCl precipitation (Cathala et al., 1983) or by CsCl density gradient sedimentation, followed by phenol/chloroform extraction and two ethanol precipitations. RNA content was determined spectrophotometrically at 260 nm after the second precipitation.

Primer extension assays were performed using a uniformly labeled single-stranded 129 base primer as previously described (Edwards et al., 1986).

Uniformly labeled S1 nuclease probes were made from M13 subclones of the cDNAs obtained from the above submaxillary gland library using

the same 22 base primer (complementary to +328 to +306) and two of four labeled nucleotides (Edwards et al., 1986). The extended primer was cleaved in the M13 polylinker and isolated after electrophoresis on a 5% polyacrylamide-urea gel. The resulting single-standed probe (50,000 cpm) was co-precipitated with 10-25 µg of a total RNA sample, heated to 85°C in 80% formamide for 15 min and hybridized overnight at 45°C. S1 nuclease digestion with 100 units of enzyme (PL Biochemicals) at 37°C for 90 minutes was followed by extraction, ethanol precipitation and electrophoresis on a 5% acrylamide-urea gel. Laser densitometry was used to quantitate the relative amounts of NGF mRNA hybridizing specifically to a given probe, and to quantitate ethidium bromide-stained non-denaturing agarose gels of the RNA samples which were then used to normalized the S1 nuclease results to total RNA (mostly rRNA).

Isolation and Characterization of Genomic Phage Clones

A mouse phage library (MboI partial digests of Balb/c embryonic DNA ligated into Bam-Charon 28, Leder #10277) was screened using nick translated probes (Maniatis et al., 1982) representing various portions of the NGF cDNA and genomic subclones. Hybridizing phage were purified, restriction mapped and subcloned into M13 vectors for sequencing. All exons and their immediate flanking DNA were sequenced by the dideoxy method (Sanger et al., 1977). The size of the intron not represented in the phage clones was estimated by hybridization of genomic DNA with probes derived from the 3' (pH3) and 5' (pH2) ends of the respective 5' and 3' phage.

Lesion Experiments

At birth and at one, two and three weeks after birth, the right whisker pad was removed from individual mice representing several

litters, using ice as anesthesia. When the mice were four weeks old, they were sacrificed; ~3x3 mm regions encompassing only the somatosensory cortex were dissected out; the dorsolateral brainstem (including the fifth nerve nucleus) was also dissected out. RNA preparation and S1 nuclease assay were performed as described above.

RESULTS

cDNA Clones Predict Different NGF Precursors

The first mRNA shown in Figure 1 (a and b) is the original, long NGF transcript (Scott et al., 1983), here termed transcript A. A cDNA representing the short transcript (B) (Fig. 1) was identified as hybridizing to an oligonucleotide complementary to the third exon of A (+205 to +188) and failing to hybridize to an oligonucleotide from the second exon of A (+145 to +124) (Edwards et al., 1986). (Transcript B was previously suspected to exist on the basis of S1 nuclease protection and primer extension experiments.) Of the six independent cDNA clones which showed this hybridization pattern, three represented NGF transcript B, two represented prematurely terminated transcript A (having part but not all of the second exon of A) and one had an entirely different sequence 5' to the junction between the second and third exons of A. This third distinct cDNA (representing transcript C, Fig. 1) predicts a novel NGF precursor with a new exon replacing the first two exons (of transcript A). The protein encoded by the new exon (IA) (Fig. 3) would replace the 20 N-terminal residues of the original, long NGF precursor protein with a totally different 27 amino acids, resulting in a prohormone that is approximately 700 daltons larger. Because of the apparent abundance of transcript C, we sequenced an additional 20 cDNA clones obtained by the same hybridization protocol. Among them we found an additional unique

FIGURE 1. A) Nucleotide sequence of NGF cDNAs A-D. Slashes denote the presence of intervening sequences in the cDNAs. In-frame AUG codons have been underlined. The presumptive initiation methionine codons are in bold letters. B) Diagrammatic representation of the predicted NGF transcripts in relation to the gene. The gene is shown above with exons as boxes and introns as lines. The size of the exons (in bp) is shown above the boxes; the size of the introns is also shown; Mature NGF is stippled. The four identified NGF transcripts are depicted below in order of decreasing submaxillary gland abundance (A,B,D, then C). The relative abundance of each transcript in the submaxillary gland (SMG) is shown at the right of the figure. Thick letters represent sequences formed in the mature RNA and the thin lines regions that are removed by splicing of a primary transcript. The presumed sites for initiation of translation are indicated (ATG).

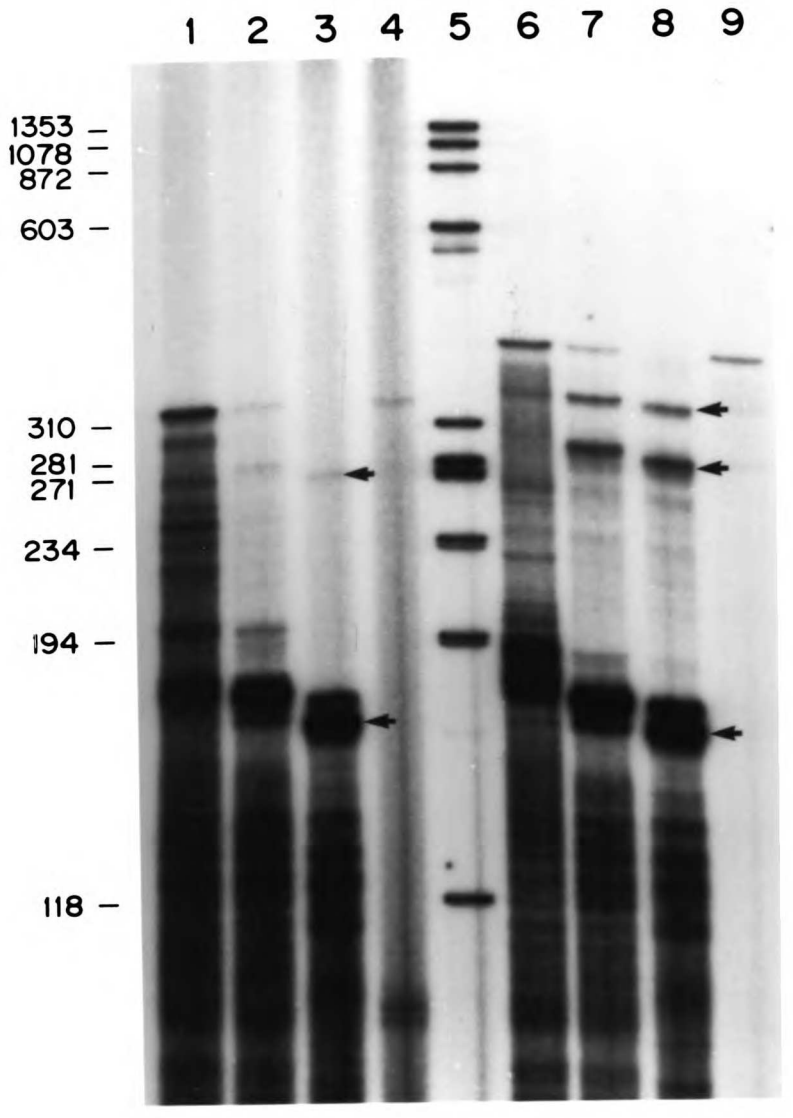
142 bp
167
33
>32 kb
~4 kb
127
161
124
~6 kb
894
ATG
ATG
ATG
NGF
A
ATG
100
B
ATG
10
D
ATG
1
C
ATG
0.1
SMG Relative Abundance

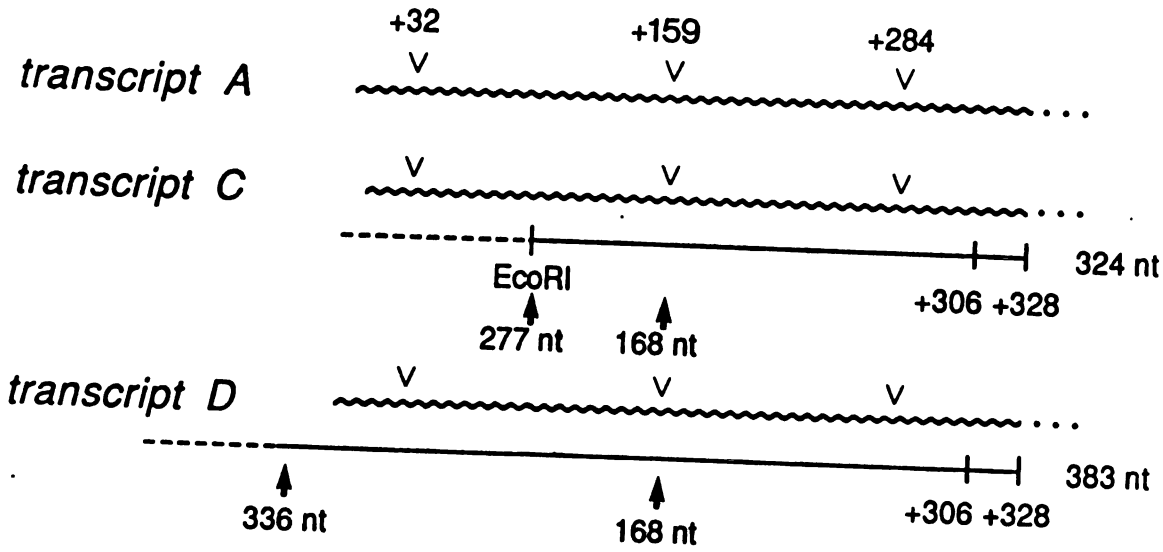
clone, cDNA D (Fig. 1). Transcript D also diverges 5' to the exon 2-3 border, but does not contain a new in-frame AUG; it thus predicts a precursor identical to that encoded by transcript B. To determine the relative in vivo levels of transcripts C and D, we performed S1 nuclease analysis on submaxillary gland RNA, using uniformly-labeled single-stranded probes derived from the respective cDNA clones (Fig. 2). Both transcripts C and D are much less abundant (by 2-3 orders of magnitude) than A or B, although D is present in greater quantities than C (Fig. 1B).

Structure of the Mouse NGF Gene Accounts for the cDNA Sequences by Using Two Promoter Elements and Alternative Splicing

To understand the structural basis for these multiple transcripts it was necessary to define the organization of the NGF gene and in particular, its 5' portion. Using NGF cDNA A as probe, we isolated a phage containing the two 3' exons (Exons III and IV, Fig. 3A). Exon IV encodes all of mature NGF (118 amino acids), as well as an additional 125 amino acids in the amino-terminal direction. Exon IIIB comprises 124 bp in the non-NGF moiety of the predicted prohormone. Exon IIIA is juxtaposed next to IIIB and accounts for all the unique 5' DNA sequences of cDNA D. The library was further screened with a fragment from the 5' end of cDNA A, this yielded a phage containing an additional 127 bp exon (exon II) in the 5' direction. To obtain the remaining 33 bp of the transcribed portion of the gene, we rescreened the library with a probe specific for this region (5' end of exon II). Additional clones identified in this way extend 2.5 kb further in the 5' direction but lacked the final 33 bp of transcripts A and B. Repeated screening of two mouse phage genomic libraries with a unique sequence in the most 5' region of

FIGURE 2. S1 nuclease analysis of submaxillary gland RNA using cDNA clones C and D as probes. A) 20 µg of male mouse submaxillary gland RNA (lanes 1-3, 6-8) was assayed using single-stranded probes derived from C (lanes 1-4) and D (lanes 6-9). One unit of S1 nuclease per reaction was used in lanes 1 and 6, 10 units in lanes 2 and 7, 100 units in 3 and 8; tRNA controls (20 µg, lanes 4 and 9) were digested with 10 units of S1 nuclease. Arrows indicate the protected fragments, with the two major transcripts represented by the lower, intense band (located at the anticipated splice junction). B) Diagram of the S1 nuclease probes derived from cDNAs C and D. An oligonucleotide complementary to +328 to +306 (of transcript A) was annealed to the appropriate M13 subclone, extended with the Klenow fragment of E. coli DNA polymerase, digested in the M13 polylinker (represented by the dashed lines) and prepared from a 5% urea-polyacrylamide gel. cDNA C has an internal EcoRI site so the relevant (3') subclone was used to generate the probe. The size of the probes is shown to the right in nucleotides (nt). The arrows below indicate the predicted sites of cleavage by S1 nuclease along with the sizes of the protected fragments.


1 2 3 4 5 6 7 8 9
1353
1078
872
603
310
281
271
234
194
118

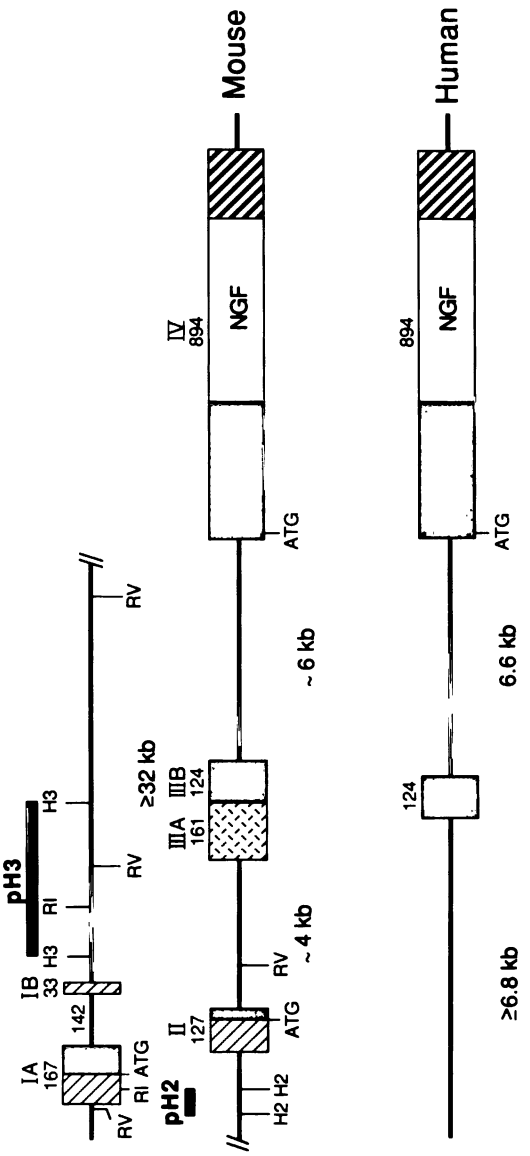

+32
+159
+284
transcript A
transcript C
EcoRI
324 nt
+306 +328
277 nt
168 nt
transcript D
383 nt
336 nt
168 nt
+306 +328

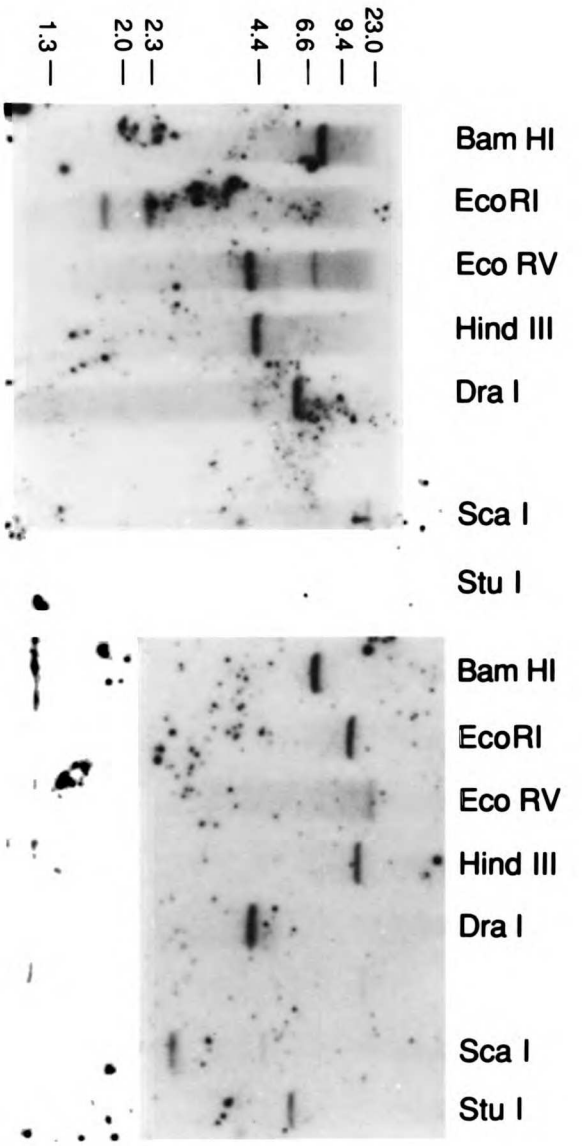
the above genomic clones failed to yield phage that had the missing exon.

The identification of cDNA C provided a new sequence with which to isolate the 5' end of the NGF gene. Using the most 5' EcoRI fragment of cDNA C as probe, we identified genomic clones possessing the exon unique to transcript C (exon IA, fig. 3A,C) and in addition, the missing 33 bp exon of transcripts A and B (exon IB, Fig. 3A,C). Only 142 bp separates these two 5' exons (Fig. 3C). Although these phage clones contain approximately 12 kb of genomic DNA in the 3' direction, they do not overlap with the previously isolated 3' phage clone. Neither probes specific for the 3' end of the 5' phage clone (pH2) nor for the 5' end of 3' phage clone (pH3) hybridizes to any single restriction fragment on a Southern blot of genomic DNA (Fig. 3b). The largest bands hybridizing to the probes were found in EcoRV-digested DNA. The pH3 DNA hybridizes to a 7.0 kb fragment extending 3' from an EcoRV restriction site in the intron between exons I and II. PH2 hybridizes to an ~23 kb fragment extending 5' from an EcoRV site in the intron between exons II and III. This indicates a minimum size of around 32 kb for the intron between exons I and II. Further mapping reveals the sizes of the second and third introns to be approximately 4 and 6 kb, respectively. Thus, the NGF gene comprises more than 43 kb. The sequences of all the exon/-intron boundaries follow the consensus signals for RNA splicing (Mount, 1982) (data not shown). A comparison of the organization of the murine NGF gene with the known portion of the human NGF gene reveals a perfect correspondence in the sizes of the two most 3' exons and in the location of the introns (Ullrich et al,. 1983; Darby et al., 1985) (Fig. 3A).

Primer extension experiments have already defined the transcription

FIGURE 3. A). The organization of the mouse and human NGF genes. The exons are denoted by boxes and introns by lines, sizes and nomenclature for the exons are given above. The dark boxes signify the precursor coding DNA, clear boxes the mature NGF, and hatched boxes the 5' and 3' untranslated regions. The location of all three potential initiator AUGs is noted below the corresponding exon. The dark bars above the introns correspond to the two probes used in the genomic Southern analyses. Noncoding exon IIIA is indicated. B). Southern blot of mouse genomic DNA. Ten ug of approximately digested genomic DNA was separated by electrophoresis in a 0.9% agarose gel. After denaturation, the gel was transferred to nitrocellulose and hybridized first with a labeled DNA fragment from the pH2 genomic subclone (right). After removing this probe by boiling in water, the same blot was rehybridized with the pH3 subclone (left). C). Sequence of the 5' end of the mouse NGF gene. The sequences of the two exons, IA and IB, are in bold type and boxed. The TATA-like (TTAAA) element for exon IB and the presumptive initiator methionine in exon IA are both underlined.


Mouse
Human
IV
894
NGF
ATG
~ 6 kb
≥32 kb
IIIB
124
IIIA
161
~ 4 kb
II
127
ATG
IB
33
142
IA
167
RI ATG
pH2
pH3
H3
RI
RV
H2 H2
6.6 kb
≥6.8 kb


23.0
9.4
6.6
4.4
2.3
2.0
1.3
Bam HI
EcoRI
Eco RV
Hind III
Dra I
Sca I
Stu I
Bam HI
EcoRI
Eco RV
Hind III
Dra I
Sca I
Stu I

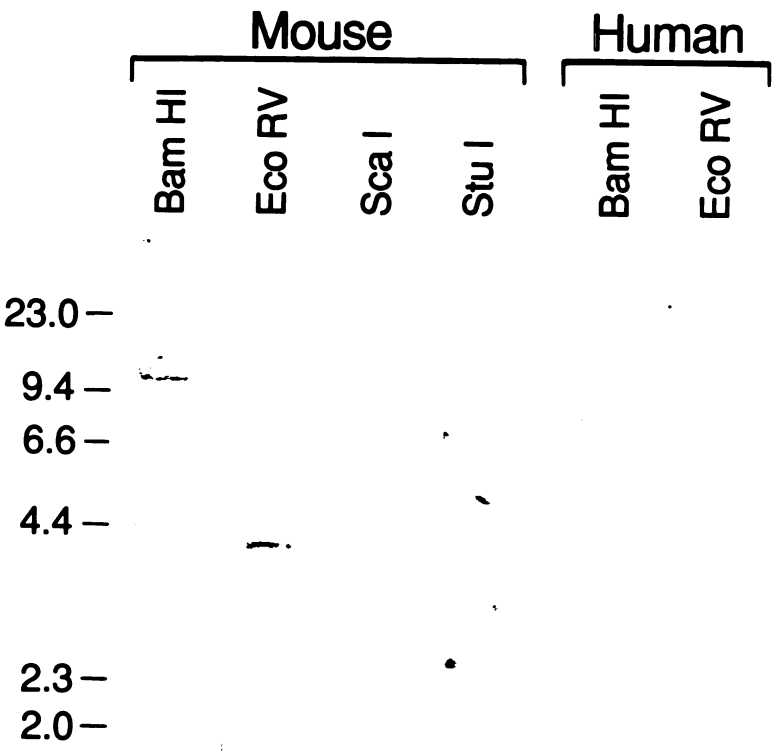

GGGGAAGAGGTAAAGGGACTCCAACTCAGCCAATCAACTCAGCCCTGCTCAGCCT
GTCTGCCCATTGGTGTTTGGAGCCAGCTACCTTGCCTGGTGCCCTAGAAAATCTCT
CTCTGCATCTGTGACCTCCCCCCACCATGCATTTTCATCTAAACCAGGGG
TTTTGAATTCTTCCCTGACTGGCTTTTCCTGGCTATGTCCCATCAACTCGG
GAGCTATCCATCCCTTGTCCCCAGGACCCTTACAACCCGGACCCCTGGG
TCTAGTCACAGCAG GTGCGGGCTGGGATTGGAGTTGGCCAGAGAGGGAGGGGT
CGGGTGAGTGGGGGGCAGGATTTGGAGAGGGTGTGACGAGCCTGGAGGAGGGGCT
AAATACAGTCAGGAAGCCTGTTAAAGAAGCTCTGTGCTCCAGCACGGCAGAGAGC
GCCTGGAGCCGGAGGGGAGCGCATCG GTGAGTCAGGCTTCTCTGAGCCGAG
CCCGATACGGGGGCA

start sites of the two major NGF transcripts (A and B) (24). The sequence upstream contains a TATA-like element, TTAAA, at -28 and an unusually GC-rich (68%) region at -50 to -140 relative to the cap site (Fig. 3C). Because of the extremely low abundance of transcript C, we could not clearly identify the transcriptional start site by primer extension or S1 nuclease analyses (using a shorter genomic probe rather than the larger, and therefore more sensitive, cDNA probe used in Figure 2). The absence of an upstream consensus splice acceptor sequence and the presence of an in-frame upstream termination codon suggest that translation initiates at the AUG indicated in Figs. 1 and 2. Examination of the genomic sequence flanking the 5' end of cDNA C does not reveal a TATA element (Corden et al., 1980) or sequences similar to the SV40 21 bp repeat that have been observed in several other genes (Reynolds et al., 1985; Valerio et al., 1985). Two CCAAT (Benoist et al., 1980) boxes are observed at positions -79 and -180 (data not shown). To determine whether transcript C is conserved, we probed a Southern blot of human genomic DNA with a 5' sequence unique to cDNA C, and detected a single hybridizing band (Fig. 4). A probe containing the 5' end of transcript D (Exon IIIA) and extending 200 bp in the 5' direction was fully protected from S1 nuclease digestion using mouse submaxillary gland RNA (data not shown). This suggests that this cDNA may represent a partially spliced intermediate in the processing of the primary NGF transcript.

Many Adult Tissues Express Both NGF Transcripts A and B

The distribution of NGF transcripts was determined in adult mouse tissues. Primer extension experiments with a high specific activity, single-stranded primer (complementary to +328 to +199 of transcript A)

FIGURE 4. Cross-species genomic hybridization: 5' end. 10 µg of digested human and mouse genomic DNAs were electrophoresed on 0.9% agarose gels and transferred to nitrocellulose paper. The filter was hybridized using standard conditions except lower formamide (35%) followed by washing at a lower temperature (45°C). The enzymes above the figure correspond to the enzyme used in that lane; the numbers on the side represent the sizes of HindIII-digested λ DNA.


Mouse
Human
Bam HI
Eco RV
Sca I
Stu I
Bam HI
Eco RV
23.0
9.4
6.6
4.4
2.3
2.0

showed that in contrast to the salivary gland, where transcript A (the longer extension product) predominates (overexposed in figure), all the peripheral organs and brain regions tested have transcript B (the short product) as the major NGF mRNA (Fig. 5). Densitometry shows ratios of approximately 3:1 (A:B) for the submaxillary gland and 1:4 for the other tissues (data not shown). Intermediate extension products presumably result from premature termination; this would produce a modest underestimation in the abundance of transcript A. It is unlikely that the long extension product represents transcripts C and D in addition to A (cDNAs C and D are virtually the same length) because they are very rare even in the submaxillary gland. There appears to be no sexual dimorphism in the NGF gene expression in the mouse brain (Fig. 5, lanes 5,6), as there does in the submaxillary gland.

NGF Transcripts Increase in Parallel Through Early Postnatal Development

To determine whether the expression of the various NGF transcripts correlates with a phase of neural development, we have used a sensitive S1 nuclease assay to distinguish the transcripts in RNA samples obtained at different times after birth. The single-stranded, uniformly labeled (^{32}P-dCTP and ^{32}P-dGTP) probe extends from +328 past the cap site of transcript A (into the M13 polylinker) and hence distinguishes transcript A from other transcripts divergent at the +159 splice junction (mostly transcript B, Fig. 6B). In the cortex, the level of NGF transcripts increases during the weeks after birth, and reaches a peak at 20 days; it then stabilizes at a slightly lower level (Fig. 6A). Similar S1 nuclease analyses of globe RNA (retina as well as iris) show a roughly 2-fold increase after birth, with a peak at 14-17 days, somewhat earlier than in the cortex (data not shown). Heart RNA also shows a

FIGURE 5. Primer extension analysis of NGF transcripts from various mouse tissues. A primer extending from +328 to +199 (of transcript A) was uniformly labeled with ^{32}P-dCTP and ^{32}P-dGTP. After annealing to the following amounts of RNA, the primer was extended with reverse transcriptase.

Lane 1: 10 µg female submaxillary gland total RNA.

Lane 2: 3 µg small intestinal rat RNA.

Lane 3: 5 µg kidney polyA+ RNA.

Lane 4: 1 µg heart polyA+ RNA.

Lanes 5,6: 3 µg of male and female brain polyA+ RNAs.

Lane 9: 3 µg deep grey matter polyA+ RNA.

Lane 10: 1 µg brainstem polyA+ RNA.

Lane 11: 1 µg cerebellum polyA+ RNA.

Lane 12: 3 µg newborn brain polyA+ RNA.

Lane 13: HaeIII digest of ϕX174 DNA.

Lane 14: 20 µg yeast tRNA.

Arrows indicate the major extension products with the predicted sizes for the two major transcripts A and B (328 and 201 bases, respectively).

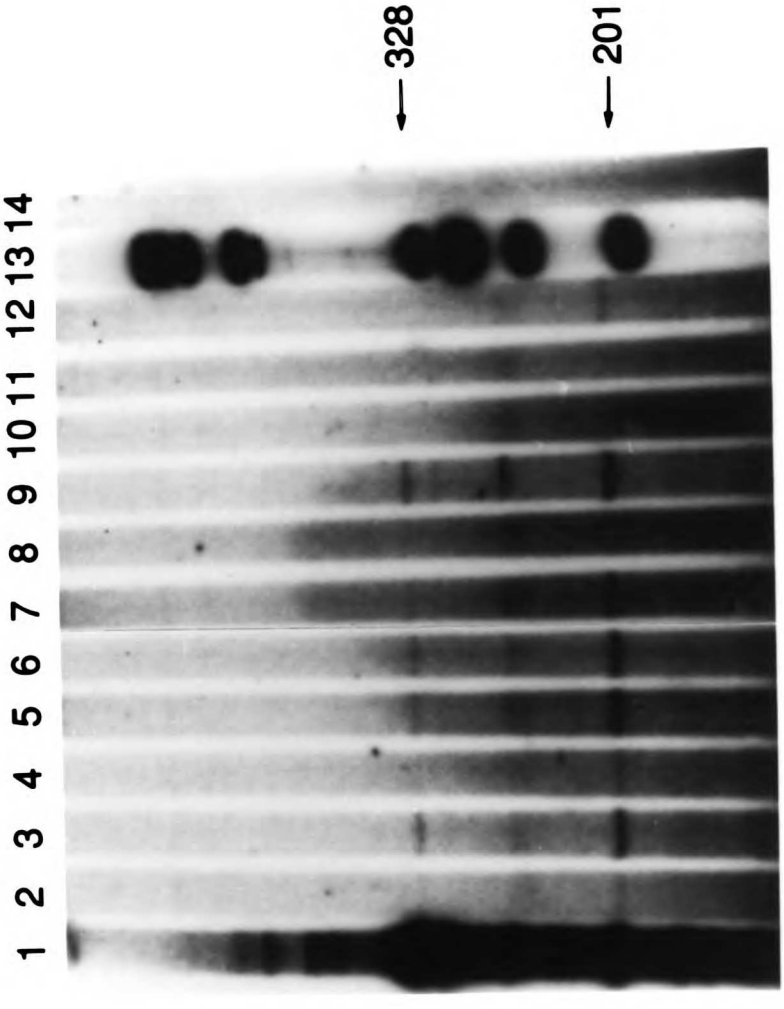
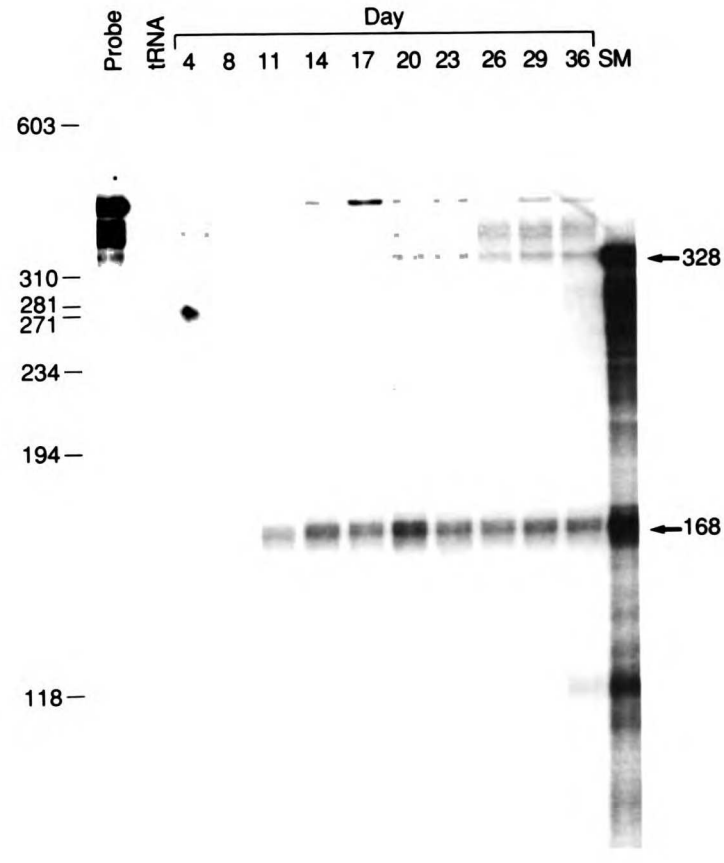
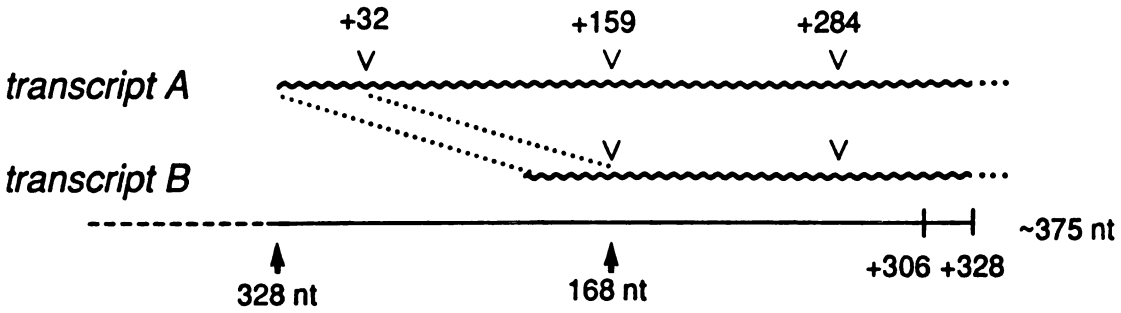

328
201
1 2 3 4 5 6 7 8 9 10 11 12 13 14

FIGURE 6.

A) S1 nuclease analysis of NGF transcripts from the developing mouse cortex. 25 µg of total RNA from mouse cortex at the postnatal days shown was hybridized to a probe representing the full 5' end of NGF transcript A. Probe fragments protected from S1 nuclease digestion by the RNA were separated on a 5% urea-acrylamide gel; the dried gel was exposed with an intensifying screen at -70°C for one week. SM denotes female submaxillary gland total RNA (less than 5 µg), and yeast tRNA was used as a negative control. Arrows indicate the sizes predicted for fragments protected by transcripts A and B. A HaeIII digest of ϕX174 DNA is marked to the left.

B) Diagram of the two major NGF transcripts (wavy lines) in relation to the S1 nuclease probe (straight line). Transcripts A and B differ only in the presence of the second exon (+32 to +159 of A). The probe was synthesized on a single-stranded M13 template (subcloned from a full-length cDNA A) using a primer complementary to +328 to +306. Arrows indicate the sites of cleavage in the probe depending on whether it is protected by transcript A (328 nt) or B (168 nt).


Probe
tRNA
Day
4 8 11 14 17 20 23 26 29 36 SM
603
310
281
271
234
194
118
328
168


+32
+159
+284
transcript A
transcript B
~375 nt
328 nt
168 nt
+306 +328

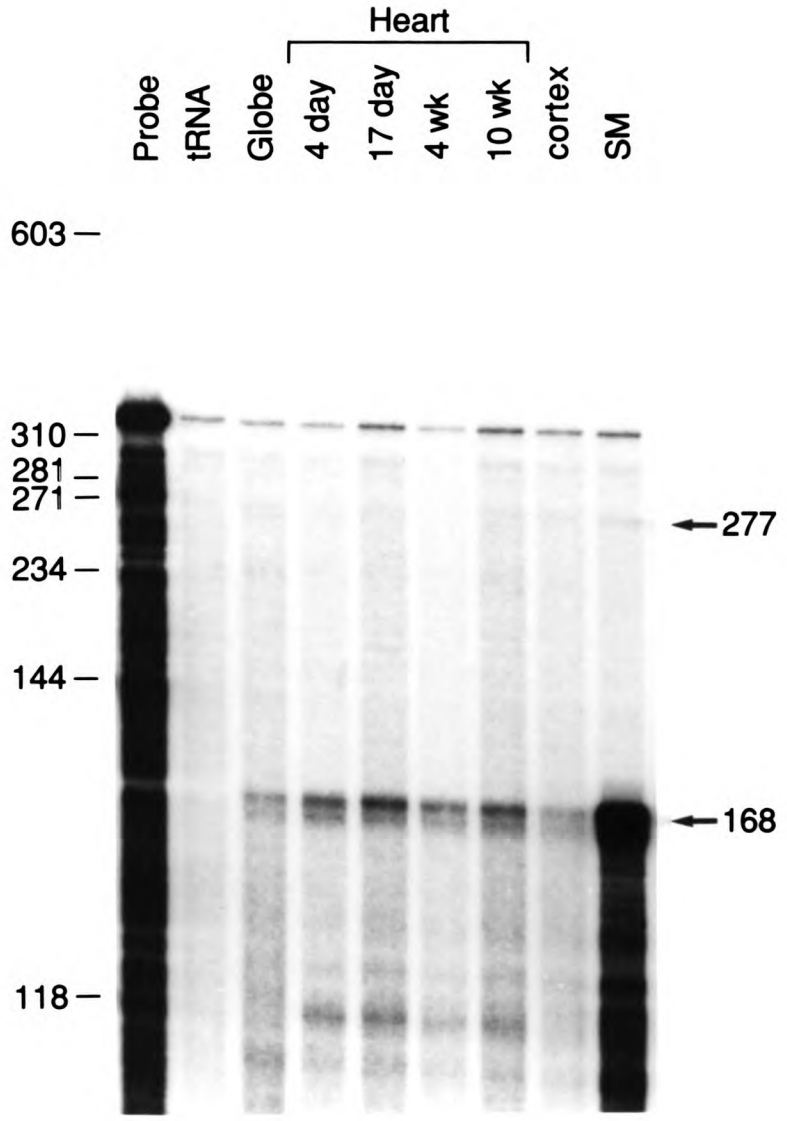
doubling in NGF expression by 17 days, which is followed by a considerable decline (data not shown). At all times transcript B predominates over A (4-8 fold as determined by densitometry) in the different tissues. The patterns remain stable from four weeks to adulthood (Fig. 6).

Using a high specific activity probe for transcript C, we detected a faint signal with the submaxillary gland, the 10 week heart and adult cortex total RNA (Fig. 7). Unlike the salivary gland where the level of C relative to B+A is very low (less than 1% of the total NGF RNA), the level of transcript C in cortex and heart have 5-10% of total NGF mRNA.

Whisker Removal And Consequent Re-Arrangement of the Sensory Pathway Does Not Alter the Levels of NGF Transcripts in the Somatosensory Cortex or Dorsolateral Brainstem

Synaptogenesis occurs at roughly the same time in development as the observed changes in the level of NGF transcripts. To determine whether synaptogenesis has a role in causing the initial rise then stabilization of NGF mRNA in the cortex, we attempted to prevent the normal formation of synapses in that brain region. For these experiments we have used the mouse whisker pad system with its convenient internal control (the opposite side) (Woolsey et al., 1981). This system has revealed the importance of sensory experience in the formation of certain neural networks. Three neurons connect whiskers to the rodent somatosensory cortex. Removal of the right whisker at birth disrupts the development of these neurons. Axons from the second neuron in the pathway (with its cell body in the left thalamus) then fail to reach their target cells (in the left cortex) as they would normally. While it

FIGURE 7. S1 nuclease analysis of transcripts corresponding to NGF cDNA C. The probe used is described in Figure 2. Standards (left) are a HaeIII digest of ϕX174 DNA. Arrows at the right denote the predicted sizes for the protected fragments.

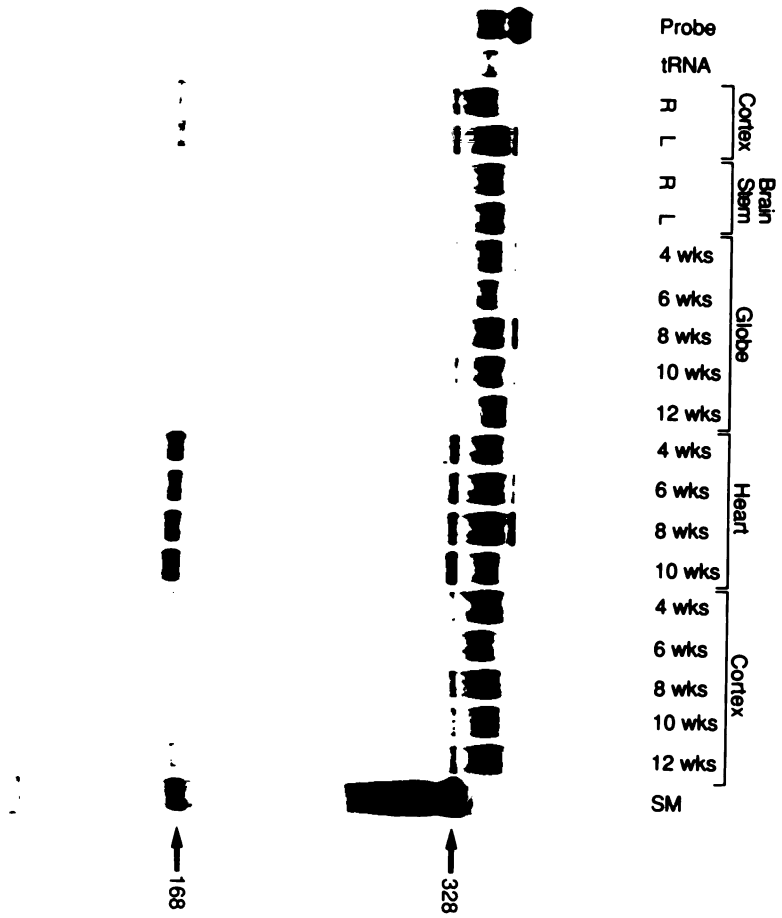

Heart
Probe
tRNA
Globe
4 day
17 day
4 wk
10 wk
cortex
SM
603
310
281
271
234
144
118
277
168

is not known whether the thalamocortical connection involves NGF, sensory deprivation at this stage of development dramatically disturbs the architecture of the somatosensory cortex (Woolsey et al., 1981). We removed the right whisker pads of mice within one day after birth. (Subsequent examination at 4 weeks after birth showed no regrowth of whiskers.) S1 nuclease protection analysis showed no difference in NGF mRNA expression, left from right (Figure 8), at 4 weeks in the relevant somatosensory cortex. Because we may have missed the critical time in which the lesion could produce an effect, we also removed whisker pads at 1,2 and 3 weeks after birth; somatosensory deprivation beginning at these times also had no effect on the levels of cortical NGF transcripts at 4 weeks after birth (data not shown). Additionally, the lesions produced no alteration in four week NGF transcript in the brainstem (where synaptic input is even more directly affected by whisker removal (Durham & Woolsey, 1984).

DISCUSSION

We have elucidated the structure of the NGF gene and its major transcripts. The NGF gene has multiple small exons in the 5' regions which were not previously evident. This organization has contributed to the difficulty in isolating the full gene (Darby et al., 1985; Ullrich et al., 1983). The four transcripts identified by cDNA cloning predict three closely related NGF precursor proteins which diverge only at their N-termini. The mature NGF encoded by the various transcripts is transcribed from the large 3' exon. Transcripts A and B differ as a result of alternative splicing at exon II; transcript C appears to derive from an independent promoter; transcript D may be a partially spliced transcript. Thus, NGF gene expression is probably regulated by transcrip-

FIGURE 8. NGF transcripts after whisker pad removal and in late development. A full-length 5' probe (as in Figure 5) was digested with S1 nuclease after hybridization to 25 µg of total RNA from the tissues shown. The right whisker pad was removed at birth and the animals sacrificed four weeks later; tRNA and SM samples are the same as in Figure 5, with arrows showing the sizes of major short and long NGF transcripts.


Probe
tRNA
R
L
R
L
Cortex
Brain Stem
4 wks
6 wks
8 wks
10 wks
12 wks
Globe
4 wks
6 wks
8 wks
10 wks
Heart
4 wks
6 wks
8 wks
10 wks
12 wks
Cortex
SM
328
168

tion initiation and by RNA processing.

Significantly, the differences in the sequence of the amino-termini of the precursors alter the position of the hydrophobic domain assumed to function as the signal peptide. Whereas the precursor predicted by transcripts B and D have this domain in the N-terminal position, transcripts A and C predict the same hydrophobic sequence is about 70 amino acids from the amino-terminus of the precursor. These differences in structure might signify functional differences of these precursors. These putative functional differences could be reflected in distinctive transcriptional patterns in different tissues and throughout development. However, we have found that all tissues, except for mouse submaxillary gland and the placenta (Edwards et al., 1986), have the same profile of NGF transcripts. Transcript B predominates; A is present at 20-30% of transcript B, and C and D are barely detectable. In mouse submaxillary gland (and placenta), transcript A is the major form; B is about 30% of A, and C and D are 0.1-1.0% the level of A. During postnatal development, A and B change in parallel. Transcript C appears at a higher level relative to A and B in cortex and heart, compared to the salivary gland. Thus its promoter may be independently regulated. This observation, however, does not necessarily imply a separate function for the C-encoded prohormone, especially in light of the very low abundance of this transcript.

These data do not allow us to conclude whether or not the precursors predicted by transcripts A and B have unique functions. The constant ratio of A to B may reflect a property common to innervated tissues that is not shared by the mouse submaxillary gland and the placenta. On the other hand, the data are compatible with the notion

that A and B are functionally equivalent. In fact, in expression experiments, both precursors give rise to mature NGF with similar facility (Edwards, Selby & Rutter, unpublished).

The quantity of total NGF mRNA changes during the postnatal period. Whittemore et al. have shown an increase in NGF mRNA in brain that plateaus 3-4 weeks after birth (1986). Large et al. have verified the rise in cortical NGF message but show a decline after the peak at 3 weeks (1986). We have confirmed the increase, decrease, then stabilization of NGF mRNA levels with respect to total RNA, but the magnitude of change (2-4 fold) is less than that seen by Large et al. (~8-fold). The difference may be attributable to the marked neonatal variation in the proportion of total brain RNA that is mRNA. The significance of these changes is unclear; they may only reflect a variable dilution of NGF transcripts by the rapidly changing RNA content of these tissues.

The changes in NGF gene expression correlate with synapse formation in the tissues examined. In the central and peripheral nervous systems, exemplified by the cortex and heart, the majority of ingrowing nerve terminals arrive at their target organ before birth. One large afferent pathway to the cortex originates in the thalamus. Its axons reach the cortex at birth and approach the cells destined for innervation shortly thereafter. However, the connections between thalamic and cortical cells continue to mature physiologically for weeks (Hubel et al., 1977; Wise & Jones, 1978). Postganglionic sympathetic neuronal processes are present in the heart before birth (Champlain et al., 1970), but do not develop varicosities until later; the connection does not become fully functional until two weeks after birth (Wekstein, 1965).

Because the period during which NGF exerts its critical effect on

neuronal survival overlaps the maturation of synaptic contacts, we attempted to ask whether innervation per se causes the developmental changes in NGF expression. Since NGF transcripts increase, then stabilize, in the developing cortex, whisker pad removal might alter this pattern of NGF expression. We analyzed whisker pad RNA at four weeks, just after the peak of NGF expression. We were unable to observe any change in the expression of the two principal NGF transcripts, A and B (Fig. 6). Because the critical window of responsiveness to sensory deprivation might occur later, we performed the same lesion at 1, 2 and 3 weeks after birth: there was no detectable change in NGF gene transcripts at four weeks. The failure to observe changes in NGF transcripts during this critical developmental period suggests that innervation does not itself regulate NGF mRNA expression in this system. Similarly, sympathetic and sensory denervation of the iris in situ also does not produce any increase in NGF mRNA (Shelton & Reichardt, 1986). These experiments do not eliminate the possibility that NGF transcripts in supporting cells (e.g. fibroblasts, Schwann cells) change during regeneration or after injury. However, there is no evidence from these experiments of a sensitive feedback regulation system operating at the level of NGF transcripts. We hypothesize that the apparent invariant NGF expression is a reflection of the inherent capacity of the target cells to be innervated.

The low level of NGF expression and the lack of feedback control is compatible with a hypothesis accounting for selective cell death in the development of NGF-responsive neuronal networks: we propose that the two cell types making NGF, neural supporting cells (e.g. Schwann cells) and innervated cells, synthesize NGF at low in an essentially unregulated

fashion. The innervated cells are the major supply at least during development. If NGF acts only in close proximity to the producing cells, a special circumstance may exist in the region of the synapse: NGF secreted by the incipient innervated cell may in part determine the extent of innervation. Indeed, NGF secretion could even be localized to the site of incipient synapse formation. The NGF receptors of developing nerve fibers may then bind the NGF and remove significant quantities via the axonal transport system. Innervating cells in the vicinity then compete for the limiting available NGF according to (a) proximity to the NGF source, (b) the number and affinity of NGF receptors on the nerve terminal, (c) some aspect of neural activity (e.g. recycling of nerve terminal membrane). The axon in the most favorable anatomical and function position eventually becomes the dominant consumer of the local supply of NGF while the availability to others is reduced. Supply to the neurons competing successfully could occur via some specialized aspect of the synapse structure. In this fashion the neurons with the appropriate connections live, the other ones receive an inadequate supply of NGF, and die. Rejection or extirpation of the neurons would result in an increase in NGF because of the absence of active axonal transport, and presumably turnover in the nerve cells.

CHAPTER V

IN VITRO ANALYSES OF THE TWO NGF PRECURSOR PROTEINS

INTRODUCTION

Two cDNAs encoding related NGF precursors of 34 kd and 27 kd have been identified (Scott et al., 1983; Edwards et al., 1986). The mRNAs for these precursors derive from alternative RNA splicing of the second exon of the NGF gene (Edwards et al., 1986; Selby et al., submitted). Both precursors encode mature NGF at their carboxy termini and are identical in the non-NGF precursor region, except that the 27 kd species is truncated at its amino terminus relative to the 34 kd molecule. The initiator methionine of the 27 kd precursor is followed by a sequence that bears similarity to signal sequences found at the N-termini of most secreted proteins (Blobel, 1983; Perlman & Halvorson, 1983). On this basis, one would predict that this precursor is secreted. Curiously, no signal like sequence appears at the N-terminus of the 34 kd precursor; rather, it is situated internally. Thus, it seemed possible that this precursor might have a different cellular function, e.g., it might be a transmembrane protein and display the functional NGF moiety on the cell surface. In lieu of an operational heterologous expression system, I opted to determine the fate of the synthetic precursors in an in vitro system.

MATERIALS AND METHODS

SP-6 Transcription

The construction of the two SP-6 subclones used here have been previously described (Edwards et al., 1986). Labeled SP-6 transcripts were generated in vitro with α^{32}P-GTP using templates linearized at HindIII or NcoI. Unincorporated ribonucleotides were removed by two rounds of ammonium acetate precipitations. The labeled RNA was dissolved in sterile water and a fraction of the counts from each sample

was added to formamide dyes, heated to 95°C for 3 minutes and loaded onto a 6% polyacrylamide-urea gel (Maniatis et al., 1983). Labeled DNA standards (ϕX174) were electrophoresed in adjacent lanes. The gel was visualized by autoradiography with an intensifying screen. Unlabeled SP-6 transcripts were generated following manufacturer's instructions.

In Vitro Translations

RNAs generated by the SP-6 system were used to program wheat germ translational extracts prepared according to Erickson and Blobel (1983). Pancreatic microsomal vesicles (ERM) were prepared as described by Walter and Blobel (1983). The reactions were analyzed on SDS polyacrylamide gels following dilution with Laemmli buffer, or after TCA precipitation. Endoglycosidase H (Endo-H) reactions were carried out overnight at 37°C with one unit of enzyme as in Perara and Lingappa (1985). Protease protection experiments used proteinase K at a final concentration of 300 µg/ml (Walter & Blobel, 1983); Triton X-100 was added to 1% in protease and detergent experiments. Translation extracts were layered on a sucrose cushion consisting of 0.5M sucrose, 0.5M KOAc and 2mM $Mg(OAc)_2$. The samples were then centrifuged at 30 psi in a Beckman airfuge for 5 minutes. The supernatants and pellets were precipitated with TCA and gel electrophoresed. For two-dimensional gel analyses, the samples and gels were prepared according to O'Farrell (1975). The wheat germ extracts, membranes and Endo-H were generous gifts of Dr. Peter Walter and Pablo Garcia.

γ Subunit Reactions

The γ subunit was generously provided by Dr. Eric Shooter. α and γ subunits were purified from the 7S complex in the saliva by following the procedures of Burton et al. (1978). γ activity was determined

fluorometrically after incubation with the synthetic substrate, D-val-leu-arg-amino-fluoro-coumarin. Translation reactions were diluted to a final Tris-HCl concentration of 100mM (1977). The γsubunit was added at the indicated amount and digestions proceeded for 20 minutes at 37°C. The reaction products were then analyzed by PAGE.

RESULTS

Hydrophobic Domains

A hydropathy plot (Kyte & Doolittle, 1982) of the amino acids of the 34 kd NGF precursor showed only one major hydrophobic peak which started near the first internal methionine residue (Fig. 1). In the 27 kd precursor, the initiator methionine flanks this hydrophobic core. Hydrophobic residues typically comprise the core of sequences signalling secretion or membrane insertion. No other region displays a similar long stretch of hydrophobic residues.

In Vitro Translation of NGF Precursor RNAs

Gel analysis of the labeled SP-6 NGF transcripts (SmaI-PstI/-HindIII) shows that they are principally full-length (Fig. 2). A truncation at the internal NcoI restriction site generates a shorter transcript as predicted. The shorter transcript presumably represent prematurely terminated species. They should, in turn, generate truncated translation products.

To test for NGF translation products and their N-glycosylation, portions of the NGF cDNA were subcloned into the SP-65 plasmid to generate two clones which would direct the expression of RNAs corresponding to the short (BstXI partial-PstI) and long (SmaI-PstI) transcripts. SDS polyacrylamide gel electrophoresis (PAGE) of wheat germ extracts programmed with the two synthetic RNAs showed translation

FIGURE 1. Schematic display of the hydrophobicity plot (Kyte & Doolittle, 1982) of the long (34 kd) precursor. The amino acid sequence beginning with the initiator methionine is presented below. The center is denoted as 0 in the plot, while the outside lines represent -2 (hydrophilic) to +2 (hydrophobic). Note the relatively wide and high peak from His 64 to Val 82 that appears 25% to the right of the left end of the diagram. This region represents the most significant hydrophobic region of the entire molecule. Most other hydrophobic regions are not as wide or as high as this one. Also note the absence of a significant hydrophobic region at the amino-terminus of the precursor.

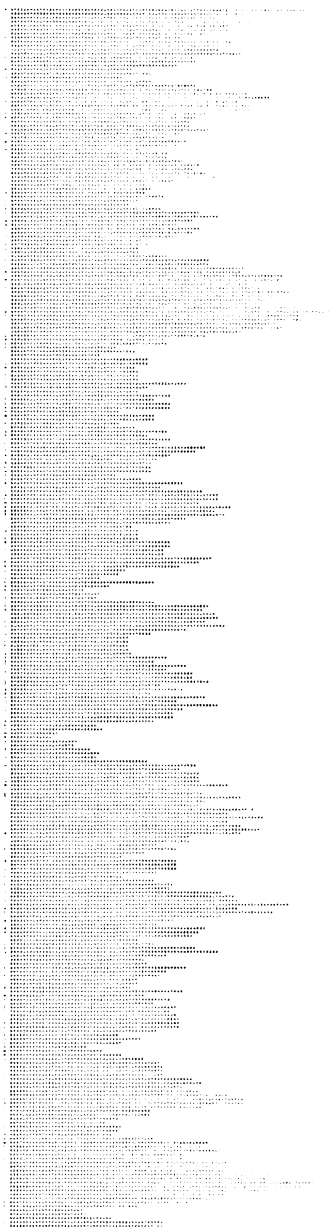
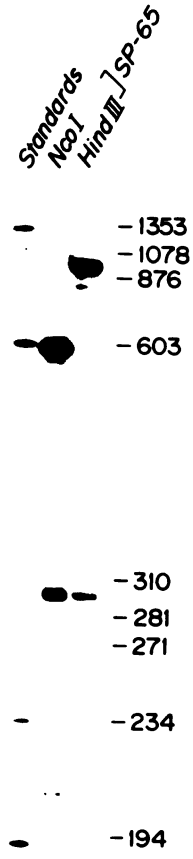


FIGURE 2. In vitro transcription analysis. An SP-65 plasmid containing the NGF cDNA (SmaI-PstI) was linearized with either NcoI (truncated transcript) or HindIII (full-length transcript) and used to program in vitro transcription reactions (using α-^{32}P-GTP) (see manufacturer's specifications). A fraction of each reaction (1 ul) was electrophoresed on an 8% sequencing gel along with labeled standards (ϕX174) and exposed to autoradiography. The numbers to the right of the figure correspond to the sizes of HaeIII-digested ϕX174 DNA. Note that the major product in each lane migrates the slowest and represents full-length transcripts. The other products represent prematurely terminated transcripts.


Standards
NaoI
HindIII
SP-65
1353
1078
876
603
310
281
271
234
194

products of roughly the predicted sizes of 27 and 34 kd (Edwards et al., (1986). Further, transcripts specifically truncated by restriction enzyme cleavage within the cDNA (ScaI and NcoI) resulted in the predicted diminution in size of the proteins (Edwards et al., 1986).

To determine if signal peptide cleavage and N-linked glycosylation occurred, different amounts of dog pancreas microsomal membranes (ERM) were added to wheat germ extracts and the synthetic precursor RNAs. PAGE analysis of the reaction products demonstrated that lower molecular weight species were generated in the presence of ERM (Fig. 3). These most likely result from signal sequence cleavage. Examination of the N-terminus of the 27 kd precursor reveals that a potential cleavage site is present between alanine 18 and glutamate 19 (von Hjeine, 1986); cleavage here would reduce the molecular weight of this precursor by almost 2 kd. Cleavage between the same residues in the 34 kd precursor would remove roughly 9 kd of protein, giving it a mobility identical with that of the processed 27 kd precursor. The 27 kd precursor is further modified for in the presence of ERM, a higher molecular weight species of ~34 kd is observed. This is apparently due to N-linked glycosylation because this species disappears following endo-H treatment. No similar modification is apparent in the longer precursor but may be masked by full-length precursor. One typically observes translocated glycosylated and non-glycosylated species in ERM; the non-glycosylated species presumably represent intermediates in glycoprotein biosynthesis (Perara & Lingappa, 1985) and/or may reflect lability of the enzyme responsible for glycosylation.

Analytical analysis of the reaction products translated in the presence of membranes showed that the two precursors are translocated

FIGURE 3. In vitro translations: Titration of ERM. Lane 1 contains no ERM, all lanes contain SP (SmaI-PstI) or BP (BstXI-PstI) RNAs. ERM was added to samples 2-6 at 1, 2, 3, 4 and 4 µl, respectively. Sample 6 was treated with Endo-H. Following incubation at 26°C, samples were TCA precipitated, resuspended in loading buffer, heat-denatured and then loaded on a 10-15% polyacrylamide gel. The gel was dried and exposed to autoradiography for 12 hours. Arrows to the left and right of the figure denote nascent precursors and their respective signal sequence-cleaved precursors. Note that in the presence of ERM, a band that migrates ~34 kd appears in the 27 kd precursor. This modification is a consequence of N-glycosylation because the band disappears after Endo-H digestion. No similar N-glycosylation is apparent in the translation/translocation of the 34 kd precursor but may be masked by the full-length 34 kd species. The principle bands in SP correspond to 34 kd and ~25 kd (after presumptive signal sequence cleavage); in BP, bands correspond to ~34 kd (glycosylated form), 27 kd (non-translocated) and 25 kd (other presumptive signal sequence cleavage). Other bands on the gel are derived from either background RNAs from the wheat germ extract or from SP-6 transcripts that are not full-length.

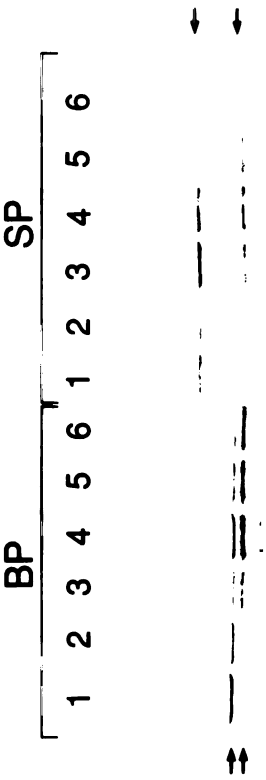

BP
1 2 3 4 5 6
SP
1 2 3 4 5 6

FIGURE 4. Analysis of in vitro translation reactions. Wheat germ translation extracts were programmed with synthetic RNAs corresponding to the two principle NGF transcripts. Figures A and B represent the 27 kd and 34 kd precursors, respectively. As a background control, translation reactions in the absence of RNA (-RNA) were carried out. In the presence of the appropriate RNA (+RNA) the precursors of roughly the predicted mobility are observed. The addition of ERM with and without subsequent Endo-H digestions confirms the previous observation of glycosylation of the 27 kd precursor and not the 34 kd precursor. Incubation with proteinase K (PK) was carried out to determine whether a translocated molecule is fully or partially protected from digestion. Both translocated precursors are fully protected from protease digestion, suggesting that they are not transmembrane proteins with some portion of the precursor exposed on the cytoplasmic face of the ERM vesicles. In the presence of non-ionic detergent, Triton X-100, no significant protection is observed, further supporting the conclusion that both precursors are fully translocated. Fractionation of the translation reactions + ERM demonstrates that translocated species cosediment with ERM (pellet) while non-translocated species are found in the soluble fraction (supernatant). Note that glycosylated 27 kd is found exclusively in the pellet as expected since N-linked glycosylation occurs within the ERM. The other bands on the gel most likely represent prematurely truncated translation products derived from truncated transcripts. The precursors in the supernatant (sup) represent those species translated independent of ERM or that leached out of the vesicles.

46K —
30K —
14K —
Standards
-RNA
+RNA
+RNA +ERM
+RNA +ERM +Endo H
+RNA +ERM +PK
+RNA +ERM +PK + Triton
+RNA +ERM (Sup)
+RNA +ERM (Pellet)

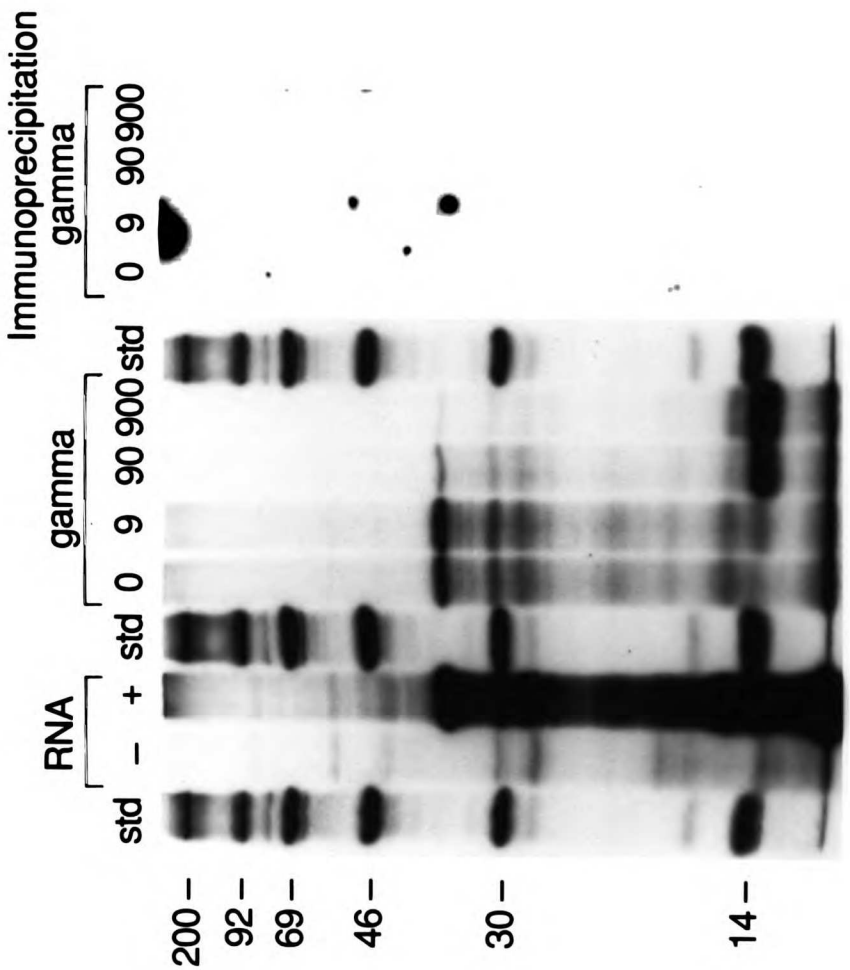
46K —
30K —
14K —
Standards
- RNA
+ RNA
+ RNA + ERM
+ RNA + ERM + Endo H
+ RNA + ERM + PK
+ RNA + ERM + PK + Triton
+ RNA + ERM (Sup)
+ RNA + ERM (Pellet)

into ERM vesicles (Fig. 4). PAGE analysis of microsomes (recovered by centrifugation of the translation extracts through sucrose cushions) showed that translocated 34 kd and 27 kd (non- and glycosylated forms) were recovered intact (see tRNA+ERM (pellet)). Since N-linked glycosylation occurs only with the ERM, no glycosylated 27 kd was found in the supernatant from the fractionation which are non-translocated species (see tRNA+ERM (sup.)). Further, both translocated precursors were completely protected from proteinase K (PK) digestion, suggesting that both precursors are fully translocated and not integrated into the membrane (see tRNA+ERM+PK). If the 34 kd species had been partially translocated (i.e., membrane spanning), protease digestion would have reduced the molecular weight of the precursor by an amount corresponding to the number of amino acid residues found outside of the membranes. The addition of non-ionic detergent (1% Triton X-100) to the protease K digestions causes complete degradation of the precursors because the integrity of the membrane is abolished. These data suggest that the precursors are translocated into the microsomes and that neither displays a significant portion of the molecule outside the membrane.

Gamma Subunit Processing of the NGF Precursors

The two NGF precursors synthesized in vitro are not capable of liberating mature NGF when incubated with the γ subunit of 7S NGF. Incubation of these NGF precursors with increasing amounts of the γ subunit using conditions described by Burger and Shooter (1977) decreased the level of precursor and increased the level of lower molecular weight products (Fig. 5). The presence of microsomal membranes did not seem to influence the cleavage pattern (Fig. 5). γ subunit digestions of carboxy-terminal truncated translation products (no NGF

FIGURE 5. γ subunit digestion of in vitro synthesized precursor. Wheat germ translation extracts including ERM were programmed with synthetic RNA corresponding to the 34 kd precursor. Under the heading "RNA" are TCA precipitations of reactions with and without synthetic RNA. Note the heavily labeled band migrating between the 46 kd and 30 kd standards; this corresponds to 34 kd precursor. Under the heading "gamma" are digests of solubilized ERM containing translocated precursor incubated with increasing amounts of γ subunit as indicated. With increasing amounts of γ subunit, the intensity of the precursor is progressively decreased while more of a lower molecular weight species (14 kd) appears. This gel was exposed to autoradiography for 2 days. Immunoprecipitation of γ subunit treated precursor with anti-NGF antiserum failed to recover the precursor itself or a distinct product derived from it. The exposure for this part of the gel was 8 days.


Immunoprecipitation
gamma
0 9 90 900
gamma
std 0 9 90 900 std
RNA
std – + std
200 –
92 –
69 –
46 –
30 –
14 –

present) showed the same lower molecular weight species as full-length translation products; thus these products are derived from the amino-terminal portion of the precursors (Fig. 6). The γ subunit independently isolated from saliva 7S NGF by R. Edwards and treated in the same way resulted in the same cleavage pattern (data not shown). To confirm that the γ subunit digestions did not generate mature NGF, two dimensional electrophoresis was carried out. The digestion products had a mobility that was substantially different from that of mature iodinated NGF (Fig. 7). Anti-NGF antibodies (Darling et al., 1983) that recognized the precursor did not immunoprecipitate these lower molecular weight species (Fig. 5). Further, the addition of purified α subunit from the saliva 7S complex (prepared by R. Edwards) along with the γ subunit did not modify the precursor cleavage pattern (data not shown).

DISCUSSION

A hydropathy plot confirms that a single major hydrophobic domain is present in both NGF precursor proteins. This domain is found at the N-terminus of the 27 kd species, and is internal in the longer 34 kd precursor. Despite this difference both are translocated across microsomal membranes. Signal sequence cleavage of both precursors was observed. The 27 kd precursor was, however, N-glycosylated as evidenced by its susceptibility to Endo-H cleavage. The signal sequence cleaved 34 kd precursor may also be glycosylated but may be masked by full-length precursor.

It is intriguing that the large N-terminal flanking region was removed from the 34 kd precursor. Further, ERM sedimentation and protease protection experiments seem to rule out membrane insertion of the 34 kd precursor. It is perhaps significant that the hydrophobic

FIGURE 6. In vitro translation of synthetic NGF RNAs. Wheat germ extracts were programmed with synthetic SP-6 RNAs encoding full-length long (34 kd, linearized SmaI-PstI construct in the 3' polylinker at HindIII; SP-H3) and short (27 kd, BstXI-PstI, also linearized at HindIII; BP-H3) precursors. Truncated RNAs were generated from templates linearized at NcoI (with the cDNA coding region; designated SP-Nco or BP-Nco). A portion of each translation was then digested with 900 ng γ subunit, followed by electrophoresis on a 12.5% polyacrylamide gel. The migration of γ subunit-digestion products is identical from both full-length and truncated products in the two constructs (SP & BP). The low molecular weight products generated from the BP constructs are not well resolved on this gel and thus appear as a smear.

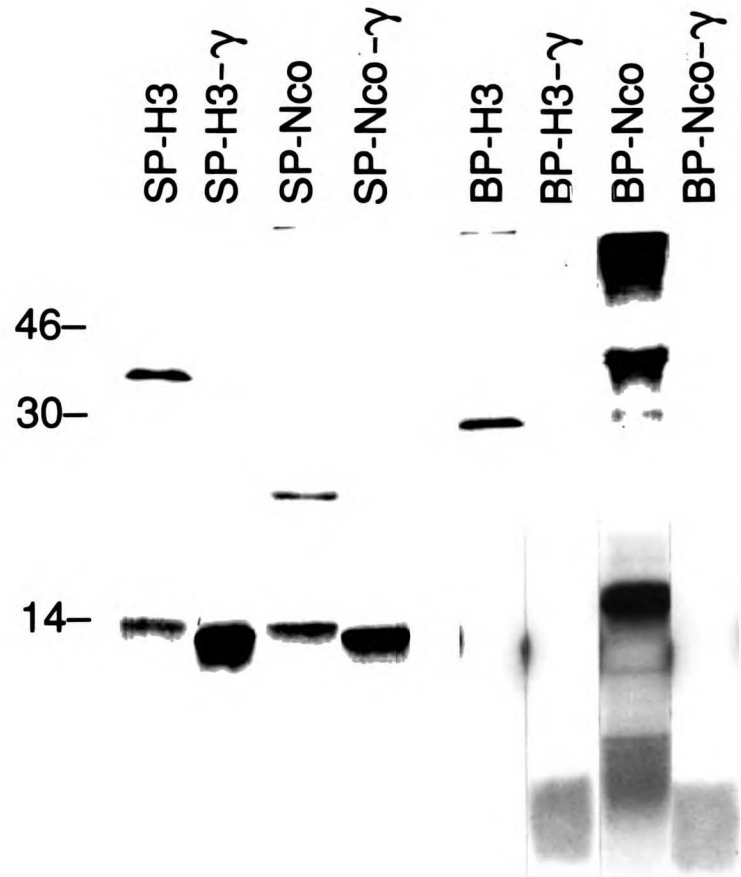
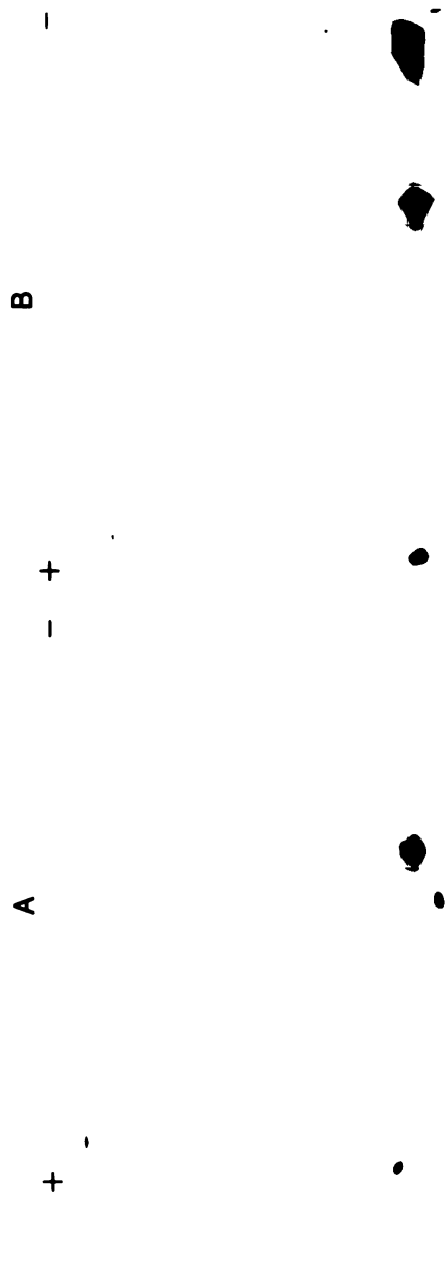

SP-H3
SP-H3-γ
SP-Nco
SP-Nco-γ
BP-H3
BP-H3-γ
BP-Nco
BP-Nco-γ
46–
30–
14–

FIGURE 7. Two-dimensional gel electrophoretic analysis of γ subunit-treated NGF precursor synthesized in vitro. The SmaI-PstI (SP) subclone encoding the 34 kd precursor was linearized at HindIII (in the 3' polylinker) to generate full-length transcripts. These transcripts were used to program a wheat germ translation extract and was followed by γ subunit digestion. The reaction products were resolved on a two-dimensional gel (A). The acidic and basic ends of the first dimension gels are denoted as + and -. Note that one major protein is observed; the minor species is presumably derived from the major species after extensive γ subunit digestion; the second two-dimensional gel (B) corresponds to the same sample but also includes iodinated NGF. Iodinated NGF migrates to a position different from that of the processed products, demonstrating that mature NGF is not produced from the precursor under these conditions.


B
+
-
A
+

region of the 34 kd precursor which might serve as a transmembrane region (Sabatini et al., 1982), is not flanked on either end by charged residues as in the membrane-spanning transferrin receptor (McClelland et al., 1984). In a similar in vitro system programmed with synthetic RNA, an analogous result has been observed: Perara and Lingappa (1985) ligated the globin coding domain (which does not have a signal peptide and is not usually secreted) to the N-terminus of preprolactin and observed that both globin and prolactin were co-translationally transferred across the membranes. Thus, it is presumed that the signal sequence of prolactin mediated the secretion of the hybrid protein. Signal sequence processing yielded the globin sequence with a hydrophobic tail as well as authentic prolactin. Additionally, the secretion of ovalbumin is probably mediated by an internal signal sequence (Lingappa et al., 1979). To decisively demonstrate the function of the internal hydrophobic domain of the 34 kd precursor one could generate a hydrophobic domain-less precursor and determine if it is translocated or not. If this domain is responsible for translocation, no translocation should occur. Also, the presumptive signal sequence of the NGF precursor could be ligated to the N-terminus of a cytoplasmic protein to determine if this protein would then be secreted into ERM vesicles.

In vitro generated precursors did not yield mature NGF on incubation with the γ subunit of 7S NGF. Only amino-terminal polypeptides were stably generated by γ subunit catalysis. Since no other products of the reaction were obvious, one perceives that the NGF moiety was degraded. The size of the lower molecular weight species suggests that they were generated by cleavage at the first set of basic amino acid residues. More extensive digestions generated smaller polypeptides

which were presumably cleaved after Arg or Lys residues. However, correct cleavage may not have been observed because the substrate may not have been in the appropriate configuration. Nascent polypeptides are glycosylated in ERM and are trimmed and modified in the Golgi. It is possible that the γ subunit recognizes only terminally-glycosylated species and not those found in the ERM. The maturation of the carbohydrates might lead to a conformational change in the precursors so that they appear as the appropriate processing substrates. It is also possible that the precursor substrates are not folded correctly due to improper disulfide bond formation. Specifically, three disulfide bonds found in mature NGF may not be formed. In a similar in vitro translation system, Scheele and Jacoby (1982) observed that serine protease zymogens were not correctly folded. Moreover, correct folding was achieved when the redox environment of the translation reaction was modified by the addition of glutathione. Since protein disulfide isomerase is sensitive to redox potential changes, it is possible that this enzyme was activated and caused the correct disulfide bond formation (Freedman, 1984). If the NGF precursors are not folded correctly, alteration of the redox environment during translation might allow for correct folding and hence, correct processing.

CHAPTER VI

PROTEOLYTIC PROCESSING OF RECOMBINANT NGF PRECURSORS

INTRODUCTION

β-Nerve Growth Factor (NGF) is a soluble 118 amino acid protein required for the development and maintenance of sympathetic and sensory nerves (Thoenen & Barde, 1980). Administration of NGF antiserum to animals during development of their sympathetic and sensory systems causes a nearly complete destruction of these neurons (Levi-Montalcini & Booker, 1960; Gorin & Johnson, 1979). Conversely, systemically-provided NGF ensures the survival of many neurons that would otherwise atrophy at lower in vivo levels (Angeletti et al., 1971). NGF, presumably provided to the nerve terminals by the innervated target tissue, is taken up by NGF receptors (Herrup & Shooter, 1973) and transported in a retrograde fashion to the cell body (Stockel et al., 1975). Although NGF is found in extremely low levels in the appropriate target organs (Korshing & Thoenen, 1983; Shelton & Reichardt, 1984), it is found in surprisingly high levels in the mouse submaxillary gland (SMG) (Cohen, 1960). NGF can be isolated from the SMG as a member of the 7S complex, along with α and γ subunits (Server & Shooter, 1977). While the γ subunit is a potent arginyl-esteropeptidase (Green et al., 1969), the α subunit is an enzymatically inactive protein (Isackson & Bradshaw, 1984) even though both are members of the kallikrein gene family (Isackson et al., 1984). Burger and Shooter (1977) observed γ subunit cleavage of an immunoprecipitated NGF precursor intermediate (22 kd) into a 13 kd species, presumably NGF. Thus, the protease to which NGF is complexed may process its precursor. To ascertain whether the γ subunit can process the two full-length NGF precursors predicted by cDNA cloning (Scott et al., 1983; Edwards et al., 1986) into mature NGF, we tested whether synthetic NGF precursors made in vivo could be cleaved by the γ subunit. We found

that precursors synthesized in vivo do liberate mature NGF upon γ subunit digestion. Trypsin also can effect the same processing events. We present data showing that the two NGF precursors generated from the two mRNAs appear to be identical by several criteria.

MATERIALS AND METHODS

NGF Iodination

Purified 2.5S NGF (generously provided by W. Mobley) (intact and β-endopeptidase clipped) was iodinated by the chloramine-T method according to Miller et al. (1978) and purified on a G-10 column equilibrated in 250mM $NaHPO_4$, pH 7.4 + 0.1% BSA. Aliquots of each 1 ml fraction were counted in a Beckman 5500 Gamma counter to identify the peak tubes.

γ and α Subunits Reactions

The γ subunit was generously provided by Dr. Eric Shooter. α and γ subunits were purified from the 7S complex in the saliva by following the procedures of Burton et al. (1978). γ subunit activity was determined fluorometrically after incubation with the synthetic substrate, D-val-leu-arg-amino-fluoro-coumarin. For reactions including the α subunit, the α subunit was incubated at 37°C for 20 min followed by the addition of γ subunit and then further incubated. The reaction products were then analyzed by SDS-PAGE.

Vaccinia Virus Infection and Labeling

Purified virus of known titer (A:2-3x10^{10} pfu/ml; B: 1x10^{10} pfu/ml; wild type (wt): 2-3x10^{10} pfu/ml) was added to PBS-M/BSA (phosphate buffered saline + 1.0mM $MgSO_4$, 0.01% BSA) at ~3x10^8 pfu/ml and then placed on the L-929 cells growing in 24-well dishes. The infection progressed at ambient temperature for 60 minutes and was followed by

incubation in Dulbecco's Modified Eagle's (DME) medium with 10% fetal bovine serum and penicillin/streptomycin for 1 hour at 37°C. The cells were depleted of intracellular methionine pools by incubating the infected cells in methionine- and cysteine-free DME for 30 min at 37°C. Radiolabeled proteins were made following the addition of ^{35}S-methionine and ^{35}S-cysteine (ICN, translabel). Radiolabelings proceeded for 30-35 minutes at 37°C and were terminated by washing the cells with ice-cold PBS. The cells were recovered by scraping and then concentrated by centrifugation in the microfuge for 2-3 minutes at 4°C. The cell pellet was resuspended in 150 µl of 100mM Tris pH 7.6 + 50mM NaCl. The cells were disrupted by sonication (Braun-Sonic 2000 micro probe; 3-times, 5 sec) or by two cycles of freezing/thawing and cleared by centrifugation for 15 min in the microfuge at 4°C. The supernatant was used for γ subunit with/without α subunit as described above.

Immunoprecipitations and Gel Electrophoresis

At the end of the incubations, the reactions were placed on ice. Four volumes of ice-cold modified RIPA buffer (1.25% sodium deoxycholate, 1.25% Triton X-100, 150mM NaCl, 50mM Tris, pH 7.4, 0.25% SDS and 20µM PMSF) were added and the samples were cleared by centrifugation at 4°C for 15 minutes in a microfuge. One microliter of rabbit NGF antiserum (generous gifts of Drs. D. Shelton and L.F. Reichardt and Drs. T. Darling and E. Shooter) (1:100-1:1000 dilution) or non-immune serum (NIS) was added to each sample and incubated overnight at 4°C. 200 µl of washed protein A-Sepharose (Sigma) in RIPA buffer (1% sodium deoxycholate, 1% Triton X-100, 150mM NaCl, 50mM Tris, pH 7.4 and 0.2% SDS) was added; the samples were rocked at 4°C for 4 hours. The affinity beads were pelleted in a microfuge for 0.5 min at 4°C and washed 4 times with RIPA buffer. 50 µl of Laemmli buffer was added to the pellet,

heated to 100°C for 5 minutes, and 25 µl was loaded on SDS-polyacrylamide gels. Following electrophoresis, the gels were fixed in 50% methanol/10% acetic acid and then enhanced with 1M sodium salicylate for 20 minutes. The gels were dried and exposed to autoradiography with an intensifying screen at -70°C. The ability of NGF antiserum to quantitatively immunoprecipitate mature NGF was demonstrated using iodinated NGF: to 25 µl of iodinated NGF was added 100 µl modified-RIPA buffer + 1 µl NGF antiserum for overnight incubation. The immunoprecipitation with protein A-Sepharose was as described above. Half of this was loaded next to a lane that contained half of the following: 25 µl of iodinated NGF + 25 µl Laemmli.

Conditions for NGF Precursor Cleavage by Trypsin

The short and long NGF precursors (A and B) were cleaved with trypsin in the same buffer used for γ subunit digestion. To ascertain if the two enzymes were catalytically comparable, the activity of each was independently determined using the artificial Arg substrate, D-val-leu-arg-amino-fluoro-coumarin; trypsin appeared 17X more potent on a weight-to-weight comparison with the same amount of this substrate. Thus, trypsin was first diluted 1:17 before use in the NGF cleavage experiment. At the end of the 30 minute incubation, the reactions were transferred to ice and then immunoprecipitated as described above.

Time-Course of Gamma Subunit Cleavage of Precursors

Labeled extracts from vaccinia virus A- and B-infected cells were digested with γ subunit at 37°C. A zero time point was taken just before the addition of γ subunit. Subsequently, 5, 10, 15, 20, 30, 40, 50 and 60 minute samples were taken and immunoprecipitated with NGF antiserum as above.

Unlabeled NGF Competition with Radiolabeled NGF in Immunoprecipitation Assays

Extracts prepared as outlined above were set up with and without added γ subunit. After the reactions were incubated for 30 minutes at 37°C, 1 µl of NGF antiserum with or without 200 ng of mature NGF. The samples were then prepared for immunoprecipitation and gel electrophoresis as above.

γ Subunit Processing of Immunoprecipitated NGF Precursor

To 12 µl of A and B extracts were added 13 µl water, 1 µl NGF antiserum and 100 µl of RIPA buffer. The samples were prepared for immunoprecipitation as above except two washes in water followed the RIPA buffer washes. The beads were finally resuspended in 25 µl of 100mM Tris, 7.6 + 50mM NaCl. Either 2 µl H_2O or 2 µl of γ subunit (1.8 µg) were added to the appropriate tubes and incubated at 37°C for 30 min. 25 µl of 2X Laemmli buffer was added to the reactions, heated and then loaded on a 12.5% polyacrylamide gel.

Detection and Characterization of Secreted NGF Protein

100 µl of medium was mixed with 15 µl 1M Tris 7.6 + 0.5M NaCl and brought to 150 µl with water. 0 or 1.8 µg γ were added and then incubated for 35 minutes at 37°C. The reactions were terminated with 1.1 ml of modified RIPA buffer (composition described above) + 1 µl of NGF antiserum. Immunoprecipitation were carried out as outlined above.

Protein Sequencing

60mm plates of L-929 cells were infected with B vaccinia virus as above. The cells were labeled with 500 µCi ^{35}S-met and 500 µCi ^{35}S-cysteine (Amersham) for 35 minutes at 37°C. Extracts were made as outlined above. For the analysis of secreted NGF, cells were labeled for

7 hours and NGF was recovered by immunoprecipitation of the media (50%). The samples were either directly immunoprecipitated or were processed and then immunoprecipitated as above. Protein A-Sepharose beads were washed 5 times in RIPA buffer and twice in double-distilled water. Then, 200-400 µl of 10mM $NaHCO_3$ containing 0.1% SDS (Ultrapure, BDH) was added to the beads, followed by heating at 100°C for 5 minutes. The beads were pelleted and the supernatant transferred to a new tube and then submitted to amino acid sequencing on a microsequencer. Fractions were quantitatively transferred to counting vials containing 4 ml scintillation fluid (Ecolite) and then counted.

Bioactivity

100 µl reactions with extracts from wt, A and B virus-infected L-cells were processed with 900 ng of γ subunit for 35 minutes at 37°C. Varying amounts of these reactions were added directly to PC-12 cells in 1 ml of DME + 15% serum that had been seeded 24 hr earlier in 24-well Costar dishes. These were compared to PC-12 cells grown with 2000, 1000, 200, 20 and 2 ng of purified NGF. We also examined the media of infected (3 hours post-infection) cells to determine the bioactivity of the secreted NGF. The media were analyzed 3 days later.

RESULTS

γ Subunit Correctly Processes In Vivo Generated NGF Precursors

Vaccinia virus vectors containing portions of mouse NGF cDNA corresponding to long (A; SmaI-PstI) or short (B; Apa-Pst) precursors were constructed. Vaccinia virus recombinants were prepared by recombination between a plasmid containing NGF cDNA inserts which bisects the TK gene and wild type vaccinia virus. Recombinants express inserts under the control of the viral 7.5 promoter. Northern blot analysis of RNA

from infected cells showed the predicted mobilities of the two RNA species (unpublished observations, R. Edwards). Upon infection of various cell lines, precursor NGFs are observed on immunoprecipitation of cell extracts. Further, in pulse-chase experiments, these precursors are converted into a 13 kd secreted species, while the level of intracellular precursors diminishes (unpublished observations, R. Edwards).

In vivo synthesized precursors can be processed in vitro into a 13 kd species. L-929 cells were infected with vaccinia virus A, B and wild type (WT) and pulse-labeled such that virtually all the immunoprecipitated products were precursors. Cleared lysates of these infected cells were treated with increasing amounts of γ subunit. This resulted in the generation of increasing amounts of a 13 kd species, presumably mature NGF (Fig. 1A). Under these conditions, both NGF precursors gave rise to 13 kd species with roughly equal facility. No proteins of similar mobility were observed on immunoprecipitation of WT infected cells (Fig. 1A). No precursors or mature NGF, with and without γ subunit digestion (Fig. 1B), were observed on immunoprecipitation of extracts with non-immune serum. The same immunoprecipitation pattern was obtained with an independent anti-NGF serum (Darling & Shooter) (Fig. 1C). A kinetic analysis showed that processing intermediates (~30 kd) of both precursors A and B appear within the first few minutes after addition of γ (Fig. 2). After one hour, most of the precursors have been converted into 13 kd species. The $t_{\frac{1}{2}}$ for cleavage is about 25 minutes.

The specificity of the processing reaction was examined by using trypsin instead of γ subunit since they both have a similar specificity for dibasic amino acid residues. The amount of enzymes used in the reactions was adjusted for their relative activities on an artificial

FIGURE 1A. Immunoprecipitation of γ subunit-processed vaccinia virus-infected cell lysates. L-929 cells were infected with either A, B or wild type (WT) vaccinia virus. After absorption and expression, the cells were pulse-labeled for 30 minutes followed by cell disruption. The lysates were aliquoted and digested with varying amounts of γ subunit for 25 minutes at 37°C. The reactions were terminated and immunoprecipitated with anti-NGF serum (Shelton & Reichardt). The designations 1-5 in the figure correspond to 0, 9, 90, 900 and 1800 ng of γ subunit added to each reaction, respectively. The designations A, B and WT refer to the virus-infected cell extract used. Note that the precursors (migrating between 46 and 30 kd standards; slower migrating species of doublet) decrease with increasing amounts of γ subunit while the ~13 kd species steadily increases. Precursor processing intermediates of ~30 kd are observed in processing of A and B extracts. No precursor or 13 kd-equivalent is observed in the WT-infected extracts. At the higher concentrations of γ, the overall background decreases, reflecting γ digestion of other proteins as well as the precursors.

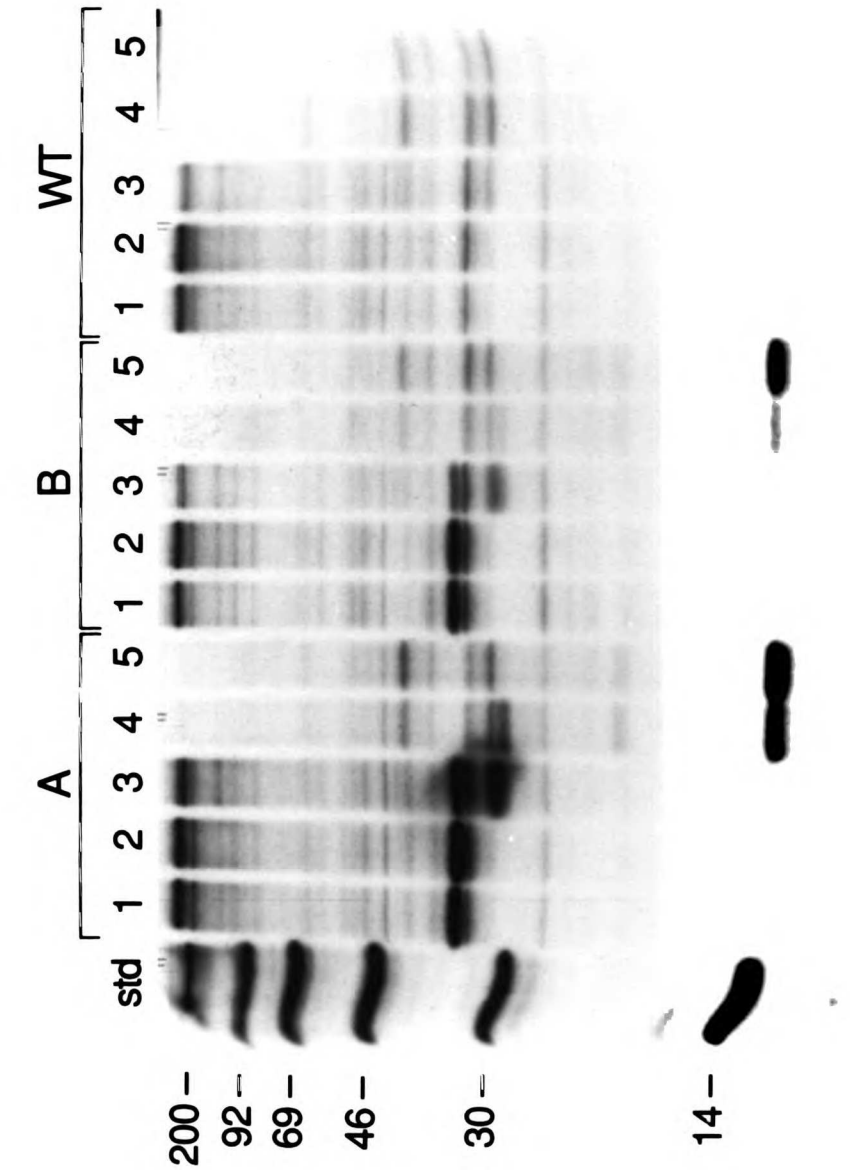

A
B
WT
std
1
2
3
4
5
1
2
3
4
5
1
2
3
4
5
200
92
69
46
30
14

FIGURE 1B. Non-immune serum precipitation of vaccinia virus-infected cells with and without γ subunit digestion. Cell extracts were made from labeled virus-infected cells (A, B or WT) followed by incubations with and without γ subunit (900 ng) and then immunoprecipitated with rabbit non-immune serum. The extracts used are designated above the figure; the addition of γ subunit to incubation is denoted by +. No products corresponding to NGF precursors or mature are observed. Moreover, all the samples appear nearly identical with or without digestion.

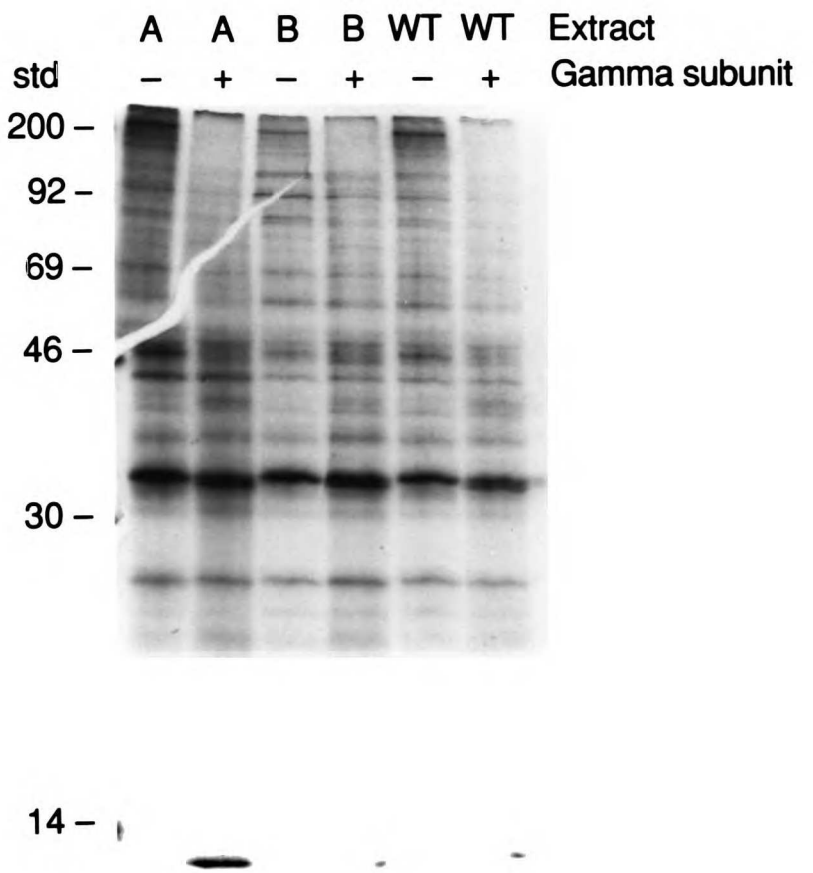

A A B B WT WT Extract
std – + – + – + Gamma subunit
200 –
92 –
69 –
46 –
30 –
14 –

FIGURE 1C. Immunoprecipitation with two different antisera of precursors and NGF from infected cell lysates. Radiolabeled (^{35}S-methionine) lysates of A and B vaccinia virus-infected cells were incubated with 0 or 900 ng of γ subunit for 30 minutes at 37°C. This was followed by immunoprecipitation with anti-NGF serum from Drs. Darling and Shooter (panels A and B, left of internal standard lane). This was compared to a previous immunoprecipitation of A infected lysates with antiserum from Drs. Shelton and Reichardt (panel A, right of internal standard lane). Further, the Darling and Shooter antiserum was used to immunoprecipitate SP-6 generated precursor and γ subunit-digested species (panel labeled SP-6). The presence of γ subunit in the reaction is denoted by a- or + above the figure. Note the precursors and mature NGF are recognized with comparable efficiencies with the two different antisera. The Darling and Shooter antiserum, like the Shelton-Reichardt serum, cannot immunoprecipitate in vitro derived precursors or products derived from them by γ digestion; this supports the conclusion that these in vitro-derived precursors do not liberate mature NGF on γ subunit treatment.

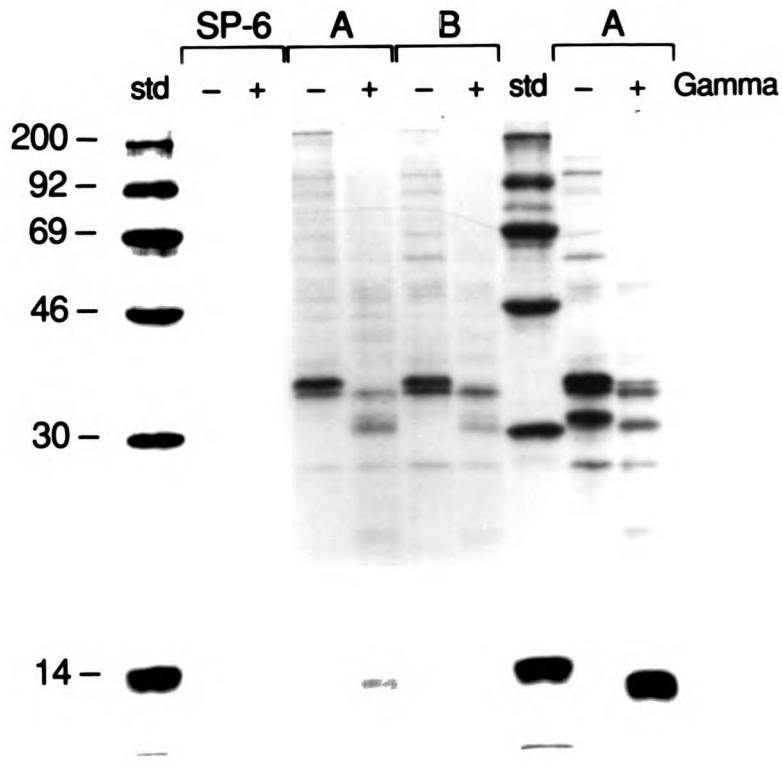
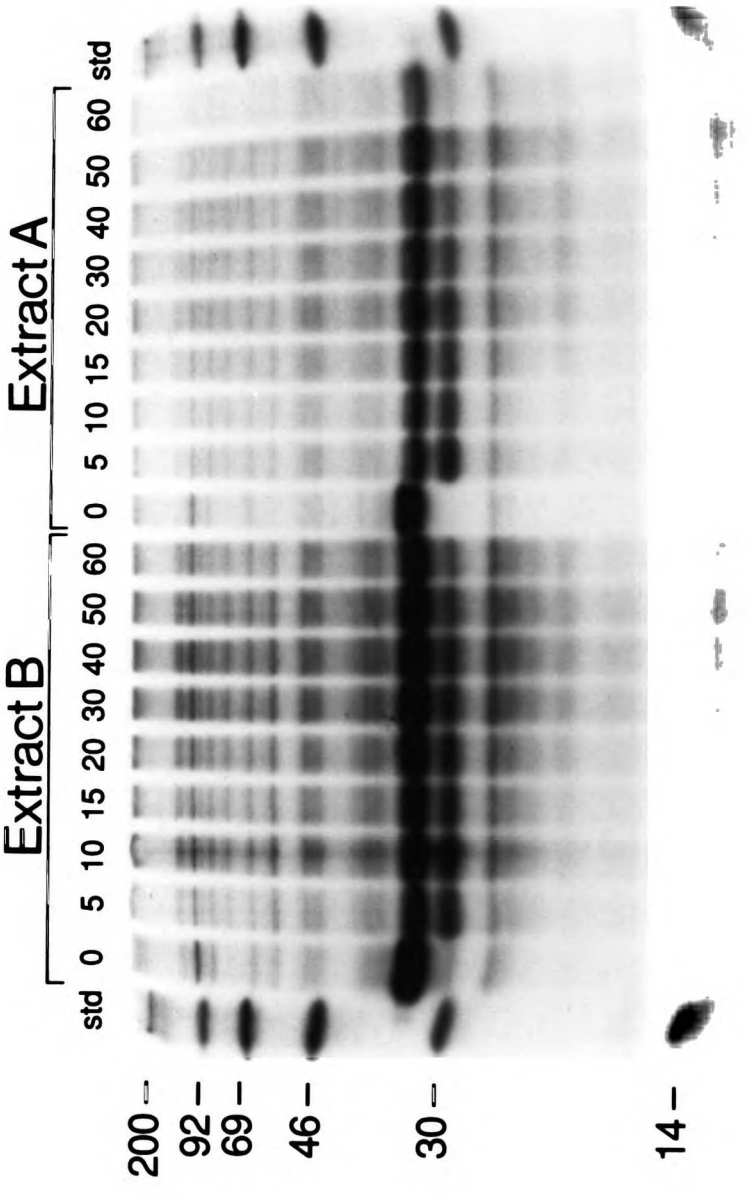

SP-6
A
B
A
std
–
+
–
+
–
+
std
–
+
Gamma
200 –
92 –
69 –
46 –
30 –
14 –

FIGURE 2. Kinetics of γ subunit digestion. Extracts (50 µl) of A and B vaccinia virus-infected cells were digested with γ subunit (3.6 µg). Samples were taken before (0 time), and at 5 or 10 minute intervals after the addition of γ subunit. Samples were removed from the reaction and transferred directly into immunoprecipitation buffer, cleared and then precipitated with anti-NGF serum. The molecular weights of the labeled protein standards appear at the left of the figure. Note the progressively decreasing precursor bands and the concomitant rise in 13 kd NGF. While the precursor is somewhat obscured by other proteins migrating in the same region as the precursor, the intermediates (~30 kd) appear early in the digestion and diminish through the incubation period. The $t_{\frac{1}{2}}$ for digestion based on the signal of the precursor is roughly 25 minutes. γ subunit digestions of extracts from A and B lysates gave qualitatively the same profile.


Extract B
Extract A
std 0 5 10 15 20 30 40 50 60 0 5 10 15 20 30 40 50 60 std
200
92
69
46
30
14

arginine substrate; 17-fold less trypsin than γ subunit (data not shown) was required to cleave the synthetic substrate. When trypsin was added to extracts of vaccinia A- and B-infected cells, a 13 kd species was was also generated (Fig. 3). The addition of only 59 ng of trypsin effected complete precursor cleavage. 900 ng of γ subunit was required to obtain comparable cleavage; thus trypsin is 15-fold more potent than the γ subunit on these precursor substrates. No progressive processing observed in γ subunit digestion is observed with trypsin treatment. This indicates that trypsin is significantly more proficient than γ subunit in converting precursor to the NGF and may have a different cleavage pattern.

The 13 kd species can also be derived from immunoprecipitated precursors. A and B precursors were immunoprecipitated from the corresponding infected cell lysates and then digested with the γ subunit. A 13 kd species was observed which co-migrated with the 13 kd species derived from immunoprecipitates of γ subunit-processed extracts (Fig. 4).

The sequence of the immunoprecipitated 13 kd species certified that this is indeed mature NGF. Extract of pulse-labeled VV B-infected cells was digested with γ subunit and immunoprecipitated. The immunoprecipitate was eluted from protein A and cleaved by sequential Edman degradation. Scintillation counting of fractions representing individual amino acid residues showed a peak of label incorporation into residue #9, the location of methionine in mature NGF (Fig. 5A). The substantial shoulder in fractions 10 and 11 probably reflects incomplete degradation and/or slight heterogeneity at the amino termini of the molecules. No significant incorporation of cysteine was observed, so uptake and incorporation were apparently very inefficient. The sequence

FIGURE 3. Trypsin digestion of NGF precursors. A and B vaccinia virus-infected cell lysates (designated A and B) were digested with 0, 0.58, 5.8 and 58 ng of trypsin (lanes 1-4) for 25 minutes at 37°C and then immunoprecipitated with anti-NGF serum. The molecular weights of protein standards appear to the left of the figure. Only digestion with 58 ng of trypsin effects complete cleavage of the precursors. Background bands are significantly reduced at this same amount of trypsin. The intensity of the two panels differs because they derive from digestions of two different extracts. These data demonstrate that trypsin at comparatively low levels liberates mature NGF from the two precursors.

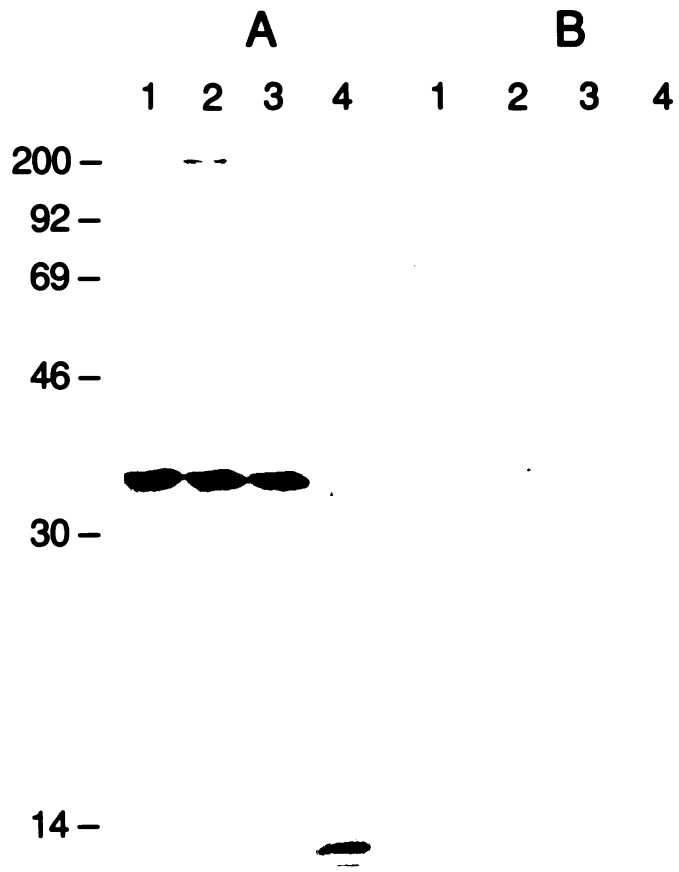

A
B
1 2 3 4 1 2 3 4
200 –
92 –
69 –
46 –
30 –
14 –

FIGURE 4. Processing of immunoprecipitated NGF precursors produced in vivo. ^{35}S-labeled extracts from A and B vaccinia virus-infected cells were immunoprecipitated with anti-NGF serum. The immunoprecipitates were washed with RIPA buffer (see text) and then water. Subsequently, Tris-saline buffer was added and followed by incubation with 0 or 1.8 µg of γ subunit. After 30 minutes, 2X Laemmli buffer was added, the samples were heated, denatured, loaded onto a 15% polyacrylamide gel and electrophoresed. The particular extract and γ subunit additions are denoted above the figure. The molecular weights of labeled protein standards are presented to the left of the figure. In the presence of γ subunit, 13,000 NGF is generated from both precursors. Aside from residual precursor, no smaller (than NGF) labeled polypeptides are observed, suggesting that the amino-terminal peptide, which contains two methionine residues (met-ser-met) is not present.

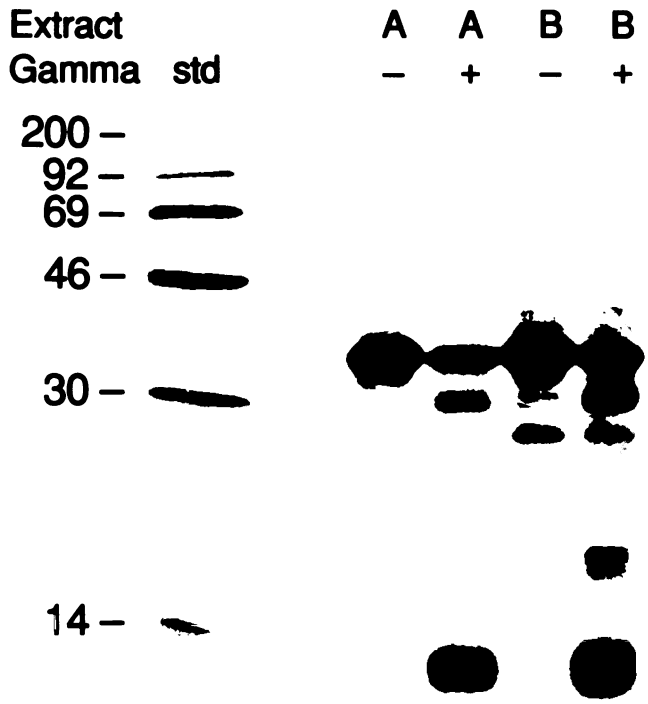
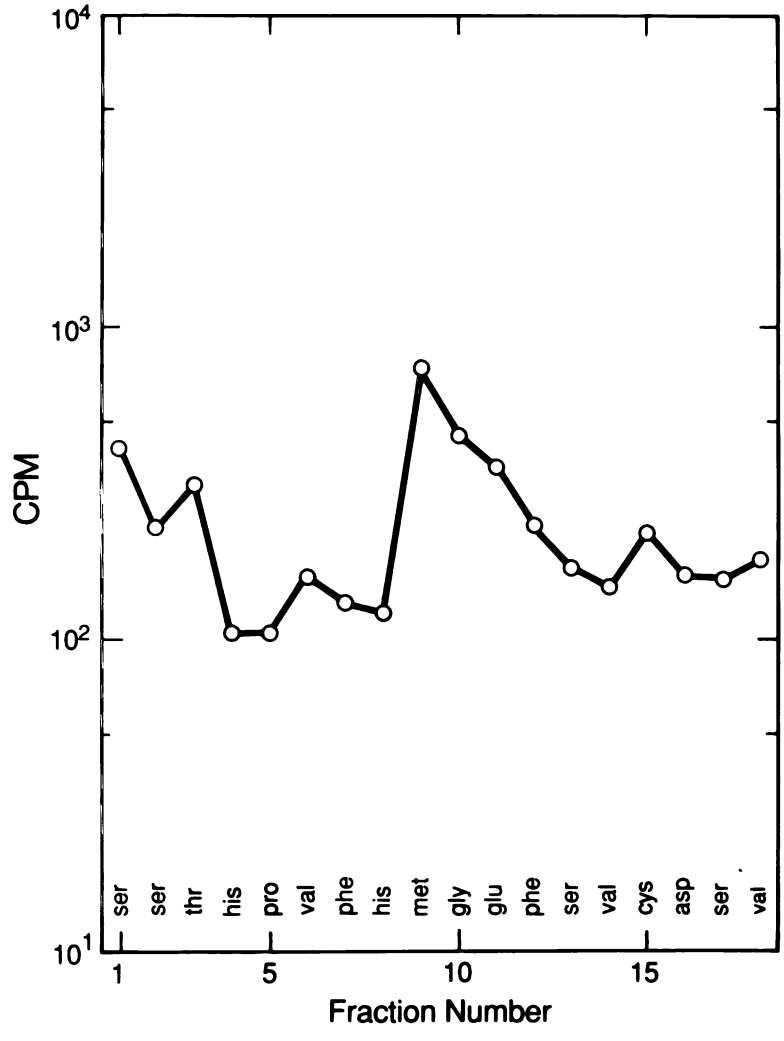
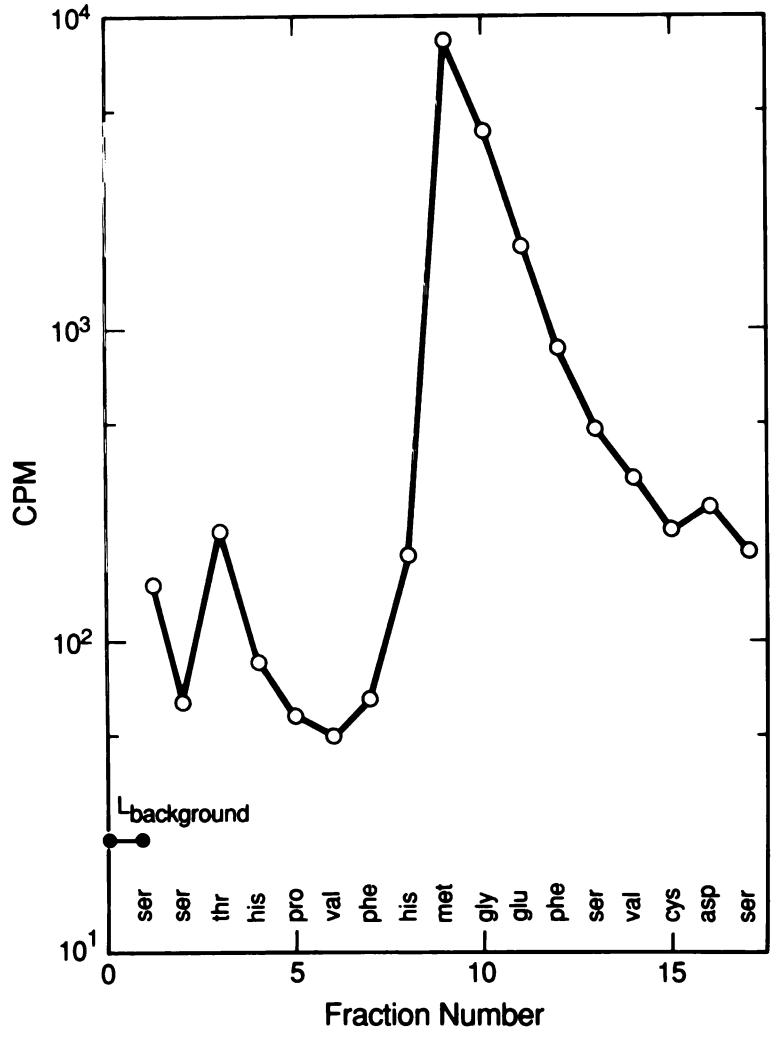

Extract
A A B B
Gamma std
– + – +
200 –
92 –
69 –
46 –
30 –
14 –

FIGURE 5. Sequential Edman degradation of putative NGF produced by γ subunit treatment of NGF precursor and by secretion from the cells. L-929 cells were infected with B vaccinia virus and labeled with ^{35}S-methionine outlined in Materials and Methods. (A) The cell lysate (75 µl) was treated with γ subunit (2.7 µg), immunoprecipitated and subjected to sequential Edman degradation. The open circles in the figure represent the cpm of the fractions from the sequenator. The amino acids of NGF are presented at the bottom of the figure along with the fraction number. Note the peak at methionine residue #9 corresponding to the met position in NGF. (B) Sequential Edman degradation of labeled and secreted mature NGF. Mature NGF was recovered by immunoprecipitation of media from A vaccinia virus-infected cells and was then submitted to sequential Edman degradation. Similar to Figure A is the peak at methionine #9, the signal is much greater here, owing to the much longer labeling period (7 hours vs 30 minutes). The trail in the peak fraction probably represents incomplete degradation and/or slight heterogeneity at the amino-terminus.


10^4
10^3
10^2
10^1
CPM
ser ser thr his pro val phe his met gly glu phe ser val cys asp ser val
1
5
10
15
Fraction Number


10^4
10^3
10^2
10^1
CPM
background
ser
ser
thr
his
pro
val
phe
his
met
gly
glu
phe
ser
val
cys
asp
ser
0
5
10
15
Fraction Number

of immunoprecipitated material from the medium of vaccinia virus A infected cells gave qualitatively the same profile (Fig. 5B). Thus, mature NGF was both secreted from infected cells and processed from a mature NGF was both secreted from infected cells and processed from a precursor synthesized in an infected cell lysate. The correct sequence is consistent with the observation that the addition of mature NGF to lysate immunoprecipitation reactions resulted in substantially reduced recovery of the corresponding precursor or mature NGF, while not affecting the background bands (Fig. 6).

Sequencing of the two precursors demonstrates that they are identical at their amino-termini. The NGF precursors were immunoprecipitated from pulse-labeled infected cells and prepared from sequential Edman degradation. When the infected cells were labeled with ^{35}S-methionine, no radioactive peaks appeared in the 22 cycles examined (data not shown); when ^{3}H-valine was used, radioactive peaks at fractions 8 and 14 were observed for both immunoprecipitated precursors (Fig. 7). These data are consistent with signal sequence cleavage after the same amino acid, namely, alanine (Ala) 84 and Ala 18 in the 34 kd and 27 kd precursors, respectively. Signal sequence cleavage removes the two amino-terminal methionines (met-ser-met--) and thus accounts for the absence of methionine radioactivity peaks in the first 22 cycles of precursor sequencing. The cleavage site between Ala and the adjacent glutamic acid residue agrees well with the cleavage site predicted by the "-3,-1" rule (von Hjeine, 1986). 18 amino acid residues are removed from the amino-terminus of the 27 kd precursor; the full 84 residues, or ~9 kd, are removed from the 34 kd species (if one assumes that the nascent precursor is actually synthesized). All sequenced samples were

FIGURE 6. Unlabeled mature NGF competes with the NGF antibody for immunoprecipitation of radiolabeled precursors and mature NGF. Extract of B vaccinia virus-infected L-929 cells were incubated with 0 or 900 ng of γ subunit as indicated (- or +). The samples were then immunoprecipitated with anti-NGF serum with or without 200 ng of mature NGF. The molecular weights of labeled protein standards appear to the left of the figure. Note the recovery of both precursor and mature NGF is diminished substantially when mature NGF is added, while the background remains unchanged. This demonstrates that the antibody is recognizes determinants on both the precursor and mature NGF. It also distinguishes NGF precursor and mature species from background bands, which are not competable.

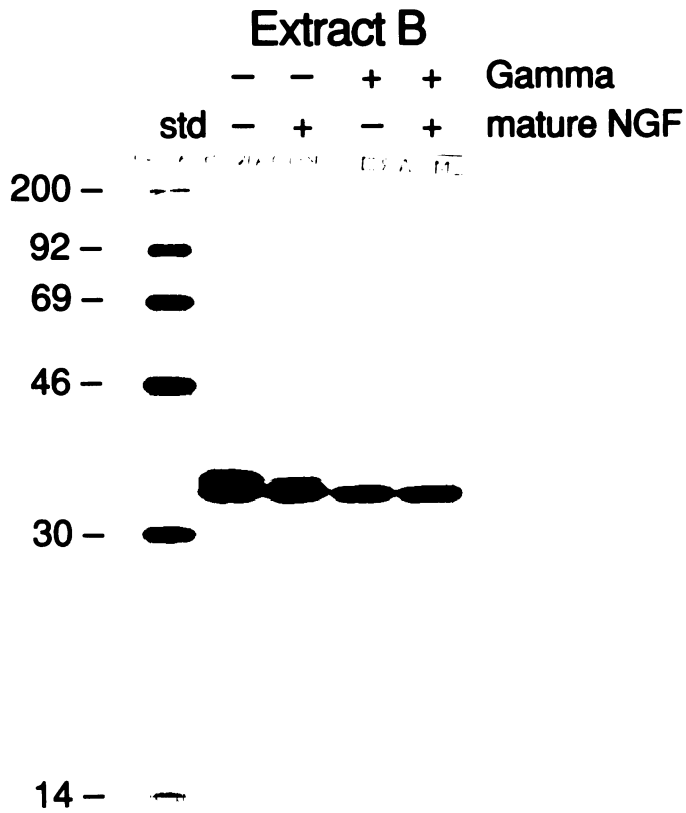

Extract B
– – + + Gamma
std – + – + mature NGF
200 –
92 –
69 –
46 –
30 –
14 –

FIGURE 7. Sequential Edman degradation of radiolabeled NGF precursors. L-929 cells were infected with A and B vaccinia viruses and labeled with ^{3}H-valine as described in Materials and Methods. The open circles in the figure represent the cpm of the fractions from the sequenator. The amino acids of the amino-terminal regions of NGF precursors (minus signal sequence) are displayed at the bottom of the graph.

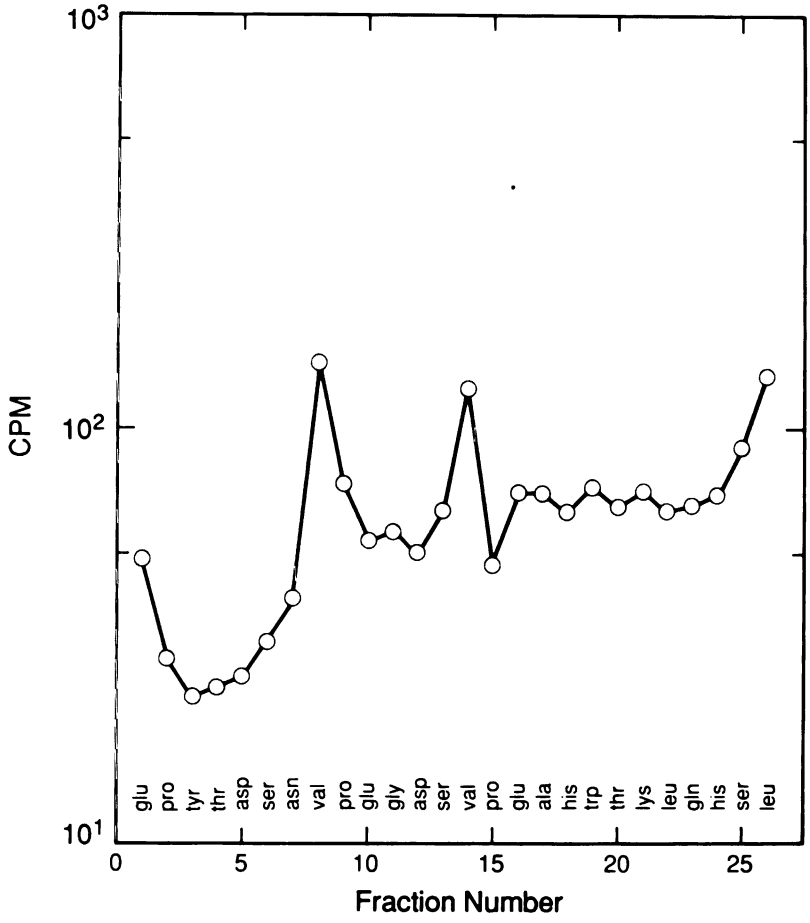
A. Graph of labeled fractions from sequencing of A vaccinia virus-infected cells. Note the two peaks that correspond to the location of valine residues after the signal sequence cleavage site.

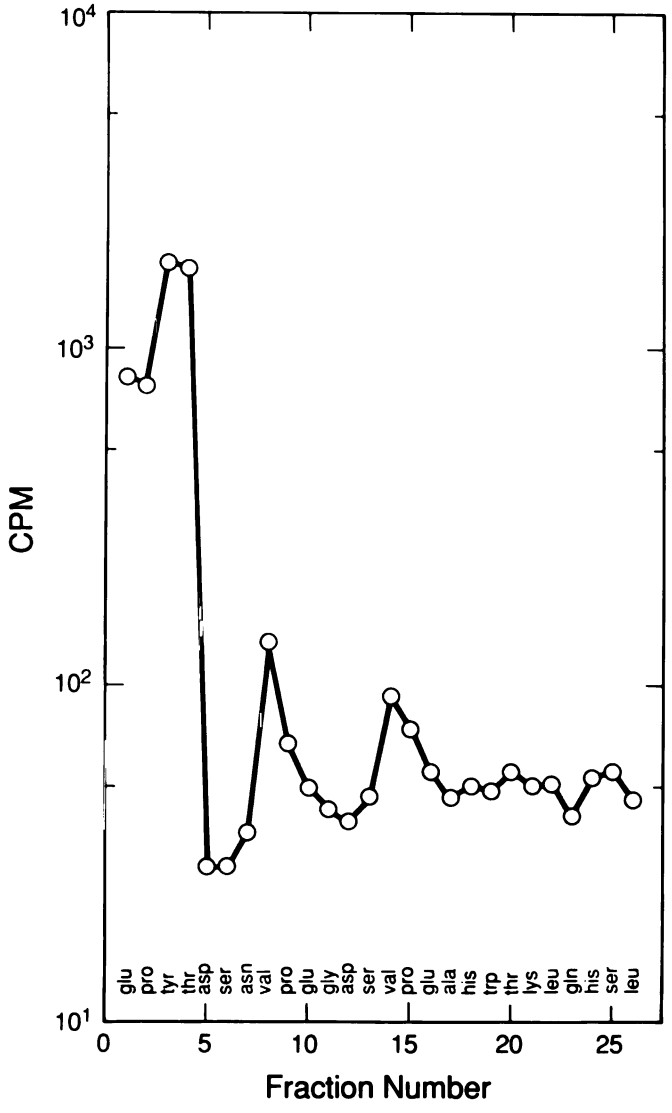
B. Graph of labeled fractions from sequencing of B vaccinia virus-infected cells. Note the same peaks that correspond to the location of valine residues. The high level of label detected in the first few samples most likely was a consequence of residual labeled products on the column (B was sequenced before A).

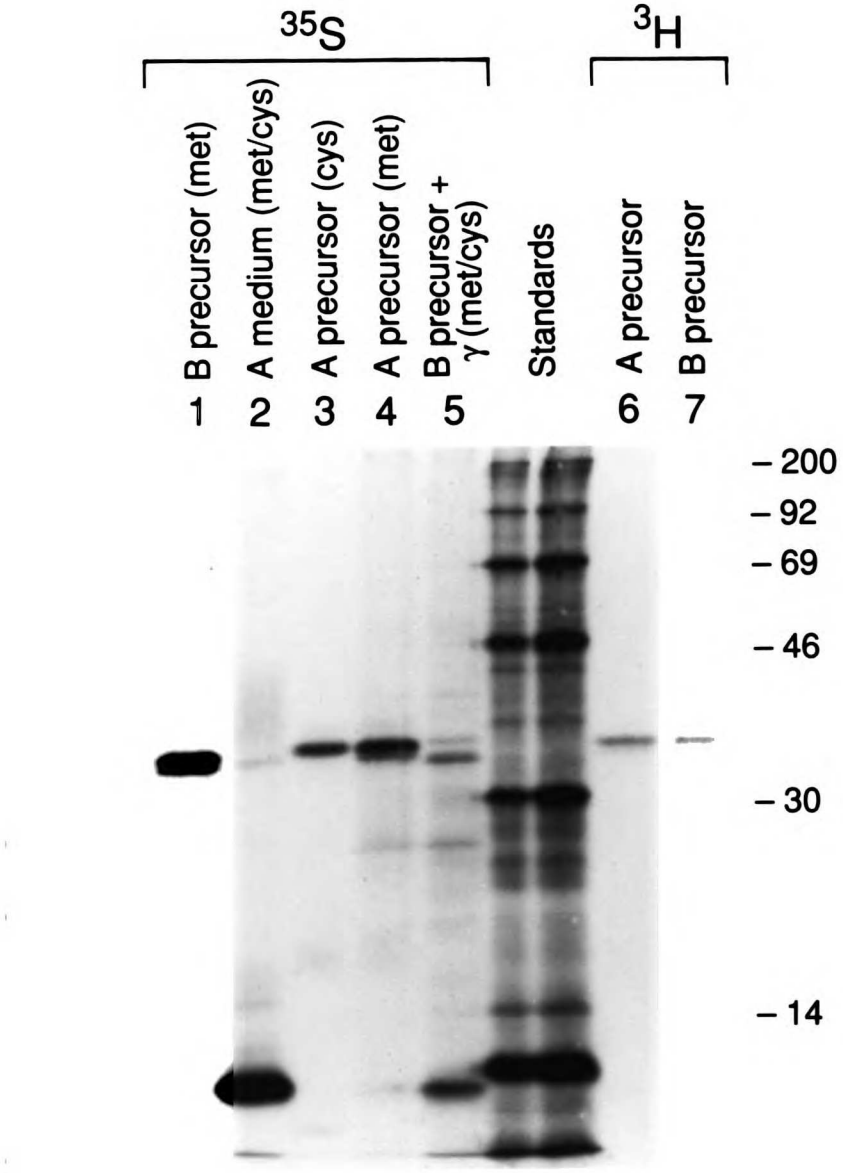
The identical pattern of peaks for the two sequenced proteins demonstrates that the precursors are identical, at least at the amino-terminal region. They are also most likely identical at their carboxy-termini because they migrate identically on SDS gels and are both N-glycosylated. Attempts to sequence ^{35}S-methionine-labeled precursors did not yield peaks of radioactivity presumably because the two methionines at the amino-terminus of the 27 kd precursor (met-ser-met) are removed by cleavage of the signal sequence.

C. Gel electrophoresis of samples for amino acid sequencing. The figure shows an autoradiograph of all the immunoprecipitated samples that were subjected to radio-sequencing. The particular isotope used in cell labeling is shown above as is the identification of the individual samples. The standards are ^{14}C-labeled (Amersham) with the corresponding

molecular weights listed at the right of the figure. Samples in lanes 1, 3, 4, 6 and 7 represent those samples used for precursor analyses; samples 2 and 5 were used for mature NGF analyses. In all samples, the principle labeled species corresponds to that which was sequenced. Only the processed precursor (lane 5) showed some background; this translated into a higher background in the radio-sequencing.


10^3
10^2
10^1
CPM
glu pro tyr thr asp ser asn val pro glu gly asp ser val pro glu ala his trp thr lys leu gln his ser leu
0
5
10
15
20
25
Fraction Number


10^4
10^3
10^2
10^1
CPM
glu pro tyr thr asp ser asn val pro glu gly asp ser val pro glu ala his trp thr lys leu gln his ser leu
0
5
10
15
20
25
Fraction Number


35S
3H
1 B precursor (met)
2 A medium (met/cys)
3 A precursor (cys)
4 A precursor (met)
5 B precursor + γ (met/cys)
Standards
6 A precursor
7 B precursor
– 200
– 92
– 69
– 46
– 30
– 14

analyzed by PAGE for purity (Fig. 7C) and were found to be substantially free of contaminating bands.

The NGF produced in this fashion causes neurite outgrowth in cultured PC-12 cells. Media from recombinant VV-infected cells, when added to NGF-responsive PC-12 cells (Greene, 1977a), elicited neurite outgrowth that was not observed in medium from WT-infected cells (Fig. 8A and B). This outgrowth was short-lived for vaccinia virus contaminating the transferred media probably caused the observed cytopathic effects. Since some recombinant VV was apparently transferred, it cannot be stated unequivocally that the effect was due solely to addition of NGF from the infected media. However, when unprocessed or γ subunit-processed precursors were added to PC-12 cell media, neurite outgrowth was also observed (Fig. 8C and D). Since no outgrowth was observed on the addition of WT extracts, the effect is attributed to the products derived from the recombinant viruses. When extract of infected cells was diluted 1:10 and added to PC-12 cells, much less sprouting was observed compared to undiluted extract (Fig. 8E and F). Continuous labeling studies (R. Edwards et al., in preparation) suggest that most of the intracellular NGF is in the precursor form, and thus a dilution of extract should significantly reduce the amount of mature NGF. Thus, the precursor itself may have neurite outgrowth-promoting activity. Somewhat more sprouting was observed in unprocessed extract than in the γ subunit-treated sample. This phenomenon can be accounted for if complexes of γ subunit and mature NGF form since the 7S complex has less biological activity than NGF (Server & Shooter, 1977). Purified mature NGF was added as positive controls and showed the characteristic sprouting (Fig. 8G). From a comparison of neurite outgrowth induced by

FIGURE 8. Bioactivity of recombinant NGF and the NGF precursor on PC-12 cells. PC-12 cells were grown in DME-H21 + 15% serum + antibiotics. The cells seeded into a 24-well plate; additions were made 24 hours later. Photographs of the cells were taken 3 days after additions.

A. and B. The media of infected cells were added to PC-12 cultures. L-929 cells, previously seeded in 60mm dishes, were infected with WT (wild type) or A (long precursor) vaccina viruses for one hour, followed by aspiration of inocula and two washings with media. The monolayers were then incubated with 2 ml fresh media for 7 hours followed by collection. 5 µl of each of these two media were then added to PC-12 cells growing in 1 ml of media (1:200 dilution of media from infected cells). In 8A, no sprouting is observed when these cells were exposed to medium from WT-infected cells. By contrast, substantial sprouting is evident in those cells exposed to medium from a recombinant (A) virus-infected cell. Identical results were observed with the other vaccinia virus recombinant, B (data not shown). This demonstrates that the recombinant product(s) is responsible for the sprouting.

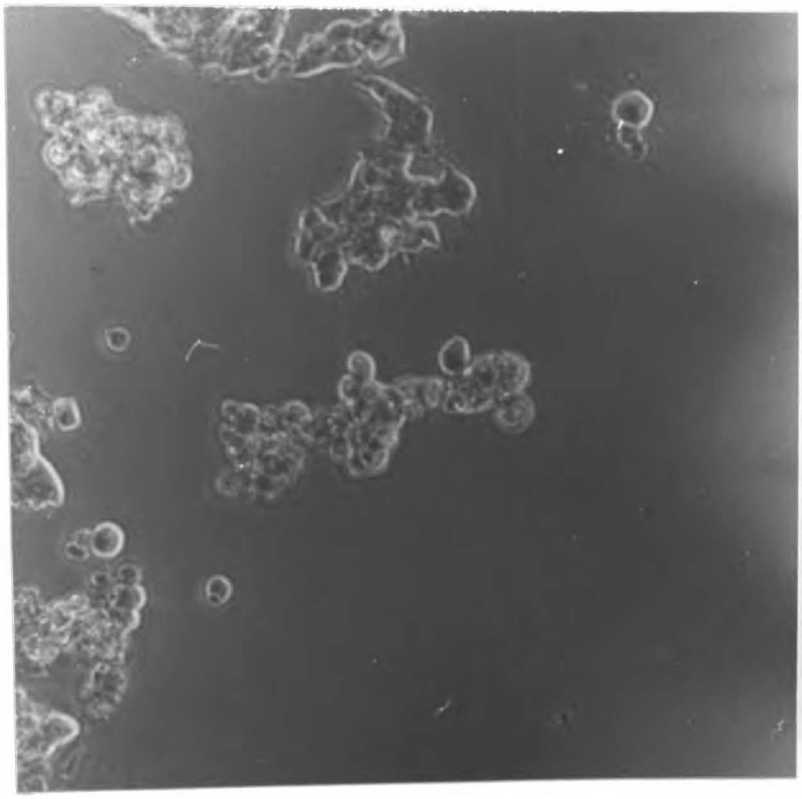
B. and C. Extracts of vaccinia virus (A)-infected L-929 cells (60mm plates) were made as outlined in Materials and Methods. One-third (50 µl) of each extract was incubated in buffer for 35 minutes at 37°C with or without added γ subunit. Aliquots were then added to the cultured PC-12 cells. In C and D, 5 µl of extract with (C) and without (D) incubation with γ subunit were added to the 1 ml of PC-12 media. Both cultures show extensive neurite outgrowth owing to the presence of recombinant NGF.

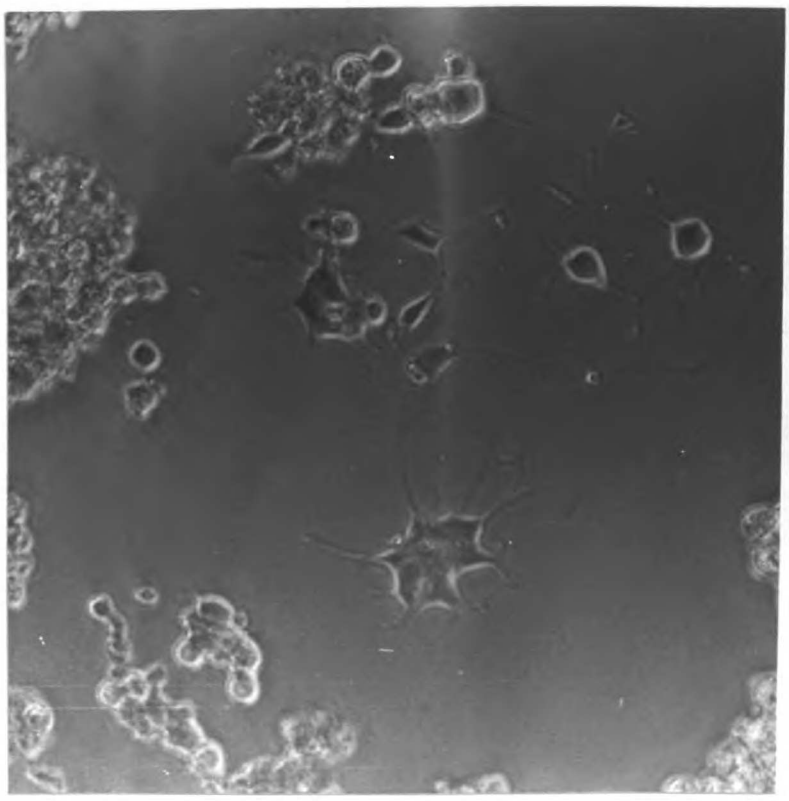
E. and F. These photographs represent cells treated as in C and D except that the extracts were first diluted 1:10 and then 1 µl was added

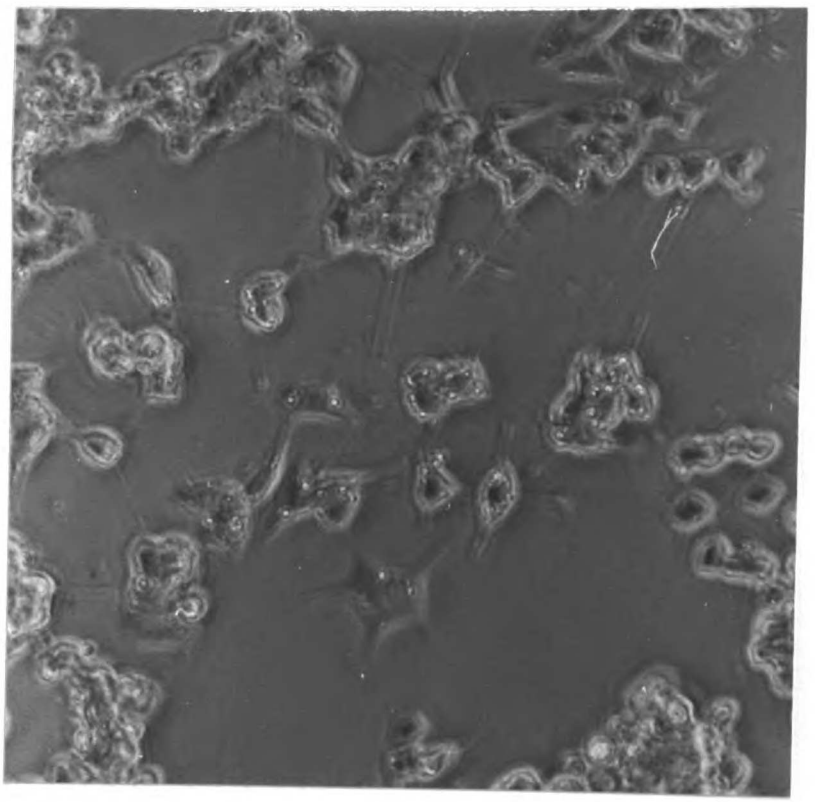
to 1 ml of media. As expected, much less sprouting is observed in both samples. Since precursor is the major NGF species in these extracts, it appears that the precursor may have NGF bioactivity since intracellular mature NGF is probably too dilute to be responsible for most of the outgrowth. There appears to be somewhat more outgrowth in the cultures treated with unprocessed extract (E) suggesting that γ subunit may complex with the small amount of NGF generated by processing to effectively remove it from the medium.

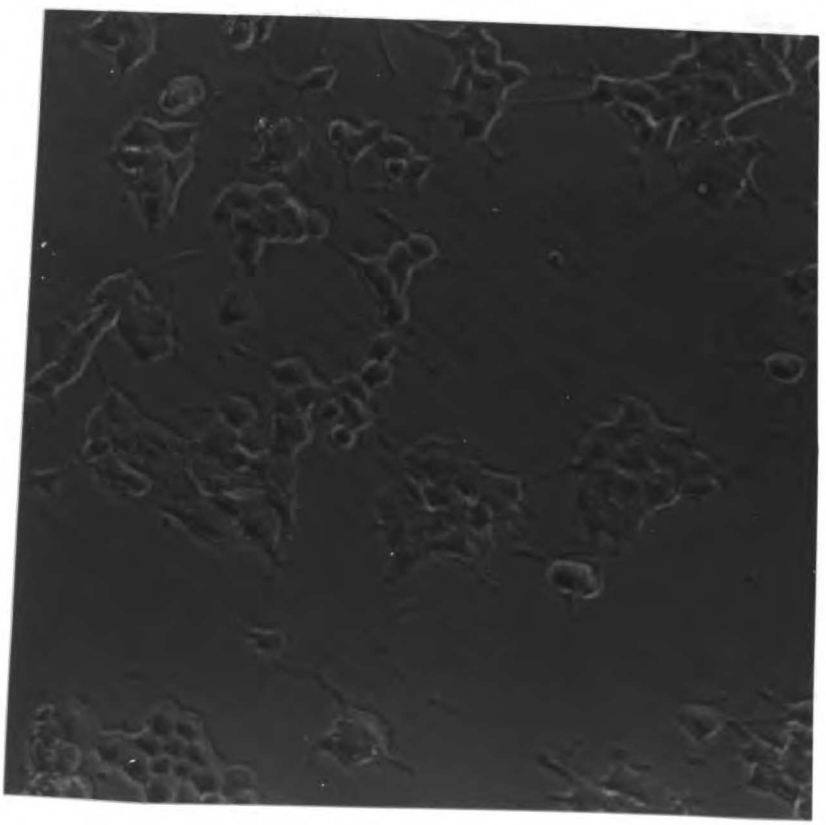
G. 200 ng of purified mature NGF (gift of W. Mobley, UCSF) was added to 1 ml of medium. Note the characteristic sprouting is apparent on virtually all the cells. Dilution of mature NGF caused correspondingly less sprouting (data not shown).

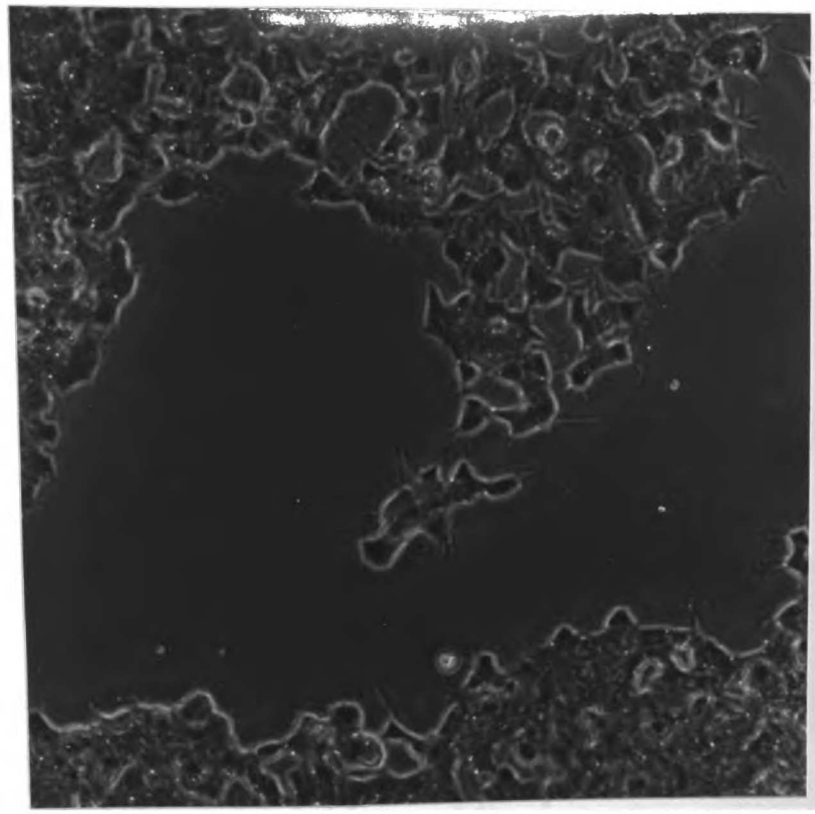
H. PC-12 cells grown without addition of extract or media. Note the absence of significant sprouting.

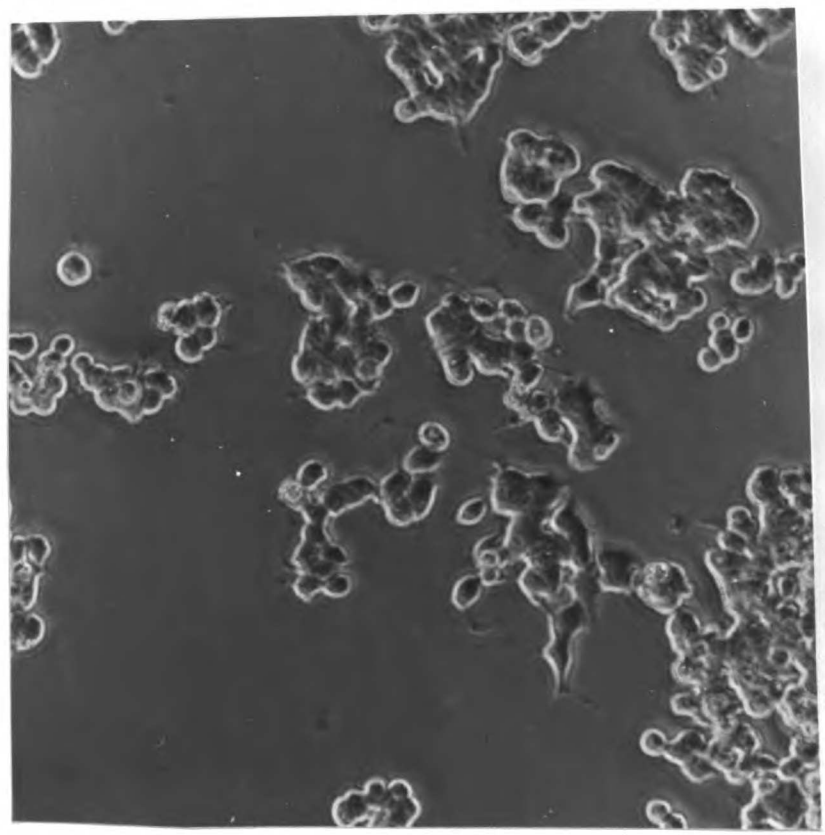


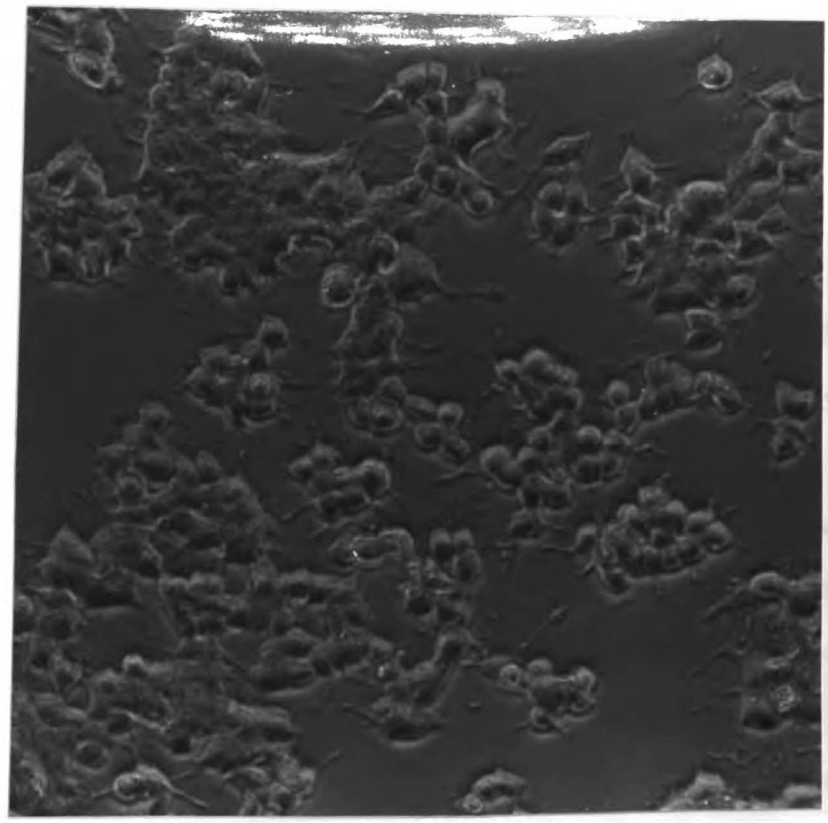












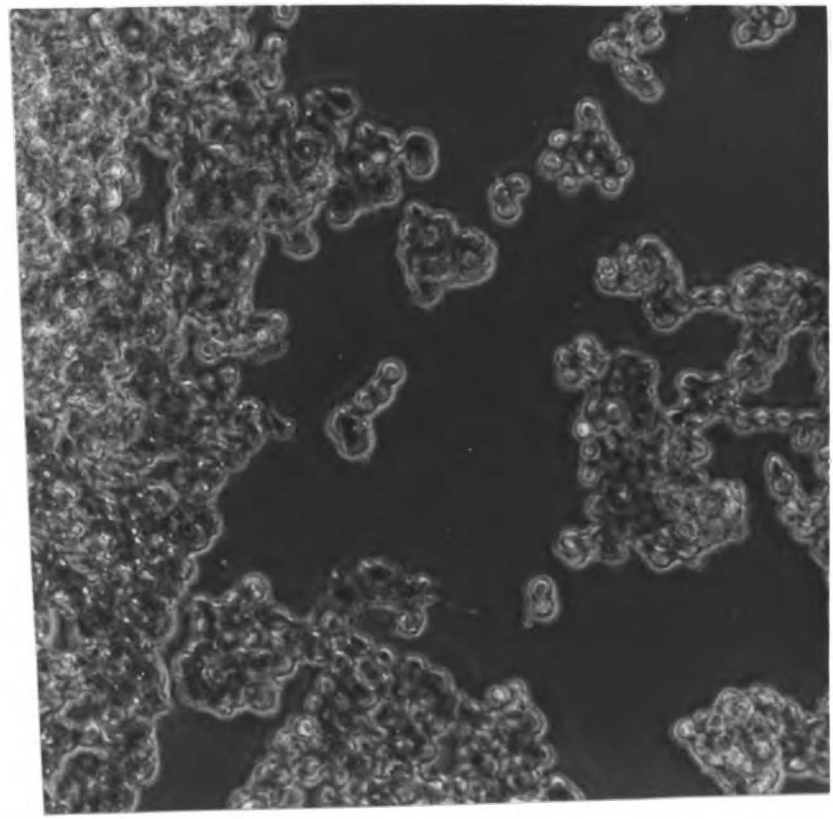


FIGURE 9. α subunit does not affect γ subunit digestion of NGF precursor. Lysates from both A (left) and B (right) vaccinia virus-infected L-929 cells were first incubated with varying amounts of α subunit (indicated above the figure), followed by incubation with a fixed amount of γ subunit (0.9 μg). To ensure that α subunit has no catalytic activity, it was incubated with in vitro-derived precursor; no cleavage was detectable, even after several hours of incubation (data not shown). Mature NGF migrates just below where the 14 kd standard migrates. No precursors are apparent; the generation of mature NGF is unaffected by preincubation with the α subunit. Thus, no role is demonstrated for the α subunit in the processing.

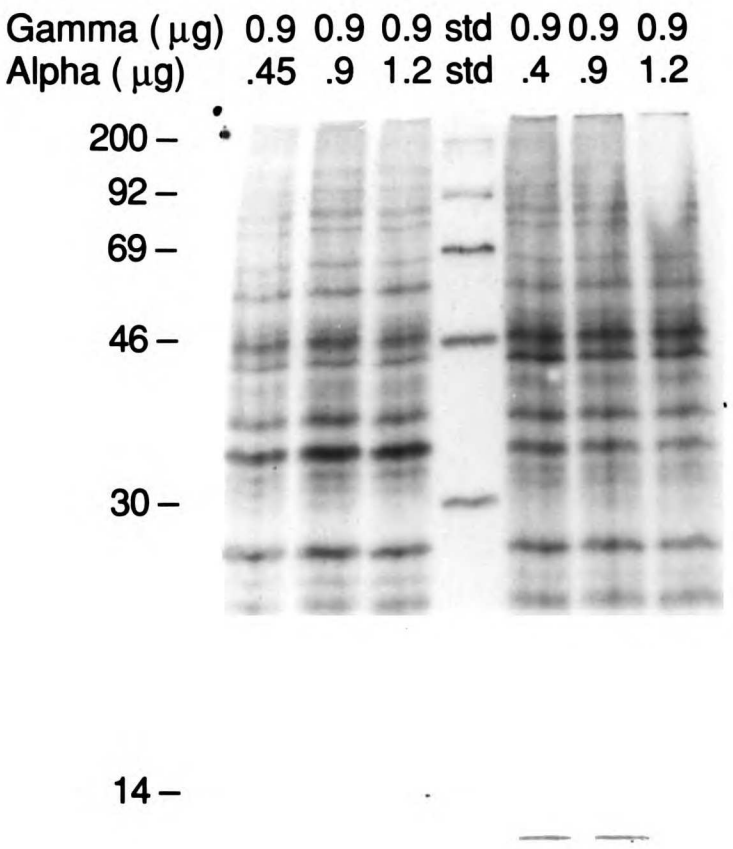
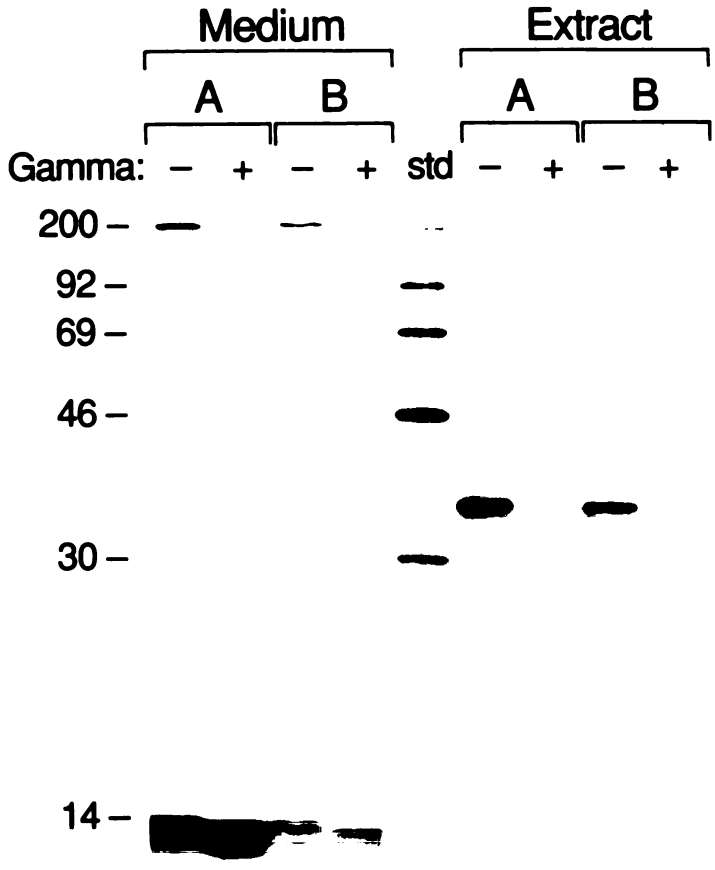

Gamma (μg) 0.9 0.9 0.9 std 0.9 0.9 0.9
Alpha (μg) .45 .9 1.2 std .4 .9 1.2
200 –
92 –
69 –
46 –
30 –
14 –

FIGURE 10. Immunoprecipitation of cell medium and cell extracts from vaccinia virus-infected HIT cells. HIT (hamster insulinoma tumor) cells were infected with A and B vaccinia viruses and labeled with ^{35}S-methionine for 4 hours. γ subunit (900 ng) was added (+) to both medium and cell lysates for 35 minutes at 37°C and compared to samples not incubated with this subunit (-) following immunoprecipitation. The presence of γ subunit in the reaction caused a minor mobility shift in secreted NGF, suggesting that excess γ subunit may be able to further cleave mature NGF. γ subunit digestion of cell extracts generates mature NGF from the two precursors. This experiment shows that HIT cells also secrete an NGF molecule, although its mobility is altered by incubation with the γ subunit.


Medium
Extract
A
B
A
B
Gamma: – + – + std – + – +
200 –
92 –
69 –
46 –
30 –
14 –

purified NGF and extracts, it can be estimated that 40 ng/μl of NGF in processed extracts are synthesized. Less than 900 ng of γ subunit complete converts precursor (10 μl extract) into mature NGF (~400 ng). Thus, a 1:1 stoichiometry of NGF to γ subunit is approached.

In an attempt to detect a role for the α subunit in processing, purified α subunit was added in advance of the γ subunit. No difference in the ability of γ subunit to effect cleavage was noted when α subunit was present (Fig 9).

The ability to process precursor NGF is not restricted to L-929 cells: NGF precursors synthesized in vaccinia virus-infected endocrine hamster insulinoma cells (HIT) were processed similarly to those from L-cells (Fig. 10). Immunoprecipitation of media from long-term labelings of infected HIT cells showed bands with a mobility of ~13.5 kd. The mobility of this band was increased following γ subunit cleavage to a position identical to that derived from γ subunit processing of infected cell extracts. This suggests that the secreted and processed NGF species differ. By contrast, NGF secreted by L-929 cells has the identical mobility to that derived from γ subunit processing of infected cell extracts (data not shown). Thus, HIT cells appear to secrete a partially cleaved NGF species. It is possible that the two amino acids residues which follow mature NGF in the precursor were not removed in the NGF secreted from infected HIT cells. Amino-terminal radiosequencing of immunoprecipitated species from the media of infected cells would indicate whether the amino-terminus is intact or not.

DISCUSSION

NGF precursors synthesized in VV-recombinant infected cells are correctly processed in vitro into mature NGF by the γ subunit. This

conclusion is supported by the following evidence. 1) The sequence of the processed product; 2) The inhibition of immunoprecipitation with mature NGF; and 3) bioactivity of the putative NGF on PC-12 cells. The kinetic data illustrate that the reaction proceeds with a $t_{\frac{1}{2}}$ of roughly 25 minutes. The titration experiment with increasing quantities of γ subunit show that large quantities of enzyme are rquired to effect a complete cleavage of substrate. From the biological assay data, we can calculate roughly the amount of NGF present in extracts. This amount roughly corresponds to the amount of γ subunit required to effect a complete cleavage of precursor substrate. Thus the stoichiometry of enzyme to substrate approximates 1:1. This observation would be consistent with the proposed role of the γ subunit in cleaving the substrate and concomitant formation of the complex (Server & Shooter, 1977).

The specificity requirements for the cleavage reaction are obviously loose because trypsin can replace the γ subunit in faithful precursor processing. Burger and Shooter reported that the 22 kd precursor intermediate could be processed into 13 kd species by the EGF binding protein as well as by γ subunit (Burger & Shooter, 1977). Further, every VV recombinant infected-cell type analyzed thus far (R. Edwards et al., in preparation) can process precursor and secrete mature NGF into the media. These results argue that several enzymes can recognize and correctly process the NGF precursors.

The presence of the α subunit failed to alter the cleavage of the precursors by the γ subunit. Thus, there is no apparent role in processing per se; however it may protect the β subunit (NGF) from further proteolytic digestion.

Radiosequencing of the two NGF precursors suggests that signal

sequence cleavage is occurring at the predicted peptide bond in both. This is consistent with the results of in vitro translation experiments using pancreatic membranes. The in vitro experiments also predict that the short (27 kd) and maybe the longer (34 kd) precursor become glycosylated; this is consistent with our results from the infected L929 cells. In contrast to the in vitro analysis, we cannot demonstrate that full-length 34 kd precursor is actually being synthesized in vivo because it would not appear due to co-translational signal sequence cleavage. It is possible, although unlikely, that the internal AUG is being utilized as the initiator AUG, thus encoding a 27 kd precursor.

The size of processing intermediates (~30 kd) is consistent with cleavage at the first set of dibasic amino acid residues in the glycosylated precursor; cleavage here would yield species with a molecular weight of ~30,000.

It is noteworthy that the NGF precursors from both mRNAs can be cleaved into mature NGF even when bound by anti-NGF antibody. This is similar to the cleavage of the immunoprecipitated 22 kd precursor intermediate observed by Burger and Shooter (1977). The lack of additional radiolabeled polypeptides following gamma subunit cleavage is consistent with the removal of the signal peptide which contains two methionine residues (met-ser-met); had the signal peptide not been removed, small molecular weight polypeptides of ~3 kd would have been observed.

It would be interesting to determine if 7S complex formation could occur in the processing reactions from infected cell extracts. Additionally, it might be of interest to determine what enzyme(s) facilitates cleavage of NGF precursors in cultured cells because to date, no γ

or γ-like enzyme has been described in L-cells (Pantazis, 1983); presumably a trypsin-like enzyme is responsible for NGF precursor processing in cultured cells. This activity must be somewhat labile or not particularly potent, for in extracts incubated without γ, very little mature NGF is generated. Infected HIT cells also synthesize and secrete precursors and mature NGF, respectively. However, γ subunit processing of secreted NGF is required to generate the same NGF species (by mobility) that is derived from processing of in vivo synthesized precursors. No similar phenomenon was observed in L-cells; thus, processing of the NGF precursors is apparently different in HIT cells than in L-929 cells.

In summary, it has been demonstrated that the γ subunit can process full-length precursor molecules and liberate mature NGF. However, those precursor molecules derived from the long and short NGF mRNA species appear identical, owing to cleavage of the precursor at a common signal peptide site and possible internal translation initiation at a common site. The γ subunit cleavage of the precursors can be mimicked with trypsin and appear unaffected by the addition of the α subunit.

CHAPTER VII

DISCUSSION

I have used molecular genetic approaches to help clarify our understanding of NGF, the gene, its protein products, and their expression. Specifically, I have helped to define the structure and relative expression of multiple NGF transcripts and the corresponding gene structure and the structure and processing of the NGF precursors. Further, the isolation and characterization of a cobra NGF cDNA and its comparison to NGF precursors from other species, have led to a better understanding of important, conserved regions of the molecule. Collectively, these data provides insight into how NGF is synthesized, both at the protein and RNA levels. Further, possible additional functional roles for the NGF gene have been examined by studying the precursors derived from multiple transcripts. These studies and experiments of others (resulting from the availability of the cDNA) have laid the groundwork for more detailed analyses of the molecular details of NGF function.

At the outset of my thesis research, I isolated and characterized an NGF cDNA clone from a male mouse submaxillary gland library which predicted that NGF (13K) is derived from a higher molecular weight precursor of 33.8K (34K). Mature NGF is located near the carboxy terminus of the precursor; it is flanked at its amino and carboxy termini by 187 and 2 amino acids, respectively. The C-terminal extension, albeit small, confirmed the prediction of Server and Shooter (1977) that an additional residue(s) was present beyond the C-terminus of NGF to account for binding of the γ subunit to the terminal Arg of mature NGF following its cleavage from the precursor. The process of formation of NGF from its precursor is likely to be similar to that of other polyprotein precursors: proteolytic processing at flanking dibasic

amino acid residues liberates the bioactive polypeptide(s) (Douglas et al., 1984). Cleavage at two other sets of dibasic residues in the non-NGF region of the precursor could liberate other polypetides. The function of the peptides, if any, in the non-NGF region is unknown. The size of the predicted precursor is larger than the 22 kd species which had been detected by immunoprecipitations from SMG slices labeled in vitro. Cleavage at the first pair of basic amino acids residues could generate a 22 kd species, thus accounting for the above observations. Darling et al. (1983) generated an anti-NGF antibody that immunoprecipitated from labeled SMG extracts a 34 kd precursor as well as presumed intermediates of precursor processing. Trypsin digestion of all these species showed tryptic fragments in common with one another and with mature NGF. Thus, their data confirmed the existence of a 34 kd precursor from which mature NGF is derived.

Stanley Cohen (1960) demonstrated that the levels of NGF protein in the SMG are sexually dimorphic: extracts from male glands produce a 6-fold higher level in neurite outgrowth than extracts from female glands. The difference is most likely due to regulation of transcription because both Northern blot (Scott et al., 1983) and primer extension (see Fig. 1, appendix of Chapter I) analyses show a corresponding sex-dependent difference in NGF RNA levels. However, it is not determined whether this difference is due to increased transcription in males or decreased RNA turnover. The higher levels in males suggest that androgen (testosterone) plays a role in mediating the increased RNA levels. In support of this notion, the development of glandular tubule cells (GTC's), to which NGF has been localized, is controlled by testosterone. This hormone may also regulate other activities: not only

is NGF more abundant in males but proteases and EGF are as well. The larger cell size and the more dense network of Golgi and endoplasmic reticulum suggest the GTC's of male and testosterone-treated female SMG's exhibit greater synthetic capacity (Bhoola et al., 1973). Hence, the increased NGF levels could be an indirect manifestation of a generalized androgen effect on the GTC's themselves and not necessarily a direct effect on the NGF gene. Thus, NGF as well as EGF may be androgen responsive only in the context of the SMG. It is difficult to address this issue by transfection experiments since no established cell line from the mouse SMG is currently available.

The transcriptional complexity of the NGF gene was perceived with the discovery of several NGF RNA transcripts (Chapters II and IV). Three of these encode related NGF precursors that have common carboxy-termini and unique amino-termini. Two of the precursors have internal hydrophobic stretches and one precursor has that same sequence at its amino-terminus by virtue of differential splicing. Why are multiple precursors synthesized? Important polypeptides may be derived from the amino-termini of the molecules in addition to NGF. Alternatively, they may serve to generate mature NGF and multiple precursors may be a reflection of different maturation pathways within the cell.

The structure of the mouse NGF gene accounts for the generation of all the transcripts. Alternative splicing (of exon II) is responsible for the synthesis of the two principle NGF transcripts. It is perhaps significant that the ratio of the two principle NGF transcripts is inverted in the SMG and placenta as compared with all other tissues. This observation suggests that some mechanism exists to preferentially select a particular splicing pathway in the SMG. How the transcription

processing machinery selects particular exons from a linear array is a vexing question. Any mechanism explaining splicing must account for this selectivity. Further, the mechanism must be flexible enough to incorporate cell- or developmental-specific splicing pathways. Obviously, some higher order control must be rseponsible for fine tuning the splicing machinery. Ribonucleoproteins (RNP), for example, facilitate the splicing mechanics and may be a good target for fine control (Padgett et al., 1986). Specifically, trans-acting factors interacting with specific RNP's may be present at different concentrations in different tissues or during development and thus account for the selectivity of splicing.

The NGF gene contains an unusually large first intron which precluded the isolation of the first exon (33 bp) of the major transcripts. This intron may, in fact, contain an unstable sequence(s) as evidenced by the inability to recover overlapping phages. Indeed, phages possessing unstable inverted repeats are non-viable in standard host strains (Leach & Stahl, 1983). However, recBC sbcB hosts can propagate phages with inverted repeats fairly well (ibid; Wyman et al., 1985; Nader et al., 1985). The absence of exoV and exoI wild type gene products may be responsible for the increased viability of otherwise unrecoverable phage; palindromic structures that would be "resolved" by these exonucleases in wild type hosts (analogous to resolution of the Holliday junction found during recombination) are stable in the modified hosts. These observations can possibly be extended to account for the absence of overlapping phage in the first intron; specifically, inverted repeats might be present in the genomic DNA unrepresented by the two NGF genomic phages. If the Leder mouse genomic phage library had been propagated

through the aforementioned host, the first two exons (IA and IB) might have been recovered after "walking" through the large first intron.

Promoters of the transcription of genes reside proximal to the 5' most exon. Some of these different elements (Myers et al., 1986) include the TATA and CCAAT boxes, which are binding sites for factors that are responsible for transcription and its regulation (see Dynan & Tjian, 1985). The flanking region for the two major NGF transcripts contains a sequence (TTAAA) at the predicted location for a TATAAA element within the first 30 bases of the transcription start site. No CCAAT element is found in the region between the TTAAA element and the proximal exon. The 5' flanking region of the third transcript does not have a TATA box but does have two CCAAT elements located at positions -79 and -195 relative to the presumptive transcription start site; the -79 location is similar to that for several other genes and may thus be functional. The NGF gene is not unique in lacking a TATA consensus sequence: other genes e.g., human chorionic somatomammotropin, has a variant such as CATAAA (Selby et al., 1984) or structurally unrelated elements such as CCGCCC (Valerio et al., 1985).

The expression level of the two major NGF transcripts is high in the SMG, thus, specific factors here may promote high level transcription or alternatively, negative regulatory factors repressing transcription might be reduced or absent in the SMG. In this regard, it might be informative to perform gel retardation experiments using the 5' flanking DNA as probe and nuclear extracts from the SMG and a control tissue, such as liver. If, for example, positive regulatory factors are present in the SMG to augment transcription, one might observe probe retardation on polyacrylamide gels. Specificity could be demonstrated by compe-

tition with specific DNA fragments from the 5' end of the gene. Footprinting would confirm that protein-DNA interactions are actually occurring and it is sequence-specific. These experiments must follow 5' deletion/expression analysis.

The presence of multiple NGF transcripts prompted us to examine the expression of the two principle transcripts during postnatal development. If one of the transcripts has a unique function during this period, a change in the ratios of transcripts might reflect this. We observed that the overall level of transcripts did change modestly, but no alteration in the ratio was observed during development through adulthood; there was a parallel early peak of expression followed by stabilization at lower levels. Significantly, the pattern of NGF RNA expression in the mouse cortex was similar to that of the rat brain, where hippocampus and neocortex NGF RNA and protein levels have been monitored (Large et al., 1986). The temporal rise in NGF expression during early postnatal development may coincide with the expression of terminal varicosities and synaptic contacts. Alternatively, NGF expression in particular cells of a given lineage could be programmed to synthesize NGF with this pattern. If this is indeed the case, then flexibility in the system might reside at another level: for example, the regulation of the affinity of the receptor (high vs low) or in the delivery of NGF to the cell surface (time, post-translational modifications, etc.).

Genomic Southern blots of various DNAs probed with mouse NGF cDNA indicate that the NGF gene is most likely a single copy gene (see Fig. 4, appendix of Chapter I; Shelton & Reichardt, 1984). Probing with the cDNA does not reveal the complexity of the first exons of the major and

minor 5' exons of the gene because these regions were not present in the clone used as probe. A mouse genomic Southern blot probed with a synthetic oligonucleotide representing the first 33 bp exon of the major transcripts (nucleotides 2-34) shows a single copy hybridization pattern (see Fig. 5, appendix of Chapter I). This hybridization was not attempted for human genomic DNA (conditions would have to be modified to correct for possible sequence divergence). However, a homologue to the larger 5' exon of the third transcript exists in the human genome because a mouse probe consisting of this exon hybridized to human genomic DNA in a single copy pattern (see Fig. 4, Chapter IV). Thus, it is likely that multiple NGF transcripts also derive from a single copy human gene.

The longer 34 kd precursor has no identifiable amino-terminal signal-type sequence but does have an internal relatively hydrophobic region (see hydrophobicity plot, Fig. 1, Chapter V). This same region is present at the amino terminus of the 27 kd precursor by virtue of a different AUG initiation codon provided by alternative RNA splicing. The absence of an amino-terminal hydrophobic domain in the larger precursor is reminiscent of ovalbumin which may also lack an amino terminal signal sequence (Lingappa et al., 1979), but apparently contains an internal equivalent. I have carried out coupled in vitro translation/translocation experiments programmed with synthetic precursor RNAs. The results suggest that both the 34 kd and 27 kd precursors are secreted intact across dog pancreas microsomal membranes with concomitant signal sequence cleavage. Based on the in vitro observations, one would predict that in cultured cells both precursors are secreted into the lumen of the endoplasmic reticulum and are then

subject to proteolytic processing. The function of the relatively hydrophobic region in translocation could be further tested by linking this region to a normally cytoplasmic protein: the location of the hybrid could then be assayed to determine if it remains soluble or is translocated into the ER.

I have shown that in vitro generated precursors are not substrates for γ subunit processing. The likely explanation for this observation is that the precursors do not assume the conformation that is recognized by the γ subunit. It is possible that the redox environment may not be appropriate for correct disulfide bond formation and thus signal sequence-free substrates may still not be correctly processed. Modifying the redox environment with oxidized and reduced glutathionine may allow for correct disulfide bond formation and subsequent γ subunit cleavage. It is also possible that the γ subunit can only cleave terminally-glycosylated species and not the non-mature form that is found in the ERM.

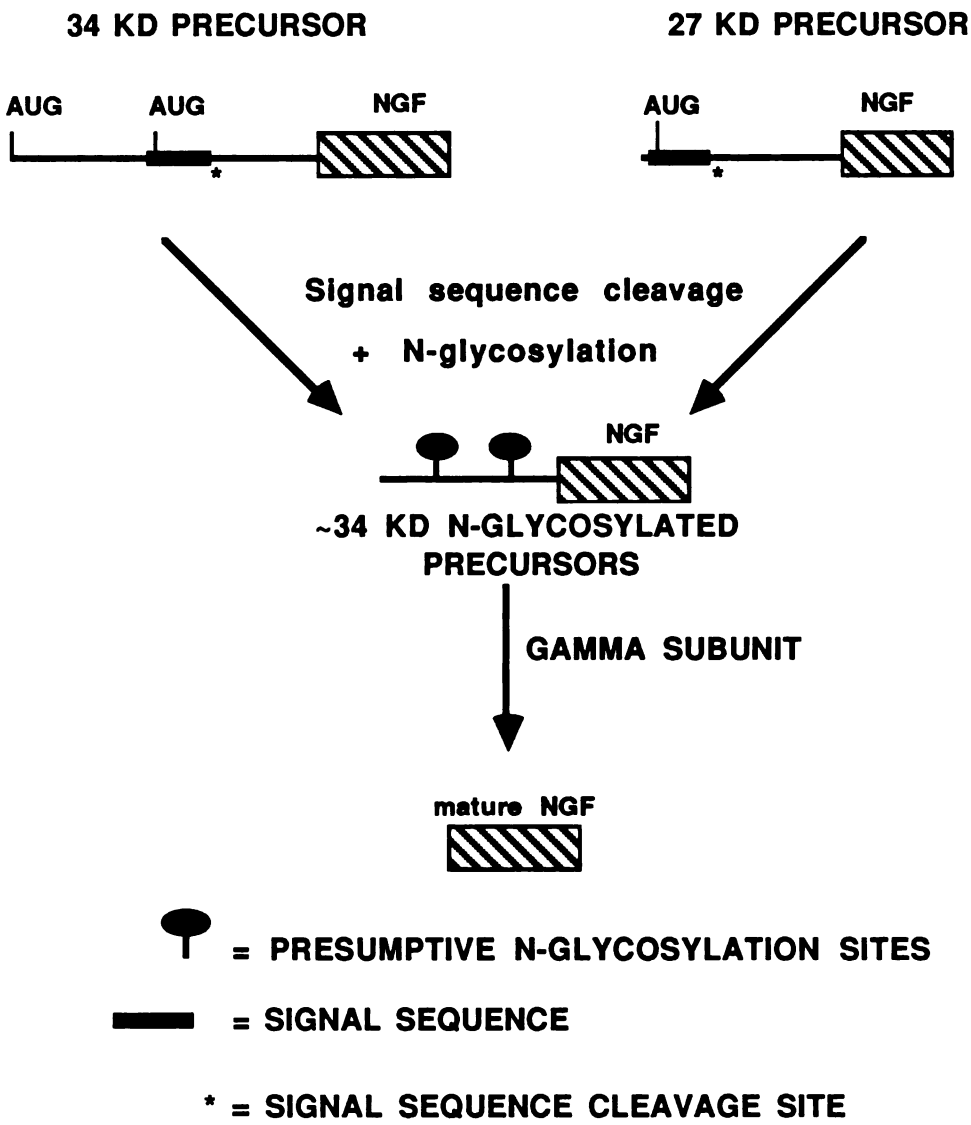
Burger and Shooter (1977) derived mature NGF from a 22 kd mouse submaxillary gland precursor by γ subunit processing. The 22 kd species can be accounted for by assuming that processing occurred only at the first pair of dibasic amino acid residues in an unglycosylated precursor. γ subunit digestion of in vivo synthesized precursors generates 13 kd NGF by gel mobility, sequence analyses, immunocompetition and bioassay. In both kinetic and γ subunit titration studies, no stable 22 kd species was observed. Further, both precursors are identical in gel mobility, glycosylation and amino-terminal sequence. This can be rationalized if one assumes that no translation initiation occur at the first AUG; initiation at the second AUG would encode a 27 kd protein

that would be modified by both signal sequence cleavage and glycosylation. This could account for the identical gel mobilities and amino-terminal sequences. Alternatively, the 22 kd species could represent a partially degraded species and not a processing intermediate. To certify that the first AUG is capable of being recognized in virus-infected cells requires site-directed mutagenesis to replace the internal AUG's (met-ser-met) such that any NGF precursor synthesized can only be derived from translation initiation at the first AUG.

Saboori and Young (1986) have isolated a 32 kd NGF intermediate that can be converted into mature NGF, as determined by isoelectric gel analysis and by bioassay. However, this 32 kd species is inconsistent with the predicted processing at dibasic amino acid residues and the amino-terminal sequence of the NGF precursors from Vaccinia virus-infected L-929 cells. The 32 kd precursor begins with Ala 28 which has no predicted site for signal peptide cleavage (von Hjeine, 1986). However, since Ala is flanked by Arg, cleavage at Arg by a trypsin-like enzyme (abundant in the submaxillary gland) and subsequent cleavage by a carboxypeptidase could generate this amino-terminus. This is difficult to rationalize (von Hjeine, 1986) with preferred cleavage at dibasic amino acid residues. On the other hand, we cannot demonstrate that the 34 kd precursor is actually synthesized in Vaccinia virus-infected cells. However, the in vitro translation in the presence of ERM does not show a 32 kd species (see Fig. 3, Chapter V).

One is left with the conclusion that both NGF messages produce the same precursor. If the 34 kd nascent precursor is indeed synthesized, then it must occur via cleavage at the internal signal petpide (see schematic Fig. 1). Both precursors, as expected, are equivalent

FIGURE 1. Schematic presentation of the fate of the two NGF precursors and subsequent γ subunit digestion. The two predicted precursors of 34 kd and 27 kd are shown with their two or one important AUG codons, respectively. The mature NGF coding region is illustrated as a stripped box. Intracellular signal sequence (illustrated as a filled-in box) cleavage between amino acid residues 18 and 19 (denoted by stars) and N-linked glycosylation (illustrated by stippled circles) generate the same precursor species. These are then converted into mature NGF by the addition of γ subunit.


34 KD PRECURSOR
27 KD PRECURSOR
AUG
AUG
NGF
AUG
NGF
Signal sequence cleavage
+ N-glycosylation
NGF
~34 KD N-GLYCOSYLATED
PRECURSORS
GAMMA SUBUNIT
mature NGF
= PRESUMPTIVE N-GLYCOSYLATION SITES
= SIGNAL SEQUENCE
* = SIGNAL SEQUENCE CLEAVAGE SITE

substrates of the γ subunit in generating mature NGF. It is curious that cells utilize two different, but related transcripts to generate the same precursors. Perhaps the amino-terminal peptide (extending from the initiator methionine of the 34 kd precursor to the signal sequence cleavage site) is a stable product and has some biological role. Alternatively, the precursors could be shuttled to different extracellular compartments where their secretion might be influenced by a particular stimulus (electrical activity; neurotransmitter release). Apparently, the NGF precursor is synthesized identically in all cell types examined; this contrasts with preproglucagon where correct processing is restricted to particular cell types (Drucker et al., 1986). Our experiments do not elucidate the reason for the two NGF transcripts. Similar experiments with recombinant vaccinia viruses in primary innervated cells may be more informative.

The identification of conserved sequences in evolutionarily distant organisms helps to identify important regions. For example, the importance of the kinase domain in the insulin receptor is suggested by its conservation between human and Drosophila (Petruzzelli et al., 1986). In order to reveal conserved and potentially important regions in the NGF precursors, I isolated a cDNA clone corresponding to the snake NGF precursor and compared its predicted amino acid sequence with that from mouse, human, cow and chicken. The cobra NGF sequence is fairly well conserved (66%) compared to mouse although it is two amino acids shorter (116 vs 118). Certain residues such as tryptophan and cysteine residues are invariant and hence may be important for activity. Antisera raised against snake or mouse NGF's show little or no cross-reactivity to one another despite the structural relationship. (Bailey et al., 1976). No

carboxy-terminal dibasic residues are found in snake NGF; this precludes complexing with a γ subunit equivalent. Chicken and bovine NGF's also do not have this terminal dibasic sequence and thus probably cannot partake in complex formation (Ebendal et al., 1986; Meier et al., 1986; Wion et al., 1986). Using the cobra cDNA as a probe one obtains the characteristic NGF hybridization pattern to restriction digested human and mouse genomic DNAs (Fig. 4, Chapter III).

The non-NGF moiety is not as well conserved (45%) as the mature NGF region. In fact, the only three well-conserved portions are the N-terminal presumptive signal sequence and two of the three dibasic cleavage sites. One pair of basic residues abuts the NGF moiety; cleavage must occur here to liberate NGF from the precursor. Cleavage at the next pair of dibasic residues would bifurcate the non-NGF moiety. Although the overall homology in the non-NGF domain is low, the number of amino acid residues from the initiator methionine to the NGF moiety is fairly well-conserved. The length conservation coupled with poor sequence homology may suggest this region has a structural rather than a biological function. It is conceivable that the purpose of the non-NGF moiety is to aid in the correct folding of the precursor so as to ensure recognition by the appropriate precursor processing enzyme(s).

In contrast to the multiple NGF transcripts in mouse, only one NGF RNA species was detected in cobra venom gland RNA. Indeed, the inability to detect a long transcript in snakes equivalent to that in mouse suggests that this translation product may have a unique function in mammals (if it has one). For example, the longer precursor predicted from the mouse RNA may be secreted via a unique secretion pathway.

One consequence of the availability of NGF cDNA and genomic probes is that the genetic basis of particular neurological defects can be investigated. Specifically, NGF has been implicated in several disorders of human peripheral nerve growth or development (Schenkein et al., 1974; Siggers et al., 1976; Schwartz & Breakfield, 1980; DeSchryver-Kecskemeti et al., 1983). This hypothesis was drawn from gross morphological observations and controversial measurements of NGF (bioassay/RIA). Restriction digests of genomic DNAs from families afflicted with either peripheral neurofibromatosis or familial dysautinomia were probed with a unique region derived from human NGF genomic phages. Restriction site polymorphisms in the NGF gene have been identified. The diseases segregate independently from the polymorphisms and thus argue against an obvious defect linked to the NGF gene (Breakfield et al., 1983; Darby et al, 1985). These data do not rule out point mutation(s), defects in precursor processing, or defects in the NGF receptor.

The correlation between the NGF levels and the density of sympathetic innervation in peripheral targets (Korshing & Thoenen, 1983) was extended to the RNA level using our NGF cDNA as a probe. Specifically, the intensity of hybridization of the probe to RNAs from various tissues parallel the density of sympathetic innervation. Further, there is a strong correlation between NGF message levels and norepinephrine content for each tissue (Shelton & Reichardt, 1984). The ganglia contain high levels of immunoreactive NGF and a relatively weak RNA signal, suggesting that most of this NGF is derived from retrograde transcript from sympathetically innervated targets. The remaining NGF must be derived from local synthesis, most likely from supporting cells (e.g.

Schwann and fibroblast cells) which synthesize NGF in culture (Rush, 1984; Brachet & Dicou, 1984).

NGF gene expression is not limited to the periphery, it has also been detected in brain. Analysis of NGF immunoreactivity demonstrated the highest levels of NGF are found in those areas containing either the cell bodies (septum, diagonal band of Broca and nucleus basalis) or the terminals of the magnocellular cholinergic neurons (hippocampus. olfactory bulb and neocortex) (Korshing et al., 1985). Moreover, Northern blot analysis show that cholinergic regions of the basal forebrain, the hippocampus and cortex, contain the highest levels of NGF message (Korhsing et al., 1985; Shelton & Reichardt, 1986a; Whittemore et al., 1986). Higher ChAT levels appeared in the striatum following exposure to NGF (Mobley et al., 1986; Martinez et al., 1985). Following fimbria fornix transection, the administration of NGF to the brain (intraventricular injection or infusion) muted the cell damage in the septal area (Hefti, 1986; Kromer, 1987). Thus, NGF is important for the cholinergic nerves of the forebrain and most likely plays some role in other brain regions since both NGF protein and NGF mRNA are detected elsewhere in brain.

Fimbria transection of neonatal rat brains leads to a modest increase (60%) in hippocampi NGF mRNA (Whittemore et al., 1987) in neonates; no change in NGF mRNA levels is detectable following the same lesion in adults (Whittemore et al., 1986; Korshing et al., 1986). From this the authors argue that different mechanisms of NGF regulation are operating in neonates and adults; for example, in neonates, the regulation may be at the level of transcription, while in adults it may be at the level of translation (Whittemore et al., 1987). However, the sample

size in the neonatal experiment was so small (two) that the reliability of this observation is questionable.

The developmental profile of NGF protein and NGF mRNA in particular regions of the rat brain was determined by Large et al. (1986). They confirmed the relatively high NGF RNA levels in the hippocampus and neocortex. The observation of low NGF mRNA and high NGF protein levels in the basal forebrain is consistent with retrograde transport of NGF to this region from the target tissues. The increase in hippocampus and neocortex NGF mRNA correlates well with the contact of basal forebrain neurons with these targets. Further, the very low levels of NGF at or just after birth argue against a role for NGF in initial guidance of cholinergic neurons to these targets. These results therefore supplement the observation of Davies and Lumsden (1984) in the developing peripheral sensory nervous sytems.

Quantitation of NGF RNA from the various peripheral and central target tissues predicts there is less than one copy of NGF mRNA per cell, thus implying that NGF RNA may be synthesized in only a subpopulation of cells in a given target (Shelton & Reichardt, 1984). Alternatively, the NGF could be derived from nerve supporting cells and from the target cells. The presence of NGF receptors at the nerve terminals is consistent with uptake from the targets, while receptors all along the nerve may take up NGF derived from supporting cells. Cultured glial and astrocyte cells (Burnham et al., 1972; Lindsay, 1979) do make NGF and induction of NGF synthesis has been observed upon culturing these cells (Shelton & Reichardt, 1986b). Immunohistochemical analysis of cultured rat irides shows NGF-positive staining of accessory cells but no staining of the nerve fiber itself or the smooth muscle target (Rush,

1984). However, given the artifactual NGF induction in culture, this conclusion may not be physiologically relevant. Korshing et al. (personal communication) have detected NGF mRNA around the site of a sciatic nerve crush, this presumably results from synthesis by the Schwann cells lining the nerves. Immunohistochemical staining of sciatic nerves with anti-NGF receptor antibody shows labeling of the endoneurium of transected nerves and not in normal nerves, suggesting that Schwann cells synthesize the receptor. Taniuchi et al. (1986) proposed that Schwann cells present NGF on their surface (via NGF receptors) to injured nerve cells and that axons regenerate. Thus the Schwann cell NGF provides chemotactic and haplotactic guidance to these neurons. In this view, supporting cells act as a physiologic source of NGF, at least for regeneration. A clear resolution of the relevant in vivo sources of NGF expression awaits in situ hybridization and the corresponding cellular definition.

In summary, the experiments presented in this thesis define the structure of multiple NGF mRNAs and the corresponding gene structure. The expression of the gene proteolytic processing of predicted NGF precursors. I have analyzed the gene in detail with regard to sequence, splice junctions and potential flanking elements. The developmental expression of the two principle NGF transcripts has been monitored through early postnatal development; they change in parallel. Further, the isolation of the evolutionarily distant cobra NGF cDNA allowed the recognition of conserved/non-conserved regions and hence, for predictions of regions of functional and/or structural importance. Also, I have shown that NGF precursors synthesized in vivo are suitable substrates for γ subunit processing, while those generated in vitro are

not. The cloning of the cDNA has also provided a tool to examine the relationship between NGF levels and sympathetic innervation as well as the possible role of NGF in particular neurologic diseases.

NGF and other polypeptide hormone, e.g., EGF, secreted by the submaxillary tubule cells, appear to be androgen regulated. It would be of interest to establish stably transformed cells which are androgen responsive to address the issue of whether or not NGF expression itself is regulated by androgens. This might be studied via a chimeric construct consisting of the NGF promoter(s) linked to a reporter gene (such as chloramphenicol acetyltransferase (CAT)). Inductions of the reporter gene following androgen treatment can be conveniently monitored by CAT assays (increased mobility of acetylated radioactive chloramphenicol substrate). If the 5' end of the gene is sufficient to ensure enhanced expression of the reporter, deletion mapping would serve to define those sequences responsible for the androgen induction, as was done with the rat insulin gene (Walker et al., 1983). If androgens increase NGF mRNA by message stabilization, no effect would be observed in these stable transformants. Analysis of message stabilization would require the expression of NGF mRNA in androgen-responsive cells. The amount of NGF message under induced or non-induced conditions could be quantitated by RNase protection experiments, for example.

The available data show a correlation between the level of NGF expression and the density of sympathetic innervation. One means to certify this directly would be to over-express NGF in a defined tissue of an animal via microinjection of NGF DNA coupled to a cell specific promoter, such as the rat insulin II or elastase promoters directing expression to the appropriate pancreatic cells. Histochemical analysis

of tissue expressing NGF from transgenic mice should determine if increased NGF expression provides for a greater degree of sympathetic innervation. There are several limitations to this approach: (1) Expression of NGF should be limited to a tissue that is sparsely innervated normally (such as the pancreas) so the prospective increase in innervation can be distinguished from the endogenous innervation. (2) The most effective approach would require the microinjection of constructs representing both NGF precursors. However, analysis of many offspring could make this a cumbersome task. If one precursor is to be selected, the "short" precursor might be the choice for its message predominates in all tissues except the salivary gland and the placenta. (3) It is possible that the expressed precursor may not be processed correctly in the particular tissue, although directing expression to cells normally innervated by sympathetic neurons should reduce this possibility. (4) Expression may not occur during development. Since NGF is most important for sympathetic neurons in postnatal development, it should be expressed at least at this point if not earlier. Expression at later periods might not be informative.

The level of NGF mRNA expression suggests that only a subset of cells in a given target tissue are expressing NGF. To better understand the physiologically relevant sites of synthesis of NGF during development, in adult neural networks and during regeneration, the actual cells synthesizing NGF mRNA must be defined. This requires in situ hybridization, but it is an awesome task since the levels of NGF mRNA are so low. Although Davies et al. (1987) published impressive in situ hybridization in the maxillary process of embryonic mice, precise cellular definition was lacking.

Another issue that needs to be addressed in this system and in many others is the role of splicing in controlling selective gene expression. To begin to analyze this, one could make transcripts in vitro that correspond to the hnRNA for a gene that is alternatively spliced. Once generated, the RNA could be incubated with nuclear extracts from specific cell types where the gene is and is not normally expressed. One would then assay for the generation of particularly spliced products. If this works, then fractionation of the extracts could begin to identify the required protein/ribonucleoprotein components. One may then be able to study the subtleties of the system that account for splice site selection.

CHAPTER VIII

FUTURE PROSPECTS

One of the important issues in understanding neural development is how nerves find and innervate their targets. Nerve targeting includes both cell guidance and selection. In nerve cell guidance, a prospective target might synthesize a chemotactic substance(s) specific for the appropriate ingrowing nerves. NGF was considered as a possible guidance molecule because cultured nerves turn towards the highest concentrations of NGF. However, the following points argue against a role for NGF in guidance in development: 1) non-physiological amounts of NGF were utilized; 2) in the experiment, thsee nerves had already interacted with targets; thus, the conditions did not mimic normal development; 3) nerve outgrowth towards naive targets was not inhibited by NGF antiserum; 4) NGF mRNA was detected concomitant with innervation and not before. To define molecules involved in neuronal cell guidance, the appropriate system must be used to search for the factors elaborated during development (developing trigeminal ganglia interactions with maxillary process target tissue) and not during regeneration. The fractionation and characterization of molecules elaborated by naive targets will be necessary to identify candidate substances. This approach is applicable to factors derived from a variety of targets to identify their specific trophic factor that might be NGF-like in function.

Selection of neural contacts is the phenomenon by which an excess number of innervating nerves or their varicosities are reduced to a particular amount. This might be subdivided into two possible mechanisms: 1) elaboration of a set quantity of guidance factor which defines the number of incoming nerves; 2) a mechanism that eliminates particular neural connections while preserving others. The amount of guidance factor synthesized could be proportional to the size of the

target cell or, alternatively, developmentally preprogrammed. Our data suggests the latter is the case. The means by which particular nerve connections are maintained may be a function of competition for limited trophic factor. Since exogenous NGF administration can increase the survival of additional neurons that would otherwise atrophy, it seems reasonable that NGF plays a role in mediating the survival of particular neurons. Those neurons that can best compete for NGF (in closest proximity to the target, for example) would then reduce the effective concentration of available NGF as a result of retrograde transport. Those not obtaining sufficient NGF would atrophy. Since the level of NGF is defined in a given target, the number of surviving neurons is also set.

Another issue of importance is the role played by NGF in nerve supporting cells, e.g., Schwann cells. NGF can be visualized immunohistochemically in the Schwann cells of nerves innervating irides; no immunoreactive material was observed in the target tissue. It is formally possible that NGF accumulation may be a consequence of binding of target or Schwann cell-derived NGF to Schwann cell NGF receptors. This may be particularly relevant during regeneration where NGF may be presented to regrowing nerves via Schwann cell NGF receptors.

The identification of NGF mRNA and protein in brain is suggestive of a role for NGF in the CNS. Further, it is specifically retrogradely transported and has a regional distribution in the brain. However, the data suggesting NGF ameliorates the deleterious effects of brain lesions is not particularly impressive, especially when compared to similar experiments in the periphery. What roles does NGF play in the brain? Since NGF protein is most abundant in the cell bodies of cholinergic

neurons, it may play a restricted role in cholinergic neuron survival. However, other regions of the brain also contain NGF protein or message, albeit at lower levels. This implies that NGF may have a role in other brain cell types. Defining the cells in the various regions that actually synthesize NGF may prove to be enlightening.

Understanding the role of NGF receptors (NGFR) in regulation of innervation, if any, has yet to be clarified. Two affinity classes (high and low) have been defined for the NGFR; the high affinity NGFR appears to be the physiologically relevant form. To date, only the low affinity NGFR has been cloned. The source of high affinity receptor is an interesting question. Northern and genomic Southern blot hybridizations indicate that only a single transcript from a single copy gene is present. This suggests that the receptor may be converted from the low to the high affinity form. Alternatively, the two receptors may be derived from unrelated genes, although low stringency hybridization experiments argue against this. At the physiological level, are NGF receptors regulated? Specifically, in the presence of ligand, do they down-regulate and if so, by what mechanism? Is new synthesis reduced, receptor recycling altered or delivery to lysosomes enhanced? Regulation at the receptor level could figure prominently into the regulation of NGF-mediated neuronal survival/maintenance.

Isolation of the Drosophila NGF equivalent might help in understanding how NGF may influence neural development. Low stringency hybridization experiments might be employed to recover the Drosophila gene from a genomic library. Analysis of the Drosophila NGF gene expression and the possibility that particular neurological phenotypes have been mapped to the same genetic locus (established by in situ

hybridization) would be revealing. In particular, it would be interesting to know the developmental consequence of the complete absence of the NGF gene.

Another issue to be addressed is how NGF is delivered to innervating nerve fibers. While it is reasonably clear that target-derived NGF is taken up by the nerve terminals, it is not apparent how NGF arrives at the synapse. Specifically, is it an active or passive process? Is NGF packaged in vesicles and then delivered when the incipient nerve fiber triggers its release? What is the trigger? Perhaps correct synapse formation as determined by electrical or chemical stimulation.

In summary, future endeavors in NGF research will undoubtedly focus on the role of NGFR, supporting cells and delivery of NGF to innervating nerves as potential regulatory mechanisms. Interest will also be directed towards understanding how NGF might act outside the peripheral nervous system. Genetic experiments in a more manipulable system might be informative.

CHAPTER IX

REFERENCES

Aloe, L., Cozzari, C. & Levi-Montalcini, R. (1981). The submaxillary salivary glands of the African rodent Promys (mastoms) natalensis as the richest available source of the nerve growth factor. Exptl. Cell Res. 133, 475-480.

Angeletti, P.U., Levi-Montalcini, R. & Caramia, F. (1971a). Analysis of the effects of the antiserum to the nerve growth factor in adult mice. Brain Res. 27, 343-355.

Angeletti, P.U., Levi-Montalcini, R. & Caramia, F. (1971b). Ultrastructural changes in sympathetic neurons of newborn and adult mice treated with nerve growth factor. J. Ultrastruc. Res. 36, 24-36.

Angeletti, R.H., Hermodson, M.A. & Bradshaw, R.A. (1973) Amino acid sequence of mouse 2.5S nerve growth factor II. Isolation and characterization of the thermolytic and peptic peptides and the complete covalent structure. Biochemistry 12, 100-115.

Bailey, G.S., Banks, B.E.C., Carstairs, J.R., Edwards, D.C., Pearce, F.L. & Vernon, C.A. (1976). Immunological properties of nerve growth factors. Biochim. Biophys. Acta 437, 259-263.

Banerjee, S.P., Snyder, S.H., Cuatrecasas, P. & Greene, L.A. (1973). Binding of nerve growth factor receptor in sympathetic ganglia. Proc. Nat. Acad. Sci. USA 70, 2519-2523.

Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980). The ovalbumin gene-sequence of putative control regions. Nucl. Acids Res. 8, 127-142.

Bhoola, K.D., Dorey, G. & Jones, C.W. (1973). The influence of anchogens on enzymes (chymotrypsin- and trypsin-like proteases, renin, kallikrein and amylase) and on cellular structure of the mouse submaxillary gland. J. Physiol. 235, 503-522.

Bjerre, B., Bjorklund, A., Mobley, W. & Rosengren, E. (1975). Short and long-term effects of nerve growth factor on the sympathetic nervous system in the adult mouse. Brain Res. 94, 263-277.

Blobel, G. (1980). Intracellular protein topogenesis. Proc. Nat. Acad. Sci. USA 77, 1496-1500.

Brachet, P. & Dicou, E. (1984). L cells potentiate the effect of the extracellular NGF activity in co-culture with PC-12 pheochromocytoma cells. Exptl Cell Res. 150, 234-241.

Breakefield, X.O., Orloff, G., Castiglione, C.M., Axelrod, F.B., Coussens, L. & Ullrich, A. (1983). Genetic linkage analysis in familial dysautonomia using a DNA probe for the beta nerve growth factor gene. In Biochemical and Clinical Aspects of Neuropeptides. ed. Koch & Richter, Academic Press, San Francisco, pp. 113-128.

Bruno-Bechtold, J.K. & Hamburger, V. (1979). Retrograde transport of nerve growth factor in chicken embryos. Proc. Nat. Acad. Sci. USA 76, 1494-1496.

Bueker, E.D. (1948). Implantation of tumors in the hind limb of the embryonic chick and the developmental response of the lumbosacral nervous system. Anat. Rec. 102, 369-390.

Burger, E.A. & Shooter, E.M. (1977). Evidence for pro-nerve growth factor, a biosynthetic precursor to nerve growth factor. Proc. Nat. Acad. Sci. USA 74, 3647-3651.

Burger, E.A. & Shooter, E.M. (1978). Biosynthesis of beta nerve growth factor in mouse submaxillary glands. J. Biol. Chem. 253, 804-810.

Burton, L.E., Wilson, W.H. & Shooter, E.M. (1978). Nerve growth factor in mouse saliva. Rapid isolation procedures and characterization of 7S nerve growth factor. J. Biol. Chem. 253, 7807-7812.

Caramia, F., Angeletti, P.U. & Levi-Montalcini, R. (1962). Experimental analysis of the mouse submaxillary salivary gland in relationship to its nerve growth factor content. Endocrinology 70, 915-922.

Carbonetto, S. & Stach, R.N. (1982). Localization of nerve growth factor bound to neurons growing nerve fibers in culture. Devel. Biol. 3, 463-473.

Cathala, G., Savouret, J.F., Mendez, B., West, B.L., Karin, M., Martial, J. & Baxter, J.D. (1983). A method for isolation of intact, translationally active RNA. DNA 2, 329-335.

Champlain, J. de, Malmfors, T., Olson, L. & Sachs, C. (1970). Ontogenesis of peripheral adrenergic neurons in the rat: Pre- and postnatal observations. Acta Physiol. Scand. 80, 276-288.

Chao, M., Bothwell, M.A., Ross, A.H., Koprowski, H., Lanahan, A.A., Buck, C.R. & Sehgal, A. (1986). Gene transfer and molecular cloning of the human NGF receptor. Science 232, 518-521.

Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. & Rutter, W.J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry 18, 5294-5299.

Chretien, M. (1977). Action of testosterone on the differentiation and secretory activity of a target organ: the submaxillary gland of the mouse. Int. Rev. Cytol. 50, 333-400.

Cohen, S. (1960). Purification of a nerve growth factor promoting protein from the mouse salivary gland and its neurocytoxic antiserum. Proc. Nat. Acad. Sci. USA 46, 302-311.

Cohen, S. and Levi-Montalcini, R. (1956) A nerve growth stimulated factor isolated from snake venom. Proc. Nat. Acad. Sci. USA 42, 571-574.

Cohen, P., Sutter, A., Landreth, G., Zimmermann, A. & Shooter, E.M. (1980). Oxidation of tryptophan-21 alters the biological activity and receptor binding characteristics of mouse nerve growth factor. J. Biol. Chem. 255, 2949-2954.

Connolly, J.L., Greene, L.A., Viscarello, R.R. & Riley, W.D. (1979). Rapid sequential changes in surface morphology of PC12 pheochromocytoma cells in response to nerve growth factor. J. Cell Biol. 82, 820-827.

Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kedinger, C. & Chambon, P. (1980). Promoter sequences of eucaryotic protein-coding genes. Science 209, 1406-1414.

Coughlin, M.D., Boyer, D.M. & Black, I.B. (1977). Embryologic developments of a mouse sympathetic ganglion in vivo and in vitro. Proc. Nat. Acad. Sci. USA 74, 3438-3442.

Darby, J.K., Feder, J., Selby, M., Riccardi, V., Ferrell, R., Siao, D., Goslin, K., Rutter, W.J., Shooter, E.M. & Cavalli-Sforza, L.L. (1985). A discordant sibship analysis between beta nerve growth factor and neurofibromatosis. Am. J. Hum. Genet. 37, 52-59.

Darling, T.J., Petrides, P.E., Beguin, P., Frey, P., Shooter, E.M., Selby, M. & Rutter, W.J. (1983). The biosynthesis and processing of proteins in the mouse 7S nerve growth factor complex. Cold Spring Harbor Symp. Quant. Biol. 48, 427-434.

Davies, A. & Lumsden, A. (1984). Relation of target encounter and neuronal death to nerve growth factor responsiveness in the developing mouse trigeminal ganglion. J. Comp. Neurol. 223, 124-137.

deChamplain, J., Mamlfors, J., Olson, L. & Sachs, C. (1970). Ontogenesis of peripheral adrenergic neurons in the rat: Pre- and postnatal

observations. Acta Physiol. Scand. 80, 276-288.

Deschryver-Kecskemeti, K., Clouse, R.E., Goldstein, M.N., Gersell, D. & O'Neal, L. (1983). Intestinal ganglioneuromatosis: A manifestation of overproduction of nerve growth factor? New Eng. J. Med. 308, 635-639.

Douglass, J., Civelli, O. & Herbert, E. (1984). Polyprotein gene expression: generation of diversity of neuroendocrine peptides. Ann. Rev. Biochem. 53, 665-715.

Drucker, D.J., Mojsov, S. & Habener, J.F. (1986). Cell specific post-translational processing of preproglucagon expressed from a metallothionein-glucagon fusion gene. J. Biol. Chem. 261, 9637-9643.

Durham, D. & Woolsey, T.A. (1984). Effects of neonatal whisker lesions on mouse central trigeminal lesions. J. Comp. Neurol. 223, 424-447.

Dynan, W.S. & Tjian, R. (1985). Control of eukaryotic mRNA synthesis by sequence-specific DNA-binding proteins. Nature 316, 774-778.

Ebendal, T., Larhammar, D., Persson, H. (1986) Structure and expression of the chicken beta nerve growth factor gene. EMBO J. 5, 1483-1487.

Edwards, R.E., Selby, M. & Rutter, W.J. (1986). Differential RNA splicing predicts two distinct nerve growth factor precursors. Nature 319, 784-787.

Ennis, M., Pearce, F.L. & Vernon, C.A. (1979). Some studies on the mechanism of action of antibodies to nerve growth factor. Neuroscience 4, 1391-1398.

Erickson, A.H. & Blobel, G. (1983). Cell-free translation of mRNA in a wheat germ system. Methods Enzymol. 96, 38-50.

Evans, B.A. & Richards, R.I. (1985). Genes for the alpha and gamma subunit of mouse nerve growth factor are contiguous. EMBO J. 4,

133-138.

Feinstein, S.C., Dana, S.L., McConlogue, L., Schooter, E.M. & Coffino, P. (1985). Nerve growth factor rapidly induces ornithine decarboxylase mRNA in PC12 rat pheochromocytoma cells. Proc. Nat. Acad. Sci. USA 82, 5761-5765.

Frazier, W.A., Hogue-Angeletti, P.A., Sherman, R. & Bradshaw, R.A. (1973). Topography of mouse 2.5S nerve growth factor. Reactivity of tyrosine and tryptophan. Biochemistry 12, 3281-3293.

Freedman, R. (1984). Native disulfide bond formation in protein biosynthesis: evidence for the role of protein disulfide isomerase. Trends Biochem. Sci. 9, 438-441.

Gnahn, H., Hefti, F., Heumann, R., Schwab, M.E. & Thoenen, H. (1983). NGF-mediated increase of choline acetyltransferase (ChAT) in the neonatal rat forebrain: evidence for a physiological role of NGF in the brain? Devel. Brain Res. 9, 45-52.

Goedert, M., Otten, U. & Thoenen, H. (1978). Biochemical effects of antibodies against nerve growth factor on developing and differentiated sympathetic ganglia. Brain Res. 148, 264-268.

Goedert, M., Otten, U., Schafer, T., Schwab, M. & Thoenen, H. (1980). Immunosympathectomy: lack of evidence for a complement-mediated cytotoxic mechanism. Brain Res. 201, 399-409.

Goedert, M., Stoeckel, K. & Otten, U. (1981). Biological importance of the retrograde axonal transport of nerve growth factor in sensory neurons. Proc. Nat. Acad. Sci. USA 78, 5895-5898.

Goldstein, M.N. & Burdman, J.A. (1965). Studies of the nerve growth factor in submandibular glands of female mice treated with testosterone. Anat. Rec. 151, 199-209.

Gorin, P.D. & Johnson, E.M. (1979). Experimental autoimmune model of nerve growth factor deprivation: effects on developing peripheral sympathetic and sensory nerves. Proc. Nat. Acad. Sci. USA 76, 5382-5386.

Green, S.H. & Greene, L.A. (1986). A single Mr~103,000 ^{125}I-β-nerve growth factor-affinity-labeled species represents both the low and high affinity forms of the nerve growth receptor. J. Biol. Chem. 261, 15316-15326.

Greene, L.A., Shooter, E.M. & Varon, S. (1969). Subunit interactions and enzymatic activity of mouse 7S nerve growth factor. Biochemistry 8, 3735-3741.

Greene, L.A. (1977a). A quantitative bioassay for nerve growth factor (NGF) activity employing a clonal pheochromocytoma cell line. Brain Res. 133, 350-353.

Greene, L.A. (1977b). Quantitative in vitro studies on the nerve growth factor (NGF) requirement of neurons. II. Sensory neurons. Devel. Biol. 58, 106-113.

Greene, L.A. & McGuire, J.C. (1978). Induction of ornithine decarboxylase by nerve growth factor dissociated from effects on survival and neurite outgrowth. Nature 276, 191-194.

Greene, L.A., Varon, S., Piltch, A.J. & Shooter, E.M. (1971). Substructure of the beta subunit of the mouse 7S nerve growth factor. Neurobiology 1, 37-48.

Griffiths, G., Pfeiffer, S., Simons, K. & Matlin, K. (1985). Exit of newly synthesized membrane proteins from the trans cisterna of the Golgi complex to the plasma membrane. J. Biol. Chem. 101, 949-964.

Grob, P.M., Ross, A.H., Koprowski, H. & Bothwell, M. (1985). Characteri-

zation of the human melanoma nerve growth factor receptor. J. Biol. Chem. 260, 8044-8049.

Gubler, U & Hoffman, B.J. (1983). A simple and very efficient method for generating cDNA libraries. Gene 25, 263-269.

Gundersen, R.W. & Barrett, J.N. (1980). Characterization of the turning response of dorsal root neurites toward nerve growth factor. J. Cell Biol. 87, 546-554.

Hamburger, V., Bechtold, J.K. & Yip. J.W. (1981). Neuronal death in the spinal ganglia of the chick embryo and its reduction by nerve growth factor. J. Neurosci. 1, 60-71.

Hamburger, V. & Yip, J.W. (1984). Reduction of experimentally induced neuronal death in spinal ganglia of the chick embryo by nerve growth factor. J. Neurosci. 4, 767-774.

Harper, G.P., Barde, Y.A., Burnstock, G., Carstairs, J.R., Dennison, M.E., Suda, K. & Vernon, C.A. (1979). Guinea pig prostate is a rich source of nerve growth factor. Nature 279, 160-162.

Hefti, F. (1986). Nerve growth factor promotes survival of septal cholinergic neurons after fimbrial transections. J. Neurosci. 6, 2155-2162.

Hendry, I.A. & Iversen, L.L. (1973). Reduction in the concentration of nerve growth factor in mice after sialectomy and castration. Nature 243, 500-504.

Hendry, I.A. (1975a). The retrograde trans-synaptic control of the development of cholinergic terminals in sympathetic ganglia. Brain Res. 86, 483-487.

Hendry, I.A. (1975b). The effects of axotomy on the development of the rat superior cervical ganglion. Brain Res. 90, 235-244.

Hendry, I.A. (1975c). The response of adrenergic neurons to axotomy and nerve growth factor. Brain Res. 94, 87-97.

Hendry, I.A. & Campbell, J. (1976). Morphometric analysis of rat superior cervical ganglion after axotomy and nerve growth factor treatment. J. Neurocytol. 5, 351-360.

Hendry, I.A. (1977). The effect of the retrograde axonal transport of nerve growth factor on the morphology of adrenergic neurons. Brain Res. 134, 213-223.

Herrup, K. & Shooter, E.M. (1973). Properties of the nerve growth factor receptor of avian dorsal root ganglia. Proc. Nat. Acad. Sci. USA 70, 3884-3888.

Herrup, K. & Thoenen, H. (1979). Properties of the nerve growth factor receptor of a clonal line of rat pheochromacytoma (PC-12) cells. Exptl. Cell Res. 121, 71-78.

Heumann, R., Schwab, M. & Thoenen, H. (1981). A second messenger required for nerve growth factor biological activity? Nature 292, 838-840.

Hogue-Angeletti, R.A., Frazier, W.A., Jacobs, J.W., Niall, H.D. & Bradshaw, R.A. (1976) Purification, characterization and partial amino acid sequence of nerve growth factor from cobra venom. Biochemistry 15, 26-34.

Hogue-Angeletti, R.A. & Bradshaw, R.A. (1978). Nerve growth factor in snake venoms. In: Handbook of Experimental Pharmacology, vol. 52, C.Y. Lee, ed., Springer, N.Y., pp. 176-294.

Hokfelt, T., Kellerth, J.O., Nilsonn, G. & Pernow, B. (1975). Substance P: Localization in the central nervous system and in some primary sensory neurons. Science 190, 889-890.

Hokfelt, T., Elde, R., Johansson, O., Luft, R., Nilsson, G. & Amimura, A. (1976). Immunohistochemical evidence for separate populations of somatostatin-containing and substance P-containing primary afferent neurons in the rat. Neuroscience 1, 131-136.

Hosang, M. & Shooter, E.M. (1987). The internalization of nerve growth factor by high-affinity receptors on pheochromocytoma PC-12 cells. EMBO J. 6, 1197-1202.

Hubel, D., Weisel, T.N. & LeVay, S. (1977). Plasticity of ocular dominance columns in the monkey striate cortex. Philos. Trans. R. Soc. Lond. (Biol.) 278, 377-409.

Huber, H. & Bode, W. (1978). Structural basis of the activation and action of trypsin. Acc. Chem. Res. 11, 114-121.

Huynh, T., Young, R. & Davis, R.W. (1985) Construction and screening cDNA libraries in λgt10 and λgt11. in Practical Approaches in Biochemistry, D. Glover, ed. (IRL, Oxford) pp. 49-78.

Isackson, P.J. & Bradshaw, R.A. (1984). The α-subunit of 7S nerve growth factor is an inactive serine protease. J. Biol. Chem. 259, 5380-5383.

Isackson, P.J., Ullrich, A. & Bradshaw, R.A. (1984). Mouse 7S nerve growth factor: complete sequence of a cDNA coding for the alpha-subunit and its relationship to serine proteases. Biochemistry 23, 5997-6002.

Ishii, D.N. & Shooter, E.M. (1975). Regulation of nerve growth factor synthesis in mouse submaxillary glands by testosterone. J. Neurochem. 25, 843-851.

Jessell, T., Tsunoo, H., Kanazawa, I. & Otsuka, M. (1979). Substance P: Depletion in the dorsal horn of rat spinal cord after section of

the peripheral processes of primary sensory neurons. Brain Res. 168, 247-259.

Johnson, E.M., Andres, R.Y. & Bradshaw, R.A. (1978). Characterization of the retrograde transport of nerve growth factor (NGF) using high specific activity ^{125}I-NGF. Brain Res. 150, 319-331.

Johnson, E.M., Jr., Gorin, P.D., Brandeis, L.D. & Pearson, J. (1980). Dorsal root ganglion neurons are destroyed by exposure in utero to maternal antibody to nerve growth factor. Science 210, 916-918.

Johnson, E.M., Jr., Taniuchi, M., Clark, H.B., Springer, J.E., Koh, S., Tayrien, M.W. & Loy, R. (1987). Demonstration of the retrograde transport of nerve growth factor receptor in the peripheral and central nervous system. J. Neurosci. 7, 923-929.

Johnson, D., Lanahan, A., Buck, C.R., Sehgal, A., Morgan, C., Mercer, E., Bothwell, M. & Chao, M. (1986). Expression and structure of the human NGF receptor. Cell 47, 545-554.

Kessler, J.A. & Black, I.B. (1980). Nerve growth factor stimulates the development of substance P in sensory ganglia. Proc. Nat. Acad. Sci. USA 77, 649-652.

Kim, S.U., Hogue-Angeletti, A. & Gonatas, N.K. (1979). Localization of nerve growth factor receptors in sympathetic neurons cultured in vitro. Brain Res. 168, 602-608.

Kondo, S., Konoshita, M., Shimizu, A., Saito, Y., Konishi, M., Sabe, H. & Honjo, T. (1987). Expression and functional characterization of artificial mutants of interleukin-2 receptor. Nature 327, 64-67.

Korshing, S. & Thoenen, H. (1983). Nerve growth factor in sympathetic ganglia and corresponding target organs of the rat: correlation with the density of sympathetic innervation. Proc. Nat. Acad. Sci.

USA 80, 3513-3516.

Korshing, S. & Thoenen, H. (1985). Treatment with 6-hydroxydopamine and colchicine decreases nerve growth factor levels in sympathetic ganglia and increases them in the corresponding target tissue. J. Neurosci. 5, 1058-1061.

Korshing, S., Auburger, G., Heumann, R., Scott, J. & Thoenen, H. (1985). Levels of nerve growth factor and its mRNA in the central nervous system of the rat correlate with cholinergic innervation. EMBO J. 4, 1389-1939.

Korshing, S., Heumann, R., Thoenen, H. & Hefti, F. (1986). Cholinergic denervation of the rat hippocampus by fimbrial transection leads to a transient accumulation of nerve growth factor without change in $mRNA^{NGF}$ content. Neurosci. Lett. 66, 175-180.

Kozak, M. (1981). Possible role of flanking sequences in recognition of the AUG initiator codon by eucaryotic ribosome. Nucl. Acids Res. 9, 5233-5252.

Kromer, L. (1987). Nerve growth factor treatment after brain injury prevents neuronal death. Science 235, 214-216.

Kyte, J. & Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157, 105-132.

Landreth, G.E. & Shooter, E.M. (1980). Nerve growth factor receptors on PC-12 cells: Ligand-induced conversion from low-to-high affinity states. Proc. Nat. Acad. Sci. USA 77, 4751-4755.

Large, T., Bodary, S., Clegg, D., Weskamp, G., Otten, U. & Reichardt, L. (1986). Nerve growth factor gene expression in the developing rat brain. Science 234, 352-355.

Leach, D.F. & Stahl, F.W. (1983). Viability of λ phages carrying a

perfect palindrome in the absence of recombination nucleases. Nature 305, 448-451.

Levi, A., Schechter, Y., Neufeld, E.J. & Schlessinger, J. (1980). Mobility, clustering and transport of nerve growth factor in embryonal sensory cells and in a sympathetic neuronal cell line. Proc. Nat. Acad. Sci. USA 77, 3469-3473.

Levi-Montalcini, R. (1948). Implantation of tumors in the hind limb of the embryonic chick and the developmental response of the lumbasacral nervous system. Anat. Rec. 102, 369-390.

Levi-Montalcini, R. (1952). Effects of mouse tumor transplants on the nervous system. Ann. N.Y. Acad. Sci. 55, 330-343.

Levi-Montalcini, R. (1966). The nerve growth factor: its mode of action on sensory and sympathetic nerve cells. Harvey Lect. 60, 217-259.

Levi-Montalcini, R., Aloe, L., Mugnaini, E., Oesch, F. & Thoenen, H. (1975). Nerve growth factor induces volume increase and enhances tyrosine hydroxylase synthesis in chemically axotomized sympathetic ganglia of newborn rats. Proc. Nat. Acad. Sci. USA 72, 595-599.

Levi-Montalcini, R. & Angeletti, P.U. (1963). Essential role of the nerve growth factor on the survival and maintenance of dissociated sensory and sympathetic embryonic nerve cells in vitro. Devel. Biol. 7, 653-659.

Levi-Montalcini, R. & Angeletti, P.U. (1966). Nerve growth factor. Pharmacol. Rev. 18, 619-628.

Levi-Montalcini, R. & Angeletti, P.U. (1968). Biological aspects of the nerve growth factor. In: Ciba Foundation Symp. Growth of the Nervous System, ed. E.N. Wolstenholme & M. O'Connor, Churchill, London, p. 126-147.

Levi-Montalcini, R. & Booker, B. (1960). Destruction of the sympathetic ganglia in mammals by an antiserum to the nerve growth factor protein. Proc. Nat. Acad. Sci. USA 46, 384-390.

Levi-Montalcini, R. & Cohen, S. (1956). In vitro and in vivo effects of a nerve growth-stimulating agent isolated from snake venom. Proc. Nat. Acad. Sci. USA 42, 695-699.

Levi-Montalcini, R. & Hamburger, V. (1951). Selective growth-stimulating effects of mouse sarcoma on the sensory and sympathetic nervous system of the chick embryo. J. exptl Zool. 116, 321-361.

Levi-Montalcini, R. & Hamburger, V. (1953). A diffusible agent of mouse sarcoma producing hyperplasia of sympathetic ganglia and hyperneurotization of viscis in the chick embryo. J. Exptl Zool. 123, 233-278.

Levi-Montalcini, R., Meyer, R.H. & Hamburger, V. (1954) In vitro experiments on the effects of mouse sarcomas 180 and 37 on the spinal and sympathetic ganglia of the chick embryo. Cancer Res. 14, 49-57.

Lindsay, R.M. (1979). Adult rat brain astrocytes support survival of both NGF-dependent and NGF-insensitive neurones. Nature 282, 80-82.

Lingappa, V.R., Lingappa, J.R. & Blobel, G. (1979). Chicken ovalbumin contains an internal signal sequence. Nature 281, 117-121.

Lumsden, A.G.S. & Davies, A.M. (1983). Earliest sensory nerve fibers are guided to peripheral targets by attractants other than nerve growth factor. Nature 306, 786-788.

Lumsden, A.G.S. & Davies, A.M. (1986). Chemotropic effect of specific target epithelium in the developing mammalian nervous system. Nature 323, 538-539.

Maniatis, T., Fritsch, E.F. & Sambrook, J. (1982). Molecular Cloning, a

Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Martinez, H., Dreyfus, C.F., Jonakait, G.M. & Black, I.B. (1985). Nerve growth factor promotes cholinergic development in brain striatal cultures. Proc. Nat. Acad. Sci. USA 82, 7777-7781.

Maxam, A. & Gilbert. W. (1980). Sequencing end-labeled DNA with base specific chemical cleavages. Methods in Enzymol. 65, 499-560.

Meier, R., Becker-Andre, M., Gotz, R. Neunamy, R., Shaw A. & Thoenen, H. (1986). Molecular cloning of bovine and chick nerve growth factor (NGF): delineation of conserved and unconserved domains and their relationship to the biological activity and antigenicity of NGF. EMBO J. 5, 1489-1493.

McClelland, A., Kuhn, L.C. & Ruddle, F.H. (1984). The human transferrin receptor gene: Genomic organization and the complete primary structure of the receptor deduced from a cDNA sequence. Cell 39, 267-274.

Menesini-Chen, M.G., Chen, J.S. & Levi-Montalcini, R. (1978). Sympathetic nerve fibers ingrowth in the central nervous system of neonatal rodent upon intracerebral NGF injections. Arch. Ital. Biol. 116, 53-84.

Milbrandt, J. (1986). Nerve growth factor rapidly induces c-fos mRNA in PC12 rat pheochromocytoma cells. Proc. Nat. Acad. Sci. USA 83, 4789-4793.

Miller, R.J., Chang, K.J., Leighton, J. & Cuatrecasas, P. (1978). Interaction of iodinated enkephalin analogues with opiate receptors. Life Sci. 22, 379-388.

Mobley, W.C., Rutkowski, J.L., Tennekoon, G.I., Buchanan, K. & Johnson,

M.O. (1986). Choline acetyltransferase activity in striatum of neonatal rats increased by nerve growth factor. Science 229, 284-287.

Mount, S.M. (1982). A catalogue of splice junctions sequences. Nucl. Acids Res. 10, 459-472.

Mowry, M.D., Jensen, R.J. & Pantazis, N.J. (1984). Immunocytochemical localization and concentrations of the alpha and gamma subunits of 7S nerve growth factor in the submandibular gland of the mouse. Cell Tissue Res. 238, 627-633.

Myers, R.M., Tilly, K. & Maniatis, T. (1986). Fine structure genetic analysis of a beta-globin promoter. Science 232, 613-618. Science 232, 613-618.

Nader, W.F., Edlind, T.D., Huettermann, A. & Sauer, H.W. (1985). Cloning of Physarum actin sequences in an exonuclease-deficient bacterial host. Proc. Nat. Acad. Sci. USA 82, 2698-2702.

O'Farrell, P.H. (1975). High resolution two-dimensional electrophoresis of proteins. J. Biol. Chem. 250, 4007-4021.

Olson, L. (1967). Outgrowth of sympathetic adrenergic neurons in mice treated with a nerve growth factor (NGF). Z. Zelforsch. 81, 155-173.

Otten, U., Goedert, M., Mayer, N. & Lembeck, F. (1980). Requirement of nerve growth factor for development of substance P-containing sensory neurons. Nature 287, 158-159.

Otten, U., Lorez, H.P. & Businger, E.F. (1983). Nerve growth factor antagonizes the neurocytoxic action of capsaicin on primary sensory neurones. Nature 301, 515-517.

Padgett, Z.A., Grabowski, P.J., Konarska, M.M., Seiler, S. & Sharp, P.A.

(1986). Splicing and mRNA precursors. Ann. Rev. Biochem. 55, 1119-1150.

Pantazis, N. (1983). Nerve growth factor synthesized in mouse fibroblast cells in culture: absence of α and γ subunits. Biochemistry 22, 4264-4271.

Pearson, J., Johnson, E.M. & Brandeis, L. (1983). Effects of antibodies to nerve growth factor on intrauterine development of derivatives of cranial neural crest and placode in the guinea pig. Devel. Biol. 96, 32-36.

Perara, E. & Lingappa, V.R. (1985). A former amino-terminal signal sequence engineered to an internal location directs translocation of both flanking protein domains. J. Cell Biol. 101, 2292-2301.

Perlman, D. & Halvorson, H.O. (1983). Putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J. Mol. Biol. 167, 391-409.

Petruzzelli, L., Herrera, R., Arenas-Garcia, R., Fernandez, R., Birnbaum, M.J. & Rosen, O.M. (1986). Isolation of a Drosophila genome sequence homologous to the kinase domain of the human insulin receptor and detection of the phoshporylated Drosophila receptor with an anti-peptide antibody. Proc. Nat. Acad. Sci. USA 83, 4710-4714.

Radeke, M.J., Misko, T.P., Hsu, C., Hezenberg, L.A. & Shooter, E.M. (1987). Gene transfer and molecular cloning of the rat nerve growth factor receptor. Nature 325, 593-597.

Reynolds, G.A., Goldstein, J.L. & Brown, M.S. (1985). Multiple mRNAs for 3-hydroxy-3-methylglutaryl coenzyme A reductase determined by multiple transcription initiation sites and intron splicing sites

in the 5' untranslated region. J. Biol. Chem. 260, 10369-10377.

Richardson, P. & Riopelle, R.J. (1984). Uptake of nerve growth factor along peripheral and spinal axons of primary sensory neurons. J. Neurosci. 4, 1683-1689.

Ross, A.H., Grob, P., Bothwell, M., Elder, D.E., Ernst, C.S., Marano, N., Ghrist, B.F.D., Slamp, C.C., Herlyn, M., Atkinson, B. & Koprowski, H. (1984). Characterization of nerve growth factor receptor in neural crest tumors using monoclonal antibodies. Proc. Nat. Acad. Sci. USA 81, 6681-6685.

Rush, R.A. (1984). Immunohistochemical localization of endogenous nerve growth factor. Nature 312, 364-367.

Sabatini, D.D., Kreibich, G., Morimoto, T. & Adesnik, M. (1982). Mechanisms for the incorporation of proteins in membranes and organelles. J. Cell Biol. 92, 1-22.

Sanger, F., Nicklen, S. & Coulson, A.R. (1977). DNA sequencing with chain terminating inhibitors. Proc. Nat. Acad. Sci. USA 74, 5463-5467.

Schechter, A.L. & Bothwell, M.A. (1981). Nerve growth factor receptors on PC-12 cells: Evidence for two receptor classes with differing cytoskeletal association. Cell 24, 867-874.

Scheele, G. and Jacoby, R. (1982). Conformational changes associated with proteolytic processing of presecretory proteins allow glutathionine- catalyzed formation of native disulfide bonds. J. Biol. Chem. 257, 12277-12282.

Schwab, M.E., Stockel, K. & Thoenen, H. (1976). Immunocytochemical localization of nerve growth factor in the submandibular gland of adult mice by light and electron microscopy. Cell Tissue Res. 169,

289-299.

Schenkein, I., Bueker, E.D., Helson, L., Axelrod, F. & Dancis, J. (1974). Increased nerve growth stimulating activity in disseminated neurofibromatosis. New Eng. J. Med. 290, 613-614.

Schwab, M.E., Otten, U.,Agidi, Y. & Thoenen, H. (1979). Nerve growth factor (NGF) in the rat CNS: Absence of specific retrograde axonal transport and tyrosine hydroxylase induction in locus coeruleus and substantia nigra. Brain Res. 168, 473-483.

Schwartz, J.P. & Breakfield, X.O. (1980). Altered nerve growth factor in fibroblasts from patients with familial dysautonomia. Proc. Nat. Acad. Sci. USA 77, 1154-1158.

Schwartz, J.P., Pearson, J. & Johnson, E.M. (1982). Effect of exposure to anti-NGF on sensory neurons of adult rats and guinea pigs. Brain Res. 244, 378-381.

Scott, J., Selby, M., Urdea, M., Quiroga, M., Bell, G.I. & Rutter, W.J. (1983). Isolation and nucleotide sequence of a cDNA encoding the precursor of mouse nerve growth factor. Nature 302, 538-540.

Seeley, J.P., Charles, K.H., Shelanski, M.L. & Greene, L.A. (1983). Pressure microinjection of nerve growth factor and anti-nerve growth factor into the nucleus and cytoplasm: Lack of effects on neurite outgrowth from pheochromocytoma cells. J. Neurosci. 3, 1488-1494.

Seiler, M. & Schwab, M.E. (1984). Specific retrograde transport of nerve growth factor from neocortex to nucleus basalis in the rat. Brain Res. 300, 33-39.

Selby, M., Barta, A., Baxter, J.D., Bell, G.I. & Eberhardt, N.L. (1984). Analysis of a major human chorionic somatomammotropin gene:

Evidence for two functional promoter elements. J. Biol. Chem. 259, 13131-13138.

Server, A.C., Herup, K., Shooter, E.M., Hogue-Angeletti, R.A., Frazer, W.A. & Bradshaw, R.A. (1976). Comparison of the nerve growth factor proteins from cobra venom (Naja naja) and mouse submaxillary gland. Biochemistry 15, 35-39.

Server, A.C. & Shooter, E.M. (1977). Nerve growth factor. Adv. Prot. Chem. 31, 339-409.

Shelton, D.L. & Reichardt, L.F. (1984). Expression of the beta nerve growth factor gene correlates with the density of sympathetic innervation in effector organs. Proc. Nat. Acad. Sci. USA 81, 7951-7955.

Shelton, D.L. & Reichardt, L.F. (1984). Expression of the beta nerve growth factor gene correlates with the density of sympathetic innervation in effector organs. Proc. Nat. Acad. Sci. USA 81, 7951-7955.

Shelton, D.L. & Reichardt, L.F. (1986a). Studies on the expression of the β nerve growth factor (NGF) gene in the central nervous system: level and regional distribution of NGF mRNA suggest that NGF functions as a trophic factor for several distinct populations of neurons. Proc. Nat. Acad. Sci. USA 83, 2714-2718.

Shelton, D.L. & Reichardt, L.F. (1986b). Studies on the regulation of beta nerve growth factor gene expression in the rat iris: the level of mRNA encoding nerve growth factor is increased in irises placed in explant cultures in vitro, but not in irises deprived of sensory or sympathetic innervation in vivo. J. Cell Biol. 102, 1940-1948.

Shine, J., Mason, A.J., Evans, E.A. & Richards, R.I. (1983). The kalli-

krein multigene family: Specific processing of biologically active peptides. Cold Spring Harbor Symp. Quant. Biol. 48, 417-426.

Siggers, D.C., Rogers, J.G., Boyer, S.H., Margolet, L., Dorkin, H., Banerjee, S.P. & Shooter, E.M. (1976). Increased nerve growth factor beta-chain cross-reacting material in familial dysautonomia. New Eng. J. Med. 12, 629-634.

Simson, J.A.V., Hazen, D., Spicer, S.S., Murphy, R.A. & Young, M. (1978). Secretagogue-mediated discharge of nerve growth factor from granular tubules of male mouse submandibular glands: An immunocytochemical study. Anat. Rec. 192, 375-388.

Stockel, K., Schwab, M.E. & Thoenen, H. (1975). Comparison between the retrograde axonal transport of nerve growth factor and tetanus toxin in motor, sensory and adrenergic neurons. Brain Res. 99, 1-16.

Stoeckel, K., Schwab, M. & Thoenen, H. (1975a). Specificity of retrograde transport of nerve growth factor (NGF) in sensory neurons: A biochemical and morphological study. Brain Res. 89, 1-14.

Stoeckel, K., Schwab, M. & Thoenen, H. (1975b). Comparison between the retrograde axonal transport of nerve growth factor and tetanus toxin in motor, sensory and adrenergic neurons. Brain Res. 99, 1-16.

Sutter, A., Riopelle, R.J., Harris-Warrick, R. & Shooter, E.M. (1979). Nerve growth factor receptors. Characterization of two distinct classes of binding sites on chick embryo sensory ganglia cells. J. Biol. Chem. 254, 5972-5982.

Taniuchi, M., Schweitzer, J.B. & Johnson, E.M., Jr. (1986). Nerve growth factor receptor molecules in rat brain. Proc. Nat. Acad. Sci. USA

83, 1950-1954.

Taniuchi, M., Clark, H.B. & Johnson, E.M., Jr. (1986). Induction of nerve growth factor receptor in Schwann cells after axotomy. Proc. Nat. Acad. Sci. USA 83, 4094-4098.

Thoenen, H., Angeletti, P.V., Levi-Montalcini, R. & Kettler, R. (1971). Selective induction of tyrosine hydroxylase and dopamine β-hydroxylase in rat superior cervical ganglia by nerve growth factor. Proc. Nat. Acad. Sci USA 68, 1598-1602.

Thoenen, H. & Barde, Y.A. (1980). Physiology of nerve growth factor. Physiol. Rev. 60, 1284-1335.

Ullrich, A., Gray, A., Berman, C. & Dull, T.J. (1983). Human beta nerve growth factor gene sequence highly homologous to that of mouse. Nature 303, 821-825.

Valerio, D., Doyvesteyn, M.G.C., Dekker, B.M.M., Van Des Voorn, L., van Ormondt, H. & Van der Eb, A.J. (1985). Adenosine deaminase characterization and expression of a gene with a remarkable promoter. EMBO J. 4, 437-443.

Varon, S., Nomura, J. & Shooter, E.M. (1967). Subunit structure of a high molecular weight form of the nerve growth factor from mouse submaxillary gland. Proc. Nat. Acad. Sci. USA 57, 1782-1789.

von Heijne, G. (1986) A new method for predicting signal sequence cleavage sites. Nucl. Acids Res. 14, 4683-4690

Walter, M.D., Edlund, T., Boulet, A. & Rutter, W.J. (1983). Cell-specific expression controlled by the 5' flanking region of insulin and chymotrypsin genes. Nature 306, 557-561.

Walter, P. & Blobel, G. (1983). Preparation of microsomal membranes for cotranslational protein translocation. Methods Enzymol. 96, 84-93.

Wekstein, D.R. (1965). Heart rate of the preweaning rat and its autonomic control. Am. J. Physiol. 208, 1259-1262.

Whittemore, S.R., Ebendal, T., Larkfors, L., Olson, L., Seiger, A., Stromberg, I. & Persson, H. (1986). Developmental and regional expression of β NGF mRNA and protein in the rat central nervous system. Proc. Nat. Acad. Sci. USA 83, 817-821.

Whittemore, S.R., Larkfors, L., Ebendal, T., Holets, V.R., Ericsson, A. & Persson, H. (1987). Increased β nerve growth factor mRNA and protein levels in neonatal rat hippocampus following specific cholinergic lesions. J. Neurosci. 7, 244-251.

Wilson, W., Gilmore, R. & Morrison, T. (1987). Translation and membrane insertion of the hemagglutinin-neuroaminidase glycoprotein of Newcastle Disease Virus. Mol. Cell. Biol. 7, 1386-1391.

Wion, D., Perret, C., Frechin, N., Keller, A., Behar, G., Brachet, P. & Auffray, C. (1986) Molecular cloning of the avian β nerve growth factor gene: transcription in brain. FEBS Letts 14, 82-86.

Wise, S.P. & Jones, E.G. (1978). Developmental studies of thalamocortical and commissural connections in the rat somatic sensory cortex. J. Comp. Neur. 178, 187-208.

Wood, W.I., Gitschier, J., Lasky, L. & Lawn, R.M. (1985). Base composition-independent hybridization in tetramethylammonium chloride: A method for oligonucleotide screening of highly complex gene libraries. Proc. Nat. Acad. Sci. USA 82, 1585-1588.

Woolsey, T.A., Durham, D., Harris, R.M., Simons, D.J. & Valentine, K.C. (1981). Somatosensory development. In: Development of Perception, Vol. I, Academic Press, New York, pp. 259-292.

Wyman, A.R., Wolfe, L.B. & Botstein, D. (1985). Propagation of some

human DNA sequences in bacteriophage λ vectors requires mutant E. coli hosts. Proc. Nat. Acad. Sci. USA 82, 2886-2884.

Yip, H.K. & Johnson, E.M., Jr. (1984). Developing dorsal root ganglion neurons require trophic support from their control processes: Evidence for a role of retrogradely transported nerve growth factor from the central nervous system to the periphery. Proc. Nat. Acad. Sci. USA 81, 6245-6249.

Yip, H.K., Rich, K.M., Lampe, P.A. & Johnson, E.M., Jr. (1984). The effects of nerve growth factor and its antiserum on the postnatal development and survival after injury of sensory neurons in rat dorsal root ganglia. J. Neurosci. 4, 2986-2992.

CHAPTER X

THE SEQUENCE OF THE MOUSE NGF GENE

The nucleotide sequence of the mouse NGF gene is presented in 5' to 3' direction. All the available restriction sites are presented below the double-stranded sequence. Exons are boxed by either complete lines or broken lines, representing coding and non-coding regions, respectively. Initiator ATG's and the TGA terminator are doubly underlined. Dinucleotide splice donor (GT) and acceptor (AG) splicing sequences are overlined and underlined, respectively. Those regions not sequenced but mapped by restriction digestions of phage and genomic DNAs are represented by broken lines; the sizes of the introns are estimated excluding the sequenced portions.

```
  1  AGATCTAGAAACATGCAGGGAGCACATACAGACACAGTTCCATCTACAGTTTCCAAGGGA
     TCTAGATCTTTGTACGTCCCTCGTGTATGTCTGTGTCAAGGTAGATGTCAAAGGTTCCCT
     ^^ ^^     ^^       ^                                   ^   ^^

     1 BglII XhoII, 2 MboI, 4 XbaI, 5 MaeI, 11 Nsp7524I, 12 NlaII
     I, 20 Bsp1286 HgiAI, 53 StyI, 57 NlaIV, 58 AvaII Sau96I,

 61  CCAGCTCTCCTGGGAGGTTTTCGCAATGCCAGTGCTAAGGCTGGCAGTGGGAGTGGGAGT
     GGTCGAGAGGACCCTCCAAAAGCGTTACGGTCACGATTCCGACCGTCACCCTCACCCTCA
       ^     ^    ^                  ^

     63 AluI, 69 EcorII ScrFI, 74 MnlI, 95 DdeI,

121  TGGGGTAATCTGAACACTAAGGTCCTCTTCCTTTGACCTCAGGCAGATATACCTCTGTCT
     ACCCCATTAGACTTGTGATTCCAGGAGAAGGAAACTGGAGTCCGTCTATATGGAGACAGA
                      ^  ^^  ^ ^          ^^            ^

     137 DdeI, 140 Eco0109 PpuMI, 141 AvaII Sau96I, 144 MnlI, 146
      MboII, 157 MnlI, 158 DdeI, 172 MnlI,

181  TGACATTCTAGTTAATTCTGCTTCTGCCTGGTTTCAGAGCAACCCCTGGCACTGGCCAAT
     ACTGTAAGATCAATTAAGACGAAGACGGACCAAAGTCTCGTTGGGGACCGTGACCGGTTA
           ^                    ^                   ^       ^^

     188 MaeI, 207 EcorII ScrFI, 225 EcorII ScrFI, 233 BalI EaeI
     HaeI, 234 HaeIII,

241  CCAGGAGTTTCCGGTGTTGTGGGTGTTCTCAGTGTTCCTTCAGCTCTCAGATTTGCAGAG
     GGTCCTCAAAGGCCACAACACCCACAAGAGTCACAAGGAAGTCGAGAGTCTAAACGTCTC
     ^          ^                    ^                ^   ^

     241 EcorII ScrFI, 251 HpaII, 268 DdeI, 282 AluI, 286 DdeI,

301  ATAGCAACACCTGATATCCTGGCCTCCAAAGAAGACTAAGTCACCCCAGGAGGCTAGCAA
     TATCGTTGTGGACTATAGGACCGGAGGTTTCTTCTGATTCAGTGGGGTCCTCCGATCGTT
                 ^    ^ ^^ ^       ^  ^ ^   ^^    ^   ^  ^^

     313 EcoRV, 318 EcorII ScrFI, 320 HaeI, 321 HaeIII, 323 MnlI,
      331 MboII, 334 TthIII1, 336 DdeI, 340 MaeIII, 341 HphI, 346
      EcorII ScrFI, 350 MnlI, 353 NheI, 354 MaeI,

361  GGCAGTGAGCCAGGGGAAGAGGTAAAGGGACTCCAACTCAGCCAATCAACTCAGCCCTGC
     CCGTCACTCGGTCCCCTTCTCCATTTCCCTGAGGTTGAGTCGGTTAGTTGAGTCGGGACG
           ^     ^  ^          ^         ^              ^       ^

     370 EcorII ScrFI, 376 MboII, 379 MnlI, 389 HinfI, 397 DdeI,
     410 DdeI, 420 DdeI,

421  TCAGCCTGTCTGCCCATTGGTGTTTGGAGCCAGCTACCTTGCCTGGTGCCCTAGAAAATC
     AGTCGGACAGACGGGTAACCACAAACCTCGGTCGATGGAACGGACCACGGGATCTTTTAG
                        ^    ^     ^   ^     ^  ^^    ^

     441 TthIII2, 446 NlaIV, 452 AluI, 456 HgiEII, 462 EcorII Scr
     FI, 465 BanI NlaIV, 466 Bsp1286, 471 MaeI,

481  TCTCTCTGCATCTGTGACCTCCCCCCACCATGCATTTTCATCTAAACCAGGGGTTTGAAT
     AGAGAGACGTAGACACTGGAGGGGGGTGGTACGTAAAAGTAGATTTGGTCCCCAAACTTA
           ^     ^   ^          ^^                  ^         ^

     488 SfaNI, 494 MaeIII, 498 MnlI, 509 NlaIII, 510 NsiI, 527 E
     corII ScrFI, 537 EcorI,
```

```
541  TCTTCCCTGACTGGCTTTTCCTGGCTATGTCCCATCAACTCGGGAGCTATCCATCCCTTG
     AGAAGGGACTGACCGAAAAGGACCGATACAGGGTAGTTGAGCCCTCGATAGGTAGGGAAC
     ^                ^                   ^     ^     ^

     541 MboII, 560 EcorII ScrFI, 579 AvaI, 585 AluI, 592 FokI,

601  TCCCCAGGACCTTACAACCCGGACCCCTGGGTCTAGTCACAGCAGGTGCGGGCTGGGATT
     AGGGGTCCTGGAATGTTGGGCCTGGGGACCCAGATCAGTGTCGTCCACGCCCGACCCTAA
       ^ ^^          ^^ ^    ^      ^  ^

     604 EcorII ScrFI, 606 Eco0109 PpuMI, 607 AvaII Sau96I, 618 N
     ciI ScrFI, 619 HpaII, 621 AvaII NlaIV Sau96I, 626 EcorII Scr
     FI, 633 MaeI, 636 MaeIII,

661  GGAGTTGGCCAGAGAGGGAGGGGTCGGGTGAGTGGGGGGCAGGATTTGGAGAGGGTGTGA
     CCTCAACCGGTCTCTCCCTCCCCAGCCCACTCACCCCCCGTCCTAAACCTCTCCCACACT
        ^^      ^   ^       ^                          ^     ^

     666 BalI EaeI HaeI, 667 HaeIII, 674 MnlI, 678 MnlI, 687 HphI
     , 711 MnlI, 717 MaeIII,

721  CGAGCCTGGAGGAGGGGCTAAATACAGTCAGGAAGCCTGTTAAAGAAGCTCTGTGCTCCA
     GCTCGGACCTCCTCCCCGATTTATGTCAGTCCTTCGGACAATTTCTTCGAGACACGAGGT
       ^   ^  ^                                  ^     ^

     725 EcorII ScrFI, 729 MnlI, 732 MnlI, 767 AluI, 773 Bsp1286
     HgiAI,

781  GCACGGCAGAGAGCGCCTGGAGCCGGAGGGGAGCGCATCGGTGAGTCAGGCTTCTCTGAG
     CGTGCCGTCTCTCGCGGACCTCGGCCTCCCCTCGCGTAGCCACTCAGTCCGAAGAGACTC
                ^^  ^  ^   ^  ^      ^ ^    ^  ^            ^

     792 HaeII, 793 HhaI, 796 EcorII ScrFI, 799 NlaIV, 803 HpaII,
      806 MnlI, 813 HhaI, 815 SfaNI, 820 HphI, 823 HinfI, 836 Dde
     I,

841  CCGAGCCCGATACGGGTGCAAGCCGGTGCGGCGGACTGGGGTTGCCGGGCTCCCTGGGAC
     GGCTCGGGCTATGCCCACGTTCGGCCACGCCGCCTGACCCCAACGGCCCGAGGGACCCTG
       ^                ^^    ^                ^ ^^    ^

     843 BanII Bsp1286, 862 BglI, 863 HpaII, 868 Fnu4HI, 885 HpaI
     I NciI ScrFI, 887 BanII Bsp1286, 888 NlaIV, 893 EcorII ScrFI
     ,

901  GGTGTGCTTGGATCCCTGTCTAGCTA----------~32KB----------------
     CCACACGAACCTAGGGACAGATCGAT----------~32KB----------------
         ^    ^^        ^ ^

     905 TthIII2, 910 BamHI BinI NlaIV XhoII, 911 BinI MboI, 920
     MaeI, 922 AluI,

961  ------------------------------------GATCCCCTTTGGCACTCAAACA
     ------------------------------------CTAGGGGAAACCGTGAGTTTGT
                                         ^               ^

     999 BinI MboI, 1015 TthIII2,

1021 CAGATTATAGATATATCATTAGGCATTACCAAGCCAATCAGCTGTTCTATATGGATTCCT
     GTCTAATATCTATATAGTAATCCGTAATGGTTCGGTTAGTCGACAAGATATACCTAAGGA
                                         ^^              ^

     1059 NspBII PvuII, 1060 AluI, 1074 HinfI,
```

1081 ATTTAGCAATACATTCTTGTAAAGCTGTTCTTCCCTGAACTCTGTGTCATGGAAGGCTCT
TAAATCGTTATGTAAGAACATTTCGACAAGAAGGGACTTGAGACACAGTACCTTCCGAGA

1103 AluI, 1109 MboII, 1128 NlaIII,

1141 TTGCCTTTAGATAGCTTATCTACAGCTCTAGTGTCCTTGTAATTAGCCGTACTTTGAAAG
AACGGAAATCTATCGAATAGATGTCGAGATCACAGGAACATTAATCGGCATGAAACTTTC

1153 AluI, 1164 AluI, 1168 MaeI, 1189 RsaI,

1201 CCTCTCTGTGGCATTGCACATTAGAGAGCTCATGAATTACTGCAGTCAGGTAAACACCAT
GGAGAGACACCGTAACGTGTAATCTCTCGAGTACTTAATGACGTCAGTCCATTTGTGGTA

1201 MnlI, 1226 BanII Bsp1286 HgiAI SacI, 1227 AluI, 1231 Nl
aIII, 1240 PstI, 1258 NlaIII,

1261 GCTGGCTGTGGCCAGCTGCACACCATAGAAGGCCACCGACAGTTACAGCAAATGCTCTTT
CGACCGACACCGGTCGACGTGTGGTATCTTCCGGTGGCTGTCAATGTCGTTTACGAGAAA

1269 BalI EaeI HaeI, 1270 HaeIII, 1273 NspBII PvuII, 1274 Al
uI, 1275 BbvI Fnu4HI, 1290 HaeI, 1291 HaeIII, 1302 MaeIII,

1321 CATTCCGAAGCTGAAGGGCAGCCCAACCACAATGGCTTCTCCTTTGATCTGCTTGGCTCT
GTAAGGCTTCGACTTCCCGTCGGGTTGGTGTTACCGAAGAGGAAACTAGACGAACCGAGA

1329 AluI, 1338 BbvI Fnu4HI, 1366 MboI, 1370 TthIII2,

1381 GGGAAAATGCCGCTGAGAGAGCATGTGTAATTTACGTCAACATCTCCCAACGAGCCATAG
CCCTTTTACGGCGACTCTCTCGTACACATTAAATGCAGTTGTAGAGGGTTGCTCGGTATC

1389 Fnu4HI, 1390 NspBII, 1393 DdeI, 1401 Nsp7524I, 1402 Nla
III, 1414 MaeII, 1416 HincII, 1440 AvaII Sau96I,

1441 GACCGATTCTCTTTAAGAAATGTGACAAAACCCTTTAGAATTGTCAGCATCGGGCCGTTT
CTGGCTAAGAGAAATTCTTTACACTGTTTTGGGAAATCTTAACAGTCGTAGCCCGGCAAA

1445 HinfI, 1462 MaeIII, 1487 SfaNI, 1492 Sau96I, 1493 HaeII
I,

1501 GGTGGAAAAAGAATAAGTTTTACTGAACCTTTATGTCTTCCTAAATCCTTGATGCCAGAT
CCACCTTTTTCTTATTCAAAATGACTTGGAAATACAGAAGGATTTAGGAACTACGGTCTA

1536 MboII, 1551 SfaNI, 1558 SfaNI,

1561 GCACTATTTTCCAACTGGATCTGCCAAAAACACTAAATAGTTAAAAGACATCAAATCGTC
CGTGATAAAAGGTTGACCTAGACGGTTTTTGTGATTTATCAATTTTCTGTAGTTTAGCAG

1577 BinI XhoII, 1578 MboI,

1621 TCCAAGTGGCCCAAACTGAAATTTGTTTAAAAGAAAAAGAACAACAAAAGAACTCTTAGT
AGGTTCACCGGGTTTGACTTTAAACAAATTTTCTTTTTCTTGTTGTTTTCTTGAGAATCA

1628 HaeIII Sau96I, 1646 AhaIII, 1675 DdeI,

```
1681  CAGCAAACCTTACACTCATATTTCTTTCTAGAAAATGACTGATTAGGCAAAGAATTAGAA
      GTCGTTTGGAATGTGAGTATAAAGAAAGATCTTTTACTGACTAATCCGTTTCTTAATCTT
                                  ^^

      1707 XbaI, 1708 MaeI,

1741  AAACAGATTGACAGCATAAAAGACCTCCAGCTCTTCTCCCTTTGTACACAGGCACAGCGT
      TTTGTCTAACTGTCGTATTTTCTGGAGGTCGAGAAGAGGGAAACATGTGTCCGTGTCGCA
                               ^   ^  ^          ^               ^

      1764 MnlI, 1769 AluI, 1772 MboII, 1784 RsaI, 1799 RsaI,

1801  ACGGAATGACAGACAGGGGGTAAAATTAAGAATGGCTATAGCAATGGAAATCACAATAGA
      TGCCTTACTGTCTGTCCCCCATTTTAATTCTTACCGATATCGTTACCTTTAGTGTTATCT

1861  GCAGTACCAAAATTCACTATACCCCAAGGACATGATCTCAATGCTAGACCTGAAATGAGA
      CGTCATGGTTTTAAGTGATATGGGGTTCCTGTACTAGAGTTACGATCTGGACTTTACTCT
        ^                       ^     ^  ^          ^

      1864 RsaI, 1884 StyI, 1891 NlaIII, 1894 MboI, 1904 MaeI,

1921  TTTCCTCACCCCAGAACTCCACCAGGTAGGAAGAGGGACATAGCTTCCTGGAATCTTTGG
      AAAGGAGTGGGGTCTTGAGGTGGTCCATCCTTCTCCCTGTATCGAAGGACCTTAGAAACC
        ^ ^                 ^       ^  ^          ^    ^   ^

      1924 MnlI, 1926 HphI, 1942 EcorII ScrFI, 1950 MboII, 1953 Mn
      lI, 1962 AluI, 1967 EcorII ScrFI, 1971 HinfI,

1981  ACAATGTAAGTTGCTTTAGGGACTGGCTCGAGCTTCATTGGAGCACGTGACATGTAAGAG
      TGTTACATTCAACGAAATCCCTGACCGAGCTCGAAGTAACCTCGTGCACTGTACATTCTC
                                ^^  ^          ^   ^ ^  ^^      ^

      2007 AvaI XhoI, 2008 TaqI, 2011 AluI, 2021 Bsp1286 HgiAI, 20
      25 MaeII, 2027 MaeIII, 2030 AflIII Nsp7524I, 2031 NlaIII, 20
      38 MnlI,

2041  GTTCTTTAAGTCAACCGTCCCCAAATCTTTGAAGAAAAATAAAGTCATCAAATATTCCTG
      CAAGAAATTCAGTTGGCAGGGGTTTAGAAACTTCTTTTTATTTCAGTAGTTTATAAGGAC
            ^                   ^                   ^

      2050 HincII, 2071 MboII, 2091 SspI,

2101  AACTGTTTAGTATCAAGCCACCATCACTCTAACATTGAATAATTTTTTTAAAGTGGCCTA
      TTGACAAATCATAGTTCGGTGGTAGTGAGATTGTAACTTATTAAAAAAATTTCACCGGAT
                                                  ^     ^^

      2147 AhaIII, 2154 HaeI, 2155 HaeIII,

2161  CAAAGTCAAAAAGAAAATCTCCTATTCACAGATGTGCGTGTATCTAGAGAGTGAGGCGTT
      GTTTCAGTTTTTCTTTTAGAGGATAAGTGTCTACACGCACATAGATCTCTCACTCCGCAA
                                                 ^^       ^     ^^

      2203 XbaI, 2204 MaeI, 2213 MnlI, 2219 AsuII, 2220 TaqI,

2221  CGAAATACAATTCCCCCCACCCCCGATATCGTCCGTGTTTTTTCATACGATGATTCGACC
      GCTTTATGTTAAGGGGGGTGGGGGCTATAGCAGGCACAAAAAAGTATGCTACTAAGCTGG
                                ^                       ^  ^

      2245 EcoRV, 2272 HinfI, 2275 TaqI,
```

2281 TGTATTTTATATTTTTCTAAGATATAGTGATTTCTTGACAATGAATGTCTCTTAGTTGCG
ACATAAAATATAAAAAGATTCTATATCACTAAAGAACTGTTACTTACAGAGAATCAACGC

2297 DdeI, 2331 DdeI, 2338 HhaI,

2341 CTCATCAAATGCACTGTCAAACTTCTCTCTCTGTGCAGAGTGACTTTGGAGCTGGCCTTA
GAGTAGTTTACGTGACAGTTTGAAGAGAGAGACACGTCTCACTGAAACCTCGACCGGAAT

2380 MaeIII, 2390 AluI, 2393 HaeI, 2394 HaeIII,

2401 TATTTGGATCTCCCGGGCAGCTTTTTGGAAACTCCTAGTGAACATGCTGTGCCTCAAGCC
ATAAACCTAGAGGGCCCGTCGAAAAACCTTTGAGGATCACTTGTACGACACGGAGTTCGG

2406 BinI XhoII, 2407 MboI, 2412 AvaI NciI ScrFI SmaI, 2413 HpaII NciI ScrFI, 2417 BbvI Fnu4HI, 2419 AluI, 2435 MaeI, 2442 Nsp7524I, 2443 NlaIII, 2452 MnlI,

2461 AGTGAAATTAGGCTCCCTGGAGGTGGGACACGGGCAGCATGGTGGGTATGTAAAGCTGGC
TCACTTTAATCCGAGGGACCTCCACCCTGTGCCCGTCGTACCACCCATACATTTCGACCG

2471 NlaIV, 2476 EcorII ScrFI, 2480 MnlI, 2494 BbvI Fnu4HI, 2498 NlaIII, 2514 AluI,

2521 TTGGTGAGTTGGCTATGTGGCCAGGGTTGAGAGTCATGGCTACAAAGGACCAGGGAGGGG
AACCACTCAACCGATACACCGGTCCCAACTCTCAGTACCGATGTTTCCTGGTCCCTCCCC

2523 HphI, 2538 BalI EaeI HaeI, 2539 HaeIII, 2541 EcorII ScrFI, 2551 HinfI, 2555 NlaIII, 2567 AvaII Sau96I, 2570 EcorII ScrFI, 2575 MnlI,

2581 GGCAGGGGCAGTGGAAAGCACAGACCAGCTGGCTAGATAGTCTGCATATTCACGCATGAA
CCGTCCCCGTCACCTTTCGTGTCTGGTCGACCGATCTATCAGACGTATAAGTGCGTACTT

2606 NspBII PvuII, 2607 AluI, 2613 MaeI, 2635 NlaIII,

2641 GGACTTACAGGGATGTCAGCCACTCCCCATGAATTACCCACAGAAGCAAGTTGCTTGGGT
CCTGAATGTCCCTACAGTCGGTGAGGGGTACTTAATGGGTGTCTTCGTTCAACGAACCCA

2651 FokI, 2668 NlaIII, 2692 TthIII2,

2701 TGTTGATTGAACTTGGGATAGTTGATTGAACTTTAACAGTTTGGCCGAAATACCTTTGAG
ACAACTAACTTGAACCCTATCAACTAACTTGAAATTGTCAAACCGGCTTTATGGAAACTC

2742 EaeI, 2743 HaeIII,

2761 ATGTAAAGAAAATAACTAGGGGTTGGCGAAGATAGGGAGACAATGCTCAGAGAGCAAACA
TACATTTCTTTTATTGATCCCCAACCGCTTCTATCCCTCTGTTACGAGTCTCTCGTTTGT

2776 MaeI, 2788 MboII, 2806 DdeI, 2815 TthIII2,

```
2821  TACAGCCTTATTTAGGATCTTCAGTAAAATCTCTCATCTCAGTTCTAGAAGACTCTGGGC
      ATGTCGGAATAAATCCTAGAAGTCATTTTAGAGAGTAGAGTCAAGATCTTCTGAGACCCG
                    ^^ ^                     ^     ^^  ^  ^

      2835 BinI XhoII, 2836 MboI, 2838 MboII, 2858 DdeI, 2864 XbaI
      , 2865 MaeI, 2868 MboII, 2871 HinfI,

2881  TACGTTGGCTCCTGGGTTCTAAGATGAGGGGACCCATGGAAGGCAGGTAGGCAGTTCACA
      ATGCAACCGAGGACCCAAGATTCTACTCCCCTGGGTACCTTCCGTCCATCCGTCAAGTGT
       ^    ^   ^       ^      ^  ^^   ^^

      2882 MaeII, 2887 NlaIV, 2891 EcorII ScrFI, 2899 DdeI, 2906 M
      nlI, 2909 Eco0109 NlaIV PpuMI, 2910 AvaII NlaIV Sau96I, 2914
       NcoI StyI, 2915 NlaIII,

2941  GGCATCAGGTCCAGTCTCTGAGTTTGTCCTCTTTGACTTGGACCTCTCACTCATATAGTG
      CCGTAGTCCAGGTCAGAGACTCAAACAGGAGAAACTGAACCTGGAGAGTGAGTATATCAC
       ^    ^         ^         ^           ^  ^

      2942 SfaNI, 2948 AvaII Sau96I, 2958 DdeI, 2968 MnlI, 2980 Av
      aII Sau96I, 2983 MnlI,

3001  TTAGCCTATTAGGCAGGGAACTCAAGTGAACAGGTACAAGAGAACTGCCAATAGCTTTAT
      AATCGGATAATCCGTCCCTTGAGTTCACTTGTCCATGTTCTCTTGACGGTTATCGAAATA
        ^                       ^                   ^

      3004 BglI, 3034 RsaI, 3053 AluI,

3061  GGCATTCTGTTAAATTGGCTTCGACATAGGTGACACAGGATGACCCAACATTGTCTATCT
      CCGTAAGACAATTTAACCGAAGCTGTATCCACTGTGTCCTACTGGGTTGTAACAGATAGA
                         ^      ^^      ^

      3081 TaqI, 3089 HphI, 3090 MaeIII, 3098 FokI,

3121  GGCTACGATCTGAAGAGATACAGAACCCAGGAACTAAGTGCTCAGACCGAGAAGCTAAAT
      CCGATGCTAGACTTCTCTATGTCTTGGGTCCTTGATTCACGAGTCTGGCTCTTCGATTTA
        ^   ^               ^     ^  ^  ^          ^

      3127 MboI, 3132 MboII, 3147 EcorII ScrFI, 3154 DdeI, 3158 Bs
      p1286 HgiAI, 3161 DdeI, 3173 AluI,

3181  TCT----------~4KB----------CCCTGGTCATCAGGCTTAGGGCAGCTGTCTTGA
      AGA----------~4KB----------GGGACCAGTAGTCCGAATCCCGTCGACAGAACT
                                 ^            ^     ^^^         ^

      3209 EcorII ScrFI, 3222 DdeI, 3228 BbvI Fnu4HI, 3229 NspBII
      PvuII, 3230 AluI, 3239 MnlI,

3241  GGTGGACCCCTTACTGAGCACTCTGTCTACATGGGCCCCCGTACACTGGTCTCTCTCATC
      CCACCTGGGGAATGACTCGTGAGACAGATGTACCCGGGGGCATGTGACCAGAGAGAGTAG
        ^         ^ ^         ^    ^  ^^       ^

      3244 AvaII NlaIV Sau96I, 3254 DdeI, 3256 Bsp1286 HgiAI, 3265
       AccI, 3270 NlaIII, 3273 ApaI BanII Bsp1286 Eco0109 NlaIV Sa
      u96I, 3274 HaeIII NlaIV Sau96I, 3281 RsaI,

3301  ACACCATACCTGTTCTCTTTATGACACGAGCAGAGTCTGAAGTTATTTATGTGCACTAGA
      TGTGGTATGGACAAGAGAAATACTGTGCTCGTCTCAGACTTCAATAAATACACGTGATCT
                                  ^                  ^    ^

      3333 HinfI, 3351 ApaLI Bsp1286 HgiAI, 3356 MaeI,
```

```
3361  CTGAGAGGTCCACAGTAGAGGTCGATACCTTGCCTGTTTAATATACTGCTGTCCTCCAGC
      GACTCTCCAGGTGTCATCTCCAGCTATGGAACGGACAAATTATATGACGACAGGAGGTCG
      ^   ^ ^          ^   ^                                    ^

      3361 DdeI, 3365 MnlI, 3367 AvaII Sau96I, 3378 MnlI, 3382 Taq
      I, 3413 MnlI,

3421  ACCAAGAGTAGTGTTGGCTCAGGAAACGAACGAACGAACGAATGAAGCACCTACATGTAT
      TGGTTCTCATCACAACCGAGTCCTTTGCTTGCTTGCTTGCTTACTTCGTGGATGTACATA
                       ^                                      ^^

      3438 DdeI, 3473 AflIII Nsp7524I, 3474 NlaIII,

3481  CATGTTGAGAGAATTAGTAGGGTTTCAGTGGCTCTAAAAGCTATAAGTTAAGGTATATTT
      GTACAACTCTCTTAATCATCCCAAAGTCACCGAGATTTTCGATATTCAATTCCATATAAA
      ^                                  ^

      3481 NlaIII, 3519 AluI,

3541  TCCCCTCATTGTGTCATCTCCATAACTTCCAATTCCTGAAATTTGATTTTTAGCGTAGAG
      AGGGGAGTAACACAGTAGAGGTATTGAAGGTTAAGGACTTTAAACTAAAAATCGCATCTC
         ^                                                      ^

      3544 MnlI, 3598 BanII Bsp1286,

3601  CCCCAAGACCCCTGTCCTGGTACGGCAAGAGGTGTAAGAGATGAGTCATGTGGTCTGTCC
      GGGGTTCTGGGGACAGGACCATGCCGTTCTCCACATTCTCTACTCAGTACACCAGACAGG
                     ^   ^         ^              ^   ^

      3616 EcorII ScrFI, 3620 RsaI, 3629 MnlI, 3643 HinfI, 3647 Nl
      aIII,

3661  CATTCCCACTAGACCTTGTCTTGTTTTCCAATCATAGAGTTGGCTTTGTTGGAATTGTTG
      GTAAGGGTGATCTGGAACAGAACAAAAGGTTAGTATCTCAACCGAAACAACCTTAACAAC
            ^  ^

      3669 MaeI, 3672 TthIII1,

3721  GGACTTTTGTATTTNNNAGTACTNNNGCTAGGGCTACTACATCTGCTCTCTGGCAGCGAG
      CCTGAAAACATAAANNNTCATGANNNCGATCCCGATGATGTAGACGAGAGACCGTCGCTC
                       ^^        ^                        ^

      3738 ScaI, 3739 RsaI, 3748 MaeI, 3773 BbvI Fnu4HI,

3781  CAAAATAATTTTCATTTTGGTTTAGGTACCAGTGGAGCATAACTCCACTGTGGATCTGGA
      GTTTTATTAAAAGTAAAACCAAATCCATGGTCACCTCGTATTGAGGTGACACCTAGACCT
                           ^^                              ^^

      3805 BanI KpnI NlaIV, 3806 RsaI, 3832 BinI XhoII, 3833 MboI,


3841  AACCACTACACTTGGTAGAGTCCTACTGTGGACCGTGTCCAGGTTGAAACTAGTAAAGGC
      TTGGTGATGTGAACCATCTCAGGATGACACCTGGCACAGGTCCAACTTTGATCATTTCCG
                       ^           ^^         ^          ^^

      3858 HinfI, 3870 AvaII Sau96I, 3871 TthIII1, 3879 EcorII Scr
      FI, 3889 SpeI, 3890 MaeI,
```

```
3901  GAGTACCATTGCAGGGCCATGCAAAACCTGAGCTCTGCAAGACTAGGGTCATGCAAGACC
      CTCATGGTAACGTCCCGGTACGTTTTGGACTCGAGACGTTCTGATCCCAGTACGTTCTGG
        ^          ^^  ^          ^ ^^             ^      ^       ^
```

3903 RsaI, 3914 Sau96I, 3915 HaeIII, 3918 NlaIII, 3928 DdeI, 3930 BanII Bsp1286 HgiAI SacI, 3931 AluI, 3943 MaeI, 3950 NlaIII, 3960 DdeI,

```
3961  TGAGCTCTGTAAGACCAGAGCTATATAAGCCTAGGACTCTGGAAATCTTGGCGACTAGAC
      ACTCGAGACATTCTGGTCTCGATATATTCGGATCCTGAGACCTTTAGAACCGCTGATCTG
       ^^                ^          ^^   ^                     ^  ^
```

3962 BanII Bsp1286 HgiAI SacI, 3963 AluI, 3979 AluI, 3990 AvrII StyI, 3991 MaeI, 3995 HinfI, 4015 MaeI, 4018 HinfI,

```
4021  TCTGTGCCTTCCTGGGCTCTAATGATGCTAAAATATTAGAACTGTGGAAATAGTGCTTGC
      AGACACGGAAGGACCCGAGATTACTACGATTTTATAATCTTGACACCTTTATCACGAACG
            ^  ^          ^        ^                          ^
```

4031 EcorII ScrFI, 4034 BanII Bsp1286, 4044 SfaNI, 4052 SspI, 4074 TthIII2,

```
4081  CTTATTGGGACATTCCTTCCTCCCTGCCTAGGGCTCAGGTGTCCGCCCATCTGCTAGGGA
      GAATAACCCTGTAAGGAAGGAGGGACGGATCCCGAGTCCACAGGCGGGTAGACGATCCCT
                      ^        ^^  ^  ^                       ^
```

4099 MnlI, 4107 AvrII StyI, 4108 MaeI, 4111 BanII Bsp1286, 4114 DdeI, 4134 MaeI,

```
4141  ATGATGGTAACCCTTATCTAACATCAGCTCTCCTTCACAGAGTTTTGGCCTGTGGTCGTG
      TACTACCATTGGGAATAGATTGTAGTCGAGAGGAAGTGTCTCAAAACCGGACACCAGCAC
        ^^                    ^                    ^^
```

4146 BstEII, 4147 MaeIII, 4166 AluI, 4186 HaeI, 4187 HaeIII,

```
4201  CAGTCCAGGGGGCTGGATGGCATGCTGGACCCAAGCTCACCTCAGTGTCTGGGCCCAATA
      GTCAGGTCCCCCGACCTACCGTACGACCTGGGTTCGAGTGGAGTCACAGACCCGGGTTAT
        ^         ^    ^^      ^       ^  ^  ^^          ^^
```

4205 BstXI EcorII ScrFI, 4215 FokI, 4220 Nsp7524I SphI, 4221 NlaIII, 4227 AvaII NlaIV Sau96I, 4234 AluI, 4237 HphI, 4240 MnlI, 4241 DdeI, 4251 ApaI BanII Bsp1286 NlaIV Sau96I, 4252 HaeIII Sau96I,

```
4261  AAGGTTTTGCCAAGGACGCAGCTTTCTATACTGGCCGCAGTGAGGTAAGTGCCTGTGGGA
      TTCCAAAACGGTTCCTGCGTCGAAAGATATGACCGGCGTCACTCCATTCACGGACACCCT
            ^     ^  ^ ^             ^^^        ^
```

4270 StyI, 4275 HgaI, 4278 BbvI Fnu4HI, 4280 AluI, 4292 EaeI, 4293 HaeIII, 4294 Fnu4HI, 4302 MnlI,

```
4321  CAGCTGCCTTCTTGGGGTGTCTAATTCTATAGCTGCTAGACTTGGAAGCTTCCCAGTGAC
      GTCGACGGAAGAACCCCACAGATTAAGATATCGACGATCTGAACCTTCGAAGGGTCACTG
      ^^^                          ^^   ^          ^^         ^
```

4321 NspBII PvuII, 4322 AluI, 4323 BbvI Fnu4HI, 4351 AluI, 4352 BbvI Fnu4HI, 4356 MaeI, 4366 HindIII, 4367 AluI, 4376 MaeIII,

```
4381  ACACAGCTTGGAAAGAATTCCTGCAGAATTCCTGC----------~6KB-----
      TGTGTCGAACCTTTCTTAAGGACGTCTTAAGGACG----------~6KB-----
          ^    ^   ^    ^   ^

      4385 AluI, 4391 XmnI, 4395 EcorI, 4401 PstI, 4406 EcorI,

4441  ----------AACTTAGGTAACCTCTATCTTTTTTTATTCCAGGTGCATAGCGT
      ----------TTGAATCCATTGGAGATAGAAAAAAATAAGGTCCACGTATCGCA
                      ^   ^^   ^                  ^

      4459 DdeI, 4463 BstEII, 4464 MaeIII, 4468 MnlI, 4486 EcorII
      ScrFI,

4501  AATGTCCATGTTGTTCTACACTCTGATCACTGCGTTTTTGATCGGCGTACAGGCAGAACC
      TTACAGGTACAACAAGATGTGAGACTAGTGACGCAAAAACTAGCCGCATGTCCGTCTTGG
            ^                ^^              ^      ^

      4507 NlaIII, 4524 BclI, 4525 MboI, 4540 MboI, 4547 RsaI,

4561  GTACACAGATAGCAATGTCCCAGAAGGAGACTCTGTCCCTGAAGCCCACTGGACTAAACT
      CATGTGTCTATCGTTACAGGGTCTTCCTCTGAGACAGGGACTTCGGGTGACCTGATTTGA
      ^                      ^

      4561 RsaI, 4589 HinfI TthIII1,

4621  TCAGCATTCCCTTGACACAGCCCTCCGCAGAGCCCGCAGTGCCCCTACTGCACCAATAGC
      AGTCGTAAGGGAACTGTGTCGGGAGGCGTCTCGGGCGTCACGGGGATGACGTGGTTATCG
                           ^      ^      ^                  ^^

      4642 MnlI, 4650 BanII Bsp1286, 4659 Bsp1286, 4678 AluI, 4679
       BbvI Fnu4HI,

4681  TGCCCGAGTGACAGGGCAGACCCGCAACATCACTGTGGACCCCAGACTGTTTAAGAAACG
      ACGGGCTCACTGTCCCGTCTGGGCGTTGTAGTGACACCTGGGGTCTGACAAATTCTTTGC
        ^    ^                        ^

      4683 AvaI, 4688 MaeIII, 4717 AvaII NlaIV Sau96I,

4741  GAGACTCCACTCACCCCGTGTGCTGTTCAGCACCCAGCCTCCACCCACCTCTTCAGACAC
      CTCTGAGGTGAGTGGGGCACACGACAAGTCGTGGGTCGGAGGTGGGTGGAGAAGTCTGTG
        ^       ^^                        ^         ^ ^

      4743 HinfI, 4751 HphI, 4752 DraIII, 4778 MnlI, 4788 MnlI, 47
      90 MboII,

4801  TCTGGATCTAGACTTCCAGGCCCATGGTACAATCCCTTTCAACAGGACTCACCGGAGCAA
      AGACCTAGATCTGAAGGTCCGGGTACCATGTTAGGGAAAGTTGTCCTGAGTGGCCTCGTT
        ^^ ^^       ^  ^  ^^   ^                  ^  ^  ^      ^

      4804 BinI XhoII, 4805 MboI, 4807 XbaI, 4808 MaeI, 4816 BstXI
       EcorII ScrFI, 4819 HaeIII Sau96I, 4822 NcoI StyI, 4823 NlaI
      II, 4827 RsaI, 4846 HinfI, 4849 HphI, 4852 HpaII, 4860 HaeII
      ,

4861  GCGCTCATCCACCCACCCAGTCTTCCACATGGGGGAGTTCTCAGTGTGTGACAGTGTCAG
      CGCGAGTAGGTGGGTGGGTCAGAAGGTGTACCCCCTCAAGAGTCACACACTGTCACAGTC
      ^    ^              ^          ^          ^      ^ ^

      4861 HhaI, 4866 FokI, 4881 MboII, 4888 NlaIII, 4900 DdeI, 49
      08 MaeIII, 4910 TthIII1,
```

```
4921  TGTGTGGGTTGGAGATAAGACCACAGCCACAGACATCAAGGGCAAGGAGGTGACAGTGCT
      ACACACCCAACCTCTATTCTGGTGTCGGTGTCTGTAGTTCCCGTTCCTCCACTGTCACGA
                                                          ^ ^^        ^

      4967 MnlI, 4969 HphI, 4970 MaeIII, 4980 EaeI,

4981  GGCCGAGGTGAACATTAACAACAGTGTATTCAGACAGTACTTTTTTTGAGACCAAGTGCCG
      CCGGCTCCACTTGTAATTGTTGTCACATAAGTCTGTCATGAAAAAACTCTGGTTCACGGC
      ^   ^ ^                    ^^

      4981 HaeIII, 4985 MnlI, 4987 HphI, 5016 ScaI, 5017 RsaI,

5041  AGCCTCCAATCCTGTTGAGAGTGGGTGCCGGGGCATCGACTCCAAACACTGGAACTCATA
      TCGGAGGTTAGGACAACTCTCACCCACGGCCCCGTAGCTGAGGTTTGTGACCTTGAGTAT
        ^                   ^   ^    ^  ^ ^   ^^

      5043 MnlI, 5064 BanI NlaIV, 5068 HpaII NciI ScrFI, 5073 SfaN
      I, 5076 TaqI, 5078 HinfI, 5082 PflMI, 5083 TthIII2,

5101  CTGCACCACGACTCACACCTTCGTCAAGGCGTTGACAACAGATGAGAAGCAGGCTGCCTG
      GACGTGGTGCTGAGTGTGGAAGCAGTTCCGCAACTGTTGTCTACTCTTCGTCCGACGGAC
            ^                   ^                       ^   ^

      5110 HinfI, 5131 HincII, 5153 BbvI Fnu4HI, 5157 EcorII ScrFI
      ,

5161  GAGGTTCATCCGGATAGACACAGCCTGTGTGTGTGTGCTCAGCAGGAAGGCTACAAGAAG
      CTCCAAGTAGGCCTATCTGTGTCGGACACACACACACGAGTCGTCCTTCCGATGTTCTTC
      ^     ^ ^^                    ^  ^                   ^   ^

      5161 MnlI, 5167 FokI, 5169 BspMII, 5170 HpaII, 5195 Bsp1286
      HgiAI, 5198 DdeI, 5217 MboII, 5220 MnlI,

5221  AGGCTGACTTGCCTGCAGCCCCCTTCCCCACCTGCCCCCTCCACACTCTCCTGGGCCCCT
      TCCGACTGAACGGACGTCGGGGGAAGGGGTGGACGGGGGAGGTGTGAGAGGACCCCGGGA
                ^ ^           ^       ^           ^  ^^   ^

      5233 PstI, 5235 BbvI Fnu4HI, 5250 BspMI, 5258 MnlI, 5270 Eco
      rII ScrFI, 5273 ApaI BanII Bsp1286 Eco0109 NlaIV Sau96I, 527
      4 HaeIII NlaIV Sau96I, 5278 MnlI,

5281  CCCTACCTCAGCCTGTAAATTATTTTAAATTATAAGGACTGCATGATAATTTATCGTTTA
      GGGATGGAGTCGGACATTTAATAAAATTTAATATTCCTGACGTACTATTAAATAGCAAAT
        ^^                ^                 ^

      5286 MnlI, 5287 DdeI, 5304 AhaIII, 5322 NlaIII,

5341  TACAATTTTAAAGACATTATTTATTAAATTTTCAAAGCATCCTGTATACCGAGCAGCTTG
      ATGTTAAAATTTCTGTAATAAATAATTTAAAAGTTTCGTAGGACATATGGCTCGTCGAAC
            ^                             ^^     ^       ^ ^    ^

      5347 AhaIII, 5377 SfaNI, 5378 FokI, 5384 AccI SnaI, 5393 Bbv
      I Fnu4HI, 5395 AluI, 5400 HinfI,

5401  AGTCCTTTCTCTCAGTAAGAGTAGCGCCACAAAGTTTGAGAGAGTGCGAAATAACTGTCC
      TCAGGAAAGAGAGTCATTCTCATCGCGGTGTTTCAAACTCTCTCACGCTTTATTGACAGG
            ^           ^^

      5411 DdeI, 5423 HaeII, 5424 HhaI,
```

5461 CCATCACTCCTGGCACCATAGCCCATCTGAAACCTTCATGTGCGGCAGCCAACCCTGCTC
GGTAGTGAGGACCGTGGTATCGGGTAGACTTTGGAAGTACACGCCGTCGGTTGGGACGAG

5469 EcorII ScrFI, 5472 BanI NlaIV, 5497 NlaIII, 5502 Fnu4HI
, 5505 BbvI Fnu4HI, 5520 EcorII ScrFI,

5521 CAGGCGCTGACTAACTAGTTTCAGACACTAACTCTACACCCACCCAACATAATTAGCGA
GTCCGCGACTGATTGATCAAAGTCTGTGATTGAGATGTGGGTGGGTTGTATTAATCGCT

5523 HaeII, 5524 HhaI, 5534 SpeI, 5535 MaeI,