Draft 10
9-19-90

ANNUAL REPORT OF THE

LABORATORY OF BIOCHEMICAL GENETICS

NATIONAL HEART, LUNG, AND BLOOD INSTITUTE

October 1, 1989 through September 30, 1990

Previously, we cloned three neighboring Drosophila homeobox genes designated NK-1, NK-3, and NK-4 that reside in the right arm of the third chromosome at bands 93E1-5, and a fourth homeobox gene, NK-2 that resides in the X chromosome at bands 1C1-5. The nucleotide sequences of approximately 3.5 kb of NK-2 cDNA and genomic DNA were determined. The deduced amino acid sequence of NK-2 protein consists of 723 amino acid residues. NK-2 is a basic protein, with regions rich in alternating acidic and basic amino acid residues, regions rich in seryl-, threonyl-, and prolyl-residues, repetitive alanine residues, PEST sequences, which specify rapid turnover of protein, an acidic domain, a homeobox domain, and an alternating repeat of histidine and alanine residues. Northern analysis revealed one species of NK-2 poly A^+ RNA, 3.2 kb in length, which is most abundant in 3-6 hr embryos. NK-2 mRNA was detected at all later stages of embryonic development and in adult flies, but less NK-2 mRNA was detected in larvae and pupae. In situ hybridization analysis showed that expression of the NK-2 gene is initiated at the cellular blastoderm stage in the ventrolateral portion of the embryo, which later in development, gives rise to the nervous system. In 12 hr and older embryos expression of the NK-2 gene was found both the central nervous system and in cells associated with the midgut.

The nucleus of a fertilized Drosophila egg divides thirteen times during the first 130 minutes of development to form the blastoderm, which is a single cell with five thousand nuclei. During the next forty minutes of development plasma membranes are laid down around each nucleus resulting in the formation of a five thousand cell embryo termed the cellular blastoderm. As the cellular blastoderm is forming, the nuclei in the ventrolateral portion of the blastoderm become committed to the neuroectodermal pathway of differentiation. This formation of the neuroectodermal anlage is the first step in the developmental program that leads to the formation of the nervous system. Both the initial time of expression of the NK-2 gene at the cellular blastoderm stage of development and the initial pattern of expression of NK-2 coincide exactly with the appearance of the neuroectodermal anlage. Since homeobox proteins are known to be regulators of gene expression, we predict that NK-2 protein activates the gene program that commits cells to the neuroectodermal pathway of development and thereby initiates the developmental program for the formation of the nervous system.

Both genomic and cDNA NK-1 clones were sequenced and approximately 7500 nucleotide residues of the NK-1 gene were determined. The NK-1 gene contains 4 exons; the last intron is located within the region of the DNA that codes for the third helix of the homeobox protein at the same site as the corresponding introns of labial, Abdominal-B, and Distal-less homobox genes. NK-1 protein is comprised of 661 amino acid residues and contains repetitive,

alternating histidyl-prolyl residues termed a paired repeat, repetitive histidyl-glutamyl residues, repetitive alanyl residues, a highly acidic domain before the homeobox, and PEST sequences. One major poly A^+ RNA transcript, 2.9 kb in length, was detected in 6-12 hr embryos and thereafter during embryonic development, but not in 0-3 hr embryos. In situ hybridization analysis of embryo sections and whole embryos showed that the expression of the NK-1 gene starts in 7.5-8.0 hr embryos in the ventral nervous system and ultimately results in a bilaterally symmetric, segmentally repeated pattern that consists of a of small subset of neurons on each side in each thoracic and abdominal segment and some unidentified cells in the head. The NK-1 gene also is expressed in a similar pattern in a small subset of striated muscle cells in each thoracic and abdomenal segment.

The nucleotide sequence of Drosophila NK-1 homeobox DNA was used to clone genomic DNA from monkey, rat, Xenopus, and salmon DNA. Nucleotide sequence analysis showed that the deduced amino acid sequences of the homeobox domains of monkey and rat are identical to that of the Drosophila NK-1 homeobox, and that salmon and Xenopus homeoboxs are novel homeoboxes that differ from NK-1 by 4 and 9 amino acid residues, respectively.

NK-3 and NK-4 are neighboring homeobox genes that are separated by only 7.5 kb. Approximately 3.6 kb of cloned NK-3 cloned cDNA and genomic DNA were sequenced. The NK-3 gene contains a short intron

and codes for a homeobox protein 374 amino acid residues in length, which contains charged regions rich in acidic and basic amino acids, and a region rich in serine, threonine, and proline residues. Northern analysis of NK-3 revealed one species of poly A^+ RNA, 1.6 kb in length, that is expressed transiently in 6 to 12 hr embryos but not thereafter. By in situ hybridization NK-3 mRNA was shown to be expressed in a bilaterally symmetric, segmentally repeated pattern in the visceral mesoderm of Drosophila embryos.

Sequence analysis of NK-4 cDNA and genomic DNA clones show that the NK-4 gene has 2 introns and that the deduced NK-4 protein is 371 amino acid residues in length. NK-4 is a basic protein with CAX repeats, a homeobox domain, glutamine repeats (M repeats), and regions rich in serine, threonine, and proline. One major species of NK-4 poly A^+ RNA was detected, 1.7 kb in length, which first appears in 3 - 6 hr Drosophila embryos, is less abundant in 6 - 12 hr embryos, and was not detected thereafter. NK-4 was found by in situ hybridization to be expressed only in cephalic, thoracic, and abdominal mesodermal tissue starting at gastrulation (3 hr) when the mesoderm first appears. NK-4 mRNA was not detected after 7.5 hr of embryonic development.

Short segments of mouse genomic DNA that correspond to homeobox regions of homeobox genes were selectively amplified by the polymerase chain reaction and were cloned. Nucleotide sequence analysis showed that some of the DNA clones correspond to novel mouse homeobox genes. A mouse genomic DNA library with inserts in

the 15 kb range then was screened using the small DNA inserts that had been identified as novel mouse homeobox genes as probes. Genomic DNA clones, were obtained that correspond to four novel mouse homeobox genes. The nucleotide sequences before and after the homeobox regions of two of the genomic DNA clones (CL-8B and CL-101) were determined. RNA transcripts of portions of CL-8B DNA then were used as probes to screen a mouse embryo cDNA library and a CL-8B cDNA clone was obtained. Further work is needed to determine the structures of these homeobox genes and the time and sites of expression of the genes in mouse embryos and adults.

A subfamily of homeobox genes has been reported that code for proteins that contain a highly conserved domain, termed a Pou-box, before the homeobox. During the past year, two novel species of Pou-box-homeobox cDNA were cloned and the Pou-box-homeobox regions were sequenced. Both of the novel cDNA's correspond to species of poly A^+ RNA that are expressed in adult mouse brain.

A method of selection for cloned mouse genomic DNA fragments that contain functional enhancer or promoter sequences was developed that is based on the observation that the synthesis of polyoma virus DNA in mouse cells is dependent upon viral enhancer or promoter sequences that also are needed for activation of transcription of polyoma genes. An E. coli-mammalian cell shuttle vector, pPyEO, was used that contains the β-lactamase gene and origin of replication from pBR322, and polyoma viral DNA with a library of mouse genomic DNA fragments inserted in the vector in

place of the polyoma enhancer region. Thus, recombinant plasmids with DNA inserts that contain enhancer or promoter sequences were expected to replicate in mouse cells, but not recombinants that lack such sequences. Neuroblastoma, glioma, myoblast, or fibroblast cell lines were transfected with plasmid DNA and incubated for several days to allow plasmid replication. Then plasmid DNA was recovered, purified, and cloned in E. coli. Thousands of recombinants were recovered from mammalian cells and all of the cloned DNA inserts that were tested had enhancer activity when ligated to a reporter gene. The results show that the enhancer selection method yields many kinds of cloned enhancer sequences. Marked differences also were found in the activities of some cloned sequences in stimulating the expression of a reporter gene in different cell types.

Previously, we showed that prolonged elevation of cyclic AMP levels of NG108-15 neuroblastoma-glioma hybrid cells and other neuroblastoma cell lines results in the appearance in cells of electrically excitable ion channels, and that storage vesicles for acetylcholine increase in abundance, and the number of synapses of NG108-15 cells with striated muscle cells increases. The increase in synapses is due to a cAMP-dependent acquisition by cells of functional L-type voltage-sensitive calcium channels that are required for stimulus-secretion coupling in NG108-15 cells. cDNA clones were obtained that correspond to mRNA from two voltage-sensitive calcium channel α_1-subunit genes from rat brain. Northern analysis revealed that poly A^+ RNA corresponding to an

α_1-subunit of a voltage-sensitive calcium channel markedly increased in abundance when NG108-15 cells were incubated for several days under conditions that increase cellular levels of cAMP.

In addition, cDNA clones were obtained that correspond to species of NG108-15 poly A^+ RNA that increase in abundance when cells are treated with compounds that elevate cellular cAMP. Six kinds of ~~the~~ cDNA clones were identified by nucleotide sequence analysis. Four of the cDNA clones correspond to different portions of mouse mitochondrial DNA; i.e., pNG-10 corresponds to RNA transcribed from the mitochondrial DNA displacement loop which contains promoters for the transcription of mitochondrial heavy- and light-chain DNA, the origin of replication for mitochondrial heavy chain DNA, and other regulatory signals. pNG-32 cDNA corresponds to mRNA transcribed from the mitochondrial ATPase 6 gene, which codes for a protein that is part of the proton pump-ATP synthase complex. pNG-37 corresponds to mRNA transcribed from the mitochondrial cytochrome oxidase I; whereas, pNG-47 corresponds to mitochondrial 12S ribosomal RNA. Treatment of NG108-15 cells with dibutyryl cAMP for 5 - 8 days results in two to five fold increases in RNA transcripts from both the mitochondrial heavy and light strands of DNA and in a two to three-fold increase in the amount of mitochondrial DNA. Similar results were obtained when adenylate cyclase of NG108-15 cells was activated with prostaglandin E1, mediated by PGE1 receptors, or by forskolin. In addition, pNG-64 cDNA was shown to correspond to secretogranin I,

a secretory vesicle protein. Poly A^+ RNA for secretogranin I was found to increase three-fold in abundance due to treatment of cells for 5 days with dibutyryl cAMP. cDNA clone pNG-57 was shown to correspond to mRNA for ribosomal protein S-10.

In previous studies cAMP and phorbol esters that activate protein kinase C were shown to synergistically elevate neuropeptide Y mRNA levels 20 to 200-fold in 4 to 24 hr in PC12 rat pheochromocytoma cells. Treatment of cells with nerve growth factor (NGF) results in 40 to 100-fold increases in neuropeptide Y mRNA within 1 to 6 days. Glucocorticoids biphasically modulate the stimulations by NGF, potentiating early (3-10 hr) effects of NGF, but inhibiting later stimulations (1-6 days). The results of nuclear run-on transcription assays showed that the apparent rate of neuropeptide Y gene transcription in PC12 cells is altered by NGF, NGF plus dexamethasone, and cAMP plus phorbol ester. The effects of these compounds on neuropeptide Y mRNA stability were less significant. These results show that neuropeptide Y gene transcription is controlled by multiple, potentially interacting regulators.

To identify regions upstream from the neuropeptide Y promoter that regulate gene expression, portions of the upstream region of the neuropeptide Y gene were ligated to plasmids that contained the chloramphenicol acetyltransferase (CAT) gene. In transient expression assays, constructs that contained 144 and 671 nucleotide residues of upstream sequence were weakly responsive to cAMP elevation by forskolin or to NGF, but no enhancement by

phorbol ester was noted. The responses were much less than were observed with the endogenous gene, suggesting the additional sequences or variables are important.

Glucocorticoids such as dexamethasone potentiate proenkephalin gene expression in several systems including rat brain. We previously showed that glucocorticoids and cAMP synergistically elevate proenkephalin mRNA levels. Nuclear run-on transcription assay results show that the proenkephalin gene transcription rate was not altered by dexamethasone alone, was weakly stimulated by cAMP, and was persistently stimulated by dexamethasone and cAMP together (3-6-fold over 1-24 hr). To search for cooperative glucocorticoid and cAMP regulatory elements, chimeric plasmids were constructed containing the CAT reporter gene under the control of upstream proenkephalin gene sequences in C6 glioma cells transfected with these constructs. Full stimulation of CAT activity by cAMP required sequences between residues −145 and −192, which are beyond the previously described human proenkephalin cAMP responsive site. Unexpectedly, dexamethasone did not stimulate expression but reduced cAMP stimulated CAT activity; the full negative effect required sequences between −435 and −1000. A DNA fragment from the 5'-upstream region of a rat proenkephalin gene was inserted into a promoterless CAT vector and a Rous sarcoma virus-transformed quail embryo retina cell line was transfected with the DNA. Treatment of the transfected cells for 24 hr with forskolin, prostaglandin E1 and theophylline, or retinoic acid resulted in 12-, 6-, and 4-fold increases in CAT

expression, respectively. In addition, the transactivation protein tax_1 of human T-cell leukemia virus I was found to activate the proenkephalin gene promoter and resulted in a 3-9-fold increase in CAT expression in C6 glioma cells transfected with the chimeric rat proenkephalin regulatory region-CAT DNA.

Ankyrin, a peripheral membrane protein that links certain integral membrane proteins to the spectrin-actin membrane cytoskeleton, was shown by immunofluorescence and immunogold labeling and electron microscopy to be concentrated at the triad junctions of striated muscle cells formed by the transverse tubules and the sarcoplasmic reticulum. Ankyrin thus may play a role in organizing the structure of the triad, or may be involved in linking the triad membranes to the muscle cytoskeleton at specific sites along the sarcomere.

Mice with the mdg/mdg genotype die at birth due to failure of excitation-contraction coupling in skeletal muscle that results from a mutation in the gene for the voltage-sensitive calcium channel α_1 subunit. Disruptions have been reported in the formation of cross-striated myofibrils and in the organization of the excitation-contraction coupling membrane system in mdg/mdg mice. Sections of striated muscle from 18 day embryo wild type mice or from homozygous or heterozygous mdg mutant mice were incubated with antibodies directed against the α_1 calcium channel subunit and examined by indirect immunofluorescence. The results

show that the expression of the α_1 subunit is sharply reduced or absent in homozygous mdg mutants; however, the spatial organization of the sarcoplasmic reticulum and myofibrils is normal, although the intensity of staining is reduced. T-tubule organization also is normal, although the intensity of staining is reduced. These results suggest that the α_1 calcium channel subunit is not required for the assembly of the excitation-coupling membrane system, and that the absence of the α_1 subunit may indirectly inhibit myotube maturation.

In cultured rat myotubes, transverse tubule membranes differ in molecular composition from plasma membranes. The precursors of transverse tubules appear to be intracellular vesicles or tubules that first appear in myoblasts. Intracellular tubular membranes with markers for both transverse tubules and sarcoplasmic reticulum are associated with the Z lines of immature myofibrils in newly formed myotubes.

Catabolite gene activator protein (CAP) in the presence of cAMP stimulates transcription from several operons in E. coli. A cAMP-independent variant, in which alanine-144 is replaced by threonine (CAP-91), is activated by adenosine, which does not activate the wild type CAP. To define the mechanism of the regulation, site-specific mutations in serine-46 of HPr were constructed and then tested for PTS activity. The results indicate that imposition of a negative charge in the region of serine-46 of HPr markedly reduces the activity of HPr as a

phosphocarrier.

Since high levels of cAMP are toxic to E. coli, difficulties have been experienced in constructing strains that overproduce adenylate cyclase which catalyzes the synthesis of cAMP. A plasmid vector suitable for the expression of lethal genes was constructed and the vector DNA was ligated to genomic DNA for adenlyate cyclase. The level of adenylate cyclase in E. coli cells transformed with the recombinant plasmid DNA increased about 7,000 fold, corresponding to 30 percent of the total protein. A relatively simple procedure was devised to purify adenylate cyclase from extracts after hyperexpression to yield nearly homogenous protein.