Final

ANNUAL REPORT OF THE
LABORATORY OF BIOCHEMICAL GENETICS
NATIONAL HEART, LUNG, AND BLOOD INSTITUTE
October 1, 1992 through September 30, 1993

NK-2 Homeobox Gene. The first known step in the zygotic development of part of the Drosophila CNS is the expression of the NK-2 gene. During the past year, proteins that regulate the expression of the NK-2 gene were identified by determining the expression of the NK-2 gene in embryos with mutations in various genes. The results show that the NK-2 gene is activated in the ventral half of the embryo, presumably by dorsal protein, which is distributed in nuclei in a ventral-dorsal concentration gradient. The NK-2 gene is activated but not expressed in the most ventral horizontal stripe of nuclei, the mesodermal anlage, due to repression by snail, a zinc finger protein, or in the adjacent horizontal stripe of nuclei, the mesectodermal anlage, due to repression by single-minded and Enhancer of split m8, which are basic, helix-loop-helix proteins. However, the NK-2 gene is expressed by nuclei in the ventral half of the ventrolateral neurogenic anlage early in Drosophila embryonic development as the nuclei undergo commitment to the neuroblast pathway of differentiation, or soon thereafter. Initially, the NK-2 gene is expressed fairly uniformly in a horizontal stripe of nuclei about 7 nuclei in width on each side that extends over 90% of the length of the embryo. During gastrulation, the horizontal stripe of cells expressing NK-2 is converted to 12 vertical stripes by repression of the NK-2 gene in some cells. Later, 26 clusters of cells that express the NK-2 gene are formed on each side, presumably by repression of the NK-2 gene in additional cells. Therefore, 2 clusters of neuroectodermal cells that synthesize NK-2 RNA are formed per hemisegment that are precursors of many neuroblasts in the ventral nerve cord.

Twenty high-affinity and 13 low-affinity NK-2 binding sites were found in 2.2 kb of DNA from the 5'-upstream region of the NK-2 gene, which suggests that NK-2 protein may be required to maintain the expression of the NK-2 gene. Putative sites for other proteins that overlap or are adjacent to the NK-2 protein binding sites were found. The conversion of neuroectodermal cells to neuroblasts is accompanied by activation of the snail gene in the neuroblasts, thereby repressing activation of the NK-2 gene by dorsal protein. The results suggest that the NK-2 gene receives and integrates information from the ventral-dorsal and anterior-posterior gradients of gene regulators that is needed to generate a pattern of clusters of neuroectodermal cells that synthesize NK-2 RNA that are precursors of different types of neuroblasts.

One of the major goals in neurobiology is to understand how the nervous system is assembled. Studies on the NK-2 homeobox gene led to some novel ideas and to a hypothesis which predicts the overall strategy of the gene program (that is the rules) for the early development of part of the CNS of Drosophila. Every aspect of the hypothesis can be tested experimentally using the NK-2 gene. With a slight modification the hypothesis also applies to the assembly of

part of the mammalian CNS.

Circular dichroism measurements and 1D NMR spectra showed that the $t_m$ for denaturation of the NK-2 homeodomain, NK-2H, is approximately 25°C at pH 4.4 and that denaturation is fully reversible. NK-2H was found to have relatively little α-helical content. No dramatic change in the CD spectra was observed on addition of an oligodeoxynucleotide with a high-affinity NK-2 binding site. NK-2H uniformly enriched with $^{15}N$ was examined by 2D and 3D NMR. The results suggest that NK-2H has a novel homeodomain secondary structure.

Genes Expressed In The Developing Nervous System. Transposition of a P-element that contains the β-galactosidase gene from 1 site in the Drosophila genome to another yielded many transgenic fly lines that express β-galactosidase only in the nervous system during embryonic development. The developmental time and location of β-galactosidase expression then is determined by regulatory signals of the genes that contain the inserted P-element DNA. DNA flanking the P-element insertion sites were cloned from 15 of the most interesting transgenic fly lines and corresponding cDNA clones were obtained and were sequenced partially. Clone 393C-2 was shown to encode Drosophila high-mobility-group protein D (HMG-D), a DNA binding protein. A homologous mammalian protein, HMG-1, recognizes DNA conformation rather than nucleotide sequence; HMG-1 binds to cruciform DNA and to DNA with axial distortion due to cisplatin. The functions of HMG-1 and HMG-D proteins have not been identified; however, the proteins are thought to play a role in chromatin structure. Also the HMG domain has been found in many DNA binding proteins that regulate transcription. We find that the HMG-D gene is expressed ubiquitously during early embryonic development but later in development is expressed exclusively in the nervous system. The homozygous P-element insertion is a lethal mutation and is accompanied by striking morphologic defects in the central nervous system.

Clone 367C-3 DNA corresponds to a gene that encodes a novel zinc finger protein that is expressed in the CNS and anterior sensory organs. The homozygous P-element insertion is a lethal mutation that results in extraordinary morphologic defects in the ventral nerve cord of developing embryos. Clone 7D3C-1 corresponds to a novel Drosophila gene that encodes a member of the kinesin heavy chain gene family. Kinesin functions as a molecular motor for axonal fast transport of organelles or cell membranes on microtubule tracks from soma of neurons towards axon tips. Clone 314-4C-2 encodes a protein that is similar to the human QM protein, an apparent suppressor of Wilm's tumor, a pediatric nephroblastoma. Sequence analysis of cDNAs from other transgenic fly lines suggest that the cDNAs correspond to novel genes expressed in the nervous system.

Mammalian Homeobox and POU-Domain Genes. Approximately 8 kb of a novel mouse homeobox gene, NKx-1, a homolog of the Drosophila NK-1 homeobox gene, was sequenced. The amino acid sequences of the NKx-1 and NK-1 homeodomains differ by only 3 of the 60 amino acid residues. Both proteins also contain an acidic domain. However, most of the other regions of the protein that have been defined

differ markedly. NKx-1 poly $A^+$ RNA was found to be most abundant in 10-day mouse embryos; the abundance progressively decreases thereafter. Northern analysis of poly $A^+$ RNA from adults revealed 1 major band of NKx-1 poly $A^+$ RNA in brain and trace bands in RNA from testes or spleen. The NK-1 gene is expressed in discrete regions of 14-day old mouse embryo mesencephalon and myelencephalon and also in spinal cord, vertebrae, and ribs.

A mouse genomic DNA library was screened with oligodeoxynucleotide probes for novel homeobox genes. Seventy-two positive recombinants were cloned. Thus far five novel homeobox genes have been found. Restriction site analysis of the 72 clones revealed additional classes of clones that have not yet been sequenced. In addition, novel POU-domain genes related to Brain-3 POU-domain cDNA were cloned from mouse and human genomic DNA and were sequenced. Two additional, novel human POU-domain genes related to Oct-3 were cloned and the POU-domain regions were sequenced. Sites of expression of Brain-1, Brain-2, Brain-4, and SKIP POU-domain genes in the mouse nervous system were determined by in situ hybridization as a function of mouse embryo developmental age and also were defined in the adult mouse. Hox 4.1 cDNA and genomic DNA were cloned and the complete Hox 4.1 open RNA reading frame was sequenced.

Regulation of a Calcium Channel $\alpha$-1 Subunit Gene. The $\alpha$-1 subunit of a voltage-sensitive calcium channel previously was shown to be inducible in NG108-15 cells and the expression of the gene was shown to control the ability of the cells to form synapses with striated muscle cells. The 5' -upstream regulatory region of the calcium channel gene was cloned and sequenced. A nucleotide sequence was found that is a powerful activator or an enhancerless chloramphenicol acetyltransferase reporter gene. A protein was found in NG108-15 nuclei that specifically binds to the activating sequence. A cDNA expression library in $\lambda$gt11 was screened for recombinants that direct the synthesis of proteins that bind to the nucleotide sequence and 35 positive clones were obtained. Seven kinds of clones were found that encode proteins that bind to oligonucleotides with appropriate sequence specificity but differ in specificity for double-stranded DNA, or (+) or (-) single-stranded DNA. Further work is needed to determine whether one or more of these proteins regulate the expression of the $Ca^{2+}$ gene channel.

Enhancer and Promotor Selection. During the past year further work has been done on the selective amplification of DNA clones that contain enhancer or promoter nucleotide sequences that activate gene expression. The method is based on the observation that the synthesis of polyoma virus DNA in mouse cells requires viral enhancer sequences that also are required for the synthesis of mRNA from polyoma genes. Mouse genomic DNA fragments were ligated to polyoma DNA that lack the enhancer region of the virus. The E. coli origin of replication and $\beta$-lactamase gene also were inserted in the polyoma coat protein gene. Promoters or enhancers in the mouse genomic DNA inserts that activate plasmid DNA synthesis in mouse cells are able to replicate and hence are selectively amplified; whereas, plasmids that lack functional enhancer sequences do not replicate. Plasmid DNA was harvested from mouse cells that had been transfected and incubated for several days. Recovered DNA then was

amplified in E. coli. The selection method is highly effective; some clones were shown to increase in abundance more than 100,000-fold. Fragments of the recovered DNA inserts were shown to bind proteins from nuclei and to activate the expression of an enhancerless chloramphenicol acetyltransferase reporter gene. Previously, cDNA clones were obtained that encode proteins that specifically bind to oligonucleotide sequences that were identified by the oligonucleotide selection method. Partial sequences of some of the cDNA clones were obtained.

Differentiation of Excitable Membranes and Myofibrils of Striated Muscle. Ventral horn neurons from fetal rats or mice were cocultured with rat striated myotubes. Neurites induce accumulation of acetylcholine receptors on the myotube surface where neurites contact the muscle cells and also reduce the concentration of acetylcholine receptors where neurites contact a pre-existing patch with a high receptor density. Steps in the assembly of functional triads in cultured skeletal myotubes were studied by calcium imaging, immunocytochemistry, and electron microscopy. Developing triads and punctate structures containing ryanodine receptors (calcium release channels) and dihydropyridine receptors were found in 3-day old myotubes. Excitation-contraction coupling was found in cells with only a few of these structures. Changes in the maximal calcium release and uptake rates were observed during the course of myotube development that were correlated with the elaboration of the sacroplasmic reticulum around the myofibrils.

Polyclonal antibodies were obtained directed against a protein from fetal pig brain that induces the formation of acetylcholine receptor aggregates on skeletal muscle cells in culture. The antibodies immunoprecipitate all acetylcholine receptor aggregating activity in crude fractions of brain and spinal cord extracts as well as in purified preparations and recognize a protein with a molecular weight of 120,000.

Genes Associated with Programmed Cell Death. During embryonic development, normal cell turnover, and other physiological and pathological processes, cells die by the activation of a mechanism of self-destruction, termed apoptosis or programmed cell death. The available information shows that apoptosis often requires gene activation and the synthesis of proteins presumably needed for cell death. Radiation and chemotherapeutic agents also elicit cell death by apoptosis. Since little is known about the genes or biochemical mechanisms involved in apoptosis, a study was initiated to identify genes involved in apoptosis in radiation-sensitive fetal rat brain neuroblasts following radiation treatment. It is anticipated that among the genes activated by irradiation will be genes required for neuroblast apoptosis, either for the triggering or for the execution of this process.

Fetal 17 day-old rats were irradiated (or sham-irradiated) with sublethal neutron/gamma irradiation in utero, and RNA isolated from brains 5 and 24 hours later was used to construct a subtractive cDNA library enriched in sequences of transcripts increased by irradiation. Clones from this library were screened and analyzed by

differential colony and Northern blot hybridization. At least 76 out of 682 analyzed clones were found to represent transcripts increased in abundance by irradiation, generally by a factor of 2-3 but much greater for at least one transcript. Sequencing of the inserts of these clones is in progress. Some clones were found to have previously unreported insert DNA sequences, while others represent ubiquitin, ferritin light chain, 12.3 protein (related to G-β proteins) and subunits of mitochondrial NADH dehydrogenase. These results indicate that a variety of transcripts are increased moderately during radiation-induced neuroblast cell death. Additional more strongly activated genes, which are expected in the triggering of cell death, are being sought currently.

Sequence and Characterization of the Hox A7 Gene. The Hox A7 gene is a mouse gene characterized by the presence of a homeobox region typical of the Antennapedia class of proteins found in Drosophila. Starting with a clone derived from mouse genomic DNA, the complete sequence of the structural gene, as well as upstream and downstream regulatory regions, were determined. A variety of studies have led to the assignment in this sequence of the probable transcription start site as well as likely regions for the interaction with proteins that regulate transcription.

Site-specific Mutagenesis of Presumptive ATP Binding Sites in Escherichia coli Adenylyl Cyclase. The region of the enzyme bounded by Alanine 190 and Arginine 197 was studied as a probable region for substrate binding. Site-directed mutagenesis was carried out on Lysine 196 by replacing this residue with a variety of other amino acids. It was found that replacement with other basic amino acids led to retention of some activity, while replacement with other classes of amino acids generally led to abolition of activity. The conclusion from these studies is that the charge in the vicinity of Lysine 196 is crucial for enzymatic activity.

Sequence and Organization of a Monocistronic ptsH Operon in Mycoplasma. A sequencing project, designed to isolate genomic clones for the enzymes of the transport system known as the phosphoenolpyruvate:sugar phosphotransferase system (PTS), was undertaken. Several clones were isolated that allowed the determination of the sequence of the gene encoding the PTS phosphocarrier protein known as HPr as well as of flanking regions upstream and that the gene encoding HPr (the ptsH gene) was located in an operon that is monocistronic. This is in marked contrast to the organization of all previously sequenced ptsH genes in other organisms where the ptsH gene is found in a polycistronic operon.