The molecular basis of mengovirus hemagglutination
by Linda Margaret Mann

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Microbiology
Montana State University

Abstract:
Mengovirus is a small neurotropic picornavirus that may also cause myocarditis and diabetes. The ability of the virus to bind and enter particular cell types is the major determinant of tissue tropism and depends'on surface features of the virus capsid and the presence of suitable cell receptors. Hemagglutination, the ability of the virus capsid to agglutinate erythrocytes, is a model system for cell-virus interactions. The goal of this thesis is to determine the molecular basis of mengovirus hemagglutination.

Mengovirus 37A is a heat stable, hemagglutination-positive variant which is virulent in mice. Two mutants 280 and 205, isolated from 37A by acriflavine mutagenesis, are unable to agglutinate erythrocytes, possess a small plaque phenotype, and are avirulent in mice. Two revertants of 205 have regained the hemagglutination-positive phenotype of 37A but retain the small plaque size and avirulence of mutant 205.

The molecular basis of mengovirus hemagglutination was determined by comparison of the nucleotide sequences determined by cDNA and consensus RNA sequencing of the capsid coding region of 37A, mutants 280 and 205, and revertants 205-A7 and 205-D2.

Two nucleotide differences were found in the 1D capsid between the hemagglutination-positive mengovirus (37A, 205-A7 and 205-D2) and the hemagglutination-negative mutants 280 and 205. These base changes result in the replacement of arginine 231 and proline 232 with lysine and serine residues in the mutants. A computer model of the mengovirus model based on the atomic structure of the M variant of mengovirus predicts that the change from proline to serine at residue 232 will alter the position of residue 231 on the capsid surface.

Two discrepancies were found between the cDNA sequence and the consensus RNA sequence of 205 indicating that sequencing a single cDNA clone is not sufficient to determine a true consensus sequence.

Consensus RNA sequencing was extended into the 5'-noncoding region of the genomic RNAs. A single base difference was found between 37A and 280 and 205, 205-A7, and 205-D2.

Eleven coding changes were identified in the 1B, 1C, and 1D capsid coding regions between 37A and the M variant. Several of these changes are in positions which may affect capsid stability.

THE MOLECULAR BASIS OF MENGOVIRUS HEMAGGLUTINATION

by

Linda Margaret Mann

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Microbiology

MONTANA STATE UNIVERSITY
Bozeman, Montana

June 1990

APPROVAL

of a thesis submitted by

Linda Margaret Mann

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

6-8-90
Date

Chairperson, Graduate Committee

Approved for the Major Department

8 June 1990
Date

Norman D Reed
Head, Major Department

Approved for the College of Graduate Studies

June 15, 1990
Date

Henry L Parsons
Graduate Dean

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute copies of the dissertation in and from microfilm and the right to reproduce and distribute by abstract in any format."

Signature Linda M. Mann

Date 6-29-1990

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES - Continued

LIST OF FIGURES - Continued

ABSTRACT

Mengovirus is a small neurotropic picornavirus that may also cause myocarditis and diabetes. The ability of the virus to bind and enter particular cell types is the major determinant of tissue tropism and depends on surface features of the virus capsid and the presence of suitable cell receptors. Hemagglutination, the ability of the virus capsid to agglutinate erythrocytes, is a model system for cell-virus interactions. The goal of this thesis is to determine the molecular basis of mengovirus hemagglutination.

Mengovirus 37A is a heat stable, hemagglutination-positive variant which is virulent in mice. Two mutants 280 and 205, isolated from 37A by acriflavine mutagenesis, are unable to agglutinate erythrocytes, possess a small plaque phenotype, and are avirulent in mice. Two revertants of 205 have regained the hemagglutination-positive phenotype of 37A but retain the small plaque size and avirulence of mutant 205. The molecular basis of mengovirus hemagglutination was determined by comparison of the nucleotide sequences determined by cDNA and consensus RNA sequencing of the capsid coding region of 37A, mutants 280 and 205, and revertants 205-A7 and 205-D2.

Two nucleotide differences were found in the 1D capsid between the hemagglutination-positive mengovirus (37A, 205-A7 and 205-D2) and the hemagglutination-negative mutants 280 and 205. These base changes result in the replacement of arginine 231 and proline 232 with lysine and serine residues in the mutants. A computer model of the mengovirus model based on the atomic structure of the M variant of mengovirus predicts that the change from proline to serine at residue 232 will alter the position of residue 231 on the capsid surface.

Two discrepancies were found between the cDNA sequence and the consensus RNA sequence of 205 indicating that sequencing a single cDNA clone is not sufficient to determine a true consensus sequence.

Consensus RNA sequencing was extended into the 5'-noncoding region of the genomic RNAs. A single base difference was found between 37A and 280 and 205, 205-A7, and 205-D2.

Eleven coding changes were identified in the 1B, 1C, and 1D capsid coding regions between 37A and the M variant. Several of these changes are in positions which may affect capsid stability.

INTRODUCTION

Picornaviruses, with their small sizes and comparably small genomes, have proved to be valuable models for a variety of biological studies. The application of classical genetic methods to the study of mutated viruses has helped elucidate how genes function. Now, with advances in molecular biology, picornavirus models can be even more informative. The actual nucleotide differences between variants can be determined by sequencing complementary DNA (cDNA) or RNA and infectious cDNAs can be constructed. Using recombinant DNA techniques or *in vitro* mutagenesis, changes can be introduced almost at will into infectious clones making no part of the viral genome inaccessible to study. Major advances have also been made in understanding the physical structure of the picornavirus virion. X-ray crystallographic studies have provided detailed information about the atomic structure of the virion providing insight into how capsids are assembled and held together, how they come apart, how they bind to cells, potential immunogenic sites for vaccine development, mechanisms and design of antiviral agents, and the evolution of biological structures.

Picornaviruses

The Picornaviridae are a large and diverse family of small, nonenveloped RNA viruses which include poliovirus, foot-and-mouth disease virus (FMDV), and human rhinoviruses (HRV), all important

pathogens of man and animals. In the last decade, over a dozen members of the family have been cloned and sequenced (21,22,25,26,91,98, 101,112,120). In the last five years, the atomic structures of several picornaviruses including poliovirus (59), human rhinovirus 14 (HRV14) (103), FMDV (1), and mengovirus (75) have been determined by X-ray crystallography providing a three-dimensional comparison of the capsid structures and surfaces.

Classification

The Picornaviridae were divided into four genera based on physical and serological properities of the virions. Now, as the genomic RNA sequences of more members become available, genomic organization has become an important determinant in their classification. Table 1 below lists the current picornavirus genera and some properties which differentiate them (105).

Table 1. Classification of the Picornaviridae.

Group	pH Stability	Buoyant Density of Virions in CsCl	Poly(C) Tract in Genome	Leader Sequences
Enterovirus	stable	1.34	-	-
Cardiovirus	stable	1.34	+	+
Rhinovirus	labile	1.39-1.42	-	-
Aphthovirus	labile	1.43-1.45	+	+

The enterovirus group contains the coxsackieviruses and echoviruses as well as poliovirus, probably the most studied and best

understood virus of man. Rhinoviruses, the most frequent cause of the common cold, are closely related to the enteroviruses but can be distinguished from them by their acid lability and disease pathology. Like the enteroviruses, cardioviruses are relatively insensitive to acid conditions. They can be differentiated from enteroviruses by the presence of a polycytidylate-rich region, the poly(C) tract, and a leader sequence in their genome. Aphthoviruses also have poly(C) tracts and leader sequences, but, unlike cardioviruses, are easily inactivated by acids. A fifth group, containing the hepatitis A virus, currently classified as an enterovirus, is likely to be added soon since its genomic organization differs considerably from the enteroviruses (128). Theiler's murine encephalomyelitis virus is also currently included in the enteroviruses although it has more sequence homology with encephalomyocarditis virus, a cardiovirus, than with any of the enteroviruses (88,92,98). Unlike the cardioviruses, it lacks a poly(C) tract in the genome. Table 2 lists some other examples from each picornavirus group.

Nomenclature

The nomenclature for the genes and gene products of picornaviruses used in this thesis are those adopted by the European Study Group on the Molecular Biology of Picornaviruses in 1983 (106). The traditional designation of the cardiovirus capsid proteins as α, β, γ, δ, and ε have been replaced by 1D, 1B, 1C, 1A, and 1AB respectively. The equivalent protein designations still often used for poliovirus are VP1 (1D), VP2 (1B), VP3 (1C), VP4 (1A), and VPO (1AB).

Using this system the proper name of the small viral protein linked to the 5'-end of the genomic RNA is 3B, however, since it is almost universally referred to as VPg (for Viral Protein, genome), I will use that term.

Table 2. Members of the Picornaviridae. (adapted from reference 105)

Group	Examples
Enterovirus	poliovirus, 3 serotypes coxsackievirus, type A, 23 serotypes type B, 6 serotypes echovirus, 32 serotypes hepatitis A virus (HAV) Theiler's murine encephalomyelitis virus
Cardiovirus	mengovirus encephalomyocarditis virus (EMCV) Maus Elberfeld virus Columbia SK virus
Rhinovirus	human rhinoviruses (HRV), 100 serotypes
Aphthovirus	foot-and-mouth disease virus (FMDV), 7 serotypes

The Picornavirus Genome

The picornavirus genome is a single-stranded RNA molecule which ranges in size from 7102 bases for HRV2 (93) to 8450 bases for FMDV (22). The RNA is plus strand, messenger-sense RNA with one long open reading frame which can be translated directly into a single, large polyprotein product. The polyprotein coding region has been divided into three regions corresponding to the first protein products seen during poliovirus replication. The P1 region contains the coding region for the 1A, 1B, 1C, and 1D capsid proteins. In cardioviruses and

aphthoviruses, a short leader sequence precedes the P1 region. The P2 and P3 regions code for the nonstructural proteins including viral proteinases and the viral polymerase.

Like cellular messages, the viral RNA is polyadenylated at the 3'-end. The length of the poly(A) region is quite variable both within and between viral variants. Its length ranges from an average length of 35 bases in EMCV (3) to 100 bases in FMDV (105). Infectivity of the viral RNA increases with increased length of the poly(A) region (63,109).

A small 2 kd protein, VPg, is covalently attached to the 5'-end of the genomic RNA via a phosphodiester bond to a phenolic hydroxyl group of a tyrosine residue of VPg (64,123). VPg is not required for infectivity of viral RNA but appears to be important in initiation of viral RNA synthesis (89) and may function in RNA packaging (90).

The 5'-noncoding or nontranslated region of picornaviruses ranges in length from 624 bases in HRV 14 to 1,199 bases in FMDV (105) and is quite long compared to 5'-noncoding regions in cellular messages. This region is highly conserved among the picornavirus groups (93) and stable stem and loop secondary structures are predicted for the region (102). This region has been shown to play a role in determining host range of picornaviruses since cellular initiation factors important in protein synthesis bind to the region (66,97,118,119). A single base change in this region increases the neurovirulence of the Sabin strain of poliovirus type 3 (45). Other point mutations can alter translation efficiency of the RNA (115).

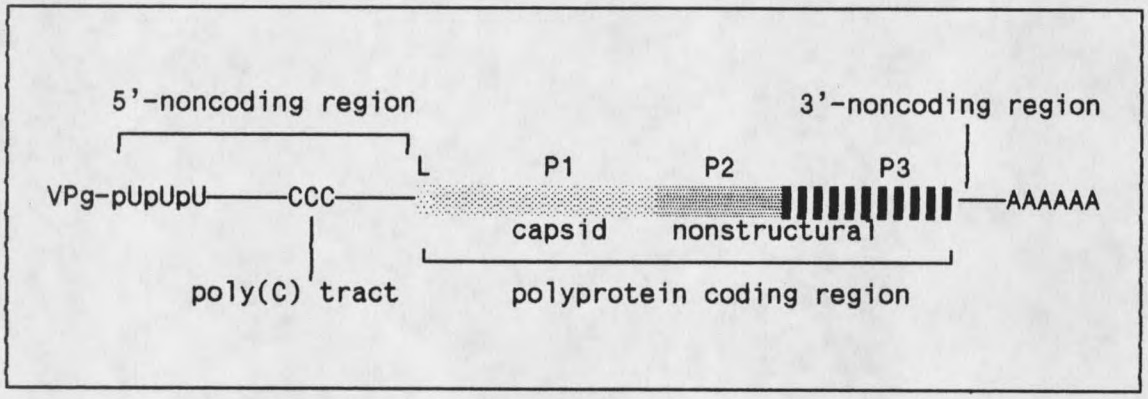
The genomes of cardioviruses and aphthoviruses contain a poly(C) tract (15,50) located about 150 bases from the 5'-end of the RNA between

VPg and the site of initiation of translation (96). The size of the poly(C) tract varies in length from 80 to 250 bases in EMCV, from 60 to 90 bases in mengovirus, and is even longer in aphthoviruses (13). The poly(C) tract is thought to function in pathogenicity since reducing the length of the tract decreases the *in vivo* virulence of infectious clones (41,86).

The 3'-noncoding region is shorter than the 5'-noncoding region and ranges in size from 47 bases in HRV14 to 126 bases in EMCV (105). The function of this region is unknown but alteration of the region can change the phenotype of the virus (110).

Figure 1 shows the basic structure of the picornavirus genome.

Figure 1. The general structure of picornavirus genomic RNA. The poly(C) tract and leader (L) sequence are found only in cardioviruses and aphthoviruses.



Translation and Processing of Viral Proteins

The translation product of the single long open reading frame of the genomic RNA is a single, large polyprotein that is post-translationally cleaved to produce the P1, P2, and P3 proteins. These initial products undergo additional processing to produce the structural

and nonstructural proteins required for capsid formation, polyprotein processing, and viral RNA replication and packaging (87,121).

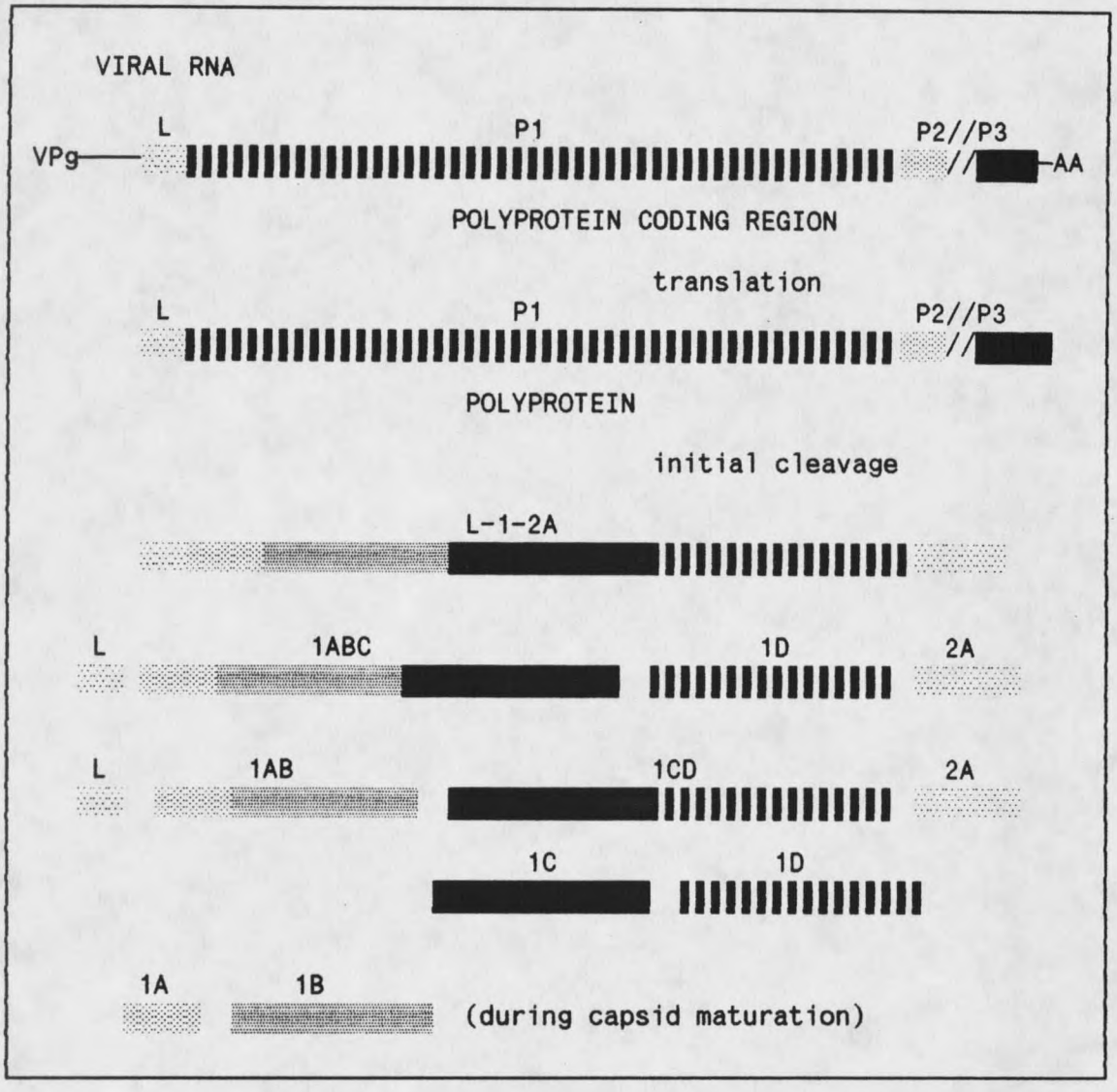
The capsid proteins 1A, 1B, 1C, and 1D are produced by cleavage of the P1 product. Figure 2 shows the posttranslation processing of the cardiovirus capsid proteins in more detail. The amino- and carboxyl-termini of the capsid proteins, except for the amino terminus of 1A which is blocked by the myristic acid residue, have been determined by partial protein sequencing of purified capsid proteins (127). Except for the carboxyl terminus of the 1D protein, which has been found to vary depending on the virion isolation procedure (14) the protein termini are not trimmed or modified after cleavage.

The leader-1A and 1C-1D cleavages occur at glutamine-glycine amino acid pairs and the 1B-1C cleavage occurs at a glutamine-serine amino acid pair. Both types of cleavages result from the activity of a viral proteinase, $3C^{pro}$. The 1D-2A cleavage occurs at a glutamic acid-cysteine amino acid pair and may involve a separate proteinase activity.

The 1A-1B cleavage at an alanine-aspartic acid amino acid pair does not occur until capsid maturation when genomic RNA is packaged (65). Analysis of the three-dimensional structure of the picornavirus capsid has led to a suggestion that a serine residue in 1B may act in conjunction with virion RNA in producing this cleavage (103).

In aphthoviruses, the cleavage of the leader appears to be an autocatalytic event (65). In the cardioviruses, this cleavage is thought to result from $3C^{pro}$ activity (95). The function of the leader sequence is still unclear.

Figure 2. Posttranslation processing of mengovirus capsid proteins.



The P2 region codes for the 2A, 2B, and 2C nonstructural proteins. In poliovirus, the 2A protein appears to have two functions. It has been identified as the proteinase responsible for inactivation of the host cell translation initiation factor eIF-4F and shutoff of host protein synthesis. The 2A protein of poliovirus is also responsible for the P1-P2 cleavage during posttranslation processing of the

polyprotein. Neither of these 2A activities have been found in the cardioviruses (74). In the aphthoviruses and cardioviruses, the initial cleavage occurs between 2A and 2B (87). The roles of the other products of the P2 region are not well understood. They are thought to function in RNA replication (74).

The P3 region codes for the 3A, 3B, 3C ,and 3D nonstructural proteins. The role of 3A is unclear. The 3B protein is VPg, the small protein covalently attached to genomic RNA. Aphthoviruses are unique in this region having 3 copies of 3B in tandem (49). The 3C protein is a cysteine proteinase, $3C^{pro}$, which is responsible for the majority of the cleavages in the processing of the polyprotein (121). Protein 3D is the RNA dependant RNA polymerase, $3D^{pol}$, which is the RNA-dependent RNA polymerase (48). In poliovirus-infected cells, the alternative cleavage products 3C' and 3D' are found (73). The role of these alternative products is unclear and they are not found in the cardiovirus-infected cells (19,20).

The Picornavirus Capsid

The picornavirus virion has a diameter of about 30 nm and is composed of a protein capsid which surrounds and protects an RNA core. Early X-ray diffraction patterns of virus crystals determined that the picornavirus capsid had icosahedral symmetry indicating that the virion was composed of identical protein subunits with fivefold, threefold, and twofold axes of symmetry (47).

The capsid is assembled from protomers composed of one copy of 1AB, 1C, and 1D. Five protomers associate to form a pentamer and 12 pentamers associate to form the procapsid. When RNA is packaged, the

final cleavage of 1AB to 1A and 1B occurs producing a mature and infectious virion (65). Sixty copies of each of the 1A, 1B, 1C, and 1D capsid proteins are found in the mature virions. The cleavage of 1AB, the precursor of 1A and 1B, is not complete and traces of 1AB, averaging two molecules per capsid, are also found (104).

The sizes of the capsid proteins of poliovirus are 33.5 kilodaltons (kd) (1D), 30.0 kd (1B), 26.4 kd (1C), and 7.39 kd (1A) but the sizes vary between the picornavirus genera (105). In aphthoviruses 1D, 1B, and 1C are nearly equal in size (114). The 1A protein of HAV is much smaller at 1.73 kd than the 1A proteins of other viruses in the family (25). The 1A and 1AB proteins have hydrophobic myristic acid residues covalently attached to their 5'-ends (23).

X-ray crystallography of virions has shown that the basic structure of the 1B, 1C, and 1D proteins are similar to each other and to the capsid protein of icosahedral plant viruses (59,103). The basic structure is a wedge-shaped β-barrel formed by the interaction of eight regions of β-sheet structure in the capsid proteins. Two conserved α-helical regions are also seen along with less conserved regions forming loops which connect the β-sheet regions. The 1D protein is the most surface exposed capsid protein while the 1A protein is buried within the virion.

The Picornavirus Multiplication Cycle

The multiplication cycle of mengovirus and other picornaviruses begins with attachment of the capsid to cell surface receptors. Initially the attachment appears to be rather loose and reversible. The attachment becomes progressively tighter causing irreversible

conformational changes in the capsid structure. Virions which are released at this stage have lost the 1A protein from the capsid and are no longer infectious (61). Receptor-mediated endocytosis appears to play a role in virus entry into cells since infection is inhibited by agents which interfere with that process (77). How the RNA leaves the capsid is still speculative but it has been suggested that the RNA is released through an opening at the vertex of a capsid pentamer with the extruded 1A molecules forming a pore through the endocytotic membrane via hydrophobic interaction of the myristate tails with the membrane lipids (59,103).

The viral RNA functions both as a template for viral RNA synthesis and as a messenger RNA for translation by the cellular protein synthesis machinery. Before replication can begin, the virus-specific proteins needed for replication of the RNA and processing of the polyprotein must be produced by the translation of the virion RNA. One product is the viral RNA dependent RNA polymerase, $3D^{pol}$, which forms a replication complex with other viral and cellular factors. The replication complex is associated with the smooth endoplasmic reticulum and initially functions to produce minus-strand copies of the RNA. From the minus-strand copies, additional plus-strand copies for translation and packaging as virion RNA are made.

As the infection continues a pool of structural proteins accumulates in the cytoplasm and self assemble into procapsids. When viral RNA covalently linked to VPg is packed into the procapsid, the final cleavage occurs in the 1AB procapsid protein producing mature virions. The virions are released by lysis of the cell.

Host Range and Tissue Tropism

The ability of a virus to infect and multiply in a particular organism determines its host range. Cell tropism refers to the ability of the virus to preferentially infect and multiply in certain cell types within the host. A major determinant in the host range and tissue tropism of a virus is the ability of the virus to adsorb to and enter particular types of cells (60,79). Binding characteristics are dependent on properties of the capsid surface of the virion and the presence of suitable cell receptors on the target cell.

The determinants of host range and cell tropism are best understood for poliovirus (30). The normal host range of poliovirus is restricted to primates but variants have been adapted to mice, rats, hamsters, and chick embryos. Poliovirus normally infects and multiplies in cells of the nasopharynx and gut of man but can readily infect and damage cells of the central nervous system. Attenuated poliovirus strains are less able to infect neural cells but retain the ability to infect and multiply in the nasopharynx and gut (100). The host range of poliovirus is determined by the presence of suitable cellular receptors on primate cells (60). A wide range of other cell types which lack the primate receptor can support a single round of poliovirus replication if the requirement for cell attachment is bypassed by transfection of viral RNA directly into the cell (62).

Alteration of the poliovirus capsid proteins can alter their ability to bind cellular receptors. The Mahoney type 1 poliovirus is virulent for man but does not cause disease in mice even when inoculated directly into the brain. Replacement of 6 amino acids in the 1D protein

with the corresponding residues found in the Lansing type 2 poliovirus variant, a neurovirulent mouse variant which is no longer virulent for man, gave the virus the ability to bind mouse cells and altered the host range and neurovirulence to include mice (85).

Alteration of capsid proteins has been found to affect the host range of rhinoviruses. When host range mutants human rhinovirus 14 (HRV14) were examined, alterations were found in the nucleotide sequence of the capsid coding region (125).

The factors that determine neurovirulence are less well understood. The study of the molecular changes responsible for the attenuation of the poliovirus vaccines have shown that neurovirulence maps to several regions of the poliovirus genome (24). Both coding and noncoding mutations are involved. Comparison of the genomic RNA sequences of the virulent Mahoney poliovirus type 1 variant and the avirulent Sabin poliovirus type 1 variant indicate that they differ at 55 nucleotide positions (91). Analysis of the effect of these changes in recombinant poliovirus clones show that while changes affecting attenuation occurred throughout the genome, the strongest attenuation was obtained in recombinants with a single base change at position 480 in the 5'-noncoding region of the genome (69).

Only ten sequence differences were found between the avirulent Sabin poliovirus type 3 and the Leon poliovirus type 3, its virulent parent. Each mutation was incorporated into an infectious cDNA and analyzed separately for its effect on virulence (122). Two strongly attenuating mutations were found. One occurred in the 5'-noncoding region at position 472. *In vivo* confirmation of the role of this

mutation in attenuation has been found in isolates isolated from cases of vaccine-associated poliomyelitis. In several cases the only back mutation found in these neurovirulent isolates is at position 472 (45). The second strongly attenuating mutation is a coding change in the 1C capsid protein which results in the replacement of a serine residue in the virulent variant with phenylalanine.

Cellular Receptors and the Canyon Hypothesis

The three-dimensional structural analysis of the surface of the HVR14 (103) and poliovirus (59) capsids reveals the presence of a canyon or cleft surrounding each of the fivefold vertices. The placement of the cellular receptor binding site in a cleft inaccessible to antibodies helped explain the conservation of the cellular receptor binding site under the pressure of the immune system. The canyon hypothesis proposed that the cellular receptor binding site was within the canyon (103). Evidence for this hypothesis has been obtained by several means. When the residues of the canyon of HRV14 are changed by site directed mutagenesis, the binding affinity to cellular receptors is changed (27). The binding of antiviral agents in a hydrophobic pocket of the canyon of HRV14 alters the canyon conformation and inhibited binding to cellular receptors (99).

The major cellular receptor for HVR14 has been identified as the intercellular adhesion molecule 1 (ICAM-1), a member of the immunoglobulin superfamily (52,113,116). ICAM-1 is able to bind to an integrin molecule on lymphocytes and is known to facilitate antigen presentation to T-cells. The receptor for poliovirus has recently been

cloned and sequenced (82) and is also a member of the immunoglobulin superfamily.

Several members of the picornavirus family including group B coxsackieviruses, echoviruses, rhinovirus, FMDV, EMCV, and mengovirus agglutinate erythrocytes (105). The significance of hemagglutination is unknown since attachment to erythrocytes is not part of the disease process.

The Mengovirus Model

Mengovirus, a small neurotropic picornavirus, continues to be a good model system for the examination of virus structure and function. Variants of mengovirus have been isolated which cause myocarditis (117) or diabetes (126). The determination of the three-dimensional capsid structure makes examination of mutations affecting capsid structure particularly rewarding. The mengovirus cellular receptor molecule has not been identified but its binding probably involves sialic acid residues of the receptor since free sialic acid binds in the putative cellular receptor binding site of the capsid (70). Mengovirus is able to agglutinate erythrocytes from several different animal species, including type O cells of man. Hemagglutination, while not required for mengovirus infection, is the result of the interaction of the viral capsid with cellular surface molecules and provides a model for virus-cell receptor interaction.

Discovery and Isolation

Mengovirus was first isolated in 1946 by Dick, Smithburn, and Haddow (36,38) from a paralyzed rhesus monkey at the Yellow Fever

Research Institute in Entebbe, Uganda. Named mengo encephalomyelitis virus after the Mengo district in which it was isolated and the disease produced in infected animals, it was recovered over the next two years from *Taeniorhynchus* mosquitos, a mongoose, another rhesus monkey and man.

Only one human case of mengovirus encephalomyelitis has been reported (37). G. W. A. Dick, one of the scientists who first isolated the virus, also contracted it either naturally or as the result of laboratory exposure. The incubation period was estimated to be between 5 and 9 days. The initial symptoms included fever, headache, and irritability that progressed to delirium which lasted several days. Late in the infection, he developed some weakness in muscles on the right side and deafness of the right ear. Aside from slight residual nerve deafness in one ear, the eventual recovery was complete.

A small survey on human sera, done at the Yellow Fever Research Institute at the time mengovirus was first isolated, showed only a small percentage (0.8%) of the population of the surrounding area to have antibodies specific for mengovirus.

Mengovirus is pathogenic for mice and was first isolated from spinal tissue from the paralyzed monkey by passage in mice. Initial isolates killed most animals in an average of 4 days when inoculated intracerebrally. The first symptom of infection was paralysis of one or more hind limbs. Later isolates, adapted by continuous passage in mice, produced more rapid paralysis and death, killing most animals within two days.

Serologically, the members of the cardiovirus group are indistinct. Despite difference in biological activity and receptor attachment, mengovirus has been considered to be closely related to EMCV (84). Infectious RNAs have been prepared from hybrid clones linking the first 299 bases from the 5'-end of mengovirus with a 7424 base EMCV sequence (41). Strains of both viruses have now been sequenced (Palmenburg and Duke, personal communication, (94) and have a 93% peptide sequence identity (93) but share less than 80% overall nucleotide identity (41).

The Atomic Structure of Mengovirus

The atomic structure of the M variant of mengovirus has been determined to a 3 angstrom (Å) resolution by X-ray crystallographic studies by Luo et al. (75). This study confirms the basic structural similarity of the capsid proteins of mengovirus to other picornaviruses such as poliovirus and human rhinovirus 14 (HRV14). The three capsid proteins 1D, 1C, and 1B share the eight-stranded antiparallel β-barrel structure. Major differences between HRV14 and mengovirus in regions of 1D not involved in the β-barrel structure increase the exposure of the 1D protein on the surface of the mengovirus virion. The addition of two regions in 1D, loop I and II, and modification of the puff region in 1B change the putative cellular receptor region from the canyon-like structure in HRV14 to a series of five pits 22 Å deep by 30 Å wide on the virion surface. Capsid protein 1A of mengovirus, which is entirely internal, has an altered relationship with the other capsid virions compared to HVR14. However, position of the carboxyl-terminus of 1A,

which is generated by the autocatalytic cleavage of 1AB to 1A and 1B during maturation of the virion, is conserved.

Relationship of the M and 37A Variants of Mengovirus

The M variant of mengovirus used in the X-ray crystallography determination of capsid structure and the 37A variant used in this work were derived from the same stock in the early 1960s but have been separated since that time. They share an ability to agglutinate human O cells but differ markedly in their virulence for mice.

The M (medium plaque size) variant of mengovirus was isolated by Ellem and Colter in the early 1960s (43) from stock obtained from Smithburn (38). The M variant was differentiated from the S (small plaque size) and L (large plaque size) variants by its intermediate plaque size, plaque morphology in L cells, and ability to agglutinate human O erythrocytes. The M variant was relative nonvirulent (28) when injected intraperitoneally into mice with a LD_{50} of 10,000 to 50,000 plaque forming units (PFU) compared to a LD_{50} of 1 PFU for the L variant by the same route. When inoculated intracerebrally, all three variants had LD_{50}s of less than 10 PFU.

The genome of the M variant of mengovirus has been sequenced (Dr. Ann Palmenberg, personal communication) and is about 7800 bases in length with a polyprotein region of 6879 bases coding for a polyprotein of 2283 amino acids. The 5'-noncoding region is 758 bases long. The poly(C) tract contains 60 C residues (41) and has the sequence $C_{50}UC_{10}$. The 3'-noncoding region is 124 bases long.

The 37A strain of mengovirus is a heat stable variant derived by Brownstein and Graham (16) also in the early 1960s from wild-type mengovirus stock obtained from Colter. Like the M variant, 37A also possesses the ability to agglutinate human O erythrocytes. Unlike the M variant it is virulent in mice with a LD_{50} of 1500 PFU when injected intraperitoneally and a LD_{50} of 7 PFU when injected intracranially (7).

Mutants and Revertants of the 37A Mengovirus Variant

Mengovirus mutants 205 and 280 were isolated from cells infected with the 37A variant of mengovirus and treated with acriflavine by Dr. M. A. Gill (personal communication). Selected for small plaque size in temperature shift analysis of infected L cell monolayers, they were originally thought to be temperature-sensitive mutants. Further studies were unable to demonstrate significant, temperature-dependent differences in virus yields or adsorption to cell monolayers between the mutants and the parental 37A variant indicating the mutation or mutations are not dependent on temperature (7).

Extensive biological characterization of 37A and mutants 280 and 205 was undertaken by Anderson and Bond (7). There are significant differences in phenotypes between the mutants and 37A. These differences are summarized in Table 3 below. The mutants have a significantly smaller plaque size than 37A and have lost the ability to agglutinate human O erythrocytes. The mutants were avirulent in mice by either the intraperitoneal or intracranial route as indicated by the difference in the LD_{50} data. Mice injected intracranially with 10^7 PFU of either 205 or 280 showed no signs of infection.

When the extent of multiplication of mutants 280 and 205 in brain tissue was determined, it was observed that mice infected with mutant 205 had detectable titers more often than mice infected with 280. Plaque isolates from the brain suspensions were tested to determine their hemagglutination phenotype. All the isolates from mice infected with 37A were hemagglutination-positive as expected. None of the isolates from 280 were hemagglutination-positive. When the isolates from mice infected with 205 were tested, 20 of the 36 isolates were hemagglutination-positive.

Table 3. Biological differences between 37A and mengovirus mutants 205 and 280 (adapted from reference 7).

Property	37A	205	280
Plaque size[a]	3.99 ± 1.08	1.56 ± 0.37	1.49 ± 0.44
Hemagglutination Titer[b]	2048	0	0
Ability to bind glycophorin[c]	+	-	-
LD_{50}[d]			
ip	1500	$> 10^7$	$> 10^7$
ic	7	$> 10^6$	$> 10^6$

[a] Mean plaque diameter in mm measured at 48 hours.
[b] Hemagglutination titers are expresses as the reciprocal of the end point dilution.
[c] As measured by a change in migration of purified virions mixed with glycophorin in sucrose density gradients.
[d] LD^{50} titers are expressed as PFU injected intraperitoneally (ip) or intracranially (ic).

The plaque sizes of the isolates from mice infected with 205, whether hemagglutination-positive or negative were significantly smaller

than 37A. Mice injected intraperitoneally with 10^4 PFU of the hemagglutination-positive isolates from mice infected with 205 showed no symptoms of disease. Mice infected by the same route with 10^6 PFU all had symptoms of infection typical of mengovirus infection by 5 days and over 80% had died by 8 days post infection. These data indicate that a partial reversion had occurred *in vivo* restoring the ability of mutant 205 to agglutinate erythrocytes without restoring the virulence or plaque size found in 37A. No revertants of mutant 280 have been isolated.

RNA synthesis follows a similar time course in all the variants but 37A produces 10-fold more RNA than 205 or 280 and 2-fold more RNA than 205-A7 or 205-D2. Synthesis of virus-specific proteins begins sooner in 37A than in 205 or 280.

The mutants and revertants also differ from 37A in their ability to bind to BHK21 cells. The absorption of the mutants 205 and 280 are only 23% and 34%, respectively, of the adsorption of 37A. The adsorption of the revertants is greater than 205 but less than 37A. The adsorption of 205-A7 and 205-D2 are 55% and 59% respectively.

Differences were also seen in the ability of mengovirus-specific ascitic fluid, produced against 37A, to neutralize the 205 and 280 mutants and the 205-A7 and 205-D2 revertants. The results of neutralization experiments are expressed in PN_{50} units. The PN_{50} unit is the reciprocal of the dilution of the antiserum able to neutralize 50% of the plaques produced by a dilution of the virus. The PN_{50} values for the mutants were significantly higher ($p \leq 0.08$) than the PN_{50} value for

for 37A and the revertants indicating that the surfaces of the mutant virions were altered.

Mengovirus Hemagglutination

EMCV, like mengovirus, can agglutinate human O erythrocytes. Burness found that glycophorin A, the major sialoglycoprotein found on the cell membranes of human erythrocytes, was the receptor on the erythrocyte surface for EMCV (5,17,18). Studies using affinity chromatography show that the binding of glycophorin to the EMCV capsid involved particular regions of the glycophorin A molecule and that the binding was a result of multiple weak, ionic interactions (5,6).

Glycophorin is also the receptor for mengovirus hemagglutination (7). When erythrocytes were treated with neuraminidase to remove sialic acid residues, they were not agglutinated by 37A in the standard hemagglutination assay suggesting that sialic acid residues like those present on glycophorin were involved in hemagglutination. Sialic acid alone has no effect on hemagglutination of 37A but when purified glycophorin isolated from human O erythrocytes was incubated with the virus prior to the hemagglutination assay hemagglutination of 37A was inhibited. Purified glycophorin was added to 37A, 280, or 205 and the mixtures analyzed on sucrose density gradients, glycophorin was found to bind to and alter the mobility of 37A but not 205 or 280 in the gradients.

To determine which of the capsid proteins was involved in glycophorin binding, glycophorin was reacted with toluene-2,4-diisocynate (TDI) a heterobifunctional crosslinking agent. The TDI-treated glycophorin was then allowed to bind to purified 37A virions and

the TDI was activated by increasing the pH to 9.6 to covalently cross-link the bound glycophorin with the nearby capsid proteins. When the capsids were disrupted with SDS and heat, and analyzed by SDS-PAGE, the glycophorin-protein complexes could not enter the gels but the amount of 1D protein resolved by the gels was substantially decreased in the lanes containing virion proteins cross-linked with glycophorin. A smaller decrease in the amount of 1C protein resolved on the gels was also observed. This indicated that the major binding site of glycophorin on the mengovirus capsid is located on the 1D protein and that the 1D protein is the major determinant of hemagglutination.

Structural Analysis of Mengovirus Proteins

A number of structural analyses of purified mengovirus capsid proteins were done to identify the location of changes in the capsid proteins between 37A and the mutants 205 and 280 (8). No changes are seen in the protein sizes on SDS-PAGE gels, sizes of tryptic peptide separated by high pressure liquid chromatography, or surface exposure of the capsid proteins by surface labeling of tyrosine residues, or in two dimensional analysis of chymotryptic peptides of 1D.

Differences are seen on two-dimensional gels when isoelectric focusing or nonequilibrium pH gradient gel electrophoresis is done in combination with SDS-PAGE separation of the capsid proteins. In gels of 37A capsid proteins, four species of 1D, differing in charge, are resolved. The 205 mutant has only three species of 1D. The 280 mutants has all four 1D species seen in 37A but the distribution of the species is different from the distribution seen in 37A.

These studies implicate the 1D protein in hemagglutination and suggest that the mutation in 205 is different from the mutation in 280.

Consensus RNA Sequencing

Viral RNA genomes are known to be heterogeneous (39,40). The heterogeneity is thought to arise during RNA replication. The frequency of errors which occur during replication of the RNA genome is estimated to be in the order of 10^{-4} per base position resulting in an estimated rate of 1 error per genome (39,71). Any population of picornavirus virions, even plaque purified stocks, would be expected to contain a large number of variants. The presence of the variants increases the ability of the virus to adapt readily to new environments and so may be of value to the virus. Of course, some of these variants will contain changes of no value and some will be lethal mutations.

DNA sequencing using the dideoxynucleotide termination method (108) is a relatively straight forward and common laboratory procedure today. There are a variety of good commercial kits and the choice of enzymes is still increasing. RNA can also be sequenced using a dideoxynucleotide termination method, however, the choice of an enzyme is limited to reverse transcriptase. To take advantage of DNA sequencing methods to sequence an RNA template, one must first produce cDNA copies of the RNA. The cloning of RNA as cDNA results in the selection of individual molecules from the population. Any errors contained in that RNA will be duplicated in the cloning process. In addition, cloning itself may introduce errors.

Direct sequencing of viral RNA is attractive since it avoids selection of individual RNA molecules and has been successfully used to sequence the viral RNA of influenza virus (4,32). If the number of errors in the genome is expected to average one per genome, many changes can be expected in any RNA population. However, the number of changes which affect any one nucleotide will be few. Minor variations in the sequence, since they affect only a few molecules of the population, will not be seen on sequencing gels among the much stronger bands produced by the majority of the molecules.

Goals and Experimental Design

The goal of my research was to investigate the molecular basis of mengovirus hemagglutination by determining the sequence differences between 37A, a hemagglutination-positive variant, and 205 and 280, hemagglutination-negative mutants derived from 37A. Initially I planned to produce cDNA clones of the capsid coding region of each virus and to sequence multiple clones of each virus to determine consensus sequences for each variant. After the cloning and sequencing process was begun, the mengovirus M variant genomic RNA nucleotide sequence and a source of relatively inexpensive synthetic oligonucleotide primers became available making it possible for me to directly sequence the viral RNAs.

During infection with 205 but not 280, hemagglutination-positive revertants were recovered from mouse brain suspensions. The capsid coding regions of two revertants of the 205 mutant, 205-A7 and 205-D2, were also sequenced in an attempt to understand the mechanism of the reversion which restored the hemagglutination-positive phenotype without

restoring virulence or plaque size of the variants. The 5'-noncoding region of the poliovirus has proved to be an important determinant in neurovirulence. When no sequence differences were found between the capsid coding regions of the revertants and 37A or between 205 and 280, additional sequencing of the 5'-noncoding regions of 37A, 205, 280, 205-A7, and 205-D2 was done.

The sequence data was used to predict the amino acid sequences of the capsid proteins and to identify changes which correlate with the ability to hemagglutinate human O erythrocytes. Analysis of the changes responsible for the hemagglutination-defective phenotype in the computer model of the mengovirus capsid predict the corresponding changes in the tertiary structure of the capsid.

METHODS AND MATERIALS

Materials

Radioisotopes were purchased from New England Nuclear Corp. Reagent grade liquid organic chemicals were obtained from J. T. Baker Chemical Co. Unless otherwise noted, all other chemicals were obtained from Sigma Chemical Co.

Unless indicated otherwise, restriction enzymes were obtained from Boehringer Mannhein and the buffers supplied by the manufacturer used for digests. Avian myeloblastosis virus reverse transcriptase XL (AMV-RT) for cloning was purchased from Life Sciences. AMV-RT used for sequencing, ribonuclease H, DNA polymerase I, exonuclease III, proteinase K, S1 nuclease, and terminal deoxynucleotide transferase (TdT) were obtained from Boehringer Mannhein. T_4 polynucleotide kinase and T_4 DNA ligase were obtained from Promega. Calf intestinal phosphatase (CIP) was purchased from Pharmacia. Ribonuclease A and lysozyme were obtained from Sigma Chemical Co.

The K/RT DNA Sequencing kits were purchased from Promega. The Sequenase DNA sequencing kits (Ver 1.0 and 2.0) were purchased from United States Biochemical Corp. The GemSeq Transcript Sequencing kit reagents used for RNA sequencing were purchased from Promega.

Phenol was purchased from Sigma Chemical Co., redistilled in the laboratory, and stored in the dark at -20°C. The working supply of phenol, stored at 4°C, was equilibrated with 37% (vol/vol) 50 mM

Tris[hydroxymethyl]aminomethane-HCl (Tris-HCl), 10 mM ethylenediamine tetraacetic acid (EDTA) (pH 8.0), and contained 7 mM 8-hydroxyquinoline. Chloroform used for phenol and chloroform extractions contained 4% (vol/vol) isopentanol (Sigma).

The water (diH_2O) used was purified by reverse osmosis (Millipore RO20) and deionized (Gilford Chemlyte).

Other reagents used are listed in Table 4.

Table 4. Reagents.

Reagent	Components
PBS (pH 7.2)	0.135 M NaCl, 2.7 mM KCl, 0.9 mM $CaCl_2$, 0.5 mM $MgCl_2$, 1.5 mM KH_2PO_4, 8 mM Na_2HPO_4
Scintillation fluid	5 g diphenyloxazole per liter xylene
Isolation and Purification of RNA	
Bu3	10 mM Tris-HCl (pH 7.4), 10 mM EDTA, 50 mM NaCl
NET	10 mM Tris-HCl (pH 7.4), 1 mM EDTA, 100 mM NaCl
Cell lysis buffer	NET with 1% (vol/vol) NP40 (Particle Data Laboratories)
10 mM TE (pH 7.4)	10 mM Tris-HCL (pH 7.4), 1 mM EDTA
Isolation and Purification of Plasmid DNA	
LB	10 g tryptone (Difco), 10 g NaCl, 5 g yeast extract (Difco) per liter diH_2O
LA	LB with 15 g Bacto-agar (Difco) per liter diH_2O

Table 4, continued.

Reagent	Components
Minimal agar	6 g Na_2HPO_4, 3 g KH_2PO_4, 0.5 g NaCl, 1 g NH_4Cl, 15 g Bacto-agar, 1 ml 1 M $MgSO_4$, 10 ml 10 mM $CaCl_2$, 10 ml 20% (wt/vol) glucose, 1 ml 10 mg/ml thiamine per liter diH_2O
TNE	10 mM Tris-HCl (pH 8.0), 10 mM EDTA, 100 mM NaCL
50 mM TE	50 mM Tris-HCl (pH 8.0), 50 mM EDTA
Lysozyme solution	5 mg lysozyme per ml 50 mM TE (prepared just prior to use)
Lysis solution	1% (wt/vol) SDS, 0.2 N NaOH (prepared just prior to use)
K acetate	29.4 g potassium acetate, adjust pH to 5.0 with glacial acetic acid, bring volume to 100 ml with diH_2O
RNase A solution	10 ug DNase-free RNase A per ml NET
Sucrose solution	20% (wt/vol) sucrose in 50 mM TE
Sephacryl column buffer	10 mM Tris-phosphate (pH 8.0) 1 mM EDTA, 1 M NaCl
PEG	13% (wt/vol) PEG (mol wt 8000) in water
Agarose Gel Electrophoresis	
Gel loading buffer	0.25% (wt/vol) bromphenol blue (Fisher Scientific), 0.25% (wt/vol) xylene cyanole FF (Eastman), 30% (vol/vol) glycerol
10X TAE	400 mM Tris-acetate pH 7.8, 20 mM EDTA

Table 4, continued.

Reagent	Components
cDNA Synthesis and Cloning	
10X Tailing buffer	2 M potassium cacodylate (pH 6.9), 10 mM DTT, 10 mM dCTP, 20 mM $MnCl_2$,
2X Annealing buffer	20 mM Tris-HCl pH 7.6, 20 mM EDTA, 300 mM NaCl
SOB	20 g tryptone, 5 g yeast extract, 0.6 g NaCl, 0.2 g KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ per liter diH_2O
SOC	SOB with 20 mM glucose (pH 6.8-7.0)
Soft agar	LB with 7.5 g Bacto-agar per liter
Analysis of cDNA Clones	
20X SSPE	3.6 M NaCl, 20 mM EDTA, 0.2 M $NaPO_4$ (pH 7.7)
20X SSC	3 M NaCl, 0.3 M sodium citrate (pH 7.0)
50X Denhardt's reagent	1% ficoll (type 400), 1% BSA 1% polyvinylpyrrolidone
Prewash solution	50 mM Tris-HCl (pH 8.0), 1 M NaCl, 1 mM EDTA, 0.1% (wt/vol) SDS
10X Kinase buffer	500 mM Tris-HCl (pH 7.5), 50 mM DTT, 10 mM $MgCl_2$, 1 mM spermidine, 0.1 mM EDTA
10 mM TE (pH 8.0)	10 mM Tris-HCL (pH 8.0), 1 mM EDTA
Low salt buffer	20 mM Tris-HCl (pH 7.4), 1 mM EDTA, 200 mM NaCl
High salt buffer	20 mM Tris-HCl (pH 7.4), 1 mM EDTA, 1 M NaCl

Table 4, continued.

Reagent	Components
Subcloning Reactions	
10X CIP buffer	0.5 M Tris-HCl (pH 9.0), 1 mM $ZnCl^2$, 10 mM $MgCl_2$, 10 mM spermidine
10X Ligation buffer	0.3 M Tris-HCl (pH 7.8), 0.1 M DTT, 0.1 M $MgCl_2$, 4 mM ATP
Exonuclease III Reactions	
10X Exonuclease III buffer	660 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1 mM 2-mercaptoethanol
10X S1 nuclease buffer	330 mM sodium acetate (pH 4.5), 500 mM NaCl, 0.3 mM $ZnSO_4$
S1 mix	15 ul 10X S1 nuclease buffer, 45 U S1 nuclease, diH_2O to 150 ul total volume
S1 Stop buffer	500 mM Tris-HCl (pH 8.0), 125 mM EDTA
10X Klenow buffer	200 mM Tris-HCl (pH 8.0), 70 mM $MgCl_2$
Klenow mix	2.5 ul 10X Klenow buffer, 2.5 U Klenow, diH_2O to 25 ul final volume
dNTP mix	125 uM dATP, 125 uM dCTP, 125 uM dGTP, 125 uM dTTP
10X Ligase buffer	500 mM Tris-HCl (pH 7.6), 100 mM $MgCl_2$
50% PEG	50% (wt/vol) PEG (mol wt 8000)
Ligation mix (for 18 time points)	100 ul 10X Ligase buffer, 100 ul 50% PEG, 10 ul 0.1 M DTT, 10 ul 0.1 M ATP, 5 U T_4 DNA ligase, diH_2O to 1000 ul

Table 4, continued.

Reagent	Components
Consensus RNA Sequencing	
TdT dilution buffer	200 mM sodium cacodylate (pH 6.5) 200 mM NaCl, 1 mM EDTA, 4 mM 2-mercaptoethanol, 50% (vol/vol) glycerol
(Components of the GemSeq Transcript Sequencing Kit)	
10X RT buffer	340 mM Tris-HCl (pH 8.3), 500 mM NaCl, 50 mM $MgCl_2$, 50 mM DTT
A mix (RT)	1.0 uM ddATP, 250 uM dCTP, 250 uM dGTP, 250 uM dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
C mix (RT)	12.5 uM ddCTP, 250 uM dCTP, 250 uM dGTP, 250 uM dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
G mix (RT)	12.5 uM ddGTP, 250 uM dCTP, 250 uM dGTP, 250 uM dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
T mix (RT)	50 uM ddTTP, 250 uM dCTP, 250 uM dGTP, 250 um dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
Chase solution	34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT, 2 mM dATP, 2 mM dCTP, 2 mM dGTP, 2 mM dTTP
Stop solution	98% formamide, 10 mM EDTA, 0.1% xylene cyanole FF, 0.1% bromphenol blue

Table 4, continued.

Reagent	Components
Sequencing Gel Preparation	
S1	38% (wt/vol) acrylamide 2% (wt/vol) bisacrylamide
S2 (10X TBE)	1 M Tris-base, 1 M boric acid 20 mM EDTA (pH 8.3)
APS	10% (wt/vol) ammonium persulfate (made fresh weekly)
1X TBE	100 mM Tris-base, 100 mM boric acid, 2 mM EDTA (pH 8.3)

Cell Culture

All propagation and assay of 37A, mutant, and revertant variants of mengovirus used in this work were done in a spontaneously transformed continuous cell line derived from baby hamster kidney, designated BHK-21 (76), and obtained from Dr. J. J. Holland.

Cells were maintained in Dulbecco's Modified Eagle's medium (Sigma) (42) containing 4.5 g glucose per liter (DME-0) and 10% (vol/vol) calf serum (Hyclone) (DME-10). Low passage stocks were stored in DME-10 with 10% (vol/vol) dimethyl sulfoxide (DMSO) in liquid nitrogen. Periodically cells were examined by the method of Del Guidice and Hopps to rule out mycoplasma contamination (34).

Virus infections were done in DME-0 containing 10 mM 3-(N-morpholino)propanesulfonic acid (MOPS), 10 mM N-tris-(hydroxymethyl)-methyl-2-aminoethanesulfonic acid (TES), 10 mM N-2-hydroxyethyl-

piperazine-N'-2-ethanesulfonic acid (HEPES), and 2% (vol/vol) fetal bovine serum (Hyclone) (DME-2).

Virus Strains

The 37A mengovirus, obtained from Dr. H. E. Swim, is a heat stable variant selected by serial passage in L cells in the laboratory of Dr. A. F. Graham (16). The 205 and 280 mutants were isolated from L cells treated with acriflavine during infection by Dr. M. A. Gill (personal communication,7). Revertants 205-A7 and 205-D2 were isolated by Dr. Kevin Anderson from brain tissue from BALB/c mice infected intracranially with mutant 205 (7,8).

All five viruses were plaque purified twice before preparation of the virus stocks used in these experiments. Stocks were prepared in BHK-21 cells by infection at a multiplicity of infection (MOI) of one plaque forming unit (PFU) per cell. Infections were allowed to proceed 24 to 48 hours until 90 to 100% lysis was achieved. Cell lysates were frozen at -70°C, thawed, clarified by centrifugation in the Sorvall RC5B in the SS34 rotor at 5000 rpm for 10 min at 4°C, and stored in 1 and 5 ml aliquots at -70°C.

Plaque Assays

Plaque assays using BHK-21 cell monolayers were done to determine the titers and plaque sizes of the viral stocks. Plastic six-well dishes (CoStar 3406) were seeded with BHK-21 cells in DME-10 and incubated at 37°C until a confluent layer of cells was obtained, usually in 24 to 48 hours. The monolayers were washed with 1 ml of warm DME-0

before addition to each well of 200 ul of an appropriate dilution of virus in DME-2. After allowing adsorption at room temperature (RT) for 30 min, the medium was removed and the cells overlaid with 2 ml per well of DME-2 with 0.75% (wt/vol) agarose (Sigma, type II).

Plaques were generally visible after 48 to 72 hours of incubation at 37°C. After fixing the monolayers at RT overnight with 0.5 ml of 2% (vol/vol) glutaraldehyde per well, the agarose overlays were removed and the wells washed with tap water and stained for 5 min with 1 ml of 0.4% (wt/vol) aqueous crystal violet per well. The plates were washed extensively with tap water and allowed to air dry.

To determine titers, the plaques from appropriate dilutions of each virus stock were counted and the results expressed as PFU/ml of the stock suspension. To determine plaque sizes, well-isolated plaques from several wells were measured and the mean diameter calculated. Results from the different variants were evaluated by the Student's *t*-test.

Hemagglutination Assays

The hemagglutination assay of Martin et al. (80), modified for U-bottom microtiter plates (Linbro), was used to determine the ability of 37A, mutants 205 and 280, and revertants 205-A7 and 205-D2 mengovirus to agglutinate human type O erythrocytes.

Human O erythrocytes from Bond or Mann were obtained by venipuncture in the laboratory, mixed with an equal volume of Alsever's solution ((83),p. 243) and stored at 4°C for no longer than two weeks before use. The cells were washed extensively in PBS with 1 mM sodium azide and resuspended in PBS-gelatin prior to use.

To perform the assay, a sample of each virus stock was diluted to 2 X 10^8 PFU/ml with DME-2. Sodium deoxycholate was added to a final concentration of 0.5% (wt/vol). Twenty-five ul of this viral suspension was diluted 2-fold, in duplicate, into successive wells of the microtiter tray containing 25 ul of Dulbecco's phosphate buffered saline (PBS) with 0.2% (wt/vol) gelatin (PBS-gelatin). Twenty-five ul of 0.3% (vol/vol) cells were added to each microtiter well. The trays were tapped to mix, covered, and then incubated overnight at 4°C.

A positive hemagglutination reaction is indicated by a diffuse layer of cells in the bottom of the well. A negative reaction results in a tight, dark pellet. The titer is expressed as the reciprocal of the end point dilution, the last dilution to show a positive hemagglutination reaction.

Isolation and Purification of RNA

Virion RNA for complementary DNA (cDNA) synthesis and consensus RNA sequencing was obtained by proteinase K digestion (58) of virions partially purified by centrifugation through a sucrose pad. RNA used for cDNA synthesis and cloning experiments was further purified by sucrose gradient centrifugation. Intracellular RNA from mengovirus- and mock-infected BHK-21 cells was labeled with [32-P] and isolated for use as probes for colony hybridization. Ribosomal RNAs from BHK-21 cells, labeled with [3-H], were prepared for use as markers in sucrose density gradients.

To prepare virion RNA, 10 cm dishes (CoStar 3100) containing confluent BHK-21 cells were infected at a MOI of three in a volume of

0.5 ml DME-2 as described above. After adsorption of virus for 30 min at RT, the virus suspension was removed and 5 ml of warm DME-2 added to each dish. Dishes were incubated at 37°C until 90 to 100% of the cells had lysed, generally 18 to 24 hours. If labeled RNA was desired, 5,6-[3-H]-uridine (NET-367) was added to several dishes to a final concentration of 10 uCi/ml at 6 hours postinfection.

The cell lysates were frozen at -70°C, thawed, pooled in 30 ml tubes (Nalgene 3119), and centrifuged in a Sorvall RC5B in a SS34 rotor at 5000 rpm for 15 min at 20°C to remove cellular debris. Sodium dodecyl sulfate (SDS) was added to a final concentration of 0.2% (wt/vol) and the lysate recentrifuged as above. The clarified virus suspension was layered over a 1 ml pad of 28% (wt/wt) sucrose in Bu3 and centrifuged in a Sorvall OTD65B ultracentrifuge in a Beckman SW41Ti rotor at 37,000 rpm for 90 min at 18°C. The pellets were resuspended in a small volume of ice cold Bu3, pooled, and frozen at -70°C.

For extraction of the RNA, the virion suspensions were thawed and transferred to 15 ml Corex tubes (Corning 8441). After adjusting the volume to 2 ml with Bu3, 0.2 ml 10% (wt/vol) SDS, 0.2 ml 4 M NaCl, 0.2 ml 0.1 M EDTA, and 10 ul proteinase K (20 mg/ml) were added. The digestion mixture was incubated at 55°C for 1 hour, extracted twice with 1 ml phenol and 1 ml chloroform, and extracted twice with 2 ml chloroform. The aqueous layer was transferred to a clean 15 ml Corex tube. To precipitate the RNA, 0.2 vol of saturated ammonium acetate (NH_4Ac) and 2.5 vol of 95% (vol/vol) ethanol (ETOH) were added and the tubes held at -20°C overnight. The RNA was pelleted in the Sorvall RC5B

in the HB-4 rotor at 10,000 rpm for 30 min at 4°C, washed with 70% (vol/vol) ETOH, and dried under a vacuum.

For sequencing, the RNA was reprecipitated twice with 0.1 volume of NH_4Ac and 2.5 volumes 95% (vol/vol) ETOH. The RNA was pelleted, washed, dried as described above, and redissolved in diH_2O. Aliquots (2 ug) were prepared in 1.7 ml microcentifuge tubes (VWR 20170-331), precipitated with 0.1 vol NH_4Ac and 2.5 volumes of 95% (vol/vol) ETOH, and stored at -70°C until use. The concentration of RNA was determined by measuring the absorbance at 260 nm estimating 1 A_{260} unit = 40 ug/ml RNA.

To purify RNA for cloning, labeled virion RNA was layered onto gradients of 10-30% (wt/wt) sucrose, 0.2% (vol/vol) SDS in Bu3 and centrifuged in a Sorvall OTD65B ultracentrifuge in a Beckman SW41Ti rotor at 40,000 rpm for 6 hours at 18°C in parallel with gradients layered with ribosomal RNA markers. Fractions (20 drops) were collected with a peristaltic pump from the bottom of the gradients and samples (10 ul) spotted onto Whatman GFC glass fiber filters. The filters were dried, placed into scintillation vials with 3 ml scintillation fluid, and counted in a Packard LSC 460CD liquid scintillation counter. The count data for the viral and marker RNAs were plotted and peak fractions of the viral RNA pooled. The RNA was recovered by ETOH precipitation, washed, dried, and quantified as described above.

Ribosomal RNA markers were prepared from six 10 cm dishes of BHK-21 cells which were 90% confluent. After removing the medium and washing the monolayers with 5 ml DME-2, 5 ml of DME-2 containing 10 uCi of 5,6-[3-H]-uridine (NET-367) per ml were added and the dishes

incubated at 37°C for 18 hours. The medium was then removed and the cell monolayers washed with 3 ml cold NET. One ml of cell lysis buffer was added to each dish and the dishes held on ice for 5 min. The pooled lysate was transferred to a 15 ml Corex tube and the dishes washed with another ml of lysis buffer. The wash was pooled with the initial lysate. The lysate was digested with proteinase K as described above except that the digest was incubated at 37°C for 1 hour.

The intracellular RNA probes used for colony hybridization were prepared by infecting seven 10 cm dishes of confluent BHK-21 cells with 37A mengovirus at an MOI of five and by mock-infecting three dishes with DME-2. After absorption of the virus and removal of the infecting medium, 4 ml of phosphate-free DME-2 was added to each dish and the dishes incubated at 37°C. At 5.5 hours postinfection, the medium was replaced with 3.5 ml of phosphate-free DME-2 with 2.5 ug actinomycin D per ml and the dishes incubated for 0.5 hour at 37°C prior to the addition of 100 ul of [32-P]-orthophosphoric acid (NEX-054) per dish. At 20 hours postinfection the intracellular RNA was isolated and proteinase K digested as described above for the ribosomal markers except that vanadyl ribonucleoside complexes (BRL) were added (100 ul per dish) during the lysis to inhibit cellular ribonucleases. Labeled RNA was separated from unincorporated [32-P] by liquid chromatography using Sephadex G-25 in disposable columns made from silconized Pasteur pipettes (VWR 14673-010). The column buffer was 10 mM TE.

Bacterial Strains and Vectors

The DH5 strain of *Escherichia coli* was used for cloning of viral cDNAs and pBR322 was used as the cloning vector. *Pst*I-digested, oligo(dG)-tailed pBR322 DNA was obtained from Bethesda Research Laboratories (BRL).

For sequencing of cDNAs, the gemini plasmid pGEM-3Z, obtained from Promega Corporation, was used. The JM109 strain of *Escherichia coli* (Promega) was used for the propagation of the pGEM-3Z parent plasmid, the pGEM205-110 plasmid, and the set of pGEM205-110-derived plasmids used for sequencing. JM 109 was grown on minimal agar plates prior to use to maintain the F factor (124).

Stock cultures of bacterial strains were prepared for storage at -70°C by resuspending the growth from 24 hr cultures in a small volume of 1% Bacto-peptone (Difco), 50% glycerol in 1.8 ml cryovials (Evergreen 222-3901-085) (78).

Plasmid DNA Isolation and Purification

Bacteria containing plasmids derived from pBR322 were grown in Luria broth (LB) with 10 ug tetracycline per ml at 37°C in a shaking waterbath until the absorbance of the medium reached an A_{600} of 0.4 to 0.5. Chloramphenicol was added to a final concentration of 170 ug/ml to amplify the plasmids and the cells were incubated overnight as described above. Bacteria with plasmids derived from pGEM-3Z were grown in LB with 100 ug ampicillin per ml as described above but without amplification.

Plasmid DNA was isolated by an alkali extraction method modified from the procedure of Birnboim and Doly (12). The exposure of the plasmid DNA to NaOH was limited to 5 to 10 min and done in an ice bath to limit nicking of the DNA.

Large scale (1 liter) bacterial cultures were done to prepare pGEM-3Z vector DNA and pGEM205-110 DNA for subcloning. The plasmid DNA isolation procedure for large scale preparations is outlined in Figure 3 and includes purification by liquid chromatography with Sephacryl S-1000 (Pharmacia). Plasmid DNA from Sephacryl-purified preparations was quantified by measuring absorbance at 260 nm estimating 1 A_{260} unit = 50 ug/ml DNA.

Plasmid DNA was isolated for sizing, slot blots, and sequencing from small scale (10 ml) cultures incubated in 20 x 150 mm screw cap tubes (Corning 99449-20X). The procedure used for small scale plasmid DNA isolation is outlined in Figure 4. Plasmid DNA for sequencing reactions was further purified by precipitation with polyethylene glycol (PEG, mol wt 8000) as recommended by Promega (Promega Protocols and Applications Guide, pp. 73-74). The quantity and quality of PEG-purified DNA was determined by agarose gel electrophoresis as described below.

Agarose Gel Electrophoresis of DNA

Agarose gel electrophoresis was used to judge the quality and quantity of DNA from small scale plasmid preparations and to estimate the size of plasmid DNAs. Low melting agarose gels were used to select

DNA fragments from restriction enzyme digests for subcloning and large cDNAs for cloning.

Figure 3. Large scale isolation and purification of plasmid DNA.

1. Transfer culture to four 250 ml bottles. Centrifuge in GSA rotor for 10 min at 5000 rpm. Decant. Resuspend and pool the pellets in 200 ml TNE. Centrifuge as above. Decant.

2. Resuspend pellet in 20 ml sucrose solution and freeze at -20°C. Thaw on ice. Add 5 ml lysozyme solution. Hold on ice 20 min.

3. Add 50 ml lysis solution. Invert to mix. Hold on ice 10 to 15 min. Add 25 ml K acetate. Invert to mix. Hold on ice 10 to 30 min.

4. Centrifuge in the GSA rotor for 20 min at 10,000 rpm. Transfer the supernatant fluid to a clean 250 ml bottle. Add 100 ml isopropanol and invert to mix. Hold at RT for at least 15 min.

5. Centrifuge in the GSA rotor at 10,000 rpm for 20 min. Decant and drain the pellet. Vacuum dry.

6. Resuspend the pellet in 4 ml RNase A solution and transfer to a 15 ml Corex tube. Hold at 37°C for 15 min, RT for 30 min, or overnight at 4°C.

7. Extract 2 to 3 times with 2 ml phenol and 2 ml chloroform. Extract twice with 4 ml chloroform. Transfer the aqueous layer to a clean 15 ml Corex tube. Precipitate with 10 ml 95% ETOH. Hold at -20°C for at least an hour to overnight.

8. Centrifuge for 20 min in the SS34 rotor at 10,000 rpm. Wash with 5 ml 70% ETOH. Centrifuge. Decant, drain, and vacuum dry the pellet.

9. Redissolve the pellet in 200 ul diH_2O and apply to Sephacryl column (1.5 x 45 cm). Rinse tube with 200 ul column buffer. Apply to column. Collect 80 40 drop fractions.

10. Measure A_{260} of fractions. Pool plasmid peak. Precipitate with 2.5 volumes of 95% ETOH. Centrifuge in the SS34 rotor at 10,000 rpm for 20 min. Wash with 5 ml 70% ETOH. Vacuum dry. Resuspend in 1 ml diH_2O.

Figure 4. Small scale miniprep isolation and purification of plasmid DNA.

1. Transfer the culture to a 15 ml Corex tube. Centrifuge for 5 min in the SS34 rotor at 5000 rpm. Discard the supernatant fluid. Resuspend the pellet in 5 ml TNE. Centrifuge as above and discard the supernatant fluid.

2. Resuspend the pellet in 1 ml 50 mM TE. Add 100 ul lysozyme solution, mix, and hold on ice 10 to 30 min.

3. Add 4 ml lysis solution. Invert to mix. Hold on ice 10 to 15 min. Add 2 ml K acetate. Invert to mix. Hold on ice 10 to 30 min.

4. Centrifuge in the SS34 rotor for 20 min at 10,000 rpm. Transfer the supernatant fluid to a clean 15 ml Corex tube. Add 7 ml isopropanol and invert to mix. Hold at RT for at least 15 min.

5. Centrifuge in the SS34 rotor at 10,000 rpm for 20 minutes. Decant and drain the pellet. Vacuum dry the pellet.

6. Resuspend the pellet in 400 ul RNase A solution and transfer to a microcentrifuge tube. Hold at 37°C for 15 min, RT for 30 min, or overnight at 4°C.

7. Extract once with 200 ul phenol and 200 ul chloroform. Extract twice with 400 ul chloroform. Transfer the aqueous layer to a clean microcentrifuge tube. Precipitate with 1 ml 95% EtOH. Hold at -20°C for at least an hour to overnight.

8. Centrifuge for 5 min. Wash with 200 ul 70% EtOH. Centrifuge. Decant, drain, and vacuum dry the pellet.

9. For sequencing, redissolve the pellet in 160 ul diH_2O. Add 40 ul 4 M NaCl. Mix. Add 200 ul PEG. Mix. Hold on ice 30 min. Centrifuge 5 min. Remove the supernatant fluid by aspiration. Wash with 200 ul 70% EtOH. Recentrifuge. decant, drain and vacuum dry the pellet.

10. Redissolve the pellet in 100 ul diH_2O. Gel 5 ul.

Three sizes of gels were run. Large 16.5 x 16.5 cm gels were run for sizing experiments. Shorter 16.5 x 9 cm gels were run to quantify

plasmids for sequencing and for preliminary sizing of uncut plasmids. Small 7 x 9 cm minigels were used for the low melting agarose gels and to check completion of restriction enzyme digests.

The concentration of agarose in most gels was 0.8% (wt/vol) agarose (Sigma, type II). When very small fragments needed to be separated 1.0 to 1.2% gels were used. For low melting agarose gels, 0.5 to 0.8% (wt/vol) Sigma type VII agarose or 2% NuSieve agarose (FMC Bioproducts) was used. All agarose gels were prepared in 1X TAE.

DNA samples (5 to 10 ul) were mixed with 3 ul gel loading buffer and heated to 65°C for 10 min before loading. All gels were run submerged in 1X TAE at voltages between 30 to 60 V. The lower voltages were used for low melting agarose gels.

Gels were stained with ethidium bromide (10 ug/ml dH_2O) for 30 min to overnight and the DNA visualized using UV illumination. A Polaroid MP-4 camera was used to photograph the UV illuminated gels.

Lambda *Hind*III DNA fragments were used as markers for sizing linearized DNA fragments. The migration distances of the unknown and marker DNAs were measured on images from a negative produced by a photographic enlarger. The GEL computer program written by Dr. Brian Fristensky, based on (111), was used to calculate the size of the DNA.

Complementary DNA Synthesis and Cloning

CDNA was synthesized using a modification of the method of Gubler and Hoffman (54) in which *E. coli* DNA ligase and NADH were omitted from the second strand reaction. Sucrose gradient-purified viral RNA (10 ug/reaction) from 37A, 205, or 280 mengovirus was used as the template

and oligo$(dT)_{12-18}$ (Collaborative Research) (3 ug/reaction) as the primer for first stand synthesis.

To determine the amount of label incorporated during the synthesis reactions, samples were removed from both first and second strand reaction mixtures for total and trichloracetic acid (TCA) precipitated counts. Total counts were made by drying duplicate samples on Whatman GFC glass fiber filters. Incorporated counts were determined by spotting the samples onto filters and placing the filters into ice cold 10% (vol/vol) TCA. The filters were batch washed twice in ice cold 5% (vol/vol) TCA and three times in 95% (vol/vol) ETOH. The filters were placed into scintillation vials and dried overnight in an oven at 100°C. After the vials cooled to RT, 3 ml of scintillation fluid were added and the filters counted.

The cDNA was prepared for insertion into the G-tailed vector by the addition of oligo(dC)-tails using TdT (35,72). The cDNA was redissolved in 100 ul of 1X C-tailing buffer containing 75 U TdT and incubated at 30°C for 15 min. Five ul of 0.2 M EDTA and 200 ul of 10 mM TE (pH 8.0) were added and the reaction mixture heated at 65°C for 5 min to stop the reaction. The mixture was extracted twice with phenol and chloroform and twice with chloroform alone, precipitated with NH_4Ac and 95% (vol/vol) ETOH, washed with 70% (vol/vol) ETOH, and dried under a vacuum.

In an effort to select for larger cDNA inserts, tailed 205 cDNAs were separated by electrophoresis in a 0.5% (wt/vol) low melting agarose minigel as described by Crouse, Frischauf, and Lehrach (29). The cDNA was divided into two fractions and run in two separate lanes on the gel.

One lane was stained with ethidium bromide while the other remained unstained. The unstained cDNA lane was used to prepare 16 slices for which the Cerenkov counts were determined by liquid scintillation spectroscopy. Seven slices from the upper portion of the gel (3-10) were selected for annealing to vector DNA.

The average volume of the gel slices, determined by weighing, was about 100 ul. To each slice 150 ul diH_2O and 250 ul 2X annealing buffer with 0.1 ng *Pst*I-cut, oligo(dG)-tailed pBR322 per ul were added. The mixtures were incubated at 65°C for 15 min to melt the agarose, then held at 58°C for 1 hour to allow the vector and inserts to anneal. The annealing mixtures were cooled in an ice bath before the addition of competent cells.

Transformation of *E. coli* strain DH5 with the annealing mixtures was done with competent cells prepared using the method of Hanahan (55). One ml of freshly prepared competent cells was added to the annealing mixture for each gel slice and gently mixed with the DNA. Each mixture was split into five 300 ul aliquots in 17 x 100 mm polypropyline tubes (Evergreen 222-2393-080), held in an ice bath for 30 min, heat shocked at 42°C for 90 sec, and cooled on ice for 10 sec. One ml of SOC, at RT, was added and the transformation mixtures incubated at 37°C for 30 min. Two ml of soft agar containing 5 ug tetracycline per ml, at 45°C, was added to each tube and the mixtures used to overlay warm LA plates containing 10 ug tetracycline per ml. After the soft agar had solidified, the plates were incubated at 37°C for 24 to 48 hours.

Known quantities of pBR322 DNA were used to transform competent cells to confirm the transformability of the cells. Competent cells

were mock-transformed and plated on LA and LA with 10 ug tetracycline per ml to check viability of the cells and to confirm the susceptibility of untransformed bacteria to tetracycline.

Analysis of cDNA Inserts

Colony hybridization using ^{32}-P labeled intracellular RNA probes from mengovirus- and mock-infected cells was done to identify transformants containing mengovirus-specific sequences. The method used was a modification of the method described by Grunstein and Hogness (53). Nitrocellulose filters (Schleicher and Schuell BA85) were placed on the surface of LA plates containing 10 ug of tetracycline per ml and inoculated with the transformants. The plates were incubated for 8 hr at 37°C when faint but visible growth was apparent on the filters. The filters were carefully lifted and transferred to plates containing LA medium with 250 ug chloramphenicol per ml and incubated overnight to amplify the plasmid DNA.

To release and denature the plasmid DNA, the filters were lifted from the plates and floated, colony site up, on 10% (wt/vol) SDS for 5 min; on 0.5 N NaOH, 1.5 M NaCl for 5 min; on 0.5 M Tris-HCl (pH 8.0), 1.5 M NaCl for 5 min; and on 2X SSPE for 5 min. After the filters had air dried, the DNA was fixed to the filters by baking at 80°C for 2 hours in a vacuum oven.

Duplicate filters were wetted with 6X SSPE and prewashed as a batch in 200 ml of prewash solution in a liter size beaker at 42°C in a shaking water bath for 3 hours changing the prewash solution once at 1.5 hours. The prewash solution was replaced with 100 ml of hybridization

solution [50% (vol/vol) formamide, 5X Denhardt's, 5X SSPE, 0.1% (wt/vol) SDS, 100 ug calf thymus DNA/ml, 100 ug tRNA/ml, and 1 ug poly(A)/ml] and the filters prehybridized as a batch overnight in the shaking waterbath at 42°. For the hybridization, the filters were placed into Seal-A-Meal bags with 20 ml of hybridization fluid, the appropriate probe was added, and the bags were incubated at 42°C overnight. The filters were washed successively at 42°C in 2X SSPE with 0.1% SDS (twice), 1X SSPE with 1% SDS (twice), and 0.2% SSPE with 0.1% SDS and mounted wet in Saran wrap. Kodak XO-MAT AR film was exposed to the filters at RT overnight and developed with GBX developer and fixer (Kodak) according to the manufacturer's instructions.

Plasmid DNA was prepared from transformants containing mengovirus-specific inserts and evaluated by agarose gel electrophoresis to determine approximate insert size. Plasmids containing the larger inserts were cut with *Pst*I and regelled for more accurate size determination.

Selected transformants with inserts greater than 2 kb in length were further analyzed by hybridization of slot blots of plasmid DNA with dephospho-oligo$(dT)_{12-18}$ (Pharmacia) and synthetic oligonucleotide probes specific for regions 3' and 5' to the capsid coding region and the 5' end of the mengovirus genome. The probes were end-labeled with [32P] using T_4 polynucleotide kinase as described in (78)(p122). Gamma-[32-P]-ATP was synthesized from [32-P]-orthophosphoric acid (NEX-054) with a Gammaprep Synthesis System (Promega) according to the manufacturer's directions.

The synthetic oligonucleotides were synthesized in the laboratory of Dr. Walter Hill (Division of Biological Sciences, University of Montana) and were based on the sequence data for the M variant of mengovirus (personal communication Dr. Ann Palmenberg). Sequences chosen for probes did not have significant homology with other portions of the mengovirus genome as determined by computer analysis with the GenePro program (Riverside Scientific Enterprises). The sequence, position, and orientation of the probes are listed in Appendix A, Table 9.

The labeled probes were purified from unincorporated label by a method using DEAE-cellulose (Dr. Andreas Luder, personal communication). One hundred ul of 10 mM TE (pH 8.0) and 900 ul of 50% (vol/vol) DEAE-cellulose in 10 mM TE (pH 8.0) were added to the labeling mixture and mixed by extensive vortexing. The DEAE-cellulose was pelleted by centrifugation in a microcentrifuge (Eppendorf 5414) for 5 min and the supernatant fluid containing unbound label discarded. The DEAE-cellulose was washed twice with 1 ml 10 mM TE (pH 8.0) and twice with 1 ml of low salt buffer before the labeled probe was eluted from the DEAE-cellulose in 1 ml of high salt buffer.

Slot blots of plasmid DNA were prepared in triplicate using the Hybri-Slot manifold (BRL 1052MM) and Zeta-Probe blotting membrane (BioRad 162-0159) (9). For each blot, 1 ug of DNA in 15 ul of diH_2O was added to 100 ul of 0.4 M NaOH and heated to 68°C for 30 min. The denatured DNAs were cooled to RT, placed in the appropriate wells of the manifold, and a vacuum from a water aspirator applied. After the wells had emptied, the manifold was disassembled and the membranes rinsed in

2X SSC and allowed to air dry. Plasmid DNAs (pBR322 and pGEM-3Z) were used as negative controls on the blots.

The blots were prehybridized as a batch in 250 ml of 6X SSPE, 0.5% (wt/vol) SDS, 10X Denhardt's, 20 ug tRNA per ml, and 50 ug fragmented, denatured salmon sperm DNA per ml at 41°C overnight. Blots were hybridized overnight at 41°C in Seal-a-Meal bags containing 10 ml of hybridization fluid (6X SSPE, 0.5% (wt/vol) SDS, 1X Denhardt's, and 200 ug tRNA per ml) and 100 ul of probe. The hybridized blots were washed twice in 2X SSPE with 0.1% (wt/vol) SDS and once in 1X SSPE with 0.1% (wt/vol) SDS for 30 min at 41°C. The blots were mounted wet in Saran wrap and used to expose X-ray film as described above. The films were developed after a 24 hour exposure.

Subcloning of pBR205-110

In order to prepare deletion mutants for sequencing, the entire cDNA insert of pBR205-110 was cut from the pBR322 vector with *Pst*I and subcloned into the *Pst*I site in the multiple cloning region of pGEM-3Z to produce pGEM205-110. The subcloning protocol was modified from the procedure described in Promega Notes (Number 6, November 1986) by the addition of low melting agarose electrophoresis selection and Elutip-d purification (Schleicher and Schuell) of the DNA fragment to be subcloned.

The insert was prepared by *Pst*I digestion of 10 ug of Sephacryl-purified pBR205-110 DNA and separated from the vector DNA on a 0.5% (wt/vol) low melting agarose minigel. The band of agarose containing the insert fragment was excised from the gel and purified using the

Elutip-d procedure for recovery of DNA from low melting agarose (Schleicher and Schull Technical Bulletin #206).

The pGEM-3Z vector DNA was prepared by *Pst*I digestion and was treated with calf intestinal phosphatase to prevent religation of the vector without the addition of the insert. An additional alkaline phosphatase digestion at 56°C of the vector DNA was included (78).

For the ligation reaction, both DNAs were extracted with phenol and chloroform, precipitated with ETOH, pelleted, washed, dried under a vacuum, and redissolved in diH_2O and quantified by measuring the absorbance at 260 nm. The ligation mixture had a final volume of 50 ul and contained 2 ug of insert DNA, 0.8 ug *Pst*I-cut pGEM-3Z, 5 ul 10X ligation buffer, and 0.5 U T_4 ligase and was incubated overnight at RT (78). A second ligation mixture without insert DNA was used as a control to check vector religation. Ten ul aliquots of the ligation mixtures were used to transform 200 ul of competent JM109 cells as described above for cDNA cloning except that pGEM-3Z DNA was used as a control for the transformation, the LA plates contained 100 ug ampicillin per ml, and the soft agar contained 100 ug ampicillin per ml, 40 ug X-Gal per ml, and 0.5 mM IPTG.

Plasmid DNA was isolated from 11 transformants and digested with *Pst*I for analysis on 0.8% agarose gels. A clone having the proper insert and vector size was selected for further work and designated pGEM205-110.

Three additional subclones of pGEM205-110, needed to complete the sequencing, were prepared by restriction enzyme digestion of pGEM205-110 and are shown in Figure 5. The pGEM205-110Xba clone was produced by

digesting the parent plasmid with *Xba*I and allowing religation of the large vector-partial insert fragment. The pGEM205-110Kpn subclone contained the 900 bp *Kpn*I insert fragment. The third subclone, pGEM205-110Bgl, contained the 500 bp *Bgl*II insert fragment. Because pGEM-3Z has no *Bgl*II restriction site, the fragment was inserted into the compatible *Bam*HI site. The fragments were isolated as described above in 0.5% low melting agarose, purified on a Elutip-d, self-ligated or ligated to restriction enzyme-digested and phosphatase-treated pGEM-3Z vector DNA, and used to transform 200 ul of competent JM109 cells. Plasmid DNA was isolated from 10 transformants from each transformation, digested with an appropriate restriction enzyme, and sized by agarose gel electrophoresis to confirm the construction.

Exonuclease III Digestion of pGEM205-110

The exonuclease III method of Henikoff (56,57) was used to create a nested set of deletion mutants of pGEM205-110 for DNA sequencing. Sephacryl column purified pGEM205-110 plasmid DNA was digested with *Sst*I (BRL) to produce a 3'-overhang to protect the vector DNA and with *Bam*HI to produce a 5'-overhang adjacent to the insert DNA. The exonuclease III procedure is outlined below in Figure 6. Uncut plasmid DNA isolated from the transformants was screened for approximate size by agarose gel electrophoresis using uncut pGEM205-110 and pGEM-3Z DNA markers. Transformant DNA running between the markers was linearized with *EcoR*I and sized more accurately by agarose gel electrophoresis as described above.

Figure 5. Restriction map of the 205-110 derived subclones. A. The restriction map and orientation of the entire cDNA insert (▒) from pBR205-110 in the multiple cloning site of pGEM-3Z. B. The pGEM205-110Kpn subclone. C. The pGEM205-110Xba subclone. D. The pGEM205-110Bgl subclone.

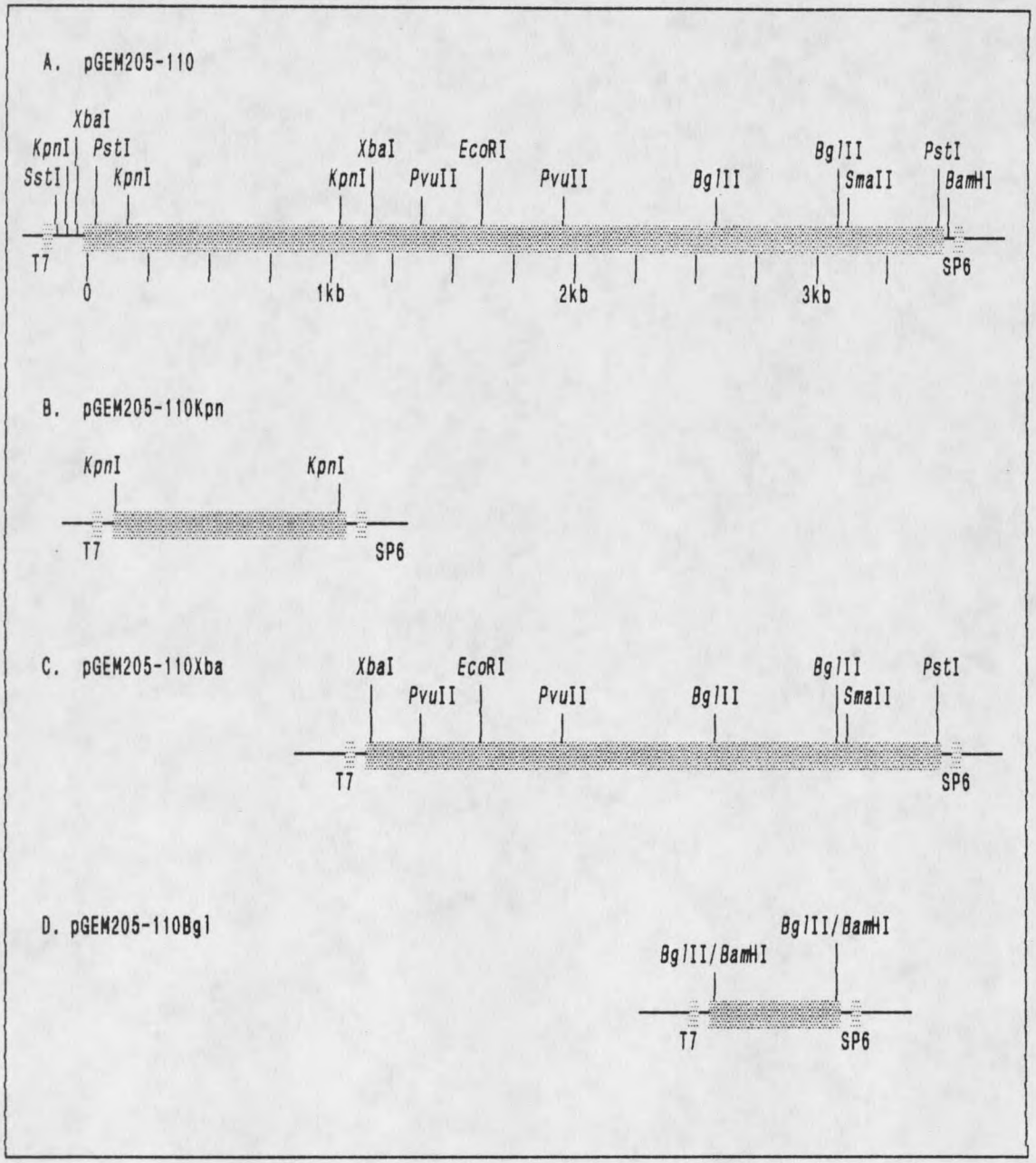


Figure 6. Exonuclease III digestion procedure for production of nested set deletions for sequencing DNA.

1. Exonuclease III digestion (for 18 time points)

   A. Redissolve restriction enzyme cut plasmid DNA in:
      44 ul diH_2O
      5 ul 10X Exonuclease III buffer.

   B. Warm to 35°C, add 1 ul (200 U) Exonuclease III, and mix with tip, incubate at 35°C.

   C. At 30 sec intervals, remove 2.5 ul of digest mixture. Add to 7.5 ul aliquots of S1 nuclease mix. Hold on ice until all samples have been taken.

2. Blunt ends with S1 Nuclease

   A. Place S1 mix tubes at RT. Incubate 30 min

   B. Add 1 ul S1 Stop. Heat at 70°C for 10 min

3. Polish ends with Klenow

   A. Place tubes at 37°C, add 1 ul Klenow mix, and incubate 2-5 min.

   B. Add 1 ul dNTP mix. Incubate 5 min

4. Ligate ends with T_4 DNA Ligase

   A. Place tubes at RT. Add 40 ul ligation mix.

   B. Incubate 30 min to overnight at RT.

5. Transformation of JM109

   A. Use 10 ul of ligation mixture to transform 20 ul competent JM 109 cells.

   B. Hold on ice 30 min.

   C. Add 900 ul SOC. Incubate at 37°C for 1 hr.

   D. Plate 100 ul to 2 LA with 100 ug ampicillin/ml plates.

   E. Incubate plates 24 to 48 hours at 37°C.

DNA Sequencing of pGEM205-110

Sequencing of the pGEM205-110-derived subclones was done using the K/RT and Sequenase (Ver 1.0 and 2.0) kits. The T7 and SP6 promoter primers supplied by Promega were used with both kits and are listed in Appendix A, Table 10.

Double-stranded plasmid DNA, isolated and purified by precipitation with PEG as described above, was denatured prior to sequencing as recommended by Promega (Promega Protocols and Applications Guide, p. 74). Two ug aliquots of plasmid DNA were pelleted, washed, and dried under a vacuum. The pellet was redissolved in 20 ul diH_2O and 2 ul 2 M NaOH, 2 mM EDTA added and the mixture incubated for 5 min at RT. Three ul 3 M sodium acetate (pH 5) and 7 ul diH_2O were added, vortexing briefly to mix. To precipitate the DNA, 100 ul 95% ETOH was added and the DNA incubated at -20°C at least 1 hour. The DNA was centrifuged in a microcentrifuge for 5 min, washed with 200 ul 70% ETOH and dried under a vacuum.

The procedures used for the kits and the components of the kits are given in Appendix B. The label used with the Promega K/RT Sequencing System was [α-32P]-dATP (NEG-034S). A label with a higher specific activity (NEG-034H) was used with the Sequenase Sequencing kit. To avoid repeated thawing of the label, aliquots were prepared and frozen at -70°C.

Consensus Sequencing of Viral RNA

Consensus sequencing of 37A, 280, 205, 205-A7, and 205-D2 viral RNAs to determine sequence differences in the capsid coding and 5'-

noncoding regions of the genomes was done using a modification of the dideoxynucleotide method of Sanger (108). RNA, not DNA, is the template for sequencing, therefore, reverse transcriptase from avian myeloblastosis virus was the enzyme used for these experiments.

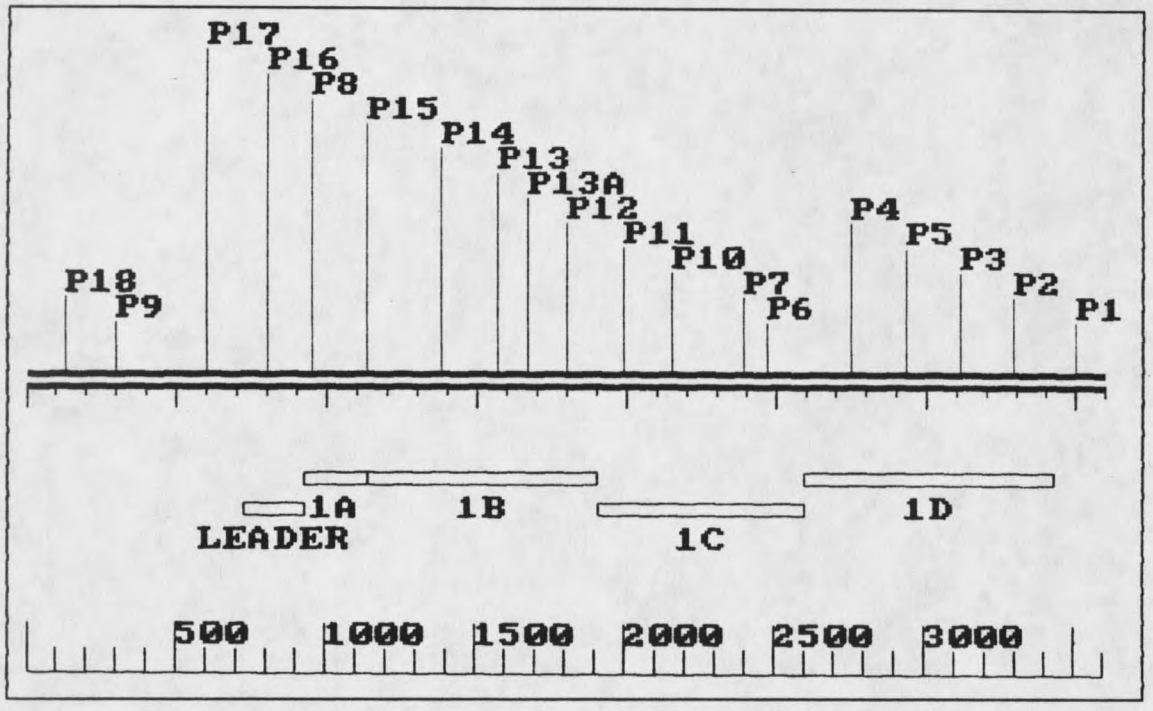
Since the template RNA was not cloned prior to sequencing, the sequence data represents the sequence of the most abundant species of virion RNA. Microheterogeneity of the population does not affect the results since a change found at any one particular base will be rare and the band generated during the sequencing of that particular template will be masked by the much darker band generated by the rest of the population.

Synthetic oligonucleotides used as primers and their sources are listed in Appendix A, Table 11. Figure 7 shows the position of the sequences complementary to the primers on the mengovirus genome. The selection of the primer sequences was based on the sequence data for both the M variant of mengovirus and pGEM205-110 and the rules suggested by Barnes (11). Sequences chosen for primers did not have significant homology with other portions of the mengovirus genome as determined by computer analysis with the GenePro program.

All the primers are 15 nucleotides in length and contain seven (P16 and P17), eight (P2-P15 and P18), or nine (P1) residues of either guanosine (G) or cytosine (C). With the exception of primer 2, the 3' terminal residue is either a C or a G. Primers were dissolved in sterile, diH_2O at a concentration of 10 ng/ml as determined by measuring the absorbance at 260 nm estimating 1 A_{260} unit = 30 ug/ml and used for

sequencing reactions without further purification. Primer solutions were stored frozen at -20°C.

Figure 7. Position of primer sites used for consensus RNA sequencing.



The RNA was prepared for sequencing as described above and was pelleted in the Sorvall RC5B in an HB-4 rotor at 10,000 rpm for 20 minutes at 4°C, washed with 200 ul cold 70% (vol/vol) ETOH, and dried under a vacuum before sequencing.

The RNA sequencing reactions were done using the reagents listed in Table 4 and supplied in the GemSeq Transcript Sequencing Kit. The procedure outlined below in Figure 8 is my modification of the Promega procedure (Promega Protocols and Application Guide pp. 65-66). To make the procedure less expensive, 5'-[α-35-S]dATP (NEG-034H) was used in place of the label recommended by Promega. This label has a higher

specific activity (1284 Ci/mmol) than the recommended label (500 Ci/mmol) so the amount of label used per reaction could be reduced. To compensate for the reduction in labeled dATP, additional cold dATP was added to the reaction mixtures.

There are regions in RNA templates which appear to stop the progression of AMV-RT at particular base residues. When this occurs, a band appears in all four lanes of the sequencing gel and the sequence information for that base is lost. One method used to minimize the loss of information due to natural stops was increasing the reaction temperature to 50°C (46). The estimated melting temperature of 15-mer primers from the RNA template is about 45°C. Therefore, the reaction protocol was altered to allow a short incubation at RT to allow elongation of the primer before a longer incubation at the increased incubation temperature.

Not all natural stops are resolved by increasing the reaction temperature. The addition of TdT to the sequencing reaction protocol also helps resolve natural stops (33). Because TdT can introduce some sequencing artifacts, two reactions, one with TdT and one without TdT, must be run. Therefore, for each reaction, the amount of template, primer, and reaction buffer used in the annealing reaction was doubled to allow the reaction to be split into two sets of termination tubes for the two reactions. For these experiments, TdT was diluted to 3 units per ul in TdT dilution buffer and 1 ul added to each reaction tube during the chase reaction.

Figure 8. Modified protocol for consensus RNA sequencing using the GemSeq Transcript Sequencing Kit.

I. Template Primer Annealing

A. redissolve 2ug RNA in 12 ul dH_2O
6 ul primer (10ng/ml)
2 ul 10X RT buffer
mix, centrifuge briefly, and incubate 2 min in a 65°C heat block

B. remove block from heat source and allow to cool slowly to <35°C, centrifuge briefly

II. Preparation of Reaction Tubes

A. thaw dNTP/ddNTP mixes, hold on ice
label 8 500 ul microcentrifuge tubes
A,A+,C,C+,G,G+,T,T+

B. add 3 ul of the appropriate mix to each tube

III. Elongation and Labeling Reaction

A. to the annealing mix add: 4 ul label
4 ul cold 20 uM dATP
mix, centrifuge briefly
add 10 U AMV-RT (1.4 ul), mix by pipeting up and down with tip

B. transfer 3 ul of the mix to each reaction tube
centrifuge briefly

C. incubate 5 min at RT, then incubate 20 min at 50°C

IV. Chase and Terminal Transferase Reaction

A. add 1 ul Chase solution to each reaction tube
add 3 U (1 ul) terminal transferase to each "+" tube

B. centrifuge briefly, then incubate 15 min at 42°C

V. Termination

A. add 5 ul Stop to each reaction tube
centrifuge briefly

B. store -20°C

Sequencing Gels

Sequencing gels of either 6% or 8% polyacrylamide were run on BRL Model SO sequencing apparatuses using wedge gel spacers (BRL 1107SC) and shark's tooth combs (BRL 1045SE) (107).

Glass plates were cleaned with Bon Ami cleanser and rinsed extensively with tap water. The inner surfaces were cleaned with 95% (vol/vol) ETOH and wiped dry with Kimwipes just before the gel plates were assembled and taped.

The formulas for the standard 6% and 8% gels are listed in Table 5 below. Gel solutions were prepared by dissolving the urea in the appropriate amount of S1, S2, and diH_2O without added heat on a magnetic stirrer. After the solution was filtered through Whatman #4 filter paper and returned to the stirrer, APS was added and allowed to mix with the solution. N,N,N',N'-tetramethylethylenediamine (TEMED) was added last and briefly allowed to mix with the solution.

Table 5. Standard 6% and 8% wedge gels formulas.

Component	Amount for 6%	Amount for 8%
Urea	75 g	75 g
S1	22.5 ml	30 ml
S2	15 ml	15 ml
diH_2O	52.5 ml	45 ml
APS	1 ml	1 ml
TEMED	80-90 ul	80-90 ul

The gel was immediately poured with the aid of a 50 ml syringe barrel while the plates were held at an angle of 20° to 30° above the bench top. A bottom spacer (BRL 1091VC) with the ends trimmed off was quickly inserted into the top of the gel to form a straight edge and clamped into place.

Gels that did not polymerize within 10 to 15 min after pouring were discarded. Gels were generally poured the day before they were used and were stored overnight at RT. If gels were used the same day they were poured and allowed to polymerize at least 1 hour.

Two gels were usually run in parallel with a starting voltage of 1050 to 1150 volts and were preelectrophoresed for 30 min before loading. Except for primer P18, reaction mixtures were run on both short and long gels. The short gels were either 6% or 8% polyacrylamide and were run until the bromphenol blue marker band reached the bottom of the gel, usually 3.5 to 4 hours. The long gels were 6% polyacrylamide and were run usually 8 to 9 hours. When bromphenol blue marker from the reaction mixes reached the bottom of the long gels, a sample of the Stop solution was loaded and the buffer changed. The gels continued to run until the xylene cyanole marker from the second loading of Stop solution reached a point 5 to 7 cm from the bottom of the gel.

Gels were fixed and the urea removed by soaking the gels in a 4 L bath of 5% (vol/vol) methanol, 5% (vol/vol) glacial acetic acid for a minimum of 2 hours. Short gels were generally held for 3 hours while long gels were fixed overnight. The gels were then transferred to Whatman 3MM paper and dried. Kodak XO-MAT AR film was exposed to the

dried gels for 72 to 96 hours and developed with GBX developer and fixer (Kodak) according to the manufacturer's instructions.

In several regions, compression of the bands on the gels did not allow resolution of the sequence. For some of these reactions, formamide gels were run in addition to the standard gels (81). To make formamide sequencing gels, the basic formula was used except that the amount of urea was reduced from 8 M to 7 M, the diH_2O replaced with formamide, and the amount of TEMED was increased fourfold. Formamide gels were run at 1500 to 1700 volts and required nearly twice the running time of standard sequencing gels. The gels were fixed and dried as described above.

Sequence Analysis

Films of the sequencing gels were read manually and the data entered into a computer file using an AsciII text editor (Wonderplus, Bourbaki) and analyzed using the GenePro program. RNA sequences are reported with T replacing U for ease of analysis. Nucleotide sequences were translated using the GenePro program. Appendix C, Table 14 lists the one letter abbreviations for the amino acids used in the protein sequence alignments.

The multiple alignments of nucleotide and protein sequences were done with the Align program (Scientific and Educational Software). Clone Version 2 (Scientific and Educational Software) was used to construct restriction maps from sequence data.

To confirm the reading of the gels and to position the sequence data, the sequence data from each reaction were aligned to the sequence

of the M variant of mengovirus and overlapping regions from adjacent primers. When differences were noted between the data, the films were reread to confirm or correct the data. Overlapping regions from short and long gels were read independently as were sequences from RNA reactions with and without TdT. If portions of the sequence could not be read on a particular film or if rereading the film did not resolve a discrepancy, the sequencing reactions were repeated and the gels rerun. Consensus sequence files were constructed from the overlapping sequences from the different primers (RNA) or clones (DNA).

Three-Dimensional Structural Analysis

The three-dimensional structural analysis of the capsid proteins was done by Dr. Ming Luo and Dr. Kevin Anderson at the University of Alabama, Birmingham. A computer model based on data files containing the coordinates of the three-dimensional structure determined by X-ray crystallography to a 3 Å resolution for the M variant of mengovirus (75) was used to predict the effect of mengovirus-specific amino acid changes found in 37A, mutants 205 and 280, and revertants 205-A7 and 205-D2. The predicted conformation allows for maximal hydrogen bonding without introduction of steric hindrance.

The three-dimensional stick images showing the position and effect of the amino acids changes were viewed and photographed from an Evans and Sutherland PS330 computer graphics system using the FRODO program (67,68). The ribbon model images were viewed and photographed from a Silicon Graphics IRIS-4D computer graphics system using the RIBBON

program developed by Mike Carson at the University of Alabama, Birmingham.

RESULTS

Plaque Sizes and Hemagglutination Titers of the Virus Stocks

The plaque sizes and hemagglutination titers of the virus stocks prepared for these experiments were determined to confirm the phenotypic identity of these stocks to the previously used stocks (7,8). The results are shown below in Table 6. As observed earlier, neither 205 nor 280 were able to agglutinate human O cells. Revertant strains 205-A7 and 205-D2 had regained that property. Both the mutants and revertants had plaque sizes significantly smaller than 37A ($p < 0.05$).

Table 6. Plaque sizes and hemagglutination titers of 37A, mutants 280 and 205, and revertants 205-A7 and 205-D2 virus stocks.

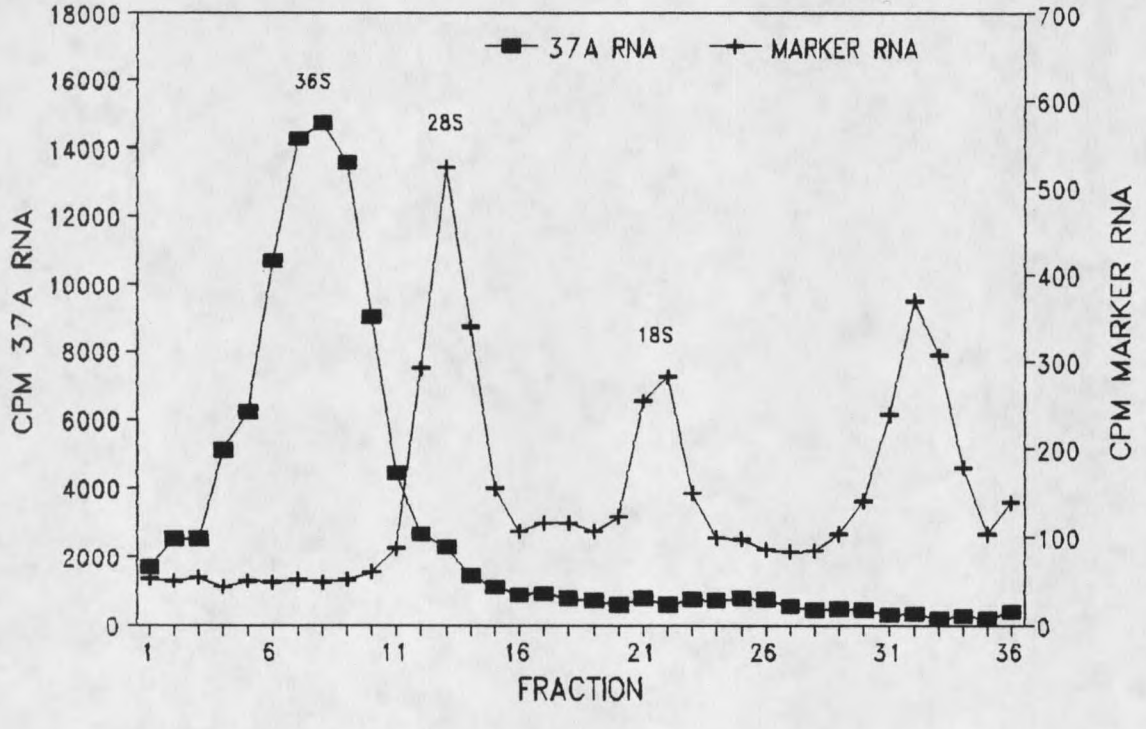
Virus	Plaque Size[a] ± STD[c]	Hemagglutination Titer[b]
37A	4.9 ± .75	1024
280	2.3 ± .53	0
205	2.2 ± .68	0
205-A7	1.7 ± .58	2048
205-D2	1.8 ± .70	1024

[a]. Mean plaque diameters in mm were calculated from 15 plaques.
[b]. Titers were done in duplicate, using equal numbers of PFU, and are expressed as the reciprocal of the end point dilution.
[c]. The standard deviation was calculated from the mean diameter and the data compared using the Student t-test.

Isolation of Virion RNA

Virion RNA was isolated by proteinase K digestion of partially purified virions as a template for cDNA synthesis and for RNA consensus sequencing. RNA used in cloning was further purified by sucrose gradient centrifugation in 10-30% sucrose gradients. A typical plot of the counts obtained from samples of [3-H]-labeled viral and cellular RNA is shown below in Figure 9. Virion RNA with a sedimentation constant of 36S was clearly separated from ribosomal 28S and 18S RNAs.

Figure 9. Separation of viral 36S RNA from cellular ribosomal 28S and 18S RNAs on 10-30% sucrose gradients centrifuged in the SW41Ti rotor at 40,000 rpm for 6 hours. The gradient profile of 37A RNA as compared to ribosomal RNA markers (28S and 18S) is shown. Fractions were collected from the bottom of the gradient.



Initial attempts to isolate viral RNA were not successful despite high virus titers in the cell lysates when the proteinase K digestion was done at 37°C. Raising the digestion temperature to 50°C was necessary to obtain viral RNA in the expected quantities.

Complementary DNA Synthesis and Cloning

The synthesis and cloning of viral cDNA was undertaken to provide DNA for sequencing. The original plan was to clone and sequence several clones from each virus to obtain the sequence of the capsid coding regions. While this work was in progress, it became possible to sequence the RNA directly. Therefore, only the cDNA insert of pBR205-110 was completely sequenced and reported here.

The yield of transformants containing mengovirus-derived cDNA inserts varied widely. Some experiments produced no successful insertions. The best yield of transformants was from an experiment that used low melting agarose for size selection of 205-derived cDNA. Nearly 3000 transformants, distributed fairly evenly between seven gel slices, were obtained.

When colonies were probed with [32-P]-labeled viral RNA, the results indicated that 93% of clones contained cDNA inserts derived from mengovirus RNA sequences. Clones were considered to be of viral origin when the clone bound the viral RNA probe and did not bind the cellular RNA probe. The clones with the strongest hybridization signals were chosen for further study.

Restriction enzyme analysis of plasmid DNA showed that most clones contained inserts of roughly 2 kb or less in size as estimated by

agarose gel electrophoresis. A total of 172 clones derived from 37A, 280, and 205 RNA with inserts of greater than 2 kb were selected for further testing. Although the primer for first stand synthesis was oligo(dT), only 30% of the clones (15 of 45 37A clones, 7 of 37 280 clones, and 29 of 90 205 clones) contained poly(A) sequences when slot blots of plasmid DNA were hybridized with oligo(dT) end-labeled with [32-P].

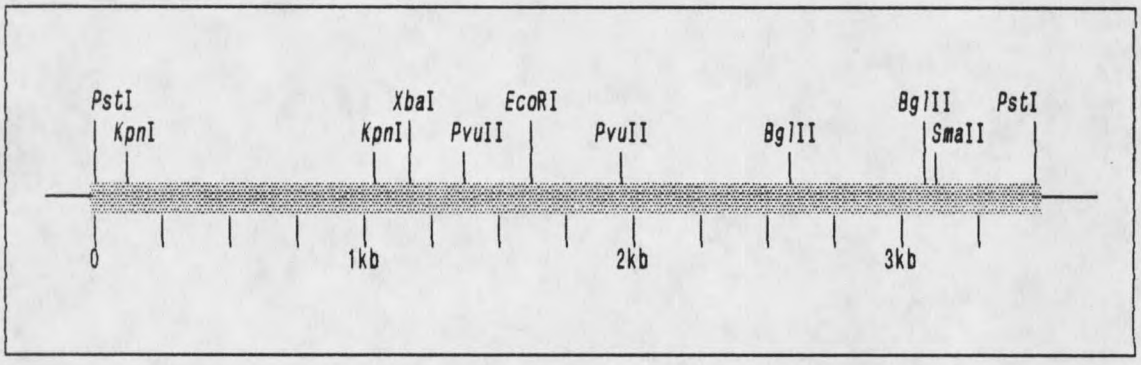
Slot blots for 32 37A clones, 24 280 clones, and 42 205 clones negative for poly(A) sequences were also hybridized with the 5'-END, 5'-CAP, and 3'-CAP probes. The results are summarized in Table 7. Although the 5'-CAP and 3'-CAP probes hybridized to a total of 32 clones, only two 205 clones were positive with both probes. Both clones were isolated from the same gel slice and were found to have the same length and restriction map. One of these clones, pBR205-110 was selected for further study.

Table 7. Results of the hybridization of plasmid DNA slot blots with synthetic oligonucleotide probes 5'-END, 5'-CAP, and 3'-CAP. The position and sequence of the probes are shown in Appendix A, Table 9.

Clones	Number Probed	Number Positive with Probe		
		5'-END	5'-CAP	3'-CAP
37A	32	5	0	0
205	42	7	4 (1 weak)	27
280	24	1	0	1

The restriction map constructed for pBR205-110 cDNA insert and confirmed by sequencing data is shown in Figure 10. The size of the insert, estimated to be about 3000 base pairs from the gels, was more correctly determined from sequencing to be 3448 base pairs. The length and positions of restriction sites are shown on the map as determined by sequencing.

Figure 10. Restriction map of the cDNA insert of pBR205-110. The insert (▒▒) is shown in its entirety. Only short, flanking sequences of the vector (——) are shown.



Sequencing of pGEM205-110

The cDNA insert of pBR205-110 containing the entire capsid coding region was subcloned into pGEM-3Z and sequenced in its entirety using both the Promega and Sequenase DNA sequencing kits. Sequence data was obtained from three sources. The ends of the pGEM205-110 subclone were sequenced using the T7 and SP6 primers. The majority of the sequence data was provided by exonuclease III-generated deletion subclones sequenced using the T7 primer. The remaining information was obtained from three additional subclones prepared from restriction enzyme fragments of pGEM205-110. The position of the sequence data provided by

these subclones, the parental pGEM205-110 clone, and deletion subclones are shown in Figure 11.

Initial attempts to produce deletion mutants of pGEM205-110 with exonuclease III for sequencing were not successful. The DNA deletions found were of random lengths and resulted from digestion at multiple sites in the DNA rather than the exclusive digestion at the unprotected 5'-overhanging end required for nonrandom deletions. The plasmid preparation procedure at that time included incubation at 55°C during the NaOH denaturation of the DNA. This can result in nicking of the DNA. Exonuclease III can initiate exonuclease activity at nicks (78). When the NaOH treatment of the plasmid DNA was limited to 10 min and done in an ice bath, plasmid DNA suitable for exonuclease III digestion was obtained.

Exonuclease III-derived subclones with minimal deletions were initially difficult to locate. Agarose gel electrophoresis of uncut plasmid DNA was used to screen transformants for deletion size and plasmids running behind uncut parental pGEM205-110 plasmid DNA markers were discarded as being too large to contain deletions. However, when several of these larger plasmids were cut with the restriction enzyme *Eco*RI for more accurate sizing, they were found to be shorter than the parental pGEM205-110 plasmid.

About 100 transformants were recovered from each timepoint in the exonuclease III digestion of the pGEM205-110 subclone. Figure 12 shows the consensus sequence obtained for the pGEM205-110 cDNA insert.

Figure 11. The position and clonal source of the sequence of pGEM205-110 obtained from T7 (▒) or SP6 (▒) primers.

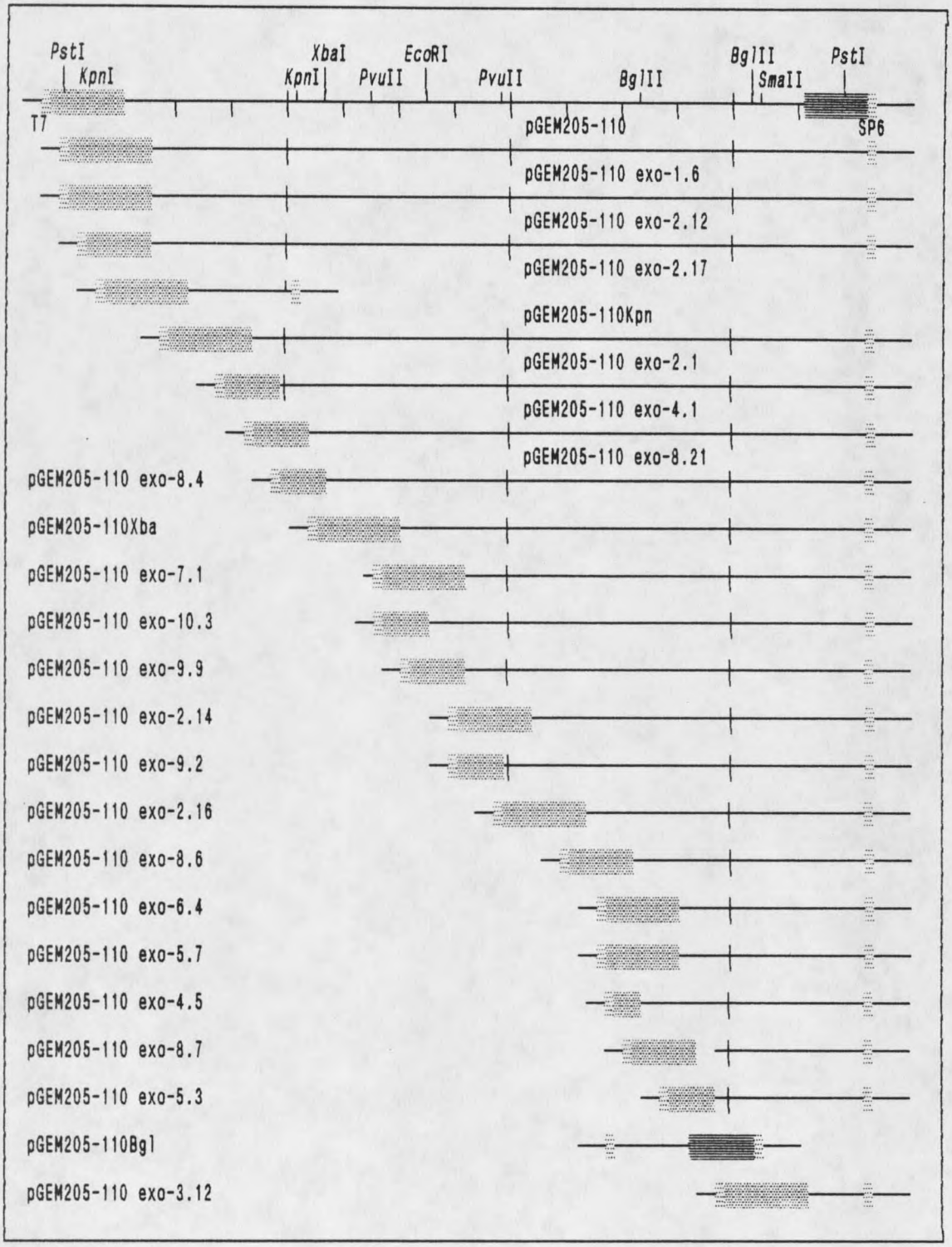


Figure 12. Nucleotide sequence of pGEM205-110.

```
TTTTCCTTTGAAAACCACGACAATAATATGGCTACAACCATGGAACAAGAGATTTGTGCT     60
CATTCCATGACCTTTGAAGAATGCCCAAAATGCTCCGCCTTACAATACAGAAATGGTTTC    120
TACCTACTGAAGTATGATGAAGAGTGGTACCCTGAGGAGTTGTTGACTGATGGTGAAGAT    180
GATGTGTTCGATCCTGATTTGGACATAGAAGTTGTGTTCGAGACACAAGGCAATTCAACC    240
TCATCCGATAAAAACAATTCTTCTTCCGAGGGTAATGAAGGAGTGATTATAAATAATTTC    300
TATTCCAATCAATACCAAAATTCAATTGATTTATCTGCCAATGCTACTGGCTCTGACCCC    360
CCAAAAACCTATGGCCAATTTTCGAATCTGTTGTCAGGAGCAGTCAATGCTTTTTCTAAC    420
ATGCTCCCCCTCCTTGCAGATCAGAATACAGAGGAGATGGAGAATCTATCAGACCGAGTG    480
TCTCAAGACACTGCCGGCAACACGGTCACAAACACCCAATCAACCGTTGGTCGTCTTGTC    540
GGATACGGAACAGTTCATGATGGGGAACATCCAGCTTCTTGTGCTGATACTGCCTCAGAG    600
AAAATTCTGGCAGTTGAAAGATATTACACTTTCAAAGTGAATGATTGGACTTCCACCCAG    660
AAACCTTTTGAATATATTCGTATACCACTACCACATGTTCTATCAGGTGAGGATGGTGGT    720
GTCTTTGGAGCTGCTCTCCGAAGACACTACCTTGTAAAAACTGGCTGGAGAGTTCAGGTC    780
CAGTGTAATGCATCCCAGTTTCATGCGGGGAGCCTCCTGGTCTTTATGGCCCCGGAGTAT    840
CCTACACTGGACGTGTTTGCAATGGACAACAAATGGTCCAAGGACAATCTGCCTAATGGA    900
ACAAGAACACAAGCCAACAGGAAGGGTCCTTTCGCTATGGACCACCAAAATTTTTGGCAA    950
TGGACTCTCTATCCGCACCAGTTTTTAAATCTAAGAACTAATACGACGGTAGATTTGGAG   1020
GTACCTTATGTAAACATTGCCCCTACCTCCTCCTGGACTCAACATGCTTCCTGGACTTTG   1080
GTGATTGCCGTGGTAGCCCCCTTGACCTATTCGACTGGTGCTTCAACAAGCCTTGACATC   1140
ACCGCCTCAATTCAACCTGTTCGTCCCGTGTTTAACGGACTACGGCATGAGGTACTGTCT   1200
AGACAGTCTCCTATCCCTGTTACCATCCGGGAGCATGCTGGGACCTGGTATTCCACTCTT   1260
CCTGATAGCACTGTTCCCATCTATGGAAAAACTCCTGTGGCCCCCGCGAATTATATGGTA   1320
GGAGAATACAAAGACTTCTTGGAGATTGCCCAAATTCCGACCTTCATTGGGAACAAAGTG   1380
```

Figure 12, pGEM205-110 sequence continued.

```
CCCAATGCTGTCCCATACATTGAGGCTTCCAACACAGCTGTCAAGACACAACCACTTGCC  1440
GTTTATCAGGTAACTCTTTCATGTTCTTGCTTGGCTAACACATTCCTAGCCGCTCTCTCC  1500
CGGAATTTTGCCCAGTACCGCGGATCACTAGTGTACACTTTTGTGTTCACTGGGACTGCG  1560
ATGATGAAGGGTAAGTTTCTTATTGCATACACCCCTCCCGGCGCCGGGAAACCGACAAGC  1620
AGAGATCAGGCCATGCAAGCCACTTACGCTATTTGGGATCTAGGTCTGAATTCTTCGTAC  1680
TCGTTTACTGTGCCTTTCATTTCTCCAACCCACTTCCGCATGGTGGGAACTGATCTAGTT  1740
AACATTACTAATGCGGATGGTTGGGTCACTGTATGGCAGTTGACCCCTCTTACTTACCCT  1800
CCAGGTTGCCCTACCTCCGCAAAAATCCTCACAATGGTTAGCGCTGGGAAGGACTTCTCT  1860
CTAAAAATGCCCATTTCTCCTGCCCCTTGGAGCCCTCAAGGCGTAGAAAATGCAGAGAAG  1920
GGTGTAACTGAAAATACAGATGCCACAGCTGATTTTGTGGCACAGCCTGTTTATTTGCCT  1980
GAGAACCAGACTAAGGTAGCCTTTTTCTATGATAGGTCTAGTCCCATTGGAGCTTTCACT  2040
GTGAAATCTGGGAGCTTAGAGTCGGGATTTGCCCCTTTCTCCAACCAGGCGTGTCCCAAT  2100
TCCGTCATTTTAACACCAGGGCCGCAGTTTGATCCGGCTTATGATCAACTGCGTCCACAG  2160
CGTTTGACCGAAATTTGGGGCAATGGGAATGAGGAAACCTCCGAAGTCTTCCCTCTAAAA  2220
ACTAAGCAGGATTATTCGTTTTGTCTCTTTTCCCCCTTTGTGTATTATAAATGTGACCTG  2280
GAAGTGACCCTAAGTCCACACACCTCCGGTAATCACGGCCTGTTGGTCCGCTGGTGCCCT  2340
ACCGGAACTCCCAACAAGCCCACCACCCAGGTGCTGCATGAGGTAAGTTCTCTCTCAGAA  2400
GGGCGAACCCCACAGGTGTACAGTGCCGGACCTGGTACTTCCAATCAGATTTCATTTGTA  2460
GTTCCTTATAATTCGCCTTTATCTGTTCTGCCTGCTGTTTGGTATAATGGGCATAAGAGA  2520
TTTGACAACACAGGCTACTTGGGAATAGCTCCCAATTCTGATTTCGGCACCCTCTTCTTT  2580
GCTGGAACGAGATCTGATATTAAATTCACTGTATATTTGAGATACAAAAACATGAGGGTC  2640
TTCTGCCCACGTCCAACTGTTTTCTTTCCTTGGCCTACTTCTGGTGACAAAATAGATATG  2700
ACTCCTAGAGCAGGAGTTCTTATGCTTGAAAGCCCCAACCCTTTGGATGTCTCTAAAACT  2760
```

Figure 12, pGEM205-110 sequence continued.

TACCCAACTCTGCACATCCTGCTACAATTCAACCACAGAGGATTGGAAGCCAGGATTTTC	2820
AGACATGGCCAGCTTTGGGCCGAAACACATGCAGAAGTGGTGCTTAGGTCAAAGACAAAA	2880
CAAATCAGCTTCCTGAGTAATGGTAGCTACCCCTCTATGGATGCCACAACACCGTTAAAC	2940
CCCTGGAAGAGCACATATCAGGCAGTTTTGCGTGCTGAACCCCATCGGGTGACCATGGAC	3000
GTGTACCACAAGAGAATAAGGCCATTTAGATTGCCCCTAGTCCAGAAAGAGTGGAGGACT	3060
TGTGAAGAGAATGTTTTTGGTCTGTATCATGTCTTCGAAACACATTATGCAGGATACTTT	3120
TCAGATCTTTTGATCCACGATGTCGAGACCAATCCCGGGCCTTTCATGTTTAAACCAAGA	3180
CAGTGGCTGGTTTTTCAGACTCAAGGAGCGGCAGTGTCATCAATGGCTCAAACCCTACTG	3240
CCGAACGACTTGGCCAGCAAAGCTATGGGATCAGCCTTTACGGCTTTGCTCGATGCCAAC	3300
GAGGACGCCCAAAAAGCAATGAAGATTATAAAGACGTTAAGTTCTCTATCGGATGCATGG	3360
GAAAATGTAAAAGGAACATTAAACAACCCGGAGTTTTGGAAACAACTCTTAAGCAGATGT	3420
GTGCAACTGATTGCCGGGATGACGATAG	3448

Consensus Sequencing of the Capsid Coding Region of the Viral RNAs

The capsid coding regions of 37A, 280, 205-A7, and 205-D2 were obtained by consensus RNA sequencing. The consensus RNA sequence of 205 was also determined and compared to the cDNA sequence of pGEM205-110.

The consensus RNA sequence data for the leader, 1A, 1B, 1C, and 1D coding regions for 37A are shown below in Figures 13, 15, 17, 19, and 21. When differences were found, the data from all five RNAs are shown. The amino acid sequences of the leader, 1A. 1B, 1C, and 1D capsid proteins predicted from the nucleotide sequence data for 37A are shown below in Figures 14, 16, 18, 20, and 22 along with the amino acid

sequence of the M variant (75). When differences were seen between 37A, mutants 280 and 205, or revertants 205-A7 and 205-D2, those data are also shown. The cleavage sites were determined from a comparison of the amino acid sequences with the amino- and carboxyl-terminal ends of the mengovirus capsid proteins (127) and by comparison with the amino acid sequence of the M variant (75).

The leader region nucleotide sequence was conserved between 37A, 280, 205, 205-D2, and 205-A7 and is shown in Figure 13. There was one area of compression and the base at position 65 in the leader region was not resolved in any of the RNA gels. The area was resolved in the DNA sequencing of pGEM205-110 and the base at position 65 is a G, the same base as found in the M variant. The amino acid sequence of the leader region is shown in Figure 14.

Figure 13. Nucleotide sequence of the leader peptide of the mengovirus 37A variant determined by consensus RNA sequencing. The DNA sequence of the 205-110 insert is shown where the RNA gels did not resolved the sequence. A "." indicates that the amino acids are the same.

```
37A        ATGGCTACAACCATGGAACAAGAGATTTGTGCTCATTCCATGACCTTTGA   50

37A        AGAATGCCCAAAATNCTCCGCCTTACAATACAGAAATGGTTTCTACCTAC  100
205-110    ..............G.................................
```

```
37A        TGAAGTATGATGAAGAGTGGTACCCTGAGGAGTTGTTGACTGATGGTGAA  150

37A        GATGATGTGTTCGATCCTGATTTGGACATAGAAGTTGTGTTCGAGACACAA  201
```

Figure 14. Amino acid sequence of the leader peptide of the mengovirus 37A variant aligned to the M variant sequence. The DNA sequence of pGEM205-110 insert is shown where the RNA gels did not resolved the sequence. A "." indicates that the amino acids are the same.

```
    37A   MATTMEQEICAHSMTFEECPKXSALQYRNGFYLLKYDEEWYPEELLTDGE   50
      M   ......................C.......................S.....
205-110   ......................C.........................

    37A   DDVFDPDLDIEVVFETQ                                   67
      M   ..........M......
```

No changes in the nucleotide sequence of the 1A region were found between 37A, 280, 205, 205-A7, and 205-D2 as shown in Figure 15. The amino acid sequence of the 1A protein is shown in Figure 16.

Figure 15. Nucleotide sequence of the 1A capsid protein of the mengovirus 37A, 280, 205, 205-A7, and 205-D2 variants determined by consensus RNA sequencing.

```
GGCAATTCAACCTCATCCGATAAAAACAATTCTTCTTCCGAGGGTAATGA      50

AGGAGTGATTATAAATAATTTCTATTCCAATCAATACCAAAATTCAATTG      100

ATTTATCTGCCAATGCTACTGGCTCTGACCCCCCAAAAACCTATGGCCAA      150

TTTTCGAATCTGTTGTCAGGAGCAGTCAATGCTTTTTCTAACATGCTCCC      200

CCTCCTTGCA                                              210
```

Figure 16. Amino acid sequence of 1A capsid protein.

37A	GNSTSSDKNNSSSEGNEGVIINNFYSNQYQNSIDLSANATGSDPPKTYGQ	50
37A	FSNLLSGAVNAFSNMLPLLA	70

The nucleotide sequence of the 1B region is shown in Figure 17. There were two areas in the 1B region for which it was difficult to determine the sequence in RNA sequencing gels although both areas were resolved in DNA sequencing gels. The first region affected position 280 in the 1B region where a single base was not resolved on the 37A, 280, and 205 RNA gels. The base was resolved as a G in the pGEM205-110 DNA sequencing gels and on 205-A7 and 205-D2 RNA gels. The second area of compression affected resolution of several bases between 366 and 381 in the 1B region. Again, these bases were resolved on the pGEM205-110 DNA sequencing gels and on some of the RNA gels. The amino acid sequence of the 1B protein is shown in Figure 18.

There were no sequence differences found in the 1C region between 37A, 280, 205, 205-A7, and 205-D2. The nucleotide sequence of the 1C region is shown in Figure 19. The amino acid sequence of the 1C protein is shown in Figure 20.

Figure 17. Nucleotide sequence of the 1B capsid protein of the mengovirus 37A variant determined by consensus RNA sequencing. The DNA sequence of the 205-110 insert is shown where the RNA gels did not resolve the sequence. A "." indicates that the sequences are the same.

```
37A       GATCAGAATACAGAGGAGATGGAGAATCTATCAGACCGAGTGTCTCAAGA    50

37A       CACTGCCGGCAACACGGTCACAAACACCCAATCAACCGTTGGTCGTCTTG   100

37A       TCGGATACGGAACAGTTCATGATGGGGAACATCCAGCTTCTTGTGCTGAT   150

37A       ACTGCCTCAGAGAAAATTCTGGCAGTTGAAAGATATTACACTTTCAAAGT   200

37A       GAATGATTGGACTTCCACCCAGAAACCTTTTGAATATATTCGTATACCAC   250

37A       TACCACATGTTCTATCAGGTGAGGATGGTNGTGTCTTTGGAGCTGCTCTC   300
280       ............................N.....................
205       ............................N.....................
205-A7    ............................G.....................
205-D2    ............................G.....................
205-110   ............................G.....................

37A       CGAAGACACTACCTTGTAAAAACTGGCTGGAGAGTTCAGGTCCAGTGTAA   350

37A       TGCATCCCAGTTTCANGCGGGGAGCCTNNTNGTCTTTATGGCCCCGGAGT   400
280       ..............N...........CC.N...................
205       ..............NN..........CC.G...................
205-A7    ..............N...........CN.N...................
205-D2    ..............N...........CC.G...................
205-110   ..............T...........CC.G...................

37A       ATCCTACACTGGACGTGTTTGCAATGGACAACAAATGGTCCAAGGACAAT   450

37A       CTGCCTAATGGAACAAGAACACAAGCCAACAGGAAGGGTCCTTTCGCTAT   500

37A       GGACCACCAAAATTTTTGGCAATGGACTCTCTATCCGCACCAGTTTTTAA   550

37A       ATCTTAGAACTAATACGACGGTAGATTTGGAGGTACCTTATGTAAACATT   600

37A       GCCCCTACCTCCTCCTGGACTCAACATGCTTCCTGGACTTTGGTGATTGC   650

37A       CGTGGTAGCCCCCTTGACCTATTCGACTGGTGCTTCAACAAGCCTTGACA   700

37A       TCACCGCCTCAATTCAACCTGTTCGTCCCGTGTTTAACGGACTACGGCAT   750

37A       GAGGTACTGTCTAGACAG                                   768
```

Figure 18. Amino acid sequence of the 1B capsid protein aligned with the M variant sequence. The amino acid sequence of the 205-110 insert is shown where the RNA gels did not resolved the sequence. A "." indicates that the bases are the same.

```
37A      DQNTEEMENLSDRVSQDTAGNTVTNTQSTVGRLVGYGTVHDGEHPASCAD    50
  M      ..................................................

37A      TASEKILAVERYYTFKVNDWTSTQKPFEYIRIPLPHVLSGEDGXVFGAAL   100
280      ..........................................X....A.
205      ..........................................G....A.
205-A7   ..........................................G....A.
205-D2   ..........................................G....A.
205-110  ..........................................G....A
  M      ..........................................G....T.

37A      RRHYLVKTGWRVQVQCNASQFXAGSLXVFMAPEYPTLDVFAMDNKWSKDN   150
280      ....................X....L..................K.....
205      ....................X....L..................K.....
205-A7   ....................X....X..................K.....
205-D2   ....................X....L..................K.....
205-110  ....................H....L..................K.....
  M      ....................H....L..................R.....

37A      LPNGTRTQANRKGPFAMDHQNFWQWTLYPHQFLNLRTNTTVDLEVPYVNI   200
  M      ........T...........................................

37A      APTSSWTQHASWTLVIAVVAPLTYSTGASTSLDITASIQPVRPVFNGLRH   250
  M      ..................................................

37A      EVLSRQ                                               256
  M      ......
```

Figure 19. Nucleotide sequence of the 1C capsid protein of the mengovirus 37A, 280, 205, 205-A7, and 205-D2 variants determined by consensus RNA sequencing.

TCTCCTATCCCTGTTACCATCCGGGAGCATGCTGGGACCTGGTATTCCAC	50
TCTTCCTGATAGCACTGTTCCCATCTATGGAAAAACTCCTGTGGCCCCCG	100
CGAATTATATGGTAGGAGAATACAAAGACTTCTTGGAGATTGCCCAAATT	150
CCGACCTTCATTGGGAACAAAGTGCCCAATGCTGTCCCATACATTGAGGC	200
TTCCAACACAGCTGTCAAGACACAACCACTTGCCGTTTATCAGGTAACTC	250
TTTCATGTTCTTGCTTGGCTAACACATTCCTAGCCGCTCTCTCCCGGAAT	300
TTTGCCCAGTACCGCGGATCACTAGTGTACACTTTTGTGTTCACTGGGAC	350
TGCGATGATGAAGGGTAAGTTTCTTATTGCATACACCCCTCCCGGCGCCG	400
GGAAACCGACAAGCAGAGATCAGGCCATGCAAGCCACTTACGCTATTTGG	450
GATCTAGGTCTGAATTCTTCGTACTCGTTTACTGTGCCTTTCATTTCTCC	500
AACCCACTTCCGCATGGTGGGAACTGATCTAGTTAACATTACTAATGCGG	550
ATGGTTGGGTCACTGTATGGCAGTTGACCCCTCTTACTTACCCTCCAGGT	600
TGCCCTACCTCCGCAAAAATCCTCACAATGGTTAGCGCTGGGAAGGACTT	650
CTCTCTAAAAATGCCCATTTCTCCTGCCCCTTGGAGCCCTCAA	693

Figure 20. Amino acid sequence of the 1C capsid protein of 37A aligned with the M variant sequence. A "." indicates that the bases are the same.

```
37A     SPIPVTIREHAGTWYSTLPDSTVPIYGKTPVAPANYMVGEYKDFLEIAQI    50
 M      ..................................................

37A     PTFIGNKVPNAVPYIEASNTAVKTQPLAVYQVTLSCSCLANTFLAALSRN   100
 M      ..................................................

37A     FAQYRGSLVYTFVFTGTAMMKGKFLIAYTPPGAGKPTSRDQAMQATYAIW   150
 M      ..................................................

37A     DLGLNSSYSFTVPFISPTHFRMVGTDLVNITNADGWVTVWQLTPLTYPPG   200
 M      ........................QA....V.....................

37A     CPTSAKILTMVSAGKDFSLKMPISPAPWSPQ                      231
 M      ............................
```

The 1D region was found to be the most variable region sequenced. The nucleotide sequence is shown in Figure 21. It was the only capsid coding region to have any differences between 37A and mutants 280 and 205. The two base changes found in mutants 280 and 205 were identical. Both sequence changes resulted in coding changes. The first base change occurred at position 692 where the A found in 37A and revertants 205-A7 and 205-D2 was replaced by a G in 280 and 205. The second change occurred nearby at position 694 where a C found in 37A and the revertants was replaced by a T in 280 and 205. The coding changes caused by these base changes occur at adjacent amino acid residues, 231 and 232, where a lysine-proline pair in 37A and the revertants is

replaced by an arginine-serine pair in 280 and 205. The amino acid sequences of the 1D proteins are shown in Figure 22.

Figure 21. Nucleotide sequence of the 1D capsid protein of mengovirus 37A, 280, 205, 205-A7, and 205-D2 determined by consensus RNA sequencing. A "." indicates that the sequences are the same.

```
37A      GGCGTAGAAAATGCAGAGAAGGGTGTAACTGAAAATACAGATGCCACAGC     50
37A      TGATTTTGTGGCACAGCCTGTTTATTTGCCTGAGAACCAGACTAAGGTAG    100
37A      CCTTTTTCTATGATAGGTCTAGTCCCATTGGAGCTTTCACTGTGAAATCT    150
37A      GGGAGCTTAGAGTCGGGATTTGCCCCTTTCTCCAACCAGGCGTGTCCCAA    200
37A      TTCCGTCATTTTAACACCAGGGCCGCAGTTTGATCCGGCTTATGATCAAC    250
37A      TGCGTCCACAGCGTTTGACCGAAATTTGGGGCAATGGGAATGAGGAAACC    300
37A      TCCGAAGTCTTCCCTCTAAAAACTAAGCAGGATTATTCGTTTTGTCTCTT    350
37A      CTCCCCCTTTGTGTATTATAAATGTGACCTGGAAGTGACCCTAAGTCCAC    400
37A      ACACCTCCGGTAATCACGGCCTGTTGGTCCGCTGGTGCCCTACCGGAACT    450
37A      CCCAACAAGCCCACCACCCAGGTGCTGCATGAGGTAAGTTCTCTCTCAGA    500
37A      AGGGCGAACCCCACAGGTGTACAGTGCCGGACCTGGTACTTCCAATCAGA    550
37A      TTTCATTTGTAGTTCCTTATAATTCGCCTTTATCTGTTCTGCCTGCTGTT    600
37A      TGGTATAATGGGCATAAGAGATTTGACAACACAGGCTACTTGGGAATAGC    650
37A      TCCCAATTCTGATTTCGGCACCCTCTTCTTTGCTGGAACGAAACCTGATA    700
280      .....................................G.T......
205      .....................................G.T......
205-A7   .....................................A.C......
205-D2   .....................................A.C......
37A      TTAAATTCACTGTATATTTGAGATACAAAAACATGAGGGTCTTCTGCCCA    750
37A      CGTCCAACTGTTTTCTTTCCTTGGCCTACTTCTGGTGACAAAATAGATAT    800
37A      GACTCCTAGAGCAGGAGTTCTTATGCTTGAA                       831
```

Figure 22. Amino acid sequence of the 1D capsid protein of 37A, 280, 205, 205-A7, and 205-D2 aligned with the M variant sequence. A "." indicates that the bases are the same.

```
37A       GVENAEKGVTENTDATADFVAQPVYLPENQTKVAFFYDRSSPIGAFTVKS   50
 M        ..............................................A...

37A       GSLESGFAPFSNQACPNSVILTPGPQFDPAYDQLRPQRLTEIWGNGNEET  100
 M        ............K.....................................

37A       SEVFPLKTKQDYSFCLFSPFVYYKCDLEVTLSPHTSGNHGLLVRWCPTGT  150
 M        ....................................A.............

37A       PNKPTTQVLHEVSSLSSEGRTPQVYSAGPGTSNQISFVVPYNSPLSVLPA  200
 M        .T..............................................

37A       VWYNGHKRFDNTGYLGIAPNSDFGTLFFAGTKPDIKFTVYLRYKNMRVFCP 250
280       ............Y.................RS..................
205       ............Y.................RS..................
205-A7    ............Y.................KP..................
205-D2    ............Y.................KP..................
 M        ............D.................KP..................

37A       RPTVFFPWPTSGDKIDMTPRAGVLMLE                       277
 M        ........................
```

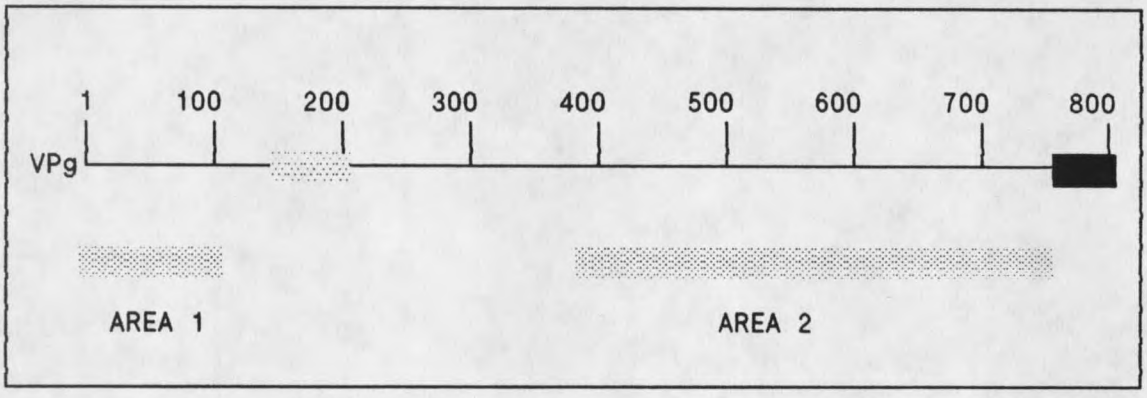
Consensus Sequences from the 5'-Noncoding Region

Since the 5'-noncoding region has been determined to be an important determinant in the neurovirulence of other picornaviruses, consensus RNA sequencing was extended into the 5'-noncoding region. Because of problems in sequencing long homopolymeric C regions, the sequence data for the entire poly(C) tract was not obtained and is not shown. Several other regions in this area were also difficult to

because of the strong secondary structure of the region. The data obtained from those areas is also not shown.

Partial nucleotide sequence data from the 5'-noncoding region of 37A, 280, 205, 205-A7, and 205-D2 were obtained. The data cover two regions of the mengovirus 5'-noncoding region as shown in Figure 23. The first region covers about the first 100 nucleotides of sequence. The second and larger region is between the poly(C) track and the site of initiation of translation.

Figure 23. Position of the sequence data obtained from the 5'-noncoding region of the mengovirus genome relative to the M variant genomic sequence. The poly(C) track is indicated by ▒▒. The polyprotein coding region is shown by ▬.



The alignment of the sequence data from area 1 is shown below in Figure 24. No differences were seen between 37A, 280, 205, 205-A7, and 205-D2. Several bases at the beginning of the RNAs were not resolved but, since the region is highly conserved among the picornaviruses, are probably T residues.

The sequence data from area 2 are shown in Figure 25. The positions shown are relative to the M variant sequence.

Figure 24. Nucleotide sequence of area 1 in the 5'-noncoding region of the 37A mengovirus genome.

```
NNTGAAAGCCGGGGGTGGGAGATCCGGATTGCCGGTCCGCTCGATATCGC      50

GGGCCGGGTCCGTGACTACCCACTCCCCCTTTCAACGTGAAGGCTACGAT      100
```

Figure 25. Alignment of the nucleotide sequence of area 2 in the 5'-noncoding region of the mengovirus genome. A "." indicates that the sequence data are identical.

```
37A       AGGAAGCAGTTCCTCTGGACGCTTCTTGAAGACAAGCAACGTCTGTAGCG   440

37A       ACCCTTTGCAGGCAGCGGAATCCCCCACCTGGTGACAGGNNCCTCTGCGG   490
280       ....................................TN.........
205       ....................................TN.........
205-A7    ....................................TG.........
205-D2    ....................................TN.........

37A       CAACCCCAGTCCGAAAGCCACGTGTGTAAGACACACCTGCAAAGGCGGCA   540

37A       GCCACGTTGTGCGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCA   590

37A       AGCGTATTCAACAAGGGGCTGAAGGATGCCCAGAAGGTACCCCACTGGCT   640

37A       GGGATCTGATCTGGGGCCTCGGTGCNCGTGCTTTACACGCGTTGAGTCGA   690
280       ......................NN..........C.........
205       ......................G...........G.........
205-A7    ......................G...........G.........
205-D2    ......................G...........G.........

37A       GGTTAAAAAACGTCTAGGCCCCCCGAANCACGGGGACGTGGTTTTCCTTT   740
280       ........................C.....................
205       ........................C............N.......
205-A7    ........................C.....................
205-D2    ........................C.....................

37A       GAAAACCACGACAATAAT                               757
```

No sequence differences were confirmed between 37A, 280, and the M variant except in region 2. The sequence of the genomes of mutant 205 and revertants 205-A7 and 205-D2 did differ from the other RNAs and the M variant at position 680 of area 2 where a C was replaced by a G. This was the only sequence change found that differentiated the revertants from 37A.

DISCUSSION

The goal of my research was to determine the basis of mengovirus hemagglutination by comparing the sequences obtained for the capsid coding regions of 37A, a hemagglutination-positive variant, and 280 and 205, two hemagglutination-negative mutants derived from 37A. The capsid coding regions of two hemagglutination-positive revertants of 205, 205-A7 and 205-D2, were also sequenced to determine the nature of the reversion which restored the hemagglutination-positive phenotype without restoring full virulence or plaque size.

The original experimental plan was to synthesize and clone cDNA from the genomic RNA of the viruses. For the 205 mutant, a cDNA clone, pBR205-110, containing the entire capsid coding region was obtained and the insert sequenced. Given the heterogeneity found in viral RNA, determination of a consensus sequence by sequencing cDNA clones requires sequencing multiple, independent clones. While this work was in progress, it became possible to sequence the viral RNA directly using synthetic oligonucleotide primers selected by comparison of the 205 cDNA and M variant sequences. Using a modification of method and the reagents from a commercial RNA sequencing kit, I determined the sequence of the capsid coding region of the RNA from 37A, mutants 205 and 280, and revertants 205-A7 and 205-D2.

Cloning and Sequencing of Mengovirus cDNA

Several thousand clones containing mengovirus-derived cDNA inserts were obtained for 37A, 280, and 205. Locating appropriate cDNA clones for this work was difficult due to the problem of obtaining sufficiently long cDNAs. The capsid coding region of mengovirus, not including the leader sequence, is 2502 bases in length. Most clones contained inserts under 2000 base pairs in length. In addition, the clones were derived from various regions of the genome. Although the primer for first stand synthesis was oligo(dT), only 30% of the clones recovered contained poly(A) sequences indicating that most of the 3' viral sequences present in the first strand of the cDNA were lost during the cloning process. Length of a clone alone could not be used to identify clones containing 5' capsid coding region sequences. One clone, pBR205-110 was found to contain the entire capsid coding region and was selected for further study.

Construction of a consensus sequence from cDNA clones requires the isolation of several different clones of the same region. Region-specific probes can be used to select clones containing the desired insert sequences but further work must be done to determine if clones arose independently. Although two 205 clones containing the entire capsid coding region were identified by the capsid region-specific probes on slot blots, they were isolated from the same transformation reaction and were found to have the same length and restriction map. This indicates that some multiplication may have occurred during incubation of the transformation mixtures prior to plating and that they were probably separate isolates of the same clone.

The insert size of pBR205-110 was estimated to be about 3000 base pairs from the gels but was determined from sequencing to be 3448 base pairs. The restriction map of the M variant covering the same region as 205-110 contains one BglII site. There were two *Bgl*II restriction sites in 205-110 but only one was identified during mapping since the sites are only 400 base pairs apart. The 400 base pair fragment was too small to be detected in the agarose gels used.

The restriction map of pBR205-110 revealed a lack of convenient sites for conventional subcloning of overlapping fragments for sequencing. Therefore, the exonuclease III procedure was used to produce a nested set of deletion subclones for sequencing. The first difficulty with the procedure was obtaining progressive digestion of the insert fragment. As noted in Results, the quality of the DNA used in the digestion was critical. The second difficulty was isolation of subclones with minimal deletions. The screening of uncut plasmids for size by agarose gel electrophoresis was found to bias the recovery of the longer inserts. Plasmids running behind uncut parental pGEM205-110 plasmid DNA in the gels were discarded as being too large to contain deletions. That larger plasmids might arise during the procedure was unexpected but could be explained by the recombination of several blunt-ended fragments during the self-ligation step. When several of these larger products were linearized by restriction enzyme digestion prior to sizing by agarose gel electrophoresis they were found to actually be shorter than the parental pGEM205-110 plasmid. It may be that the particular deletions in these plasmids decreased the degree of

supercoiling possible for the plasmids changing their mobility in agarose gels and leading to errors in estimation of their size.

When exonuclease III deletion mutants were not easily found to cover three small regions of the 205-110 insert, three subclones were constructed to provide sequence data to fill the gaps and complete the sequence of the insert.

Two commercial sequencing kits were used for DNA sequencing. The amount of sequence data provided per reaction was greater using the Sequenase kits, Ver. 1 or 2. In addition, there were fewer unreadable bands using the Sequenase kit. The entire sequence of pGEM205-110 was resolved. Band compression and natural stops were not as much of a problem with DNA sequencing as they were with RNA sequencing. However, much more labor was involved in producing, locating, and preparing cDNA clones for sequencing.

When the nucleotide sequence of the pGEM205-110 cDNA insert was aligned with the sequence of the M variant, no deletions or insertions were detected and the overall nucleotide homology was 98.93%. There were a total of 37 base changes which led to 18 coding changes. Twenty-five of the base changes were transitions where a purine was replaced by a purine or a pyrimidine was replaced by a pyrimidine. Fourteen of the transitions resulted in amino acid changes. Only five of the 12 transversions, where a purine is replaced by a pyrimidine or a pyrimidine is replaced by a purine, resulted in an amino acid change. One codon change was the result of two adjacent substitutions, one transversion and one transition. Two base changes in the 1D region introduced a *Bgl*II restriction site not present in the M variant. In

summary, there were seven coding changes found in the 1D region making that region the most variable region in the capsid. There were three coding changes in both the 1B and 1C regions, two coding changes in the leader region, and no coding changes in the 1A region.

Consensus RNA Sequencing

The sequence data for the 37A, 280, 205-A7, and 205-D2 mengoviruses were determined by direct sequencing of partially purified virion RNA. Further purification of the RNA was not required since synthetic oligonucleotide primers specific for viral sequences were used. The genomic RNA sequence of 205 capsid coding region was also determined and compared to the pGEM205-110 cDNA sequence previously obtained.

Primers for consensus RNA sequencing were selected by comparing the sequence data obtained from pGEM205-110 with the sequence data of the M variant, kindly provided by Dr. Ann Palmenberg, to locate conserved regions. Prospective primer sequences were analyzed for homology to other regions of the mengovirus genome using the GenePro program. Those primers with the least homology with other regions, especially in the last 5 bases at the 3'-end of the primer, were selected for use.

All but one of the primers, primer 13 (P13), efficiently primed DNA synthesis as expected producing readable sequence data. With primer P13, readable bands were only obtained after prolonged exposure of the gels. Surprisingly, the data obtained from this primer were not sequences immediately downstream from the primer. P13 is complementary

to bases 1564 to 1578 of the M variant of mengovirus. The sequence data obtained from this primer covered the region from bases 900 to 760, approximately 600 bases 3' to the expected sequences. The sequence used for synthesis of P13 was reevaluated, the primer resynthesized, and the sequencing reactions repeated with the same results.

Computer analysis of the P13 primer did not indicate any unusual homology with other regions of the mengovirus genome and sequence data from pGEM205-110 and RNA gels from P12 confirmed the presence of the P13 primer site in the RNA. Presumably, the secondary structure of the RNA in this area interfered with the binding of P13. The weak bands found on prolonged exposure of the films can be explained in several ways. If a small fraction of the primer molecules were able to bind to viral RNA molecules at the primer site and initiate DNA synthesis, a region of strong secondary structure could have produced a stem and loop between the priming site and the downstream region over which the polymerase may have jumped. If the base of the stem and loop structure involved the primer site itself, a small proportion of the primer molecules may have spanned the stem and been able to initiate DNA synthesis downstream of the stem and loop. Another and perhaps the most likely explanation for the weak banding pattern is that the primer bound very weakly to a downstream site and was occasionally able to prime DNA synthesis. Normally the sequence bands resulting from priming at a weak primer site would be masked by the much stronger bands resulting from the priming at the primary binding site. If the primary primer site was blocked, the bands resulting from the priming at the secondary primer site might be visible in long exposures.

Another region was selected about 100 bases downstream from P13 as a site for primer 13A (P13A). Priming from this site was normal and provided the sequence data needed to cover the region between P12 and P14.

The major problem in obtaining the RNA sequences was not related to problems with the primers but with the presence of natural stops and with compression of bands on the sequencing gels. Natural stops are sequences at which the polymerase prematurely terminates synthesis prior to insertion of a dideoxynucleotide. Increasing the temperature of the sequencing reactions to 50°C (46) decreased the number of stops observed on the gels but did not solve the problem. When TdT was added to the RNA sequencing reactions the number of stops was markedly decreased (33). Rarely, additional TdT (6 units versus 3 units per reaction tube) had to be added to resolve a particularly difficult stop.

Sequencing reactions with TdT must be run in parallel with reactions without TdT. Occasionally TdT erroneously altered the appearance of bands where there had been no stop in the regular sequencing reaction. Comparison of the TdT reaction set with the reaction set without TdT identified these areas and provided the correct data. Often the erroneous bands were much darker than the other bands in the region.

The problem with band compression on the sequencing gels was not totally solved. Compressions, the failure of bands to separate correctly on sequencing gels, were, as expected, not affected by the addition of TdT. Weak compressions were often resolved when the gel was repeated possibly because of increased gel temperatures due to changes

in environmental conditions. The electrophoresis apparatuses used did not have heat distribution plates so that the temperature of a gel also varied at different positions on the gel. A heat distribution plate may help prevent compression of bands since it would allow the gels to be run at higher overall temperatures without risking breakage of the glass plates holding the sequencing gel. Some compressions were resolved on formamide gels. All the regions were resolved in the DNA sequencing gels of pGEM205-110 and some were resolved on some of the RNA gels.

Cloning Errors

The comparison of the pGEM205-110 cDNA insert sequence with the 205 consensus RNA sequence revealed two discrepancies between the data which were confirmed by reexamination of both sets of sequencing gels. The changes which were seen in pGEM205-110 may represent selection of a rare RNA during cloning and reflect the heterogeneity of the viral RNA population (40) or may result from the introduction of transcription errors in the cloning process. Since there was not a problem resolving the RNA sequences at these positions, the changes found in pGEM205-110 were not present in a major population of the genomic RNA of mutant 205. The positions of these cloning errors are shown in Figure 26 below. Neither of the errors resulted in coding changes in the amino acid sequence of the polyprotein.

Figure 26. Discrepancies between the sequence of the cDNA insert of pGEM205-110 and the consensus 205 RNA sequence. A "." indicates that the bands were the same.

```
Error 1      1B Coding Region

                   530        540        550        560        570
                    *          *          *          *          *
205 RNA      ATGGACTCTCTATCCGCACCAGTTTTTAAATCTTAGAACTAATACGACGG
205-110      ..................................A.................

Error 2      1D Coding Region

                   320        330        340        350        360
                    *          *          *          *          *
205 RNA      TCCCTCTAAAAACTAAGCAGGATTATTCGTTTTGTCTCTTCTCCCCCTTT
205-110      ........................................T.........
```

Analysis of the Leader and Capsid Coding Region Sequences

Analysis of the consensus RNA sequencing data for the leader and capsid coding regions revealed no deletions or insertions in the genomic RNA of 37A, 280, 205, 205-A7, or 205-D2 compared to the M variant. There were no sequence differences seen between 37A, 280, 205, 205-A7, and 205-D2 in the leader, 1A, 1B, and 1C regions. The nucleotide and amino acid homologies in the leader and capsid coding regions between 37A and mutants 280 and 205 were 99.99% and 99.78% respectively. The nucleotide and amino acid homologies between 37A and the M variant were slightly lower at 98.93% and 98.56% respectively.

In the leader region, there were two sequence differences seen between 37A and the M variant. Both of these changes were transitions and both resulted in amino acid changes. The first coding change in the leader region was a conservative change with a nonpolar isoleucine

residue in 37A replacing the nonpolar methionine residue found in the M variant. The second coding change resulted in the replacement of a serine residue found in the M variant with a nonpolar and larger leucine residue. Since the function of the leader peptide is not yet known, the significance of these changes is unknown.

There were two base changes in the 1A region, both transitions, between 37A and the M variant. Neither of these base changes resulted in a coding change in the 1A region.

The changes occurring in the 1B, 1C, and 1D regions will be discussed in more detail below in relation to the hemagglutination phenotype and conservation of capsid structure. In summary, there were a total of 25 base changes in the 1B, 1C, and 1D regions which resulted in 11 coding changes between 37A and the M variant. In addition, there were two base changes found in the 1D region between 37A and mutants 280 and 205. Both changes resulted in amino acid changes in the 1D capsid protein.

Conservation of the β-Barrel Structure of Mengovirus Capsid Proteins

The 1D, 1B, and 1C capsid proteins of the picornavirus capsid all share a similar wedged-shaped eight-stranded antiparallel β-barrel structure (1,59,75,103). The major structural elements of the β-barrel are the eight β-sheets regions designated β_B through β_I. At the narrow end of the wedge are the four corners. Between the β-sheet regions of the protein chain are three areas of alpha helical structure, designated α_A, α_B and α_{AO}. The other regions of the proteins are more variable and form various structures designated as loops and puffs. Changes, which

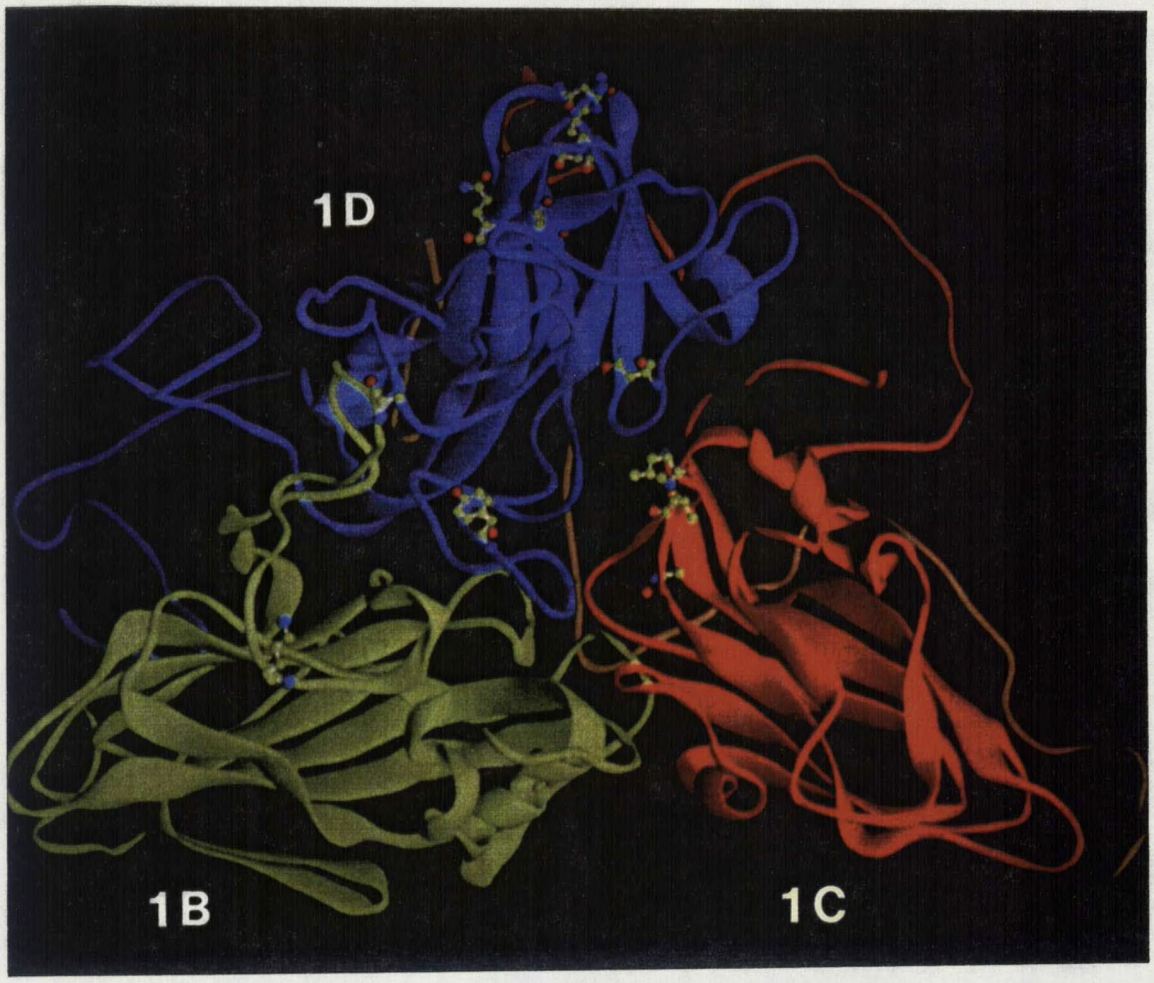
could affect the basic β-barrel structures formed by β^{B-I} regions of the capsid proteins, would be those changes which occur in the β-sheet regions. Changes in corners, loops, or puffs will not have a major effect on the overall structure.

Figure 27 shows the position of the amino acid changes found between 37A and the M variant in the 1B, 1C and 1D capsid proteins. These changes are summarized in Table 8. The position of these changes in relation to structural features in the tertiary structure of the mengovirus structure as determined by X-ray crystallography (75) are also listed in Table 8. Figure 28 shows the alignment of the structural elements of the eight-stranded antiparallel β-barrels of the 1B, 1C, and 1D capsid proteins of 37A, 205, and the M variant.

Table 8. Summary of amino acid changes between the 37A and M mengovirus variants.

Protein	Base Position	Structural Feature	37A	M
1D	47	β_B	threonine	alanine
	63	1st corner	glutamine	lysine
	138	β_E	asparagine	alanine
	152	puff	asparagine	threonine
	213	FMDV loop	tyrosine	aspartic acid
1B	99	α_A	alanine	threonine
	145	β-puff	lysine	arginine
	159	extra puff	alanine	threonine
1C	177	FMDV loop	leucine	glutamine
	178	FMDV loop	valine	alanine
	183	FMDV loop	alanine	valine

Figure 27. A RIBBON drawing of the mengovirus protomer viewed from outside of the capsid showing the amino acid changes found in 37A compared to the M variant. The carbon backbones of 1D, 1B, and 1C are shown in blue, green, and red respectively. The amino acid residues which were found to be changed are shown in molecular detail. Carbon atoms are shown in green, nitrogen atoms in blue, and oxygen atoms in red.



Only two of the amino acid changes between 37A and the M variant resulted in changes in β-sheet regions. Both changes were in the 1D protein and occurred in the β_B and β_E regions as shown in Figure 28, frames B and C. The replacement of the alanines seen in the M variant with the threonine in the β_B region and the asparagine in the β_E region

are conservative substitutions. The threonine and asparagine replacements, although bulkier, may act to strengthen the structure since they both may form hydrogen bonds with surrounding residues.

Several of the changes affect residues which line the putative cellular receptor binding site or "pit". All three coding changes seen in the 1C protein between 37A and the M variant occur in the FMDV loop and affect the pit as shown in Figure 28, frame E. Two of the changes involve substitution of valine for alanine or alanine for valine. While alanine is smaller than valine, both residues are nonpolar so these substitutions are conservative. The third change results from the substitution of leucine for a glutamine found in the M variant, and should increase the hydrophobicity of the area and perhaps affect receptor binding. The substitution of threonine with asparagine that occurs in the puff region of 1D, shown Figure 28, D, also involves a pit residue but the substitution is a conservative one.

Some of the other changes between 37A and the M variant occurred in regions which may affect the antigenic determinants of the mengovirus capsid. A replacement of the lysine residue found in the M variant at the first corner of the 1D protein with a glutamine residue, as shown in Figure 28, frame B, is probably significant since the change results in the loss of a positively-charged residue. A second major change occurred in the 1D protein at the FMDV loop as shown in Figure 28, frame E. The aspartic acid residue found in the M variant has been replaced by a tyrosine residue in 37A. Not only is the negative charge of aspartic acid lost but the tyrosine residue which replaced it is bulky and hydrophobic.

Figure 28. Structural alignment of the mengovirus capsid proteins 1D, 1B, and 1C showing the conservation of the β-barrel regions. The amino acid sequence shown for 37A is the same as 205-A7 and 205-D2. The amino acid sequence shown for 205 is the same as 280. A "." indicates that the amino acids are the same. A "-" indicates a gap or deletion. Residues in the pit are underlined. Figure is adapted from (75).

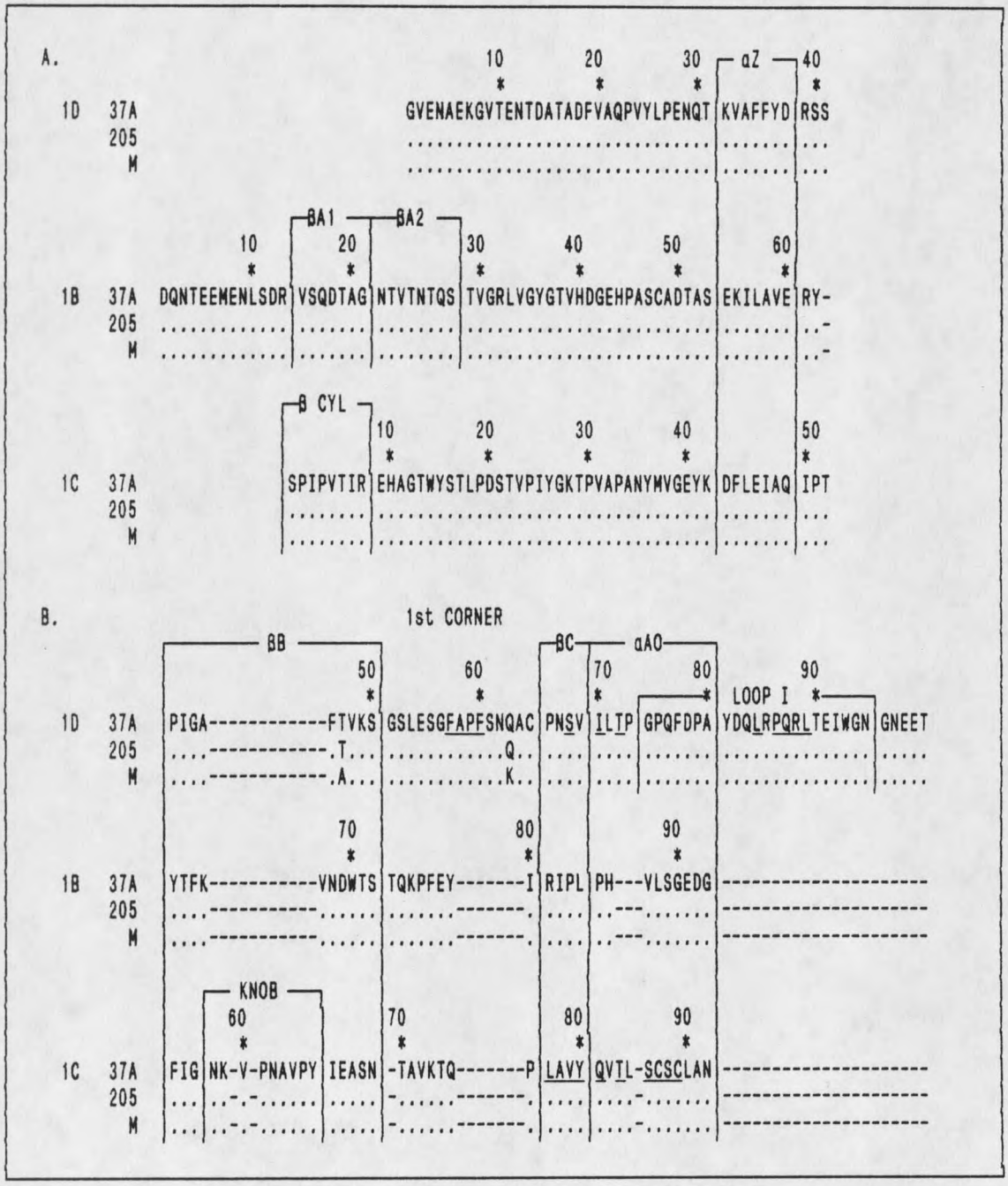


Figure 28, Structural alignment of the mengovirus capsid proteins 1D, 1B, and 1C, continued.

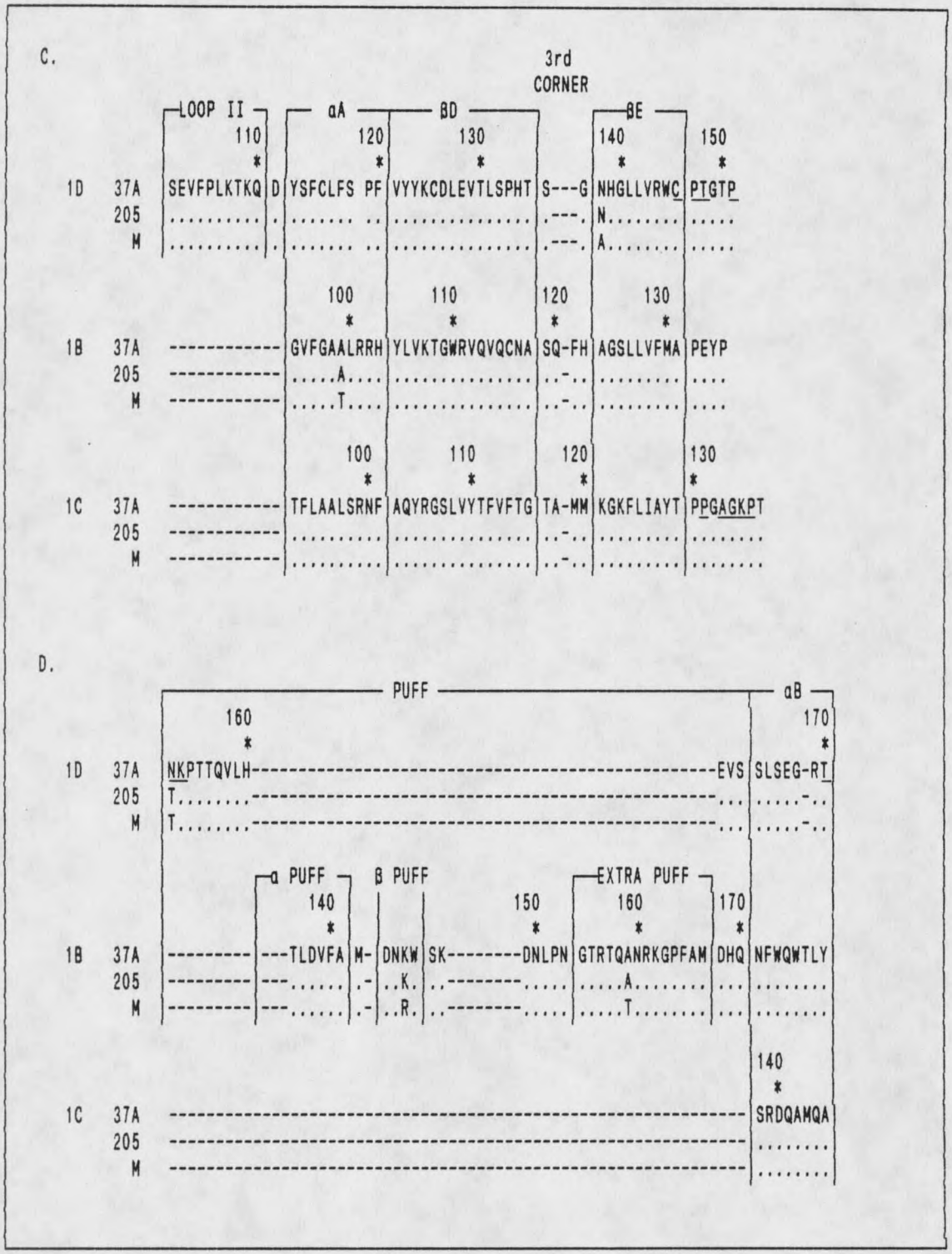
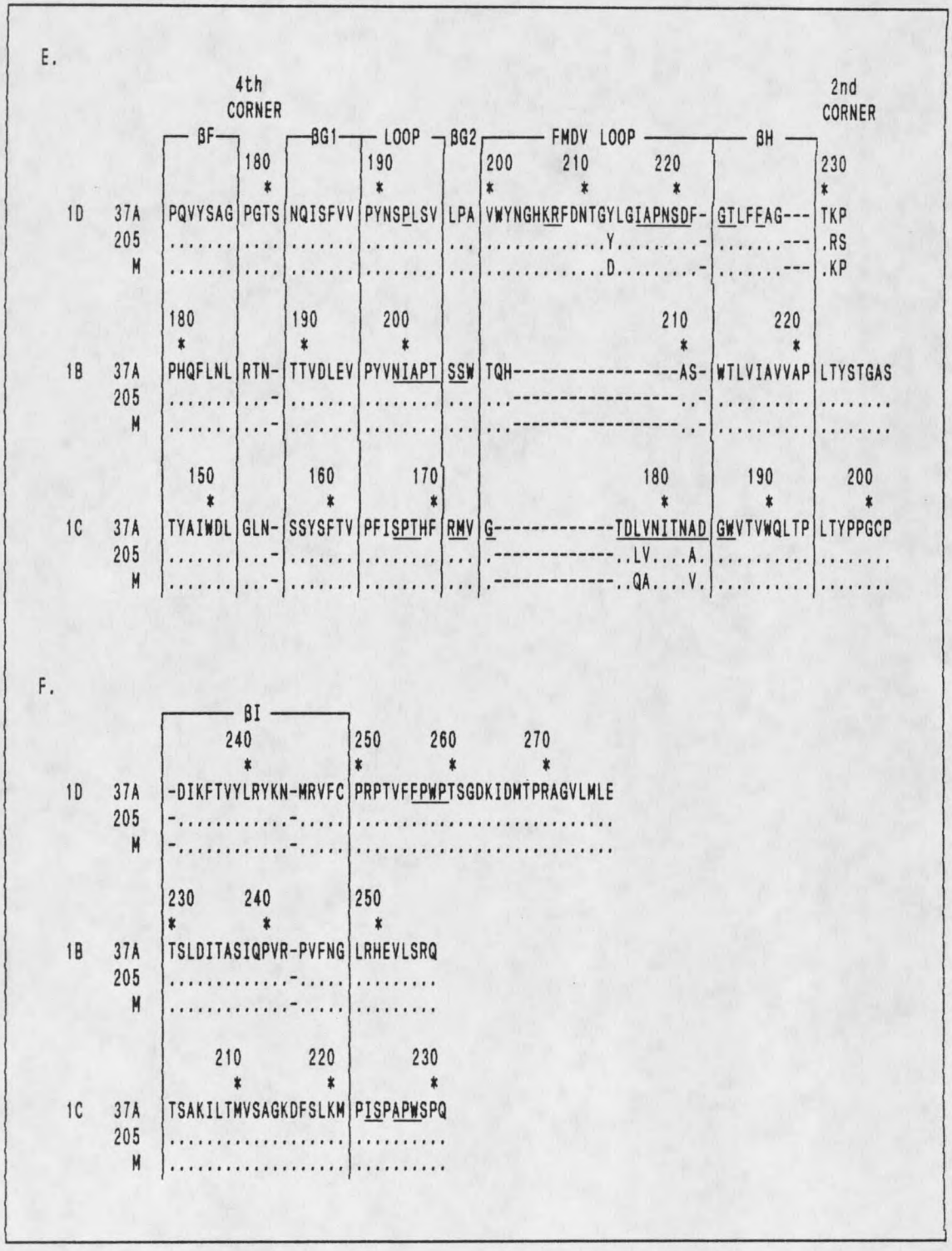


Figure 28, Structural alignment of the mengovirus capsid proteins 1D, 1B, and 1C, continued.



Capsid Changes Associated with the Hemagglutination Phenotype

Only two base differences were found in the capsid coding region of the viral RNAs between 37A and mutants 205 and 280. These changes resulted in amino acid changes at positions 231 and 232 in the 1D protein. The lysine-proline amino acid pair found in 37A was replaced by a arginine-serine pair in 280 and 205. In revertants 205-A7 and 205-D2, these positions were occupied by lysine and proline.

Figure 29 shows the positions of these two residues in relation to the putative cellular receptor binding site. It can be seen that the residues implicated in hemagglutination are not part of the pit and lie some distance away from it.

Figure 30 shows the positions of these residues as one looks down the fivefold axis of symmetry to the five 1D proteins forming the vertex of the pentamer. The proximity of the amino terminus of the 1C protein (in red) to the 231 and 232 residues of the 1D protein can perhaps explain the observation that glycophorin A bound to the 1C protein and as well as the 1D protein in cross-linking experiments by Anderson and Bond (7).

Figure 31 shows a side view of the vertex formed by five 1D molecules at the fivefold axis. The five lysine 231 residues point in towards the vertex and are part of a relatively flat platform formed by the five 1D molecules at the fivefold axis.

Figure 29. A RIBBON drawing of the mengovirus protomer showing the relationship of the amino acid residues implicated in hemagglutination to the amino acid residues forming the putative cellular receptor "pit". The two amino acid residues associated with hemagglutination (HA) are shown in molecular detail at 12 o'clock and lie at some distance from the residues shown in molecular detail at 3 o'clock which form the "pit". The carbon backbones of 1D, 1B, and 1C are shown in blue, green, and red respectively. Carbon atoms are shown in green, oxygen atoms in red, nitrogen atoms in blue, and sulfur atoms in yellow.

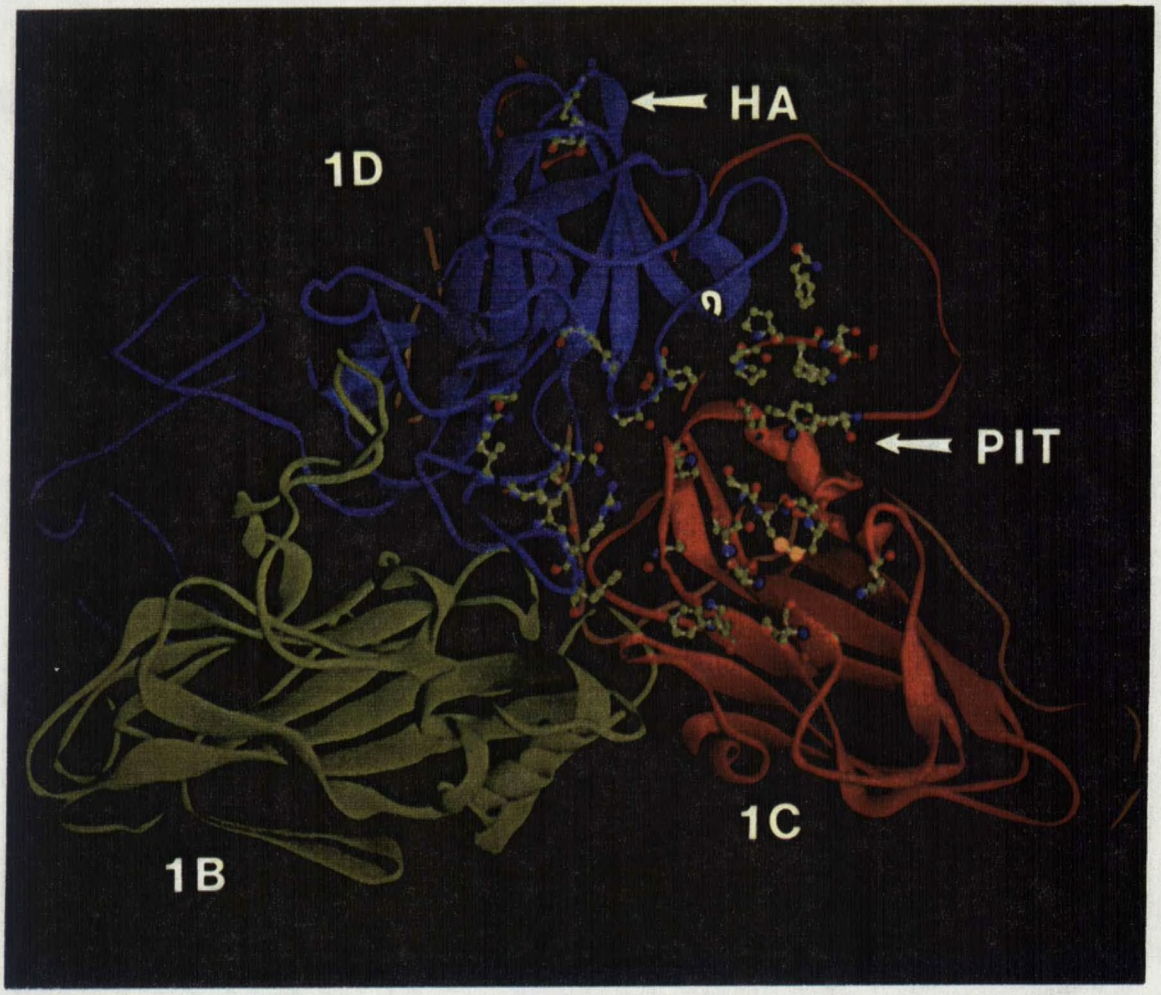


Figure 30. A line drawing looking down the fivefold axis of symmetry at the vertex (V) formed by the interaction of five 1D proteins showing the relationship of the amino acid residues implicated in hemagglutination to the vertex. Five 1D carbon backbones are shown in blue. Single 1B and 1C proteins are shown in green and red, respectively. The amino acid residues implicated in hemagglutination are shown in yellow.

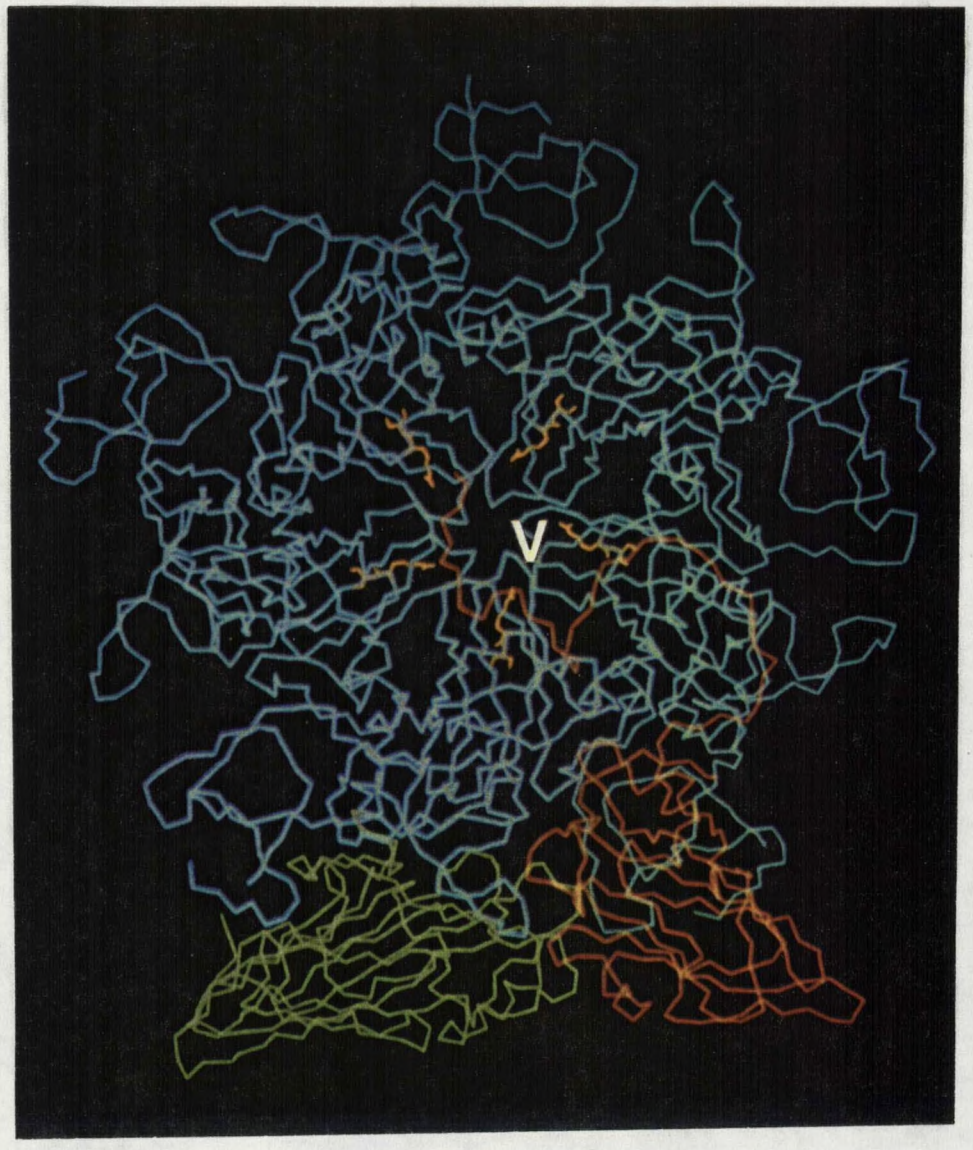
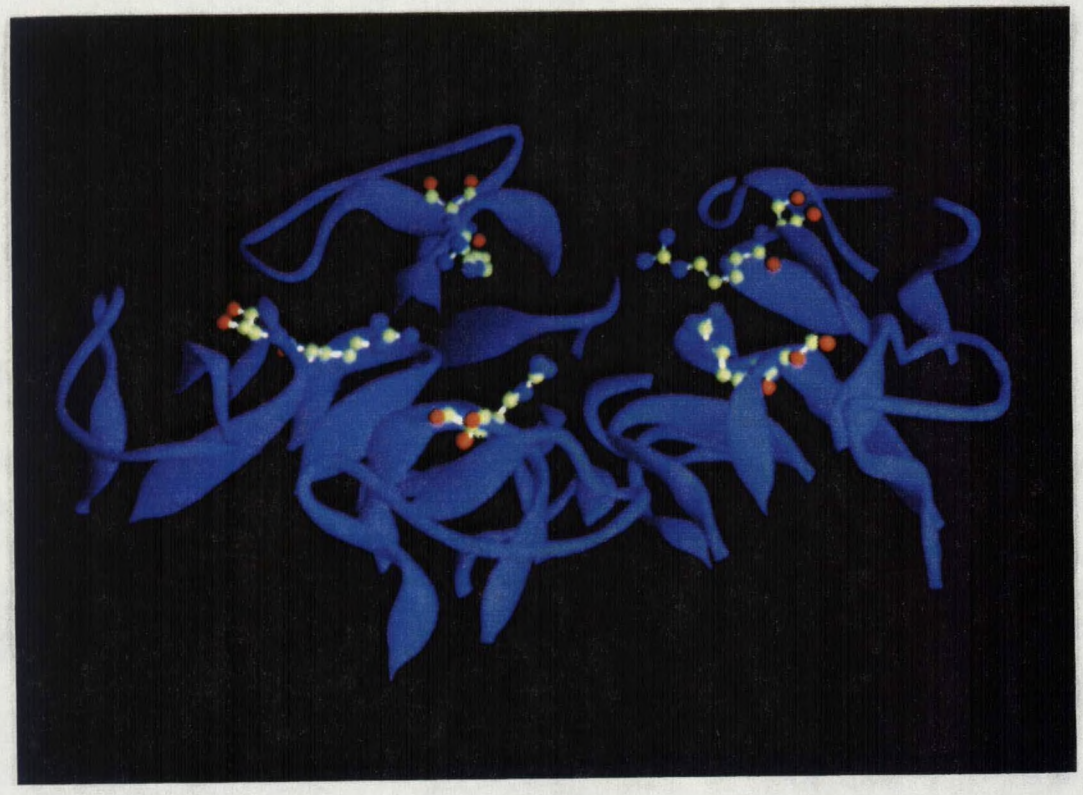


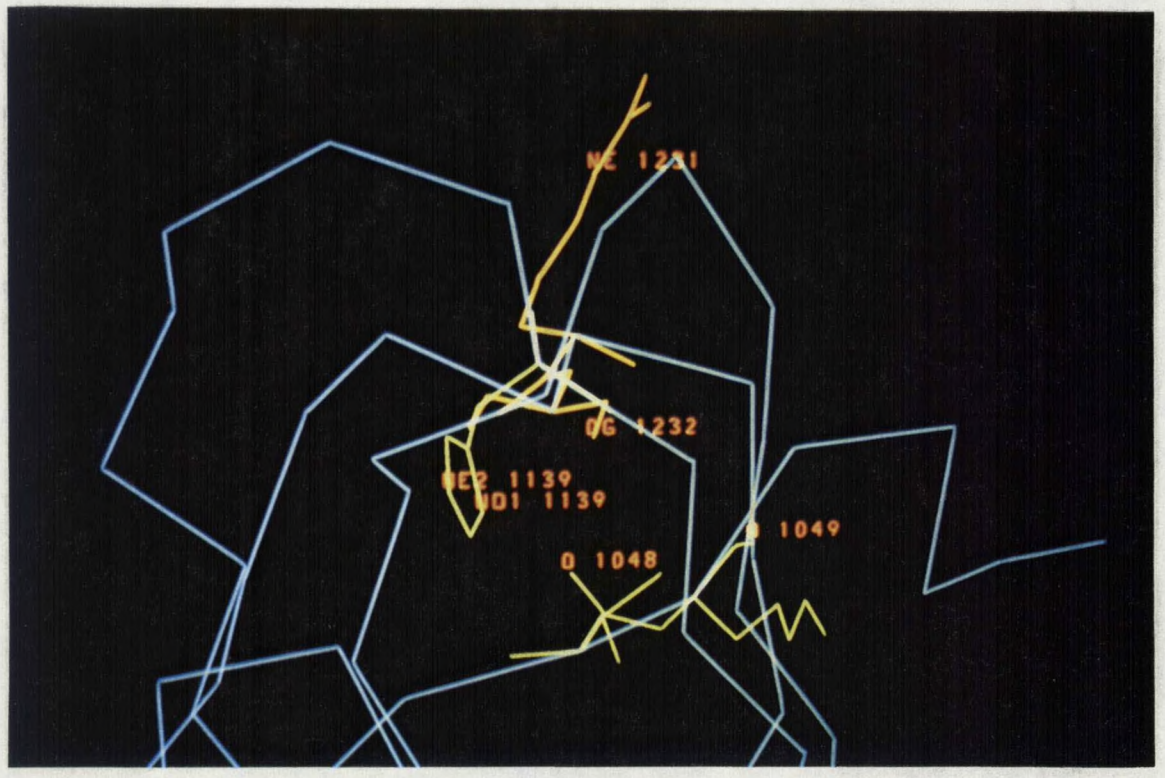
Figure 31. A RIBBON drawing of the side view of the platform formed by the five 1D molecules at the fivefold axis of symmetry. The carbon backbones of the 1D molecules are shown in blue. The two amino acid residues of 1D implicated in hemagglutination are shown in molecular detail. Carbon atoms are shown in green, nitrogen atoms in blue, and oxygen atoms in red.



Although the lysine to arginine replacement at position 231 was a fairly conservative amino acid substitution, the replacement of proline by serine at position 232 should cause a significant alteration in the secondary structure of the amino acid chain at that position since the "kink" produced by proline was removed when serine was substituted at that position. In addition, an hydroxyl group was introduced with the substitution of the serine increasing the potential for hydrogen bond formation in the region. Computer analysis of the model predicted that

the serine residue could form hydrogen bonds with three residues of the 1D protein: histidine 139; serine 48; or aspartic acid 49. The interaction with serine 48 was predicted to be the most likely. Figure 32 shows the molecular neighborhood of serine 232 and arginine 231 in relation to the possible hydrogen bond forming residues.

Figure 32. A line drawing showing the molecular neighborhood of the amino acid residues implicated in mengovirus hemagglutination. The carbon backbone of one 1D molecule is shown in blue. The amino acids implicated in hemagglutination are shown in detail in yellow (1231 and 1232). Residues with possible interactions with the serine residue found at position 232 in mutants 205 and 280 are shown in green [histidine 139 (1139), serine 48 (1048), aspartic acid 49 (1049)].



The additional flexibility of the carbon backbone, produced by replacement of the proline residue, should allow the serine 232 residue

to be pulled towards the serine 48 residue leading to a conformational change that twists the adjacent arginine 231. The conformation change occurring at the fivefold vertex could affect interaction with the erythrocyte membrane changing the hemagglutination phenotype of the virus.

The changes found between 37A and mutants 280 and 205 affect only a small portion of the capsid surface and may seem to be too minor to affect the interaction of the capsid with the erythrocyte membrane. Examination of intermolecular interaction in other systems indicate that very small changes can affect binding. When lysozyme and antilysozyme antibody complexes were studied by X-ray crystallography the interaction was determined to involve about 15 amino acid residues (31). A change in just one residue could prevent binding.

The revertants have regained both the proline 232 and lysine 231 residues of 37A implying that perhaps both changes are necessary for restoration of the hemagglutination-positive phenotype. Further studies should include analysis of the effect of single mutations in this area. When an infectious cDNA becomes available, the effects of the single substitutions of serine for proline 232 or arginine for lysine 231 can be tested.

Although the revertants regained the hemagglutination-positive phenotype, they retained the small plaque size of the mutants and their virulence was not fully restored indicating that the loss of hemagglutination ability alone is not responsible for the loss of virulence of the mutants. It is possible, however, that hemagglutination ability can play a role in virulence and that the

partial restoration of virulence in the revertants was due to the back mutation in the 1D protein. When six residues of the 1D protein were replaced in the Mahoney type 1 poliovirus, a variant avirulent in mice, by the residues found in the Lansing type 2 poliovirus, a virulent mouse variant, the recombinant virus was virulent for mice (85). These residues are part of the loop forming the first corner of the 1D protein and lie near the vertex of the virion pentamer (59). An example of a picornavirus with the hemagglutination-positive phenotype which has an increased ability compared to hemagglutination-negative virus to infect certain cell lines *in vitro* has also been found. Hemagglutination-positive variants of group B coxsackievirus are selected for by a rhabdomyosarcoma cell line infected with a hemagglutination-negative stock (30). When the hemagglutination-positive and negative variants are compared in HeLa cells the variants infect and multiply equally well. It has been suggested that the selection of hemagglutination-positive virus is due to its ability to bind to a receptor molecule present only on rhabdomyosarcoma cells. The ability to attach to the erythrocyte membrane may be an indication of an ability to attach to secondary receptors on particular cells and may improve the ability of the virus to productively infect a particular cell.

Analysis of the 5'-Noncoding Region Changes

Because of the importance of the 5'-noncoding region in the neurovirulence of poliovirus and other picornaviruses, RNA sequencing was extended from the capsid coding region into the 5'-noncoding region. Although the sequence of the entire region was not obtained, a base

change that distinguished 205 and its revertants 205-A7 and 205-D2 from 37A and mutant 280 was found. The change occurred at position 680 relative to the M variant sequence. A C found in 37A and 280 was replaced by a G in 205, 205-A7, and 205-D2. This was the only sequence change found in either the 5'-noncoding or the capsid coding regions that differentiated the revertants from 37A. No sequence differences in the 5'-noncoding region were found between 37A and 280.

Specific changes in the 5'-noncoding region of poliovirus have been shown to be associated with neurovirulence (69,122). One of the changes which differentiate the diabetogenic EMCV variant (EMC-D) from a related nondiabetogenic variant (EMC-B) occurs in the 5'-noncoding region at a position only 15 nucleotides removed from the change found in the 205 mutants and revertants (10). Secondary structure predictions of the region indicate that these changes occur in the same stem and loop structure (Palmenberg and Duke, personal communication). The effect of this base change on virulence could be tested best using recombinant virus progeny from an infectious cDNA clone of 37A in which this 5'-noncoding region base change could be isolated from other mutations in the genome. It should be noted that the 280 mutant, which does not have a mutation at position 680, was also nonvirulent.

One other possibility should be mentioned. The mutation implicated in the change in hemagglutination phenotype is a double mutation. The mutations that restored the hemagglutination-positive phenotype to the revertants occurred at the same sites as the original mutation. Both mutations are reversed in revertants 205-A7 and 205-D2 indicating that both changes were required to restore the phenotype. A

second explanation could be that the revertants resulted from a recombination event resulting from a double infection with 37A and 205. Recombination is common in RNA viruses and is known to occur *in vivo* (2,71) We have no evidence for this but it is possible and should be considered.

Further Studies

Additional sequencing of 37A, mutants 280 and 205, and revertants 205-A7 and 205-D2 should be undertaken to complete the genomic sequence and to determine if other sequence changes have occurred between these viruses. Mutation of other regions of genome may be important in the determination of virulence and plaque size of the mutants. The partial reversion of the revertants may be explained by a mutation at a second site in the genome which was not restored or suppressed by changes in the revertants.

One technique I have not tried but which may help solve sequencing problems due to compressions or natural stops is to sequence uncloned cDNA (51). This technique would take advantage of the flexibility in the choice of enzymes for DNA sequencing but would avoid the selection of single molecules as templates for sequencing from the population. It would require the synthesis of another set of synthetic oligonucleotide primers for the reactions since the template for the sequencing reactions would not be complementary to the current set of primers. The first step in this reaction is to prime first strand cDNA synthesis using a mengovirus-specific primer which binds to the virion RNA at a

region 3' to the problem area. A second primer specific for the cDNA sequence would be used to sequence the problem area.

Construction of an infectious cDNA of 37A would be of tremendous value in separating the effects of the specific nucleotide changes found between the viruses. As mentioned above, the effect of the change found in the 5'-noncoding region needs to be tested as a single change as do the two mutations implicated in hemagglutination. An infectious cDNA for 37A would also allow study of the effect of changes found between 37A and the M variant.

Conclusions

These studies have determined the probable site of interaction of the mengovirus capsid with the human O erythrocyte receptor by the identification of two amino acid changes at positions 231 and 232 in the 1D capsid protein which lead to the loss of the hemagglutination phenotype in mutants 205 and 280. When the amino acids were reverted to those seen in 37A, the ability to agglutinate human O erythrocytes was regained. Computer analysis of the effect of these changes on the atomic structure of the mengovirus capsid suggest that the replacement of proline 232 with serine results in a conformational change which could affect erythrocyte binding. The erythrocyte binding site was separate and distinct from the cellular receptor binding site.

A nucleotide change in the 5'-noncoding region suggest that the revertants arose from mutant 205 and are true revertants since the change was found only in 205, 205-A7, and 205-D2. However, a recombination event cannot be entirely ruled out. Completion of the

sequencing of the genomic RNA of the viruses could help answer this question as well as help define the molecular basis of virulence and plaque size.

Several changes have been identified between the 37A and M variants which may affect the structure and stability of the capsid by introducing residues with additional hydrogen bonding potential.

Two cloning errors were detected in the cDNA clone of the capsid coding region of mutant 205. Direct sequencing of the viral RNAs was a more efficient method of determining the consensus sequence in comparison with to cDNA synthesis and cloning since it did not require identification and sequencing of multiple independent clones.

LITERATURE CITED

1. Acharya, R., E. Fry, D. Stuart, G. Fox, D. Rowlands and F. Brown. 1989. The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. Nature **337**:709-716.

2. Agut, H., K.M. Kean, C. Bellocq, O. Fichot and M. Girard. 1987. Intratypic recombination of polioviruses: Evidence for multiple crossing-over sites on the viral genome. J.Virol. 61:1722-1725.

3. Ahlquist, P. and P. Kaesgerg. 1979. Determination of the length distribution of poly(A) at the 3' terminus of the virion RNA of EMC virus, poliovirus, rhinovirus, RAV-61 and CPMV and of mouse globin mRNA. Nucleic Acids Res. 7:1195-1204.

4. Air, G.M. 1979. Nucleotide sequence coding for the "signal peptide" and N terminus of the hemagglutinin from Asian (H2N2) strain of influenza virus. Virology 97:468-472.

5. Allaway, G.P. and A.T.H. Burness. 1986. Site of attachment of encephalomyocarditis virus on human erythrocytes. J.Virol. 59:768-770.

6. Allaway, G.P. and A.T.H. Burness. 1987. Analysis of the bond between encephalomyocarditis virus and its human erythrocyte receptor by affinity chromatography on virus-Sepharose columns. J.Gen.Virol. 68:1849-1856.

7. Anderson, K. and C.W. Bond. 1987. Biological properties of mengovirus: Characterization of avirulent, hemagglutination-defective mutants. Arch.Virol. **93**:31-49.

8. Anderson, K. and C.W. Bond. 1987. Structural and physiological properties of mengovirus: Avirulent, hemagglutination-defective mutants express altered alpha (1D) proteins and are absorption-defective. Arch.Virol. **93**:13-29.

9. Ausubel, F.M., Brent, R., Kingston, R.E., et al. *Current Protocols in Molecular Biology*, New York:Wiley and Sons, 1989.

10. Bae, Y.S., H.M. Eun and J.W. Yoon. 1989. Genomic difference between the diabetogenic and nondiabetogenic variants of encephalomyocarditis virus. Virology **170**:282-287.

11. Barnes, W.M. 1987. Sequencing DNA with dideoxyribonucleotides as chain terminators: Hints and strategies for big projects. Methods Enzymol. **152**:538-556.

12. Birnboim, H.C. 1983. A rapid alkaline extraction method for the isolation of plasmid DNA. Methods Enzymol. **100**:243-255.

13. Black, D.N., P. Stephenson, D.J. Rowlands and F. Brown. 1979. Sequence and location of the poly C track in aphtho- and cardiovirus RNA. Nucleic Acids Res. 6:2381-2390.

14. Boege, U. and D.G. Scraba. 1989. Mengo virus maturation is accompanied by C-terminal modification of capsid protein VP1. Virology 168:409-412.

15. Brown, F., J. Newman, J. Stott, A. Porter, D. Frisby, C. Newton and P. Fellner. 1974. Poly(C) in animal viral RNAs. Nature 251:342-344.

16. Brownstein, B. and A.F. Graham. 1961. Interaction of mengo virus with L cells. Virology 14:303-311.

17. Burness, A.T.H. and I.U. Pardoe. 1981. Effect of enzymes on the attachment of influenza and encephalomyocarditis viruses to erythrocytes. J.Gen.Virol. 55:275-288.

18. Burness, A.T.H. and I.U. Pardoe. 1983. A sialoglycopeptide from human erythrocytes with receptor-like properties for encephalomyocarditis and influenza viruses. J.Gen.Virol. 64:1137-1148.

19. Butterworth, B.E. 1973. A comparison of the virus-specific polypeptides of encephalomyocarditis virus. Virology 56:439-453.

20. Butterworth, B.E., L. Hall, C.M. Stoltzfus and R.R. Rueckert. 1971. Virus-specific proteins synthesized in encephalomyocarditis virus-infected HeLa cells. Proc.Natl.Acad.Sci.USA 68:3083-3087.

21. Callahan, P.L., S. Mizutani and R.J. Colonno. 1985. Molecular cloning and complete sequence determination of RNA genome of human rhinovirus 14. Proc.Natl.Acad.Sci.USA 82:732-736.

22. Carroll, A.R., D.J. Rowlands and B.E. Clarke. 1984. The complete nucleotide sequence of the RNA coding for the primary translation product of foot-and-mouth disease virus. Nucleic Acids Res. 12:2461-2472.

23. Chow, M., J.F.E. Newman, D. Filman, J.M. Hogle, D.J. Rowlands and F. Brown. 1987. Myristylation of picornavirus capsid protein VP4 and its structural significance. Nature 327:482-486.

24. Cohen, J.I. 1989. Toward understanding the molecular basis for attenuation of picornaviruses. Adv.Virus Res. 36:153-180.

25. Cohen, J.I., J.R. Ticehurst, R.H. Purcell, A. Buckler-White and B.M. Baroudy. 1987. Complete nucleotide sequence of wild-type hepatitis A virus: Comparison with different strains of hepatitis A virus and other picornaviruses. J.Virol. 61:50-59.

26. Cohen, S.H., R.K. Naviaux, K.M. Vanden Brink and G.W. Jordan. 1988. Comparison of the nucleotide sequences of diabetogenic and nondiabetogenic encephalomyocarditis virus. Virology 166:603-607.

27. Colonno, R.J., J.H. Condra, S. Mizutani, P.L. Callahan, M.E. Davies and M.A. Murcko. 1988. Evidence for the direct involvement of the rhinovirus canyon in receptor binding. Proc.Natl.Acad.Sci.USA 85:5449-5453.

28. Colter, J.S., J.B. Campbell and L.R. Hatch. 1965. The pathogenicityto mice of three variants of Mengo encephalomyelitis virus. J.Cellular Comp.Physiol. 65:229-236.

29. Crouse, G., A. Frischauf and H. Lehrach. 1983. An integrated and simplified approach to cloning into plasmids and single-stranded phages. Methods Enzymol. **101**:78-89.

30. Crowell, R. and B. Landau. 1983. Receptors in the initiation of picornavirus infections. Comp.Virol. **18**:1-42.

31. Davies, D.R., S. Sheriff and E.A. Padlan. 1988. Antibody-antigen complexes. J.Biol.Chem. **263**:10541-10544.

32. DeBorde, D.C., A.M. Donabedian, M.L. Herlocher, C.W. Naeve and H.F. Maassab. 1988. Sequence comparison of wild-type and cold-adapted B/Ann Arbor/1/66 influenza virus genomes. Virology **163**:429-433.

33. DeBorde, D.C., C.W. Naeve, M.L. Herlocher and H.F. Maassab. 1986. Resolution of a common RNA sequencing ambiguity by terminal deoxynucleotidyl transferase. Anal.Biochem. 157:275-282.

34. Del Guidice, R.A. and Hopps, H.E. Microbiological methods and fluorescent microscopy for the direct determination of mycoplasma infection of cell cultures. In: *Mycoplasma infection of cell cultures*, edited by McGarrity, G.J., Murphy, D.G. and Nichols, W.W. New York: Plenem, 1978, p. 57-70.

35. Deng, G.-R. and R. Wu. 1983. Terminal transferase: Use in tailing of DNA and for *in vitro* mutagenesis. Methods Enzymol. **100**:96-104.

36. Dick, G.W.A. 1948. Mengo encephalitis virus: Pathogenicity for animals and physical properties. Brit.J.Exp.Path. 29:559-577.

37. Dick, G.W.A., A.M. Best, A.J. Haddow and K.C. Smithburn. 1948. Mengo encephalomyelitis: A hitherto unknown virus affecting man. Lancet 255:286-289.

38. Dick, G.W.A., K.C. Smithburn and A.J. Haddow. 1949. Mengovirus encephalomyelitis virus: Isolation and immunological properties. Brit.J.Exp.Path. 29:547-558.

39. Domingo, E. and Holland, J.J. High error rates, population equilibrium, and evolution of RNA replication systems. In: *RNA Genetics: Volume III*, edited by Domingo, E., Holland, J.J. and Ahlquist, P. Boca Raton: CRC Press, 1988, p. 3-36.

40. Domingo, E., E. Martinez-Salas, F. Sobrino, J.C. de la Torre, A. Portela, J. Ortin, C. Lopez-Galindez, P. Perez-Brena, N. Villaneuva and J.J. Holland. 1985. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: Biological relevance - a review. Gene 40:1-8.

41. Duke, G.M. and A.C. Palmenberg. 1989. Cloning and synthesis of infectious cardiovirus RNAs containing short, discrete poly(C) tracts. J.Virol. 63:1822-1826.

42. Dulbecco, R. and G. Freeman. 1959. Plaque production by the polyoma virus. Virology 8:396-397.

43. Ellem, K.A.O. and J.S. Colter. 1961. The isolation of three variants of mengovirus differing in plaque morphology and haemagglutination characteristics. Virology 15:340-347.

44. Evans, D.M.A., P.D. Minor, G.C. Schild, and J.W. Almond. 1983. Critical role of an eight amino acid sequence of VP1 in the neutralization of poliovirus type 3. Nature 304:459-462.

45. Evans, D.M.A., G. Dunn, P.D. Minor, G.C. Schild, A.J. Cann, G. Stanway, J.W. Almond, K. Currey and J.V.Jr. Maizel. 1985. Increased neurovirulence associated with a single nucleotide change in a noncoding region of the Sabin type 3 poliovaccine genome. Nature 314:548-550.

46. Field, K.G. 1989. AMV reverse transcriptase sequencing of ribosomal RNAs. Promega Notes 22:4.

47. Finch, J.T. and A. Klug. 1959. Structure of poliomyelitis virus. Nature 183:1709-1714.

48. Flanegan, J.B. and D. Baltimore. 1977. Poliovirus-specific primer-dependent RNA polymerase and RNA replicase have the same viral polypeptide. Proc.Natl.Acad.Sci.USA 74:2677-2680.

49. Forss, S. and H. Schaller. 1982. A tandem repeat gene in a picornavirus. Nucleic Acids Res. 10:6441-6450.

50. Frisby, D.P., C. Newton, N.H. Carey, P. Fellner, J.F.E. Newman, T.J.R. Harris and F. Brown. 1976. Oligonucleotide mapping of picornavirus RNAs by two-dimensional electrophoresis. Virology 71:379-388.

51. Geliebter, J. 1987. Dideoxynucleotide sequencing of RNA and uncloned cDNA. Focus 9:5-8.

52. Greve, J.M., G. Davis, A.M. Meyer, C.P. Forte, S.C. Yost, C.W. Marlor, M.E. Kamarck and A. McClelland. 1989. The major human rhinovirus receptor is ICAM-1. Cell 56:839-847.

53. Grunstein, M. and D. Hogness. 1975. Colony hybridization: A method for the isolation of cloned DNAs that contain a specific gene. Proc.Natl.Acad.Sci.USA 72:3961-3965.

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

55. Hanahan, D. 1983. Studies on transformation of Escherichia coli with plasmids. J.Mol.Biol. 166:557-580.

56. Henikoff, S. 1984. Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28:351-359.

57. Henikoff, S. 1987. Exonuclease III generated deletions for DNA sequence analysis. Promega Notes No. 8:1-3.

58. Hilz, H., U. Wiegers and P. Adamietz. 1975. Stimulation of proteinase K action by denaturing agents: Application to the isolation of nucleic acids and the degradation of masked proteins. Eur.J.Biochem. 56:103-108.

59. Hogle, J.M., M. Chow and D.J. Filman. 1985. Three-dimensional structure of poliovirus at 2.9 A resolution. Science 229:1358-1365.

60. Holland, J.J. 1961. Receptor affinities as major determinants of enterovirus tissue tropism in humans. Virology 15:312-326.

61. Holland, J.J. 1962. Irreversible eclipse of poliovirus by HeLa cells. Virology 16:163-176.

62. Holland, J.J., L.C. McLaren and J.T. Syverton. 1959. Mammalian cell-virus relationship. III, Poliovirus production by non-primate cells exposed to poliovirus ribonucleic acid. Proc.Soc.Exptl.Biol.Med. 100:843-850.

63. Hruby, D.E. and W.K. Roberts. 1976. Encephalomyocarditis virus RNA: Variations in polyadenylic acid content and biological activity. J.Virol. 19:325-330.

64. Hruby, D.E. and W.K. Roberts. 1978. Encephalomyocarditis virus RNA. III. Presence of a genome-associated protein. J.Virol. 25:413-415.

65. Jacobson, M. and D. Baltimore. 1968. Morphogenesis of poliovirus. I. Association of the viral RNA with the coat protein. J.Mol.Biol. 33:369-378.

66. Jang, S.K., M.V. Davies, R.J. Kaufman and E. Wimmer. 1989. Initiation of protein synthesis by internal entry of ribosomes into the 5' nontranslated region of encephalomyocarditis virus RNA *in vivo*. J.Virol. 63:1651-1660.

67. Jones, T.A. 1978. A graphics model building and refinement system for macromolecules. J. Appl. Crystallogr. 11:268-272.

68. Jones, T.A. 1985. Interactive computer graphics: FRODO. Methods Enzymol. 115:157-171.

69. Kawamura, N., M. Kohara, S. Abe, T. Komatsu, K. Tago, M. Arita and A. Nomoto. 1989. Determinants in the 5' noncoding region of poliovirus Sabin 1 RNA that influence the attenuation phenotype. J.Virol. 63:1302-1309.

70. Kim, S., U. Boege, S. Krishnaswamy, I. Minor, T.J. Smith, M. Luo, D.G. Scraba and M.G. Rossman. 1990. Conformation variability of a picornavirus capsid: pH-dependent structural changes of mengovirus related to its host receptor attachment site and disassembly. (In preparation)

71. King, A.M.Q. Genetic recombination in positive strand RNA viruses. In: *RNA genetics, Volume II*, edited by Domingo, E., Holland, J.J. and Ahlquist, P. Boca Raton: CRC Press, 1988, p. 149-165.

72. Land, H., M. Grez, H. Hauser, W. Lindenmaier and G. Schutz. 1883. Synthesis of ds-cDNA involving addition of dCMP tails to allow cloning of 5'-terminal mRNA sequences. Methods Enzymol. 100:285-292.

73. Lee, C.-K. and E. Wimmer. 1988. Proteolytic processing of poliovirus polyprotein: Elimination of $2A^{pro}$-mediated, alternative cleavage of polypeptide 3CD by *in vitro* mutagenesis. Virology 166:405-414.

74. Lloyd, R.E., M.J. Grubman and E. Ehrenfeld. 1988. Relationship of p220 cleavage during picornavirus infection to 2A proteinase sequencing. J.Virol. 62:4216-4223.

75. Luo, M., G. Vriend, G. Kamer, I. Minor, E. Arnold, M.G. Rossmann, U. Boege, D.G. Scraba, G.M. Duke and A.C. Palmenberg. 1987. The atomic structure of mengovirus at 3.0 Å resolution. Science 235:182-191.

76. MacPherson, I. and M. Stoker. 1962. Polyoma transformation of hamster cell clones: An investigation of genetic factors affecting cell competence. Virology 16:147-151.

77. Madshus, I.H., S. Olsnes and K. Sandvig. 1984. Mechansim of entry into the cytosol of poliovirus type 1: Requirement for low pH. J.Cell.Biol. 98:1194-1200.

78. Maniatis, T., Fritsch, E.F. and Sambrook, J. *Molecular Cloning: A Laboratory Manual*, Cold Springs Harbor:Cold Springs Harbor Laboratory, 1982.

79. Marsh, M. and A. Helenius. 1989. Virus entry into animal cells. Adv.Virus Res. 36:107-151.

80. Martin, E.M., J. Malec, S. Sved and T.S. Work. 1961. Studies on protein and nucleic acid metabolism in virus-infected mammalian cells. Biochem.J. **80**:585-596.

81. Martin, R. 1987. Overcoming DNA sequencing artifacts: Stops and compressions. Focus 9:8-10.

82. Mendelsohn, C.L., E. Wimmer and V.R. Racaniello. 1989. Cellular receptor for poliovirus: Molecular cloning, nucleotide sequence, and expression ao a new member of the immunoglobulin superfamily. Cell 56:855-865.

83. Merchant, D.J., Kahn, R.H. and Murphy, W.H. *Handbook of cell and Organ Culture*, Minneapolis:Burgess, 1964.

84. Morishima, T., P.R. McClintock, G.S. Aulakh, L.C. Billups and A.L. Notkins. 1982. Genomic and receptor attachment differences between mengovirus and encephalomyocarditis virus. Virology **122**:461-465.

85. Murray, M.G., J. Bradley, X.-F. Yang, E. Wimmer, E.G. Moss and V.R. Racaniello. 1988. Poliovirus host range is determined by a short amino acid sequence in neutralization antigenic site I. Science **241**:213-215.

86. Naviaux, R.K., S.H. Cohen, K.M. Vanden Brink and G.W. Jordan. 1990. Construction and characterization of two infectious molecular clones of encephalomyocarditis virus. J.Virol. **64**:913-917.

87. Nicklin, M.J.H., H. Toyoda, M.G. Murray and E. Wimmer. 1986. Proteolytic processing in the replication of polio and related viruses. Bio/Tech. **4**:33-42.

88. Nitayaphan, S., D. Omilianowski, M.M. Toth, G.D. Parks, R.R. Rueckert, A.C. Palmenberg and R.P. Roos. 1986. Relationship of Theier's murine encephalomyelitis viruses to the cardiovirus genus of picornaviruses. Intervirology **26**:140-148.

89. Nomoto, A., B. Detjen, R. Pozzatti and E. Wimmer. 1977. The location of the polio genome protein in viral RNAs and its implication for RNA synthesis. Nature **268**:208-213.

90. Nomoto, A., N. Kitamura, F. Golini and E. Wimmer. 1977. The 5'-terminal structures of poliovirus RNA and poliovirus mRNA differ on in the genome-linked protein VPg. Proc.Natl.Acad.Sci.USA 74:5345-5349.

91. Nomoto, A., T. Omata, H. Toyoda, S. Kuge, H. Horie, Y. Kataoka, Y. Genba, Y. Nakano and N. Imura. 1982. Complete nucleotide sequence of the attenuated poliovirus Sabin 1 strain genome. Proc.Natl.Acad.Sci.USA **79**:5793-5797.

92. Ozden, S., F. Tangy, M. Chamorro and M. Brahic. 1986. Theiler's virus genome is closely related to that of encephalomyocarditis virus, the prototype cardiovirus. J.Virol. **60**:1163-1165.

93. Palmenberg, A.C. Comparative organization and genome structure in picornaviruses. In: *Positive Stranded RNA Viruses*, edited by Brinton, M.A. and Rueckert, R.R. New York: Alan R. Liss, Inc., 1987, p. 25-34.

94. Palmenberg, A.C., E.M. Kirby, M.R. Janda, N.L. Drake, G.M. Duke, K.F. Potratz and M.S. Collett. 1984. The nucleotide and deduced amino acid sequences of the encephalomyocarditis viral polyprotein coding region. Nucleic Acids Res. **12**:2969-2985.

95. Parks, G., G.M. Duke and A.C. Palmenberg. 1986. Encephalomyocarditis virus 3C protease: efficient cell- free expression from clones which link 5' noncoding sequences to the P3 region. J.Virol. 60:376-384.

96. Perez-Bercoff, R. and M. Gander. 1977. The genomic RNA of mengovirus. I. Location of the poly(C) tract. Virology **80**:426-434.

97. Perez-Bercoff, R. and R. Kaempfer. 1982. Genomic RNA of mengovirus. V. Recognition of common features by ribosomes and eucaryotic initiation factor 2. J.Virol. **41**:30-41.

98. Pevear, D.C., M. Calenoff, E. Rozhon and H.L. Lipton. 1987. Analysis of the complete nucleotide sequence of the picornavirus Theiler's murine encephalomyelitis virus indicates that it is closely related to cardioviruses. J.Virol. **61**:1507-1516.

99. Pevear, D.C., M.J. Fancher, P.J. Felock, M.G. Rossmann, M.S. Miller, G. Diana, A.M. Treasurywala, M.A. McKinlay and F.J. Dutko. 1989. Conformational change in the floor of the human rhinovirus canyon blocks adsorption to HeLa cell receptors. J.Virol. **63**:2002-2007.

100. Racaniello, V.R. 1988. Poliovirus neurovirulence. Adv.Virus Res. **34**:217-246.

101. Racaniello, V.R. and D. Baltimore. 1981. Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the virus genome. Proc.Natl.Acad.Sci.USA **78**:4887-4891.

102. Rivera, V.M., J.D. Welsh and J.V. Maizel. 1988. Comparative sequence analysis of the 5' noncoding region of the enteroviruses and rhinoviruses. Virology **165**:42-50.

103. Rossmann, M.G., E. Arnold, J.W. Erickson, E.A. Frankenberger, J.P. Griffith, H.J. Hecht, J.E. Johnson, G. Kramer, M. Luo, A.G. Mosser, R.R. Rueckert, B. Sherry and G. Vriend. 1985. Structure of a human common cold virus and functional relationship to other picornaviruses. Nature **317**:145-153.

104. Rueckert, R.R. 1976. On the structure and morphogenesis of picornaviruses. Comp.Virol. **6**:131-213.

105. Rueckert, R.R. Picornaviridae and their replication. In: *Virology*, edited by Fields, B.N., Knipe, D.M., Chanock, R.M., Hirsch, M.S., Melnick, J.L. and Roizman, B. New York: Raven Press, Ltd., 1990, pp. 507-548.

106. Rueckert, R.R. and E. Wimmer. 1984. Systematic nomenclature of picornavirus proteins. J.Virol. **50**:957-959.

107. Sanger, F. and A.R. Coulson. 1978. The use of thin polyacrylamide gels for DNA sequencing. FEBS Letters **87**:107-110.

108. Sanger, F., S. Nicklen and A.R. Coulson. 1977. DNA sequencing with chain terminating inhibitors. Proc.Natl.Acad.Sci.USA **74**:5463-5467.

109. Sarnow, P. 1989. Role of 3'-end sequences in infectivity of poliovirus transcripts made *in vitro*. J.Virol. **63**:467-470.

110. Sarnow, P., H.D. Bernstein and D. Baltimore. 1986. A poliovirus temperature-sensitive RNA synthesis mutant located in a noncoding region of the genome. Proc.Natl.Acad.Sci.USA **83**:571-575.

111. Schaffer, H.E. and R.R. Sederoff. 1981. Improved estimation of DNA fragment lengths from agarose gels. Anal.Biochem. **115**:113-122.

112. Stanway, G., P. Hughes, R. Mountford, P. Minor and J. Almond. 1984. The complete nucleotide sequence of a common cold virus: human rhinovirus 14. Nucleic Acids Res. **12**:7858-7875.

113. Staunton, D.E., V.J. Merluzzi, R. Rothlein, R. Barton, S.D. Marlin and T.A. Springer. 1989. A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. Cell **56**:849-853.

114. Suzuki, Y., H. Kato, C.W. Naeve and R.G. Webster. 1989. Single-amino-acid substitution in an antigenic site of influenza virus hemagglutinin can alter the specificity of binding to cell membrane-associated gangliosides. J.Virol. **63**:4298-4302.

115. Svitkin, Y.V., T.V. Pestrova, S.V. Maslova and V.I. Agol. 1988. Point mutations modify the response of poliovirus RNA to a translation iniation factor: A comparison of neurovirulent and attenuated strains. Virology **166**:394-404.

116. Tomassini, J.E., D. Graham, C.M. DeWitt, D.W. Lineberger, J.A. Rodkey and R.J. Colonno. 1989. CDNA cloning reveals that the major group rhinovirus receptor on HeLa cells is intercellular adhesion molecule 1. Proc.Natl.Acad.Sci.USA **86**:4907-4911.

117. Tomioka, N., C. Kishimoto, A. Matsumori and C. Kawai. 1988. Mural thrombi in mice with acute viral myocarditis. Jap.Circ.J. **49**:1277.

118. Trono, D., R. Andino and D. Baltimore. 1988. An RNA sequence of hundreds of nucleotides at the 5' end of poliovirus RNA is involved in allowing viral protein synthesis. J.Virol. 62:2291-2299.

119. Trono, D., J. Pelletier, N. Sonenberg and D. Baltimore. 1988. Translation in mammalian cells of a gene linked to the poliovirus 5' noncoding region. Science 241:445-448.

120. van der Werf, S., F. Bregegere, H. Kopecka, N. Kitamura, P.G. Rothberg, P. Kourilsky, E. Wimmer and M. Girard. 1981. Molecular cloning of the genome of poliovirus type I. Proc.Natl.Acad.Sci.USA 78:5983-5987.

121. Wellink, J. and A. van Kammen. 1988. Proteases involved in the processing of viral polyproteins. Arch.Virol. 98:1-26.

122. Westrop, G.D., K.A. Wareham, D.M.A. Evans, G. Dunn, P.D. Minor, D.I. Magrath, F. Taffs, S. Marsden, M.A. Skinner, G.C. Schild and J.W. Almond. 1989. Genetic basis of attenuation of the Sabin type 3 oral poliovirus vaccine. J.Virol. 63:1338-1344.

123. Wimmer, E. 1982. Genome-linked proteins of viruses. Cell 28:199-201.

124. Yanisch-Perron, C., J. Vieira and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.

125. Yin, F.H. and N.B. Lomax. 1983. Host range mutants of human rhinovirus in which structural proteins are altered. J.Virol. 48:410-418.

126. Yoon, J.W., T. Morishima, P.R. McClintock, M. Austin and A.L. Notkins. 1984. Virus-induced diabetes mellitus: Mengovirus infects pancreatic beta cells in strains of mice resistant to the diabetogenic effect of encephalomyocarditis virus. J.Virol. 50:684-690.

127. Ziola, B.R. and D.G. Scraba. 1976. Structure of the mengo virus. IV. Amino- and carboxyl-terminal analysis of the major capsid polypeptides. Virology 71:111-121.

128. Zurbriggen, A. and R.S. Fujinami. 1989. A neutralization-resistant Theiler's virus variant produces an altered disease pattern in the mouse central nervous system. J.Virol. 63:1505-1513.

APPENDICES

APPENDIX A

SYNTHETIC OLIGONUCLEOTIDE PROBES AND PRIMERS

Table 9. Synthetic oligonucleotide probes used for analysis of mengovirus cDNA inserts [a].

Probe	Position[b]	Orientation[c]	Probe Sequence
3'-CAP	3499-3513	-	GCAGG ATGTG CAGAG
5'-CAP	1118-1132	+	TCTGT TGTCA GGAGC
5'-END	1-15	+	TTTGA AAGCC GGGGG

[a]. These probes were obtained from Dr. Walter Hill, Division of Biological Sciences, University of Montana.
[b]. Position relative to the sequence of the M variant of mengovirus.
[c]. A "+" indicates that the probe sequence is identical to the RNA sequence. A "-" indicates that the probe sequence is the inverted complement of the RNA sequence.

Table 10. Promega primers used for DNA sequencing.

Primers	Description
T7 Promoter primer	20mer 5'-d(TAATACGACTCACTATAGGG)-3'
SP6 Promoter primer	19mer 5'-d(GATTTAGGTGACACTATAG)-3'

Table 11. Synthetic oligonucleotide primers used for consensus sequencing of viral RNAs.

Primer	Complementary Sequence Position[a]	Sequence	Source
PRIMER 1	3499-3513	GCAGG ATGTG CAGAG	A
PRIMER 2	3288-3490	GGTGC CGAAA TCAGA	B
PRIMER 3	3104-3118	CTTAC CTCAT GCAGC	B
PRIMER 4	2742-2756	GGGAC TAGAC CTATC	B
PRIMER 5	2924-2938	ACTTC GGAGG TTTCC	C
PRIMER 6	2563-2577	CAGCG CTAAC CATTG	C
PRIMER 7	2386-2400	TCAGA CCTAG ATCCC	C
PRIMER 8	949-964	TTGCC TTGTG TCTCG	C
PRIMER 9	282-296	ACATC GTCAC AGACG	C
PRIMER 10	2146-2160	GTGTC TTGAC AGCTG	C
PRIMER 11	1988-2013	AGATG GGAAC AGTGC	C
PRIMER 12	1797-1811	CAAAG TCCAG GAAGC	C
PRIMER 13	1564-1578	GTGTA GGATA CTCCG	C
PRIMER 13A	1665-1679	TTGGT GGTCC ATAGC	C
PRIMER 14	1379-1393	TTCTG GGTGG AAGTC	C
PRIMER 15	1128-1142	AGCAT TGACT GCTCC	C
PRIMER 16	796-810	CTTCA AAGGT CATGG	C
PRIMER 17	592-606	CCTTG TTGAA TACGC	C
PRIMER 18	128-142	TGGTT TTGGG TTGGC	C

[a]. Primer positions are given relative to the M variant of Mengovirus.

[b]. Sources of the primers were (A) Dr. Walter Hill, Division of Biological Sciences, University of Montana, Missoula, MT; (B) Genetic Designs, Inc., Houston, TX; and (C) Dr. Michael White, Veternary Molecular Biology Laboratory, Montana State University, Bozeman, MT.

APPENDIX B

DNA SEQUENCING KIT PROCEDURES AND COMPONENTS

Figure 33. K/RT Sequencing System procedure used for sequencing the pGEM205-110 derived plasmids.

I. Template Primer Annealing

A. redissolve in 6 ul diH_2O
3 ul primer (10 ng/ul)
1 ul 10X Klenow or RT buffer
mix and centrifuge briefly

B. incubate at 37°C for 15 to 120 min

II. Preparation of the Reaction Tubes

A. thaw A,C,G, and T mixes, hold on ice

B. label 4 500 ul microcentrifuge tubes "A,C,G,T"
add 3 ul of the appropriate mix to each tube

III. Sequencing Reactions

A. to the annealing mixture add 4 ul [α-^{35}S]dATP
(NEG-034S),centrifuge briefly

B. add 5U Klenow or 5U AMV-RT to the annealing mixture
mix with pipet tip

C. transfer 3 ul to A,C,G, and T tubes
centrifuge briefly

D. incubate at 37°C (Klenow) or 42°C (AMV-RT) for 20 min

IV. Chase and Termination

A. add 1 ul Chase solution, centrifuge briefly

B. incubate at 37°C (Klenow) or 42°C (AMV-RT) for 15 min

C. add 5 ul Stop, centrifuge briefly
store at -20°C

Table 12. Components of the Promega K/RT Sequencing System Kit.

Component	Composition
10X RT buffer	340 mM Tris-HCl (pH 8.3), 500 mM NaCl, 50 mM $MgCl_2$, 50 mM DTT
A mix (RT)	1.0 uM ddATP, 250 uM dCTP, 250 uM dGTP, 250 uM dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
C mix (RT)	12.5 uM ddCTP, 250 uM dCTP, 250 uM dGTP, 250 uM dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
G mix (RT)	12.5 uM ddGTP, 250 uM dCTP, 250 uM dGTP, 250 uM dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
T mix (RT)	50 uM ddTTP, 250 uM dCTP, 250 uM dGTP, 250 um dTTP, 34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT
10X Klenow buffer	100 mM Tris-HCl (pH 7.5), 500 mM NaCl
A mix (Klenow)	100 uM ddATP, 33 uM dCTP, 33 uM dGTP, 33 uM dTTP, 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT
C mix (Klenow)	16.7 uM ddCTP, 1.66 uM dCTP, 33 uM dGTP, 33 uM dTTP, 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT
G mix (Klenow)	16.7 uM ddGTP, 33 uM dCTP, 1.66 uM dGTP, 33 uM dTTP, 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT
T mix (Klenow)	117 uM ddTTP, 33 uM dCTP, 33 uM dGTP, 1.66 uM dTTP, 10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT
Chase solution	34 mM Tris-HCl (pH 8.3), 50 mM NaCl, 6 mM $MgCl_2$, 5 mM DTT, 2 mM dATP, 2 mM dCTP, 2 mM dGTP, 2 mM dTTP
Stop solution	98% formamide, 10 mM EDTA, 0.1% xylene cyanole FF, 0.1% bromphenol blue

Figure 34. Sequenase procedure used for sequencing the pGEM205-110 derived plasmids.

I. Primer Template Annealing

A. redissolve in 7 ul diH_2O
1 ul primer (10 ng/ul)
2 ul 5X reaction buffer
mix and centrifuge briefly

B. incubate at 65°C for 2 min, remove block from heat source
allow to cool slowly to RT

II. Preparation of the Termination Tubes

A. label 4 500 ul microcentrifuge tubes "A,C,G,T"
thaw Termination mixes, hold on ice

B. add 3 ul of the appropriate mix to each tube
hold on ice

III. Labeling Reaction

A. to the annealing mix add 1 ul 0.1 M DTT
2 ul Labeling Mix (1:5 dilution)
0.5 ul [α-^{35}S]dATP (NEG-034S)
centrifuge briefly

B. add 2 ul Sequenase (1:8 dilution)
mix with pipet tip

C. incubate at RT for 2 to 5 min

IV. Termination Reactions

A. prewarm termination tubes at 37°C for 1 min

B. add 3.5 ul of the labeling reaction to each termination tube, centrifuge briefly

C. incubate at 37°C for 3 to 5 min

D. add 4 ul Stop, centrifuge briefly, store at -20°C

Table 13. Components of the Sequenase (Ver.2) Sequencing Kit.

Component	Composition
5X Reaction buffer	200 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$, 250 mM NaCl
DTT solution	0.1 M DTT
5X Labeling (dGTP) mix	7.5 uM dCTP, 7.5 uM dGTP, 7.5 uM dTTP
ddA Termination (dGTP) mix	8 uM ddATP, 80 uM dATP, 80 uM dCTP, 80 uM dGTP, 80 um dTTP, 50 mM NaCl
ddC Termination (dGTP) mix	8 uM ddCTP, 80 uM dATP, 80 uM dCTP, 80 uM dGTP, 80 um dTTP, 50 mM NaCl
ddG Termination (dGTP) mix	8 uM ddGTP, 80 uM dATP, 80 uM dCTP, 80 uM dGTP, 80 um dTTP, 50 mM NaCl
ddT Termination (dGTP) mix	8 uM ddTTP, 80 uM dATP, 80 uM dCTP, 80 uM dGTP, 80 um dTTP, 50 mM NaCl
Enzyme Dilution buffer	10 mM Tris-HCl (pH 7.5), 5 mM DTT, 0.5 mg BSA/ml
Stop solution	95% formamide, 20 mM EDTA, 0.05% bromphenol blue, 0.05% xylene cyanole FF

APPENDIX C

ONE LETTER ABBREVIATIONS FOR AMINO ACIDS

Table 14. One letter abbreviations for the amino acids used in the protein alignments.

Amino Acid	One-letter Symbol
Alanine	A
Arginine	R
Asparagine	N
Aspartic Acid	D
Cysteine	C
Glutamine	Q
Glutamic acid	E
Glycine	G
Histidine	H
Isoleucine	I
Leucine	L
Lysine	K
Methionine	M
Phenylalanine	F
Proline	P
Serine	S
Threonine	T
Tryptophan	W
Tyrosine	Y
Valine	V
Undetermined	X