Order Number 9009755

Determination of mutation rates of RNA viruses and characterization of the receptor-destroying enzyme of bovine coronavirus

Leider, Jason Mark, Ph.D.

City University of New York, 1989

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

DETERMINATION OF MUTATION RATES OF RNA VIRUSES AND CHARACTERIZATION OF THE RECEPTOR-DESTROYING ENZYME OF BOVINE CORONAVIRUS

by

Jason M. Leider

A dissertation submitted to the Graduate Faculty in Biomedical Sciences
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy,
the City University of New York.

1989

This manuscript has been read and accepted for the Graduate Faculty in Biomedical Sciences in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

8/1/89
Date

Peter Palese, Ph.D.
Chair of Examining Committee

8/5/89
Date

Terry A. Krulwich, Ph.D.
Executive Officer

Richard Peluso, Ph.D.

Frances I. Smith, Ph.D.

8-1-89

Toon Stegmann, Ph.D.

Lu-Hai Wang, Ph.D.

Supervisory Committee

The City University of New York

ABSTRACT

DETERMINATION OF THE MUTATION RATES OF RNA VIRUSES AND CHARACTERIZATION OF THE RECEPTOR-DESTROYING ENZYME OF BOVINE CORONAVIRUS

by

Jason M. Leider

Adviser: Dr. Peter Palese

The extreme variability of some RNA-containing animal viruses is an important factor which limits the effectiveness of vaccine development against viral infection. For this thesis, two studies were performed which led to the determination of the mutation rates of three RNA viruses. First, the mutation rates of influenza A/WSN/33 virus and the Mahoney strain of poliovirus type 1 were measured. The rate of mutation for the nonstructural (NS) gene of influenza A virus and for the VP1 gene in poliovirus type 1 was assayed by direct sequence analysis. Each gene was repeatedly sequenced in over 100 viral clones which were descended from a single virion in one plaque generation. We obtained values of 1.5×10^{-5} and less than 2.1×10^{-6} mutations per nucleotide per infectious cycle for the mutation rates of influenza A virus and poliovirus, respectively.

In our second study, the newly developed technology of denaturing gradient gel analysis was applied in order to determine the mutation rate of the retrovirus Rous sarcoma virus (RSV). Progeny descended from a single virion were collected after one replication cycle, and seven regions of the genome were analyzed for mutations by denaturing gradient gel electrophoresis. In all, 65,250 nucleotides were screened,

yielding nine mutations, and the RSV mutation rate was calculated as 1.4×10^{-4} mutations per nucleotide per replication cycle. These results indicate that RSV is an extremely mutable virus.

Extensive virus variability, as described in these first two studies, is in stark contrast to the need of a virus to preserve a conserved pocket at its receptor–binding site and the active sites of any enzymes it may contain (such as polymerases and receptor–destroying enzymes). In our third project, we described the activity of the receptor–destroying enzyme of bovine coronavirus (BCV), and the effect on viral replication following its inhibition. We identified the E3 protein of BCV as possessing serine esterase activity. Furthermore, treatment of BCV with the serine esterase inhibitor diisopropylfluorophosphate dramatically reduced its infectivity in a plaque assay. It is assumed that the esterase activity of BCV is required in an early step of virus replication, possibly during virus entry or uncoating.

ACKNOWLEDGMENTS

Thanks to my beautiful wife Jolinda, whose love, support, and chicken salad sandwiches sustained me through the best and worst of times. I love you very much and look forward to spending the rest of our lives together.

Thanks to my parents, who always gave me the encouragement, love, sardines, "Ted Koppel" updates, and trips to NJ when I really needed them. Ditto to my brother and sisters (Steven, Darlene, Lorraine, and Heidi) who kept me in line and gave me the necessary diversions to keep me "reasonably" sane. If you think you've had it tough, they've been stuck with me for all these years and don't even get paid for it.

Thanks to Dr. Palese for providing excellent graduate training and terrific guidance over these three years. Also thanks for shelling out the dough for the awesome dissertation celebration at Windows on the World. (You mean that's not going to happen?)

Thanks to Dr. Frances Smith, who was instrumental in my scientific training during the retrovirus work (she has the battle scars to prove it). Also thanks to the other fine Sinai faculty members who contributed to my scientific education.

Thanks to Wayne Pan who, despite being on the wrong end of some of my jokes and squash shots, nonetheless remained a great friend and was always there to commiserate when science hit rock-bottom. So when are we going bowling?

Thanks to Wayne Pan's library for giving me that critical reference when I really needed it and for maintaining that spring-time freshness and crispness of the pages of your books.

Thanks to Terri Latham and my many friends on the 16th floor for their terrific companionship. Terri's brownies made my blood glucose cup runneth over while the rest of the gang transformed me into a "volleyball" maven.

Thanks to my other friends for all the video games, lost weekends, and ugly on-the-towners that are needed for any decent graduate career.

Thanks to the NIH and Mt. Sinai for supporting me during my studies. Also, thanks to Dr. T. Krulwich for keeping our program running at full-speed ahead.

Finally, thanks to Ms. Pacman for just being you and showing me that 1841 Pacman tokens bi-weekly don't go as far as they used to.

This thesis is dedicated to my sister Cheryl, who I wished had lived to see its completion. We all miss you very much and G-d bless you.

FORMAT OF THESIS

This thesis is prepared according to the new guidelines of the City University of New York which permit the direct incorporation of published research articles as chapters. The thesis has a general introduction, and chapters may have specific introductory statements. Materials and methods and results sections are in each individual chapter. Also, each chapter has a specific discussion section, and there is a general discussion as the final chapter of the thesis. The references for all chapters are pooled, in order to avoid redundancy.

Copyright permission for each chapter has been obtained from the publishers.

TABLE OF CONTENTS

PUBLICATIONS

Parvin, J.D., Moscona, A., Pan, W.T., Leider, J.M. and Palese, P. (1986). The mutation rate of animal viruses: Influenza A virus and poliovirus type 1, J. Virol. **59:** 377–383.

Leider, J.M., Palese, P. and Smith, F.I. (1988). Determination of the mutation rate of a retrovirus, J. Virol. **62:** 3084–3091.

Vlasak, R., Luytjes, W., Leider, J., Spaan, W. and Palese, P. (1988). The E3 protein of bovine coronavirus is a receptor–destroying enzyme with acetylesterase activity, J. Virol. **62:** 4686–4690.

MEETINGS

Cell Biology of Virus Entry, Replication and Pathogenesis: A Glaxo–UCLA Symposium. February 28–March 5, 1988; Taos, New Mexico, in attendance.

American Society for Microbiology 1988 Annual Meeting, New York, New York; Title of Presentation: Determination of the Mutation Rate of a Retrovirus (May 17, 1988).

American Society for Virology 1988 Annual Meeting. June 12–16, 1988; the University of Texas at Austin, Austin, Texas; Title of Presentation: Determination of the Mutation Rate of a Retrovirus (June 14, 1988).

TABLES

FIGURES

I. INTRODUCTION

A. VIRAL VARIATION AND RECEPTOR-DESTROYING ENZYMES

Extensive viral variation has been observed for influenza viruses and retroviruses. For influenza A virus, the genes coding for surface and non-surface proteins have been estimated to evolve at a rate which is a millionfold greater than for eukaryotic genes in nature (Buonagurio et al., 1986; Palese, 1986; Saitou and Nei, 1986). The variability associated with the human immunodeficiency virus in nature (Hahn et al., 1986; Saag et al., 1988; Fisher et al., 1988) may allow this virus to escape neutralization by the host immune response and also affect other aspects of viral pathogenesis. As a consequence of the variability of these viruses, development of effective vaccines has been severely hampered. While there are many reports of a high degree of variation for these viruses, quantitative evidence of their inherent mutability is lacking. Thus, we believe it is extremely important to precisely determine virus variation during a limited number of replication cycles. The outcome will be important to the understanding of virus evolution and will be useful in the development of appropriate vaccine strategies.

It had been postulated by Holland et al. (1982) that RNA viruses have increased mutation frequencies, in part because their polymerases lack proofreading exonucleases to remove misincorporated bases from newly synthesized strands. It is a major goal of this thesis to test whether some RNA animal viruses have significantly higher mutation rates than others. We wish to know if these differences correlate with the speed of evolution of these viruses. Also, while a primary source of variation is obviously mutation, we desire to see if the extreme variability of influenza A viruses and retroviruses may be due to unusually high mutation rates.

This thesis also focuses on characterizing the receptor-destroying enzyme of bovine coronavirus (BCV). Presently, there is a dearth of information about the steps of viral attachment to and inactivation of cellular receptor units. The regions of viral proteins

needed for these activities may be under much greater structural/functional constraints than other regions of the virus and, therefore, may be more amenable to treatment with antiviral agents than more variable sites. By understanding the role of these proteins during infection, effective antiviral agents may be developed which can be used to block specific stages of viral infection.

B. MORPHOLOGY OF THE INFLUENZA VIRION AND PROTEIN FUNCTION

Influenza A viruses are generally spherical or ovoid, enveloped, and about 80-120 nm in diameter. An influenza virion consists of an inner ribonucleoprotein core, or nucleocapsid, containing the single-stranded RNA segmented genome. The virion RNA has negative polarity and is transcribed by the virion-associated polymerase to produce mRNAs coding for the eight proteins that form the virus particle (PB1, PB2, PA, HA, NA, NP, M1, and M2) and for two nonstructural proteins which are only present in infected cells (NS1, and NS2) (Zebedee and Lamb, 1988; for review, see Lamb, 1983).

The three largest RNA segments each encode one of the polymerase molecules (PB1, PB2 or PA) which, along with the nucleoprotein NP and the viral RNA, form the nucleocapsid. Transcription begins with the cleavage by PB2 of host mRNA 10-13 nucleotides from the cap (Plotch et al., 1981). PB1 protein is associated with transcription and chain elongation (Ulmanen et al., 1981, 1983). The function of the PA protein is unclear, although during transcription the three polymerase proteins are seen as a complex (Braam et al., 1983). In the nucleocapsid, the NP protein is seen interacting with itself, the three P proteins and the RNA. It is assumed that the NP is part of the transcription and replication complex, however it is not known whether it has a catalytic or only structural function (Lamb and Choppin, 1983). Re-constitution of the influenza virus transcriptase by renaturing the NP and P proteins suggest the NP may be able to bind to vRNA to form a functional helical template. A complex of the three P

proteins bound to the 3' end of these helical structures then initiates and elongates mRNA chains (Szewczyk et al., 1988). Dissociation of the NP protein from virion nucleocapsids results in limited mRNA chain elongation (Kawakami and Ishihama, 1983; Kato et al., 1985). Thus the NP-associated virion nucleocapsids may be needed for the removal of secondary structure from vRNA which may otherwise inhibit chain elongation. Free NP molecules not associated with nucleocapsids may be required for antitermination during template vRNA synthesis and for the elongation of viral RNA chains (Beaton and Krug, 1986).

The major surface proteins of the virion are the hemagglutinin and neuraminidase. Hemagglutinin is an integral membrane protein associated with receptor-binding and fusion activities of the virus (Wilson et al., 1981). The three-dimensional structures of hemagglutinin complexed to sialic acid analogues shows sialic acid as filling the conserved pocket and surrounded by antibody-binding sites (Weis et al., 1988). Neuraminidase is a receptor-destroying enzyme that is important for the release of budding virus particles and the prevention of aggregation of virions (Palese et al., 1974). The surface proteins of influenza virus are also the major sites recognized by neutralizing antibodies. By X-ray crystallography, the three-dimensional structure of a complex between influenza virus neuraminidase and the Fab fragments of monoclonal antibodies has recently been described (Colman et al., 1987). Conformational changes in both the antibody and the antigen were observed and the interaction displays features of a handshake rather than the inflexible 'lock and key' model proposed for antigen-antibody binding.

The remaining influenza proteins are all synthesized from the mRNAs of two genomic RNAs. The M1 protein is the most abundant protein in the virion and forms a shell beneath the lipid bilayer (Compans et al., 1970). In vitro studies have shown that the influenza viral M1 protein inhibits viral RNA transcription (Zvonarjev and Ghendon,

1980). In infected cells, the viral M1 proteins may selectively interact with nucleocapsids containing vRNAs to inhibit transcription of vRNA into mRNA. The M2 protein is abundantly expressed at the plasma membrane of influenza A virus-infected cells and may be necessary for organization of virion particles (Lamb et al., 1985). Further evidence that the M2 has a role during influenza virus replication is demonstrated by the linkage of influenza A virus resistance to amantadine chloride and single nucleotide substitutions in the membrane-spanning domain of the M2 protein (Hay et al., 1985). It has now been shown that the M2 protein is present in 14 to 68 copies per virion and that an antibody to the extracellular N-terminal domain of this protein can restrict replication of some strains of influenza A virus (Zebedee and Lamb, 1988). Still unresolved are the functions of the NS1 and NS2 proteins.

C. INFLUENZA VIRUS REPLICATION

Infection begins with the attachment of the virus particle via hemagglutinin to sialylated determinants on the cell membrane. Virus then enters the cell via receptor-mediated endocytosis (White et al., 1982). Within the endosome, a drop in pH mediates a conformational change in the hemagglutinin, exposing its hydrophobic core. The virus then fuses with the endosome membrane and the viral core is released into the cytoplasm. Transcription and replication of the genome occurs in the nucleus. Using host-derived caps as a primer, the viral polymerase synthesizes mRNA until about 15 to 20 nucleotides from the 5' end of the vRNA template where termination and polyadenylation of the mRNA occur (Robertson et al., 1981; Beaton and Krug, 1986). The mRNA is then translated to make viral proteins. Full-length positive strand cRNA is also made, which functions as a template for synthesis of progeny genomes (Beaton and Krug, 1986; Hay et al., 1977).

About four hours after infection, several events occur. Virion M protein becomes

associated with the inner surface of the plasma membrane. Discrete patches of the membrane thicken and incorporate hemagglutinin and neuraminidase molecules. With the transport of nucleocapsids to the membrane, viral particles bud and form (for review, see Lamb and Choppin, 1983).

D. INFLUENZA A VIRUS EPIDEMIOLOGY

Influenza A viruses have been shown to undergo rapid and extensive evolution. They are unusual among RNA viruses in that they have two distinct forms of antigenic variation: antigenic shift and antigenic drift. In antigenic shift, reassortment of the hemagglutinin and neuraminidase genes produces viral strains with surface antigens that are immunologically remotely related to prior strains. Antigenic shift has in the past been accompanied by pandemics as novel viruses appeared against which the human population was highly susceptible. In antigenic drift, accumulation of mutations in the hemagglutinin and neuraminidase genes leads to viruses with minor antigenic changes, against which the human population has partial immune protection due to exposure to previously circulating strains. The accumulation of mutations in a subtype of influenza A virus leads to rapid evolution along a single lineage (for review, see Palese, 1986).

The properties of antigenic shift and drift in influenza viruses have presented a severe challenge to the development of vaccines that can provide lasting immunity to influenza infection (Wiley and Skehel, 1987). Existing vaccines provide only limited protection as the appearance of antigenic variants via antigenic drift can still cause epidemics. The potential of a major influenza pandemic as the result of emergence of a variant that has had antigenic shift provides a unique test to current vaccine regimens. Reports of high rates of evolution for the hemagglutinin, neuraminidase and nonstructural genes of influenza virus (Wiley et al., 1981; Webster et al., 1982; Raymond et al., 1986; Buonagurio et al., 1986) suggested that the virus evolves according to immunological

selection.

E. ANALYSIS OF MUTATION RATES OF RNA VIRUSES

1. QUANTITATIVE EVIDENCE OF THE ROLE OF MUTATION RATE IN VIRUS EVOLUTION

What is the source of influenza virus variation? Is it due solely to immune selection or is the virus also highly mutable? It was our goal to study a fundamental property which may be governing the rapid rate of evolution of influenza A virus, its mutation rate. While influenza vaccines are of only limited efficacy, vaccines against poliovirus have been used effectively for several decades. Reports of variability for poliovirus (Emini et al., 1983; Pincus et al., 1986) on the order of that seen for influenza virus (Lubeck et al., 1980) seemed paradoxical considering the success of poliovirus vaccination. Thus, in chapter II, we have calculated the mutation rates of these two viruses. The results suggest an inverse relationship between viral mutation rates and vaccination success.

2. POLIOVIRUS STRUCTURE AND VARIATION

Poliovirus is a member of the Picornaviridae, a large family of RNA viruses containing such human pathogens as Coxsackie viruses, rhinoviruses, and hepatitis virus A. Three immunologically defined serotypes exist for poliovirus (P1, P2, and P3) as well as many different subtypes (i.e., P1/Mahoney and P1/Brunhilde). The poliovirion is roughly spherical with no lipid envelope. Its protein shell is an icosahedral protein capsid consisting of sixty protomers arranged in twelve pentameric units. Each protomer contains one copy of each of the viral proteins VP1, VP2, VP3, and VP4. The three dimensional spatial arrangement of the capsid has recently been solved (Hogle et al., 1985) revealing that VP1, VP2, and VP3 are exposed on the virion surface where they can react with antibodies, while VP4 is internal to the shell and is probably in close

association with the viral RNA core. The VP1 protein appears not only to be a major antigenic site, but also involved in adsorption of the virion to cells in culture because antibodies to this protein neutralize infectivity efficiently in vitro (Chow and Baltimore, 1982).

In recent studies of the type 1 poliovirus vaccine strain, sequence analysis of the parental virus and its vaccine derivative indicated that the latter differed by 57 nucleotide base substitutions, resulting in 21 amino acid changes (Kitamura et al., 1981; Nomoto et al, 1982; Racaniello and Baltimore, 1981). Thirty-three percent of these base changes occurred in the VP1 coat protein gene, which accounts for only 12% of the viral genome. While aggregation of mutations in this coat protein suggest that attenuation results from a change in its surface properties, mutations in other viral genes may also contribute to loss of virulence.

3. MEASUREMENT OF THE MUTATION RATES OF INFLUENZA A VIRUS AND POLIOVIRUS

In chapter II, in a project in which the primary investigator was J. D. Parvin, we sequenced the NS and VP1 genes in randomly selected viral clones of influenza A virus or poliovirus type 1, respectively, to directly measure the mutation rates of the two viruses. The mutation rate of a virus can be defined as the probability that, in a single cycle of replication, a particular nucleotide is mutated (Smith and Inglis, 1988). By examining these two genes, which were previously reported to be extremely variable, it was our aim to see if influenza A virus is highly mutable, which may explain its ability to undergo extensive variation and evade immune surveillance. We measured a higher mutation rate for influenza A virus than for poliovirus. The time over which mutations could occur was confined to a single plaque generation (about five infectious cycles), and we measured a neutral mutation rate for each virus (all lethal and deleterious mutations were not observed). This difference in mutation rates of the two viruses

correlates with their evolutionary rates and is inversely related to the degree of success achieved by vaccination against them.

F. RETROVIRUS STRUCTURE

Retroviruses are a family of RNA viruses which characteristically replicate through a DNA intermediate and cause a chronic infection in susceptible host cells. The virions share a similar morphology, being 80 to 130 nm in diameter and containing an electron-dense core surrounded by a lipoprotein envelope. The viral envelope is a host-derived lipid bilayer through which viral envelope (env) gene products are inserted. The smaller, anchoring env C-terminal protein (gp37) is associated with the lipid bilayer and interacts with the env spikes (gp85), which contain the receptor-binding site as well as the major antibody-binding sites (for review, see Coffin, 1986). The core of the virus is composed of an icosahedral capsid (gag gene products). Inside the core lies reverse transcriptase (pol gene product) and the viral genome. The core RNA consists of two identical single-stranded RNA molecules held as a dimer by noncovalent bonds at their 5' ends. The RNA molecules are 3' polyadenylated and, depending on the retrovirus, 3.5 to 9 kilobases in length. Also found noncovalently-bound near the 5' end of the genome is a cellular tRNA molecule.

G. RETROVIRUS TRANSCRIPTION AND REPLICATION

The binding of retroviruses to host cells is mediated by a specific viral envelope glycoprotein-host cell receptor interaction. The most thoroughly characterized retroviral receptor is that of the human immunodeficiency virus type 1 (HIV-1), the CD4 receptor. The receptors for other retroviruses await further investigation. Virus entry is then by either direct fusion with the plasma membrane or through receptor-mediated endocytosis, with conflicting reports supporting each mode (Anderson and Nexo, 1983; Dale and

Hanafusa, 1972; Lenard and Miller, 1982; Miyamoto and Gilden, 1971). For HIV-1 and simian immunodeficiency virus, entry into $CD4^+$ cells apparently occurs via direct fusion at the plasma membrane. After entry, the viral contents are released into the cytoplasm. In the cytoplasm, the reverse transcriptase uses the viral RNA as a template and the host tRNA as a primer to make a single-strand of DNA. The RNA is then degraded by an RNase H activity (degrades the RNA from RNA:DNA hybrids) of the reverse transcriptase as it synthesizes the second strand of DNA (for review, see Stoltzfus, 1988). The double-stranded DNA migrates to the nucleus, where it is integrated at random into the cellular DNA by the endonuclease domain of reverse transcriptase (Alexander et al., 1987; Panganiban and Temin, 1984). The viral DNA is then transcribed into full-length RNA. In replication-competent viruses, about half of this RNA is used for virion RNA, and the other half is processed to form mRNA. Newly made viral proteins, viral RNA, and host tRNA assemble and bud outward (for review, see Ho et al., 1987).

H. MUTATION RATE ANALYSIS OF RETROVIRUSES

1. ROUS SARCOMA VIRUS AS A MODEL FOR RETROVIRUS MUTATION RATE ANALYSIS

Rous sarcoma virus (RSV) has been one of the most thoroughly analyzed of the retroviruses. It is a type C avian sarcoma virus, signifying it is found as an extracellular virus, 80 to 110 nm in diameter, with a centrally placed core, and develops as it buds from the cell's plasma membrane. Its life-cycle encompasses the key aspects of a retrovirus infection as it replicates through a DNA intermediate and causes a chronic infection of its host that is not cytocidal. It accomplishes both replication and transformation in its host cell without the aid of a helper virus. Found in its genome are complete copies of the three viral genes needed for competent replication: gag (an

acronym for group-specific antigen), which encodes the core proteins; pol (polymerase), which codes for the reverse transcriptase; and env (envelope), which encodes the viral envelope proteins gp 37 and gp 85. In addition, RSV contains a full-length copy of the oncogene src (sarcoma), and is thus able to transform fibroblasts. The potential for RSV to competently replicate and transform cells in tissue culture makes the virus ideally suited for mutation rate analysis. The rate due solely to RSV, without the confounding factor of a helper virus, can be studied and the cloning of the virus simplified through selection for the transformed phenotype.

2. RETROVIRUS VARIATION

Many reports on retroviruses suggested the existence of extensive genetic diversity. It was seen that five days after passage of a cloned RSV preparation through chick embryo fibroblasts, host range mutants could be detected (Zarling and Temin, 1976). In another experiment, involving the repeated passage of RSV from one cell culture to another followed by oligonucleotide fingerprint analysis, a mutation frequency on the order of 3×10^{-4} bases per passage of virus was calculated for a particular G to A change (Coffin et al., 1980). Due to the high number of virus passages, however, they could not calculate a precise mutation rate. In another study of spontaneously arising retrovirus mutants, Darlix and Spahr (1983) demonstrated by direct sequence analysis of isolated oligonucleotides that many different point mutations occurred when a cloned stock of RSV was grown in tissue culture. Regions of hypervariability have also been demonstrated for the gp 85 portion of the env protein by molecular analysis and sequence comparison of natural variants of members of the avian sarcoma and leukosis virus family (Bova et al., 1988; Dorner and Coffin, 1986). A spleen necrosis virus-based vector, containing less than one kilobase of retrovirus sequences and two dominant selectable phenotypic markers, has recently used to estimate a rate leading to expression of a suppressed gene of 5×10^{-3} per base pair per replication cycle

(Dougherty and Temin, 1987). The nature of the mutations leading to the expression of the suppressed gene were not characterized and could have been point mutations, insertions, or inversions. More recently, Dougherty and Temin (1988) have described determination of retrovirus mutation rates for base pair substitutions and for insertions. They found the mutation rate for a single base pair substitution during replication is 2 x 10^{-5} per base pair per replication cycle for a single site and the insertion rate is 10^{-7} per base pair per replication cycle for a single locus. Although the authors do not comment on the discrepancy in their calculation of the mutation rates in the two studies, their target size for mutation in the first study was poorly characterized and may have led to an erroneous estimate of the mutation rate.

For HIV-1, three recent publications present further evidence for rapid evolution for retroviruses. It has now been shown that the env and gag genes from sequential virus isolates of persistently infected individuals have rates of evolution that are a millionfold greater than for DNA genomes (Hahn et al., 1986). Sequential virus isolations from two chronically infected individuals and analysis of recombinant genomes at a molecular level indicated HIV-1 variation in vivo is rapid and variants evolve in parallel and coexist during chronic infection (Saag et al., 1988). The analysis of a series of proviral clones of HIV-1 originating from a single patient showed that hybrid genomes (in which the envelope region of six viral clones were separately substituted into a prototype HIV-1 genome) generated viruses with a range of differing growth capabilities in human T cell lines and suggested the existence of extensive biological variation in vivo within an infected individual (Fisher et al., 1988).

The genomic heterogeneity of retroviruses has been largely attributed to their unique replication. Replication of the viral genome uses three enzymes: reverse transcriptase to synthesize the proviral DNA intermediates; cellular DNA polymerase to replicate the proviral DNA along with the host genome; and RNA polymerase II for the

synthesis of new genomes. High fidelity has been seen for proviral replication by DNA polymerase (Jolly et al., 1986). The error rate for RNA polymerase II is unknown, but since this enzyme lacks proofreading capabilities, there is no reason to believe it to be very low. Much of the variability of retroviruses, however, has been assigned to reverse transcriptase. In a study on src deletion mutants in six transformation defective (td) mutants of RSV, the significant number of nucleotide differences accounting for the divergence of the genomic sequence during replication of the td virus was believed to occur via a reverse transcriptase-mediated mechanism (Parvin and Wang, 1984). Several in vitro experiments (Loeb and Kunkel, 1982; Robertson et al., 1989a; Preston et al., 1989; Robertson et al., 1989b) have estimated error frequencies for purified reverse transcriptases obtained from different retroviruses (avian myeloblastosis virus, Moloney murine leukemia virus, and HIV-1, respectively) as ranging from 10^{-3} to 10^{-4} per site. While this high error rate may be partially due to the in vitro conditions employed, the implication is that the marked variation seen for retroviruses may be due to the fidelities of their replicating enzymes.

Despite these many reports of extensive retrovirus variation, direct experimental evidence of a definitive estimate of the mutation rate of a retrovirus had not been done and the source of retrovirus diversity was still equivocal. One explanation for retrovirus variation proposed that variation existed in the virus population for a long time and that high multiplicity infection of divergent viruses provided the chance for evasion of the immune system by a complex combination of phenotypic mixing and recombination events (Coffin, 1986). In chapter III, we describe a study in which we have determined the precise mutation rate of RSV for a single cycle of replication. Our results indicate that retroviruses are highly mutable, and that the origin of their heterogeneity may be from the low fidelity of either reverse transcriptase or RNA polymerase II or both.

3. MUTATION DETECTION METHODOLOGY

In our study on the mutation rate of RSV, we employed denaturing gradient gel electrophoresis, a method that allowed the easy and rapid screening of several thousand nucleotides for mutation. The keystone of this approach is the ability to detect single base mutations by observing the altered mobility of a mutant nucleic acid hybrid with respect to that of a wild-type hybrid during electrophoresis in a polyacrylamide gel containing an increasing concentration of denaturant. Basically, hybridization of a radioactively-labelled single-stranded DNA or RNA probe that is complementary in sequence to a molecule of interest is performed. The hybrid is then treated with nuclease S1 to digest non-complementary, unpaired regions. Following digestion, the duplex, now containing nucleic acid strands identical in length, is electrophoresed in a direction parallel to that of an ascending concentration of denaturant in the gel matrix. Within the wild-type duplex DNA, melting domains exist. The melting of these domains can be predicted by a theoretical melting profile, computed using the statistical-mechanical algorithms for sequence-specific helix-disorder transitions in DNA described by Lerman et al. (1984). Use of denaturing gradient gel analysis is contingent upon knowing the sequence of the molecule of interest for generating the theoretical melting profile as well as for producing complementary single-stranded probes. The profile reveals domains of the duplex in which cooperativity is strong enough between adjoining bases so as to permit long stretches of contiguous helix to form single strands at a particular critical temperature (Tm). The instability in the helix caused by a mutation in one of these domains lowers the domain's Tm and causes premature melting. Experimentally, this instability can be demonstrated on a parallel denaturing gradient gel. If in one lane of the gel a wildtype sample is loaded and in another lane a mutant sample is present, then two bands (one band in each lane) will be seen on an autoradiogram of the gel with the mutant duplex lagging behind the migration of the wild-type duplex at the critical melting temperature of the wild-type domain.

A major advantage of denaturing gradient gel analysis over conventional sequencing practices is the expedition of screening several thousand nucleotides without having to individually sequence each base. A prior limitation of denaturing gradient gel analysis had been that only low-melting regions of the genome, in which such a domain abutted a high melting domain (HMD), could be analyzed for mutations. The inability to analyze HMDs was due to the total strand dissociation, and consequent loss of resolving power, that occurred when these highest-temperature domains melted (Smith et al., 1986). Subsequent to the completion of our RSV mutation rate project, several new mutation detection methodologies have been described.

To increase the number of mutations that can be detected by denaturing gradient gel analysis, Sheffield et al. (1989) used the polymerase chain reaction (PCR) to attach a G+C-rich sequence (GC-clamp) to one end of amplified DNA for mutation detection in B-globin promoter mutants and genomic DNA from people with hemoglobinopathies. Using the clamp, they were now able to analyze HMDs. The solution melting method, which uses a step-wise salt or formamide gradient followed by gel electrophoresis to monitor for strand dissociation, has now been used for analysis of HMDs of RNA heteroduplexes differing by a single base pair (Smith et al., 1988) as well as that of DNA heteroduplexes (Latham and Smith, 1989). An alternative method for mutation detection in DNA:RNA and RNA:RNA hybrids is by ribonuclease A cleavage at mismatches (Myers et al., 1985; Winters et al., 1985). This method has recently been combined with in vitro gene amplification by PCR to examine somatic mutational activation of c-K-ras genes in human pancreatic carcinomas (Almoguera et al., 1988). A chemical cleavage method, employing hydroxylamine and osmium tetroxide to react with mismatched cytosine and thymine, respectively, potentially offers the ability to detect all single-base-pair mismatches in a heteroduplex (Cotton et al., 1988). These new methodologies, in association with denaturing gradient gel analysis, provide powerful

approaches for mutation detection in virtually any sequence.

4. RSV MUTATION RATE

In chapter III, we offer quantitative evidence that retroviruses are highly mutable. We applied denaturing gradient gel analysis to rapidly screen 65,250 nucleotides of information, with the detection of nine mutations. Also, this approach enabled us to examine seven regions of the genome, thus obtaining an unbiased estimate of the RSV mutation rate and gaining preliminary evidence of differential variability of regions of the retrovirus genome, correlating with previous reports.

I. ANALYSIS OF THE RECEPTOR-DESTROYING ACTIVITY OF BOVINE CORONAVIRUS

1. CORONAVIRUS STRUCTURE AND REPLICATION

Coronaviruses are large, positively-stranded RNA viruses that cause diseases in humans and domestic animals (Sturman and Holmes, 1983). They generally possess three structural proteins: A phosphorylated nucleocapsid protein (N) of 50K to 60K, a glycosylated matrix-like protein (M or E1) of 23K to 29K, and a large glycosylated peplomeric protein (P or E2) of 150K. The peplomer E2 protein forms the large surface projections that are arranged in a petal-shape in the virion envelope, from which the virus' name derived. Recently, a fourth structural protein, E3, has been found for coronaviruses that hemagglutinate: bovine coronavirus (BCV), (King and Brian, 1982), the human respiratory coronavirus OC43 (HCV OC43) (Hogue and Brian, 1986), and the porcine hemagglutinating encephalomyelitis virus (Callebant and Pensaert, 1980). In BCV, the E3 has been characterized as a glycoprotein dimer of apparent M_w of 140K, consisting of two disulphide-linked subunits of apparent M_w of 65K (King et al., 1985). Coronaviruses which lack the E3 protein, such as mouse hepatitis virus A59, do not hemagglutinate (Hogue et al., 1984).

It has now been demonstrated that BCV and HCV OC43 recognize O-acetyl sialic

acid-containing receptors similar to those of influenza C viruses (Vlasak et al., 1988). The binding appears to be mediated through the E2 and/or the E3 protein. The exact details of coronavirus entry following binding to its receptor are unclear. Recently, cell fusion by BCV was shown to be maximal at pH 7.5 and 8.0, suggesting BCV may be able to penetrate the host cell by direct fusion with the plasma membrane similar to paramyxoviruses (Payne and Storz, 1988). The process of initial translation of the approximately 20 kilobase long genomic RNA is unknown, although implications are that a large unidentified protein with polymerase activity is synthesized from it. The genomic RNA would then be transcribed by this enzyme to form a complementary, full-length minus-strand RNA (Lai et al., 1982; Sawicki and Sawicki, 1986). The negative strand RNA then serves as a template for synthesis of genomic RNA and subgenomic mRNAs. All mRNAs have a common leader sequence encoded by the 3' end of the negative RNA template. The helical nucleocapsid of coronaviruses is formed in the cytoplasm by interaction of newly synthesized genomic RNA with N proteins. Virions are formed by budding at membranes of the RER and the Golgi apparatus, but not at the plasma membrane. At the budding sites, strands of nucleocapsid align in an orderly array in areas that contain viral glycoproteins. Spherical budding virions that have incorporated a complete nucleocapsid are released into the lumen of the Golgi and the RER (Dubois-Dalcq et al., 1984). After migration through the Golgi apparatus, virions are transported into smooth-walled vesicles. The vesicles migrate to the cell periphery and virions are released from the cell by fusion of the vesicles with the plasma membrane (Sturman and Holmes, 1983).

2. VIRAL RECEPTOR-DESTROYING ENZYMES

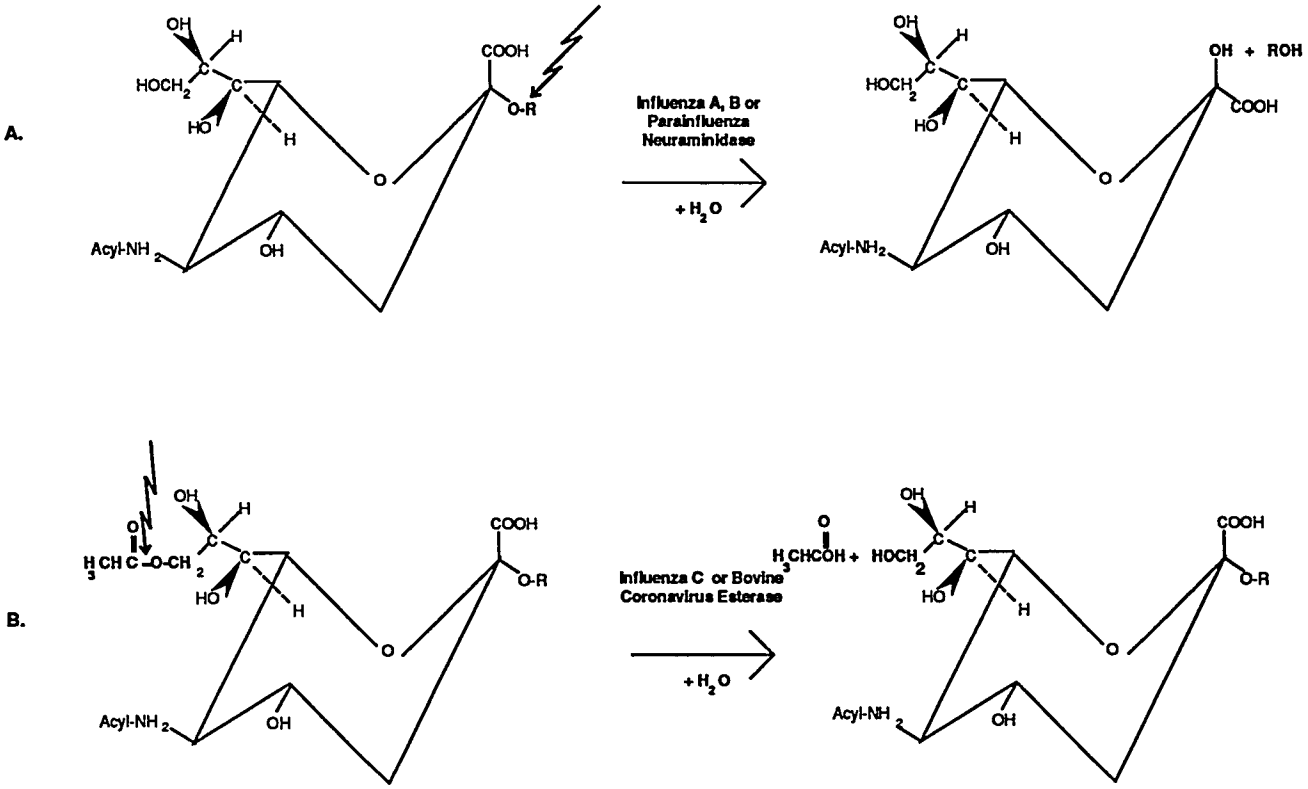
Receptor-destroying enzyme (RDE) activities have now been described for members of three families of enveloped RNA viruses (i.e. Orthomyxoviridae, Paramyxoviridae, and Coronaviruses) (Gottschalk, 1957; Scheid et al., 1972; Herrler et al., 1985; Vlasak et al.,

Figure 1. RECEPTORS FOR AND RECEPTOR-DESTROYING ACTIVITIES OF ORTHOMYXOVIRUSES, PARAMYXOVIRUSES, AND BOVINE CORONAVIRUS.

(A) The neuraminidase of influenza A and B viruses and paramyxoviruses cleave the α-ketosidic linkage joining the potential keto group of a terminal N-acylated neuraminic acid to an adjacent sugar residue in a disaccharide, trisaccharide or polysaccharide.
(B) The hemagglutinin-esterase of influenza C and the E3 protein of bovine coronavirus are specific for recognition of 9-O-acetylated sialic acid moieties and cleavage of this sugar's acetyl group.
R = monosaccharides, oligosaccharides, glycoproteins, glycolipids, aliphatic or aromatic alcohols.

FIGURE 1.

RECEPTORS FOR AND RECEPTOR-DESTROYING ACTIVITIES OF ORTHOMYXOVIRUSES, PARAMYXOVIRUSES, AND BOVINE CORONAVIRUS



1988). In paramyxoviruses and influenza A and B viruses, the RDE is a neuraminidase found in the form of a surface glycoprotein in the virion envelope. This RDE is capable of removing sialic acids from cellular receptors (Scheid and Choppin, 1974b; Rogers et al., 1986; Wiley and Skehel, 1987; Vlasak et al., 1988) (see Fig. 1A). An important consequence of the presence of neuraminidase activity in the envelope of the influenza A and B viruses and the paramyxoviruses is that no sialic acid can be detected in the viral particles (Klenk et al., 1970a; Klenk et al., 1970b). The lack of sialic acid in the viral envelope prevents the formation of large aggregates at the cell surface through inter-viral binding of hemagglutinin to envelope sialyl groups as seen with neuraminidase defective influenza A mutants (Palese et al., 1974). Thus, neuraminidase appears to be necessary for the final release and detachment of virus. Neuraminidase mediates these functions by prevention of particle aggregation and by destruction of cellular neuraminic acid. In paramyxoviruses, the neuraminidase activity is located on the hemagglutinin-neuraminidase (HN) protein, which also possesses receptor-binding/hemagglutinin activity (Scheid et al., 1978). For the influenza A viruses, the neuraminidase (NA) is a surface glycoprotein consisting of a single uncleaved polypeptide spike in a tetrameric form in the viral envelope (Varghese et al., 1983). Studies of the amino acid sequence of the tetrameric head released from viral membranes by protease treatment demonstrated that despite variation in the sequences of the N2, N1, and influenza B virus NAs, there is conservation of sequence at critical sites, strongly suggesting a common structure for these proteins (Varghese et al., 1983). The four catalytic sites of the enzyme have been identified as being on the distal surface of the tetrameric head (Varghese et al., 1983). These sites comprise a large pocket on the molecular surface, characteristic of the catalytic sites of enzymes that catalyze removal of end groups (Lamb and Choppin, 1983). The active site of neuraminidase is distally removed from its antibody-binding site. This was shown by the observations that antibody could only inhibit neuraminidase

activity if large substrates, such as fetuin, were employed, and not when smaller substrates, like sialyllactose, were used. The dependency of enzyme inhibition by antibody on substrate size suggested that the catalytic site of the enzyme was not antigenic and that inhibition of enzyme activity by antibody was mediated by steric hindrance. Structural studies demonstrate that each of the segments of the polypeptide chain identified as potential antigenic sites are sufficiently close to the enzyme's active site (for review, see Lamb and Choppin, 1983).

In influenza C viruses (Vlasak et al., 1987; Herrler et al., 1985), the RDE is an acetylesterase, removing acetyl groups from O-acetylated sialic acids (see Fig. 1B). The influenza C virus hemagglutinating-esterase (HE) protein, like the paramyxovirus HN protein, has both receptor-binding as well as receptor-destroying activity on the same molecule (Vlasak et al., 1987). The activated influenza C glycoprotein HE has an approximate molecular mass of 88,000 (88kD) (Nakada et al., 1984) and is derived by tryptic cleavage from an approximately 110 kD precursor (gpI) (Pfeifer and Compans, 1984). The mature HE is composed of two disulfide-bonded subunits, gp65 (HEI) and gp30 (HEII). Under non-reducing conditions the major peptide labeled by [^{3}H]DFP (a serine esterase inhibitor) has a Mr of 89 KD (corresponding to HE). In the presence of a reducing agent, the size of the labelled peptide shifted to 63 kD (corresponding to gp65), and thus the DFP-binding site was localized to gp65 (HEI) (Muchmore and Varki, 1987). Evidence from the microsequencing of [^{3}H]DFP-labelled HE has identified the active serine of the HEI (Vlasak et al., 1989; Herrler et al., 1989), with the enzyme probably utilizing a charge relay GLY-ASP-SER.

3. THE RECEPTOR-DESTROYING ACTIVITY OF BCV

In chapter IV, we characterize the receptor-destroying activity of BCV. In experiments, in which the primary investigator was R. Vlasak, we show that the BCV esterase activity resides on the E3 glycoprotein as seen by [^{3}H]DFP labelling.

Enzymatic activity is inhibited by DFP, indicating that the BCV RDE is a serine esterase. Furthermore, inhibition of the BCV RDE by DFP inhibits viral replication, as seen by lower virus titers in plaque assays. We suggest that the presence of an active viral esterase is essential for virus entry.

SPECIFIC AIMS

A. We sought to determine the mutation rates of influenza A virus and poliovirus. Also, we wanted to see if there may exist a relationship between the evolutionary stability and mutation rates of influenza A virus and poliovirus. Determination of the influenza A virus mutation rate should help in the development of a model which can explain the source of new variants which survive immune surveillance, cause severe illness, and serve as progenitors for future generations of virus.

B. The rate of evolution of retroviruses has been reported to be among the highest yet observed and exceeds that of their DNA-based hosts by a millionfold (Gojobori and Yokoyama, 1985; Hahn et al., 1986). We wanted to see if these viruses have intrinsically high mutation rates which would contribute to their rapid rates of evolution. For our analysis, we studied the mutation rate of RSV in tissue culture. Also, we wanted to examine several sites throughout the retrovirus genome for mutations and compare regions of the virus to one another for differential variability. Analysis of several regions of the genome should provide a truer estimate of the overall mutation rate of the virus. Finally, we wanted to apply a new method of mutation detection, denaturing gradient gel analysis, and demonstrate its value as an easier and more rapid approach to mutation detection than standard sequencing regimens.

C. We characterized the receptor-destroying activity of BCV. This included identifying the protein containing the enzymatic activity, its substrate specificity, and the nature of its catalytic activity. In addition, our results suggested that this enzyme is necessary for a productive infection and provided important insights as to the step(s) of the virus life-cycle at which it may function.

II.

MEASUREMENT OF THE MUTATION RATE OF ANIMAL VIRUSES: INFLUENZA A VIRUS AND POLIOVIRUS TYPE 1

Jeffrey D. Parvin, Anne Moscona, Wayne T. Pan, Jason M. Leider and Peter Palese

Department of Microbiology
Mount Sinai School of Medicine of CUNY
One Gustave L. Levy Place
New York, N.Y. 10029

Originally published in J. Virol. **59:** 377-383 (1986). Reprinted with permission from the American Society for Microbiology.

A. INTRODUCTION

The variability of animal viruses is a well-recognized phenomenon (reviewed in Holland et. al., 1982). The best studied example is the influenza A virus in which antigenic nature is continually changing by reassortment (antigenic shift) and by mutation (antigenic drift). The property of antigenic drift allows virus from a single subtype to persist in the human population in spite of immunity to strains from previous years (Webster et al., 1982). It has been observed form sequence analyses that the genes coding for surface as well as non-surface proteins of the influenza A virus evolve at a rate which is approximately a million-fold greater than that of eukaryotic genes in nature (Buonagurio et al., 1986; Palese, 1986; Saitou and Nei, 1986). Although a high mutation rate of influenza A viruses is most often implicated as the cause of the extensive variation (Hayashida et al., 1985; Saitou and Nei, 1986), a direct measurement of this parameter has never been obtained in vivo.

Mutation rates have been estimated for several bacteriophages. Based on the growth kinetics of spontaneous revertants from a deleterious point mutation, a mutation rate was indirectly calculated for bacteriophage QB (Batschelet et al., 1976) and mutation frequencies determined by growth under selective conditions have been estimated for bacteriophage lambda and bacteriophage T4 (Drake, 1969). Attempts have also been made to assess the mutability of animal viruses. For example, the substitution frequency of a specific 5' terminal nucleotide in the vesicular stomatitis virus (VSV) genome was determined (Steinhauer and Holland, 1986). Similarly, the frequency of variants which have lost the recognition site for a neutralizing monoclonal antibody has been used as a measure of mutation frequency. This method has been used to analyze mutability of RNA and DNA containing viruses including influenza viruses, poliovirus, VSV, and herpesviruses (Emini et al., 1983; Holland et al., 1983; Lubeck et al., 1980; Portner et al., 1980). Alternatively, reversion frequencies of viral mutants have

been used to assess mutation frequencies. For example, the accumulation of thymidine kinase deficient mutants of herpes simplex virus was employed to analyze the role of mutator and anti-mutator phenotypes of this virus (Hall et al., 1984).

In this study we attempted to measure directly the mutation rates of two animal viruses by sequencing genes in randomly selected viral clones. These viral clones were all descended from a single virion after only one plaque generation, which confined the time over which mutations could occur to about five infectious cycles. The mutation rate we measured was a neutral mutation rate, because all lethal and deleterious mutations were not observed. We found that the mutation rate of influenza A virus was higher than that of poliovirus, and we speculate that this difference correlates with the speed of evolution and the lack of success of vaccination against influenza A viruses.

Accurate in vivo measurements of mutation rates should allow a greater understanding of the evolution of viruses in nature and may also provide a new way to test different evolutionary theories with respect to animal viruses. Knowledge of the inherent mutability of a virus may also clarify the mechanism by which variants emerge that are resistant to the immunity conferred by vaccination. The probability of successful vaccination against a particular virus may be determined by the mutation rate of that virus, and thus knowledge of this parameter may influence the strategy of vaccine design. Finally, an understanding of the mutability of a virus may help in studying the development of drug resistance as well as the changes in virulence observed with different RNA- and DNA-containing viruses.

B. MATERIALS AND METHODS

1. VIRUSES AND CELLS

The influenza virus A/WSN/33 was derived from an uncloned viral stock preparation grown in Madin-Darby canine kidney (MDCK) cells in the presence of Eagle minimal

essential medium containing 1 ug of trypsin per ml (Brand and Palese, 1980). Virus was then plaque passaged twice in MDCK cells, and virus from plaque passage 2 was used for the experiment described. The agar overlay for plaquing contained 0.6% agar (Oxoid Ltd.), minimal essential medium, 0.2% bovine albumin, 0.01% DEAE dextran, and 1 ug of trypsin per ml.

The Mahoney strain of poliovirus type 1 and the HeLa cell line used for its passage were kindly provided by V. Racaniello. Again, the virus was plaque passaged twice before being used in the experiment. Viral passage was done as described previously (Bernstein et al., 1985).

2. PREPARATION OF INFLUENZA A VIRUS CLONES AND PURIFICATION OF VIRAL RNA

Confluent MDCK cells were infected with serial 10-fold dilutions of plaque-purified virus preparation. At one hour postinfection (p.i.), the inoculum was removed by aspiration, and the standard agar overlay was added. At 48 h p.i., a well-isolated plaque was identified, and the overlay above the plaque was gently aspirated with a Pasteur pipette. Virus was eluted from the agar plug into 0.5 ml of phosphate-buffered saline containing 0.2% bovine albumin. The cell monolayer from which the plaque was picked was stained with 0.1% crystal violet in 20% methanol to demonstrate that the plaque was indeed discrete.

To clone the individual virions in the plaque, a second plaque passage was done with the virus yield from the first plaque. This time many dishes were prepared to allow the isolation of several hundred discrete plaques. The total PFU in each plaque were about 10^6. One-fifth of the yield from each plaque was used to infect 2 x 10^6 MDCK cells with a liquid overlay which contained the same components as the agar overlay except for the agar and dextran. The supernatant from the completely lysed monolayer was harvested at 24 to 30 h p.i. The supernatant was then diluted in phosphate-

buffered saline containing 0.2% bovine albumin to infect (multiplicity of infection [MOI], of about 0.2) seven or eight dishes of MDCK cells, each containing 2 x 10^7 cells. The medium (65 ml) was harvested approximately 30 h p.i., after complete lysis of the cell monolayer.

The medium containing virus harvested form lysed cells was precleared by centrifugation at 8,000 x g for 30 min. The supernatant was then layered over a 3-ml 20% sucrose cushion, and the virus was pelleted by centrifuging at 25,000 rpm for 2 h in an SW27 rotor (Beckman Instruments, Inc.). The virus pellet was suspended in 4 ml of 100 mM NaCl-10mM Tris hydrochloride (pH 7.4)-1 mM disodium EDTA. The RNA from the suspended virus was extracted (approximately 100 ug) (Palese and Schulman, 1976).

3. PREPARATION OF POLIOVIRUS TYPE 1 CLONES AND OF VIRAL RNA

Confluent HeLa cell monolayers were infected with serial dilutions of the plaque-purified poliovirus preparation. At 1 h p.i., the virus inoculum was removed by aspiration, and the standard agar overlay containing 1% agar, Dulbecco modified Eagle medium, 5% horse serum, and 0.01% DEAE dextran was added. At 36 h p.i., a well-isolated plaque was identified, and the agar above it was aspirated as before. The monolayer was stained with crystal violet solution to confirm that the plaque was discrete. Virus was eluted from the agar plug into phosphate-buffered saline containing 0.2% horse serum. The total yield from the plaque was 5.2 x 10^6 PFU.

To clone individual virions from the virus yield of the plaque, a second round of plaque passage was done as before, except that the time allowed for plaque formation was 48 h. One-fifth of the virus yield from each plaque was used to infect 2 x 10^6 HeLa cells with a liquid overlay which contained the same components as the agar overlay except for the agar and dextran. At 24 h p.i., the medium above the lysed cells was harvested. This supernatant was then diluted for the infection (MOI, 0.5) of five

large dishes containing 2 x 10^7 HeLa cells each. At 24 h p.i., the yield form the lysed cells was harvested.

The medium containing virus (40 ml) was precleared by centrifugation at 9,000 x g for 30 min. The supernatant was centrifuged at 130,000 x g for 90 min. The pelleted virus was suspended in 4 ml of buffer containing 100 mM NaCl, 10 mM Tris hydrochloride (pH 7.4), and 1 mM disodium EDTA, and the RNA (approximately 100 ug) was extracted was done for the influenza virus.

4. SEQUENCING OF VIRAL RNA

The viral RNA was sequenced directly by the method of Sanger et al. (1977). The hybridization mixture contained 10 ug of viral RNA and 200 ng of a specific oligonucleotide as primer. Three primers were used to sequence each gene. The primers used for the NS gene were complementary to the viral RNA at positions 10 to 29, 293 to 312, and 593 to 612 (Buonagurio et al., 1986). The VP1 gene primers were complementary to positions 3422 to 3403, 3136 to 3117, and 2812 to 2793 on the viral RNA (Kitamura et al., 1981). The hybridization mix was heated in boiling water for 5 min, cooled to room temperature, and aliquoted to four reaction tubes. The A reaction contained 250 uM each dGTP and dTTP, 100 uM dATP, 12.5 uM dideoxy ATP, and 10 uM dCTP. The G reaction contained 250 uM each dATP and dTTP, 100 uM dGTP, 12.5 uM dideoxy GTP, and 10 uM dCTP. Both the A and G reactions had 15 uCi of [alpha-^{32}P]dCTP (3,000 Ci/mmole) or [alpha-^{35}S]dCTP (600 Ci/mmole) added to each reaction. The C reaction contained 250 uM each dGTP and dTTP, 100 uM dCTP, 12.5 uM dideoxy CTP and dGTP, 100uM dTTP, 12.5 uM dideoxy TTP, and 10 uM dATP. The T reaction contained 250 uM each dCTP and dGTP, 100 uM dTTP, 12.5 uM dideoxy TTP, and 10 uM dATP. To both the C and T reactions 15 uCi of [alpha-^{32}P]dATP. To both the C and T reactions 15 uCi of [alpha-^{32}P]dATP (3,000 Ci/mmole) or [alpha-^{35}S]dATP (600 Ci/mmole) was added. The reactions were started by addition of 5

U of avian myeloblastosis virus reverse transcriptase (Molecular Genetics Resources, Inc.) and incubated at 42°C for 30 min. The total reaction volume was 10 ul. A cold-chase solution containing 2 mM concentrations of each nucleotide was added (1.5 ul). After a second 30-min incubation at 42°C, the reactions were terminated by the addition of 11 ul of a formamide dye mix. Before gel electrophoresis, the reactions were heated in boiling water for 5 min and then rapidly cooled in an ice water bath.

As a standard procedure, two different gel systems were used to analyze the sequencing results. Gel 1 was a 6% polyacrylamide buffer-gradient gel (Biggin et al., 1983) or a 6% polyacrylamide gel which was used for a long electrophoresis of the reaction products. Gel 2 generally permitted the resolution of nucleotides from 160 to 320 or more bases away from the primer. Gels were fixed in a 10% acetic acid-10% methanol solution, dried, and exposed to Cronex 4 film (E. I. du Pont) for 3 to 10 days.

With this method, bands occasionally appeared in more than one lane at a given position. The secondary bands proved not to be troublesome since their patterns were uniform for positions among the different RNAs.

For the influenza virus NS gene, 91,708 nucleotides were sequenced. Of these, 50 nucleotides could not be positively identified because the bands in the expected lanes were too faint. However, since no other band appeared in any of the other three lanes at the same level, it was assumed that no changes occurred at these positions. Of the 95,688 sequenced nucleotides for the VP1 gene, 323 could not be positively identified for the same technical reason. Again, none of these positions was scored for mutations because bands were not observed in any of the other three lanes.

5. COMPARISON OF GROWTH KINETICS OF NS GENE MUTANTS

To determine the relative growth kinetics of each influenza virus NS gene variant, a multicycle infection was done for each mutant, the parental virus, and two randomly picked controls whose NS genes had the wild-type sequence. Samples of each virus

were titered in advance, and each was used to infect a dish of MDCK cells at an MOI of 1,000 PFU/3 x 10^6 cells. At 1 h p.i., the inocula were removed from the dish, each monolayer was washed with phosphate-buffered saline, and the standard liquid overlay was added. At 10, 19, 24, 28, 32, 37, 42, and 48 h p.i., 0.3 ml of medium was removed from each dish and used for plaque titrations. Cytopathic effect was first observed at 37 hours.

6. CONVERSION OF THE NS GENE MUTATION RATE FOR COMPARISON

The mutation rate of the NS gene as defined in this paper was calculated as the amount of change which would occur with each cell burst. Drake (1969) and Koch and Drake (1973) calculated the mutation rates from mutation frequencies. By this definition of mutation rate, the number of replications reflected the number of times the genome was copied. Replication number was approximately equal to population size. The formula was: mutation rate = 0.4343 x mutation frequency/log(population size). With this formula, the NS gene mutation frequency in a plaque was seven changes per 91,708 nucleotides, and the population size was equal to the total nucleotide pool (1.2 x 10^6 PFU x 850 nucleotides analyzed per NS gene), which resulted in the converted mutation rate of 3.7 x 10^{-6} mutations per nucleotide per replication. Alternatively, a correction factor substituting population size with population size times mutation rate could be made since no mutations would be expected until the populations size equaled the mutation rate (J. W. Drake, personal communication). The corrected mutation rate was then 8.4 x 10^{-6} mutations per nucleotide per replication. If the mutation rate ceiling for the poliovirus was similarly converted, it would have been 4.7 x 10^{-7} mutations per nucleotide per replication or, including the correction factor, 1.2 x 10^{-6} mutations per nucleotide per replication. By this procedure, the values for the mutation rates of bacteriophage lambda, bacteriophage T4, Salmonella typhimurium, Escherichia coli, and Neurospora crassa were 2.4 x 10^{-8}, 1.7 x 10^{-8}, 2.0 x 10^{-10}, 2.0 x 10^{-10}, and 0.7 x 10^{-11},

respectively (Drake, 1969).

C. RESULTS

1. DETERMINATION OF THE NS GENE MUTATION RATE

The influenza virus clones used to determined the mutation rate in the NS gene were derived from a single plaquing experiment (Fig. 2). Plaque formation occurred under standard tissue culture conditions in the absence of any new selective pressure. Cell monolayers were infected at a high dilution with plaque-purified A/WSN/33 virus, and at 48 h p.i., a single, well-isolated plaque was picked. This plaque, descended form a single virion, contained 1.2×10^6 PFU. To determine the number of mutant NS genes among these viral particles, individual clones were obtained, and viral RNA was prepared after only two amplification steps. The MOI during the amplification steps was maintained at 0.1 for step 1 and 0.2 for step 2 to minimize the enrichment of new mutants. Since RNA obtained from amplified viral preparations was sequenced by using the primer extension protocol, the sequence represented a consensus sequence which should have been identical to that of the individual clone. Any mutants which may have arisen in the amplification protocol would not have been detected.

The NS genes of 108 different clones were sequenced, producing 91,708 nucleotides of information. Comparison of the NS gene sequences revealed seven mutants. The relevant sequencing gels of two of these mutants are shown in Fig. 3. The sequences for five different clones are presented for positions 151 to 181 (Fig. 3A). As indicated by the arrows, a G to T transversion in clone 014 was observed at position 169. The sequences for five other clones are shown for positions 751 to 788 (Fig. 3B). On this gel, the G reactions of the five clones were loaded side-by-side as were the A, T, and C reactions. This latter arrangement facilitated the rapid detection of mutants. The A to G transitions at position 772 in clone 024 can easily be observed in this gel

Figure 2. EXPERIMENTAL DESIGN FOR MEASURING MUTATION RATES IN VIRAL GENES.

A single virion (the parental virus) formed a plaque after sufficient time for about five infectious cycles. The many progeny virions contained within the plaque carried genes with parental sequence (●), and a fraction of the virions carried genes with a point mutation (◆ and ■). The individual virions were cloned by plaquing and amplified under conditions that minimized the effect of new mutations on the consensus sequence of the gene in the clone. RNA obtained from purified virus was directly sequenced. The mutation rate was determined by dividing the number of observed point mutations by the number of nucleotides analyzed and by the number of infectious cycles.

FIGURE 2.

EXPERIMENTAL DESIGN FOR MEASURING MUTATION RATES IN VIRAL GENES

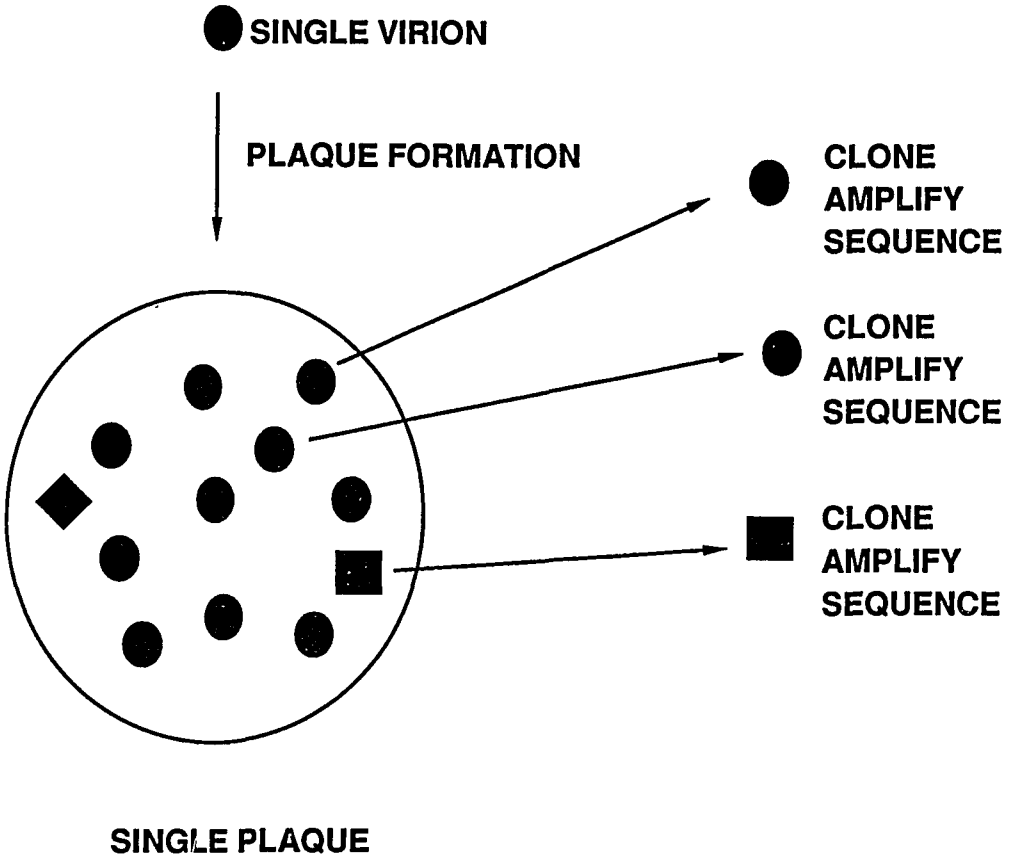
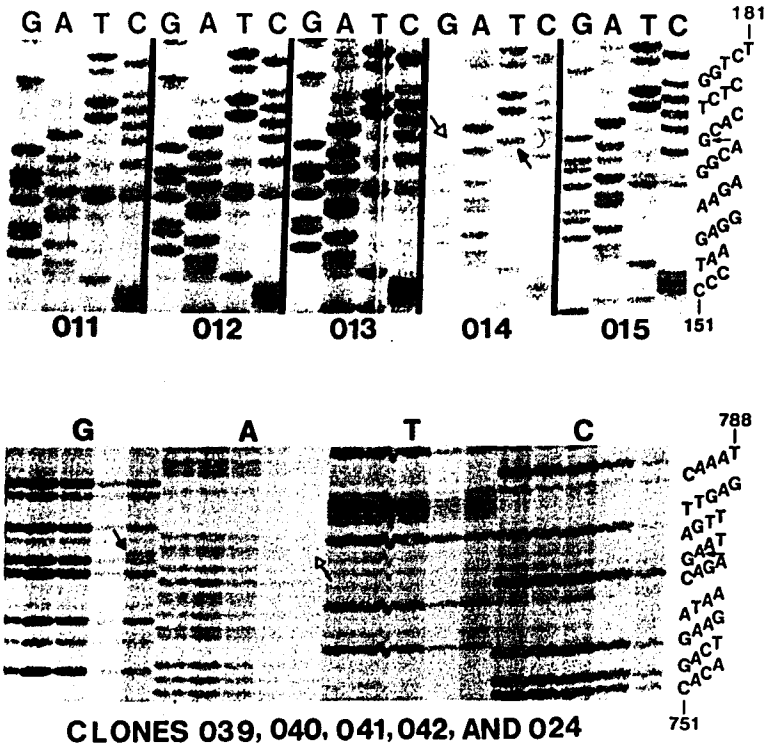


Figure 3. DETECTION OF POINT MUTATIONS IN NS GENES OF INFLUENZA VIRUS CLONES 014 AND 024.

(A) In the top half of the figure, the sequencing reactions of the NS genes of clones 011, 012, 013, 014 and 015, were electrophoresed in a 6% polyacrylamide buffer-gradient gel. Positions 151 through 181 are shown and the G to T change at position 169 in clone 014 is indicated by the arrows. In the bottom half of the figure, the sequencing reactions for the NS genes of clones 039, 040, 041, 042 and 024 were electrophoresed in a buffer-gradient gel. The G reactions for five clones, 039, 040, 041, 042 and 024 (left to right) were loaded side-by-side as were the A, T and C reactions. Positions 751 to 788 are shown, and the A to G transition at position 772 in clone 024 is indicated by arrows.

FIGURE 3.


G A T C G A T C G A T C G A T C G A T C
181
151
011
012
013
014
015
G
A
T
C
788
751
CLONES 039, 040, 041, 042, AND 024

(arrows). All seven point mutations are listed in Table 1 along with the predicted amino acid coding changes. As can be seen, the mutations were evenly distributed along the gene, and there was only one mutation per variant gene. Two mutants, 050 and 069, which were derived from clones on separated dishes were found to have the same point change. This finding was not unexpected since point mutations which occur early in the development of the plaque would represent a higher fraction of the variants than ones which occur late. Six of the seven variants were found to encode amino acid changes in the NS gene protein products. This observed frequency of coding changes is in agreement with the predicted frequency of coding changes in a randomly mutated gene coding for overlapping NS1 and NS2 proteins.

It is assumed that back-mutations were negligible and that the mutants had a growth fitness similar to that of the wild-type virus (see below). Based on these assumptions, dividing the number of variants by the number of infectious cycles accounted for the effect of mutations occurring over several replications and of mutations being sampled twice.

The mutation rate was calculated by using the formula: mutation rate = 7 mutations/91,708 nucleotides/5 infectious cycles = 1.5×10^{-5} mutations per nucleotide per infectious cycle.

Although plaque growth dynamics are complex and although synchrony in burst cycles may be lost, the number of growth cycles was most probably five. If all the sampled virus clones were derived from only four infectious cycles, the mutation rate would be underestimated by 26%. Similarly, if all the sampled clones resulted from six burst, the mutation rate would be overestimated by 17% in our calculations. Thus, the error contributed by clones that had different numbers of growth cycles was small.

2. CHARACTERIZATION OF NS GENE VARIANTS

It was assumed in calculating the mutation rate that the mutants had no significant

TABLE 1.

POINT MUTATIONS DETECTED AMONG NS GENES OF 108 INFLUENZA A/WSN/33 VIRUS CLONES

Mutant	Mutation*	Amino acid change	Position
014	G -> U	Ser -> Ile in NS1	169
024	A -> G	Asn -> Asp in NS2	772
031	U -> C	Trp -> Arg in NS1	72
049	U -> C	Silent	68
050	G -> A	Arg -> Lys in NS1	379
069	G -> A	Arg -> Lys in NS1	379
082	C -> U	Leu -> Phe in NS2	856

*Mutations detected by RNA sequencing are presented in the message orientation. The parent NS gene sequence is according to Buonagurio et al., 1986.

selective advantage or disadvantage during plaque formation. If the mutations were phenotypically neutral, the percentage of a given mutant would remain constant as the total virus population expanded in the plaque.

To test whether the NS gene mutations were neutral, the rate of virus production was assayed in a multicycle infection. The time course of virus production in multicycle infections is very sensitive to small changes in virus fitness. The results are diagrammed in Fig. 4. Included in the study were the parent clone and two randomly selected virus clones whose NS genes did not contain mutations. Samples from the liquid overlay taken at the indicated times were titered for the concentration of PFU. In the diagram, lines connect the time points in the controls. As can be seen, there was no significant difference between the wild-type and variant clones. One virus appeared to lag behind the others at some intermediate time points, but it was the variant carrying the silent mutation, and it is very unlikely that this NS gene mutation was responsible for the small lag. All seven variants were therefore indistinguishable from the wild-type virus, and thus the mutations appeared to be neutral.

3. DETERMINATION OF VP1 GENE MUTATION RATE

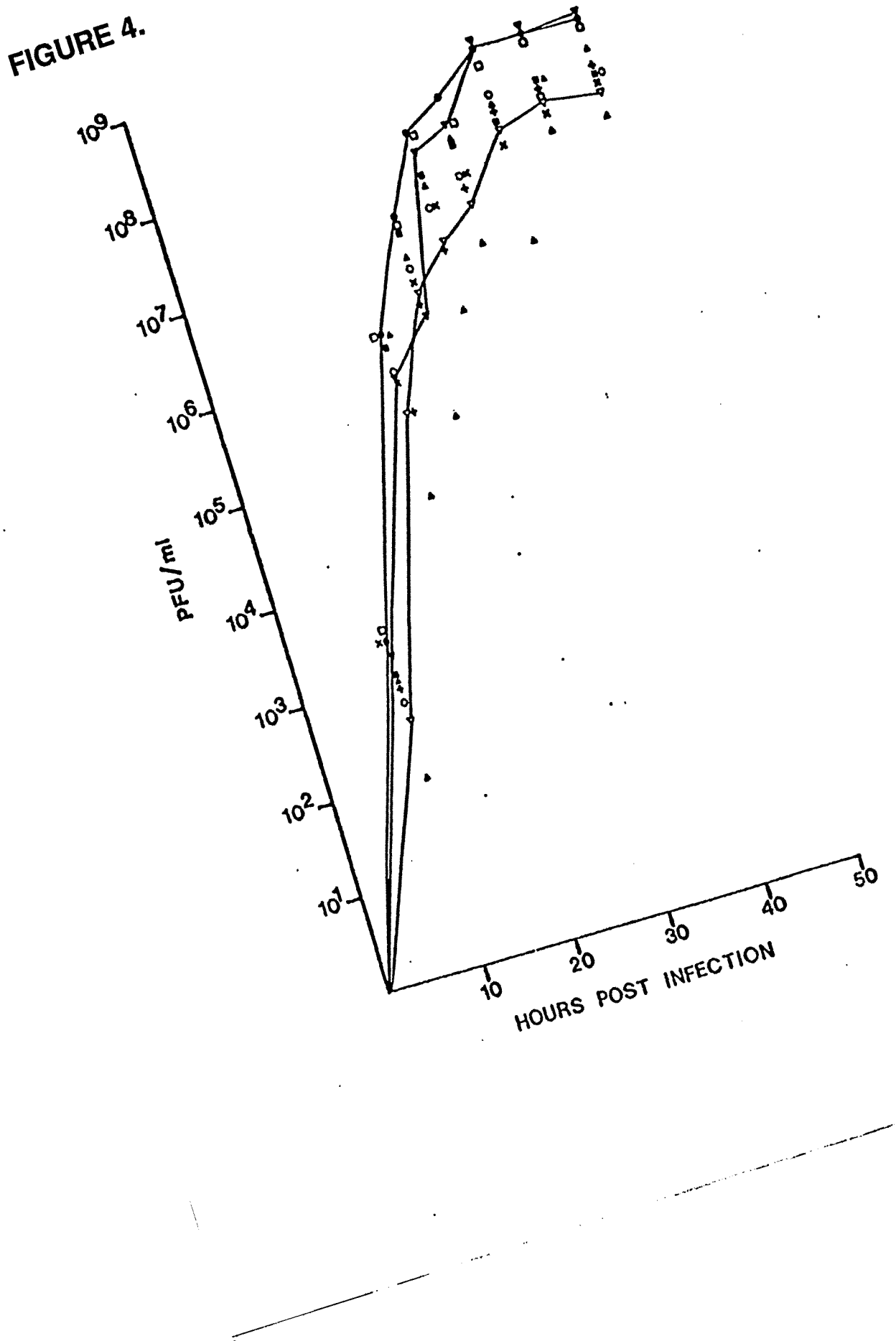
Plaque passage of the Mahoney strain of poliovirus type 1 was done as for the influenza virus experiment. Virus which had been plaque purified twice was infected onto cell monolayers, and at 36 h p.i., a well-isolated plaque containing 5.2 x 10^6 infectious particles was picked. Since the generation time of poliovirus is about 7 h, this allowed for about five infectious cycles. Again, individual clones from among the 5.2 x 10^6 PFU were obtained from this plaque, and viral RNA was prepared following a two-step amplification process.

The sequence of the Mahoney strain VP1 gene is the same as that presented by Racaniello and Baltimore (1981) except for three point changes at positions 2837 (A to G), 3139 (C to T), and 3151 (A to T). (Numbering is according to Kitamura et al.

Figure 4. COMPARISON OF GROWTH KINETICS OF INFLUENZA VIRUS NS GENE VARIANTS IN MDCK CELLS.

The time course of production of virus after infection at an MOI of 3 x 10^{-4} were determined for the parental virus, two clones which had the parental NS gen sequence, and the seven variants. The PFU per ml of the parental virus (●) and clones 007 (▼), 014(■), 024(□), 031(○), 049(△), 050(▲), 069(×), 082(✚), and 086(▽) are shown. The data points for the viruses containing the parental NS gene sequence (the parental virus and clones 007 and 086) are connected by lines for clarity. The PFU per ml of supernatant medium were determined for different time points by plaque titration on MDCK cells.

FIGURE 4.


10^9
10^8
10^7
10^6
10^5
10^4
10^3
10^2
10^1
PFU/ml
10
20
30
40
50
HOURS POST INFECTION

(1981).) These three positions also differ from the sequence presented by Kitamura et al. (1981) and probably represent the changes which randomly occur after repeated passage. Interestingly, two of the three changes were associated with amino acid changes in the VP1 protein. When the sequences of the 105 VP1 genes were compared, no point mutations were observed. Since no mutations occurred in the VP1 genes of the clones analyzed, the upper limit of the mutation rate can be calculated as follows: mutation rate < 1 mutation/95,688 nucleotides/5 infectious cycles < 2.1×10^{-6} mutations per nucleotide per infectious cycle.

Analysis of the statistical significance of the difference between mutation rates in the two experiments was performed by a chi-square test with a two-by-two contingency table of the data. It was found that the mutation rate of the VP1 gene was significantly lower than that of the NS gene ($P < 0.025$). In addition, the difference was statistically significant by a Student t test applied to the binomial standard error of the influenza virus mutation rate ($p < 0.01$). However, since each mutation rate was only measured once, we cannot predict the variation in mutation rates had more than one mutation experiment been performed for each virus.

D. DISCUSSION

1. COMPARISON OF INFLUENZA A VIRUS AND POLIOVIRUS TYPE 1 MUTATION RATES

The present study was undertaken to determine the precise mutation rates of two different RNA viruses. The specific strains were chosen because of the technical ease of plaque passaging at 37°C and because of the excellent growth characteristics which facilitated the sequencing of template preparations. The genes were elected based on prior demonstrations that they were capable of variation in nature or in the laboratory. The NS gene of influenza A virus has been shown to evolve very rapidly in nature

(Buonagurio et al., 1986; Krystal et al., 1983) and the VP1 gene of poliovirus type 1 was shown to tolerate mutations (Emini et al., 1983; Nomoto et al., 1982). The influenza A virus mutation rate was found to be significantly higher than that of poliovirus type 1. This result was surprising in the light of earlier data involving the selection of antigenic variants with neutralizing monoclonal antibody preparations. The frequency of variants, as measured by the plaque reduction assay, was as high for poliovirus type 1 as it was for influenza A virus in many cases (Emini et al., 1983; Lubeck et al., 1980). However, this assay may not reflect the true mutation frequency of the viral genomes, since the result may be affected by various other parameters. For example, the avidity and discriminating capabilities of the monoclonal antibody preparations may differ. Also, the genomic target size for the antibody-combining site is unknown. Further, it has been shown with poliovirus that even mutations outside of the antibody-binding site can inhibit neutralization by the antibody (Blondel et al., 1986). It is thus difficult to predict a precise mutation rate from data on the frequency of antigenic variants.

Is the lower mutation rate in the poliovirus VP1 gene the result of high replicase fidelity or is it due to very high constraints against lethal or deleterious mutations? The data on the selection of antigenic variants (Emini et al., 1983) demonstrates that there must be regions along the gene which have relaxed constraint, and comparison of the VP1 sequences of the Mahoney and the Sabin 1 strains reveals that many coding and silent changes are allowed in the VP1 gene. The VP1 gene, which represents 12% of the genome, contains 9 of the 57 total point changes and 7 of the 19 coding differences between the Mahoney and Sabin strains (Nomoto et al., 1982). It is thus suggested that the absence of poliovirus VP1 mutants among the 105 clones sequenced may not be due to constraints alone but may also be the result of increased fidelity of the viral replicase.

Based on the calculated mutation rates, the following prediction can be made regarding the frequencies of mutants present in influenza virus and poliovirus populations. If all the influenza A virus genes are under the same constraints as the NS gene and thus have the same mutation rate, then every virus in a plaque would average one point change per genome (mutation rate x 13.6-kilobase genome x 5 cycles). An influenza virus population therefore represents a quasispecies (Eigen, 1971; Eigen and Schuster, 1977) in which there is a consensus sequence, but each individual is unique. A similar calculation suggests that a poliovirus plaque would contain less than 8% mutant infectious particles. Therefore, the diversity in the poliovirus population is not as great as that predicted in the influenza virus population.

2. COMPARISONS OF MUTATION RATE ESTIMATES IN OTHER SYSTEMS

Mutation rates and mutation frequencies have been estimated in a variety of other systems. Recently, a study of the vesicular stomatitis virus substitution frequency revealed a surprisingly high in vitro rate of misincorporation of an extracistronic nucleotide (Steinhauer and Holland, 1986). The frequency of misincorporation was about 10^{-4} substitutions per base incorporated at the site. The same site was analyzed in vivo, and the substitution frequency was again found to be about 10^{-4}. Caution must be exercised in comparing this substitution frequency with the mutation rates determined for influenza and polioviruses, since this specific site may not be characteristic for the genome. In this respect, it should be noted that mutation frequencies of specific single bases are often idiosyncratic. In the phage T4rII locus, 10^4-fold differences in mutation frequencies have been observed for single nucleotides at different sites in the gene (Salts and Rosen, 1971). The indication that VSV might have a high mutation rate would suggest that a direct comparison with the mutation rates measured in this study by using the same methods would be very interesting. The direct analysis of clones derived from a single plaque passage, as was done in the present study, examines the

mutation rate of an entire gene and thus appears to give more reliable data.

Mutation frequencies have also been estimated for virus populations after serial uncloned tissue culture passage by using T1 oligonucleotide mapping analysis. After 28 passages, foot-and-mouth disease virus was shown to accumulate virus that consists of particles with two to eight mutations per genome (Sobrino et al., 1983), and bacteriophage QB was similarly shown to have one or two changes per genome after 30 or more passages (Domingo et al., 1978). These estimates, based on multiple passages with an unknown number of replication cycles, do not allow us to calculate mutation rates (mutation frequency per infectious cycle for these viruses, and thus these values cannot be directly compared with the data obtained for influenza and polioviruses during a single plaque generation.

Other estimates of mutation rates for bacteriophages, bacteria, and a fungus were derived from experiments in which the yield of test organisms grown under selective conditions was compared with the yield grown under permissive conditions. Although the molecular basis of the mutation allowing growth under selective conditions was unknown, it was assumed to be a point mutation, and the gene was assumed to have a standard length if the length was unknown (Drake, 1969). To compare our data with these results, the NS gene mutation rate must be converted to fit the definitions of Drake (1969). The mutation rate of the NS gene after this conversion is approximately 10^2 times higher than the rates of S. typhimurium and E. coli, and approximately 10^5 times higher than that of N. crassa. (For detailed calculations see Materials and Methods.)

3. ROLE OF MUTATION RATE IN VIRAL EVOLUTION

There is epidemiologic and genetic evidence that influenza A viruses evolve more rapidly than other viruses in humans (Webster et al., 1982; Bull. WHO, 1985, 63, 479-484). Specifically, vaccine strains used against influenza A viruses have to be changed

frequently (at least every 2 to 3 years) to protect against an evolving virus population. In contrast, vaccines against poliovirus (and most other human viruses) are based on strains which have been used for the last several decades, ostensibly without loss in efficacy. One might therefore speculate that the higher mutation rates found for influenza A viruses provide a molecular basis for this difference. High mutation rates would be necessary to generate the great amount of diversity required for the extraordinarily rare event of producing a beneficial mutant which would be selected in the host environment. It should be noted, however, that extensive variation has been observed in poliovirus populations circulating in nature (Nottay et al., 1981), and thus, the above model for the role of mutation rate must remain but one hypothesis for explaining the phenomenon of the rapid evolution of influenza A viruses.

Furthermore, one would like to know whether other viruses, for example, the acquired immune deficiency syndrome virus or herpesviruses, show comparably high mutation rates which may then affect the successful use of vaccines. The precise measurement of the mutation rates of other viruses may thus help to dissect the factors which determine the complex genetic interactions of viruses with the natural environment.

III.

DETERMINATION OF THE MUTATION RATE OF A RETROVIRUS

Jason M. Leider, Peter Palese, and Frances I. Smith

Department of Microbiology
Mount Sinai School of Medicine of CUNY
Fifth Avenue and 100th Street
New York, N.Y. 10029

Originally published in J. Virol. **62:** 3084-3091 (1988). Reprinted with permission from the American Society for Microbiology

A. INTRODUCTION

Viral variation may be the most important factor which limits the effectiveness of vaccine development against viral infection. Recent data have suggested that the amount of variation of a virus in nature can be correlated with its mutation rate (Parvin et al., 1986). Because of the emerging importance of retroviral infections in humans, it was of interest to determine the mutation rate of a representative retrovirus and to compare this value with that of other human RNA viruses.

Previous workers have studied variation of retroviruses (Coffin et al., 1980; Dougherty and Temin, 1986) and, although a precise mutation rate at the nucleotide level could not be calculated from these studies, the results obtained suggested that extremely high variability could arise in a population derived from a cloned virus over a relatively short period. In this study, we attempted to measure directly the mutation rate of the retrovirus Rous sarcoma virus (RSV) by screening for mutations arising over a single replication cycle. Denaturing-gradient gel electrophoresis, originally developed by Lerman et al. (1984) and subsequently modified to detect point mutations in RNA molecules by Smith et al. (1986), was used to screen for mutations. The ease and rapidity of this new technique allowed us to analyze several thousand nucleotides distributed over seven portions of the RSV genome. The results indicate that retroviruses are among the fastest-mutating viruses.

B. MATERIALS AND METHODS

1. CELLS

Chicken embryo fibroblasts (CEFs) were prepared from 10-day-old fertilized chicken eggs according to published methods (Hanafusa, 1969). CEFs susceptible to the Prague C strain of RSV (RSV Pr-C) were derived from eggs that were virus negative, chicken helper factor negative, and avian leukosis virus group-specific antigen negative

(v^- chf^- gs^-; SPAFAS, Inc., Norwich, Conn.). CEFs genetically-resistant to RSV Pr-C infection were derived from fertilized chicken eggs that were the kind gift of W.E. Briles, Department of Biological Sciences, Northern Illinois University. The culture medium employed consisted of Hams F-10 medium (Hazelton Research Products, Inc.) supplemented with 10% tryptose phosphate broth (TPB), 5% calf serum, 6.6 mM sodium bicarbonate, 100 U of penicillin per ml, and 100 ug of streptomycin per ml.

2. PLASMIDS

pATV-8, a recombinant plasmid containing the RSV Pr-C genome cloned into pBR322 (Katz et al., 1982), kindly provided by J. Coffin of the Tufts University School of Medicine, was the source of RSV Pr-C sequences. Various regions of the RSV Pr-C genome were subcloned into plasmids containing RNA polymerase promoters for the purpose of production of riboprobes. pGEM-4 (Promega Biotec, Madison, Wis.) and pIBI 31 (IBI) were used as cloning vectors. The following RSV Pr-C genomic fragments were subcloned from pATV-8; 256 to 520 into the SmaI-SphI site of pGEM-4 (pgag1); 520 to 1006 into the SmaI-SphI site of pGEM-4 (pgag3); 4068 to 4997 into the AccI-KpnI site of pGEM-4 (ppol1); 5257 to 5563 into the PstI-SmaI site of pGEM-4 (penv4); 5564 to 5837 into the PstI-SmaI site of pGEM-4 (penv1); 6440 to 6865 into the SacI-PstI site of pGEM-4 (penv3); 6983 to 7631 into the SalI-AvaI site of pGEM-4 (psrc1); 9238 to 9291 and 1 to 2318 into the EcoRI site of pIBI 31 (pGG-AMT3); 6285 to 9238 into the EcoRI-HindIII site of pIBI 31 (pE3S-MT7); and 2870 to 6085 into the HindIII site of pIBI 31 (pPE14-AMT7). All plasmids were made by standard techniques, transfected into Escherichia coli HB101 or DH5-alpha, selected by hybridization with radiolabeled nick-translated restriction fragments of pATV-8, and expanded, and the plasmid DNA was purified by the alkaline lysis method (Maniatis et al., 1982).

3. PREPARATION OF RNA PROBES

Plasmid DNA was digested with EcoRI or HindIII restriction enzymes, depending on

the desired sense of the transcripts. Linear templates (100 ug/ml) were then transcribed by SP6 RNA polymerase (7.5 U/ug of template DNA; Promega Biotec), T7 RNA polymerase (7.5 U/ug of template DNA; Promega Biotec), or T3 RNA polymerase (10 U/ug of template DNA; IBI) in the presence of [alpha-^{32}P] CTP, as described by Melton et al. (1984). Typically, 3 to 5 ug of single-stranded RNA were produced per microgram of input plasmid DNA.

4. PREPARATION OF DOUBLE-STRANDED (ds) RNA HYBRIDS

^{32}P-labeled minus-sense RNA was suspended in 30 ul of hybridization buffer (80% deionized formamide, 40 mM PIPES [piperazine-N,N'-bis(2-ethanesulfonic acid)], pH 6.7, 0.4 M NaCl, 1 mM EDTA), mixed with 1 ug viral RNA, and incubated for 12 to 16 hours at 45°C. To remove unhybridized RNA, 300 ul of S1 digestion buffer (30 mM sodium acetate, [pH 4.4], 280 mM NaCl, 5 mM $ZnCl_2$, containing 1,000 U of nuclease S1 per ml; Sigma Chemical Co., St. Louis, Mo.) was then added, and the reaction was incubated at 37°C for 1 h. This step was followed by a phenol-chloroform extraction, and the [^{32}P]RNA was ethanol precipitated, dissolved in 50 ul of sterile water, and stored at -20°C.

The wild-type control duplexes were made by hybridizing ^{32}P-labeled SP6- or T7-derived transcripts from the various pGEM-4 plasmids with T7- or T3-derived transcripts of the RSV Pr-C genomic fragments in the pIBI 31 plasmids. The latter plasmids contained regions complementary to the SP6 or T7 transcripts as well as additional retroviral sequences at the 5' and the 3' ends. It was necessary to employ such a strategy to make the control heteroduplexes in order to avoid the complication of hybridization due to polylinker sequence. Adjacent hybridized polylinker would alter the length and perhaps the melting properties of the hybrids. Conditions for hybridization and nuclease S1 digestion were the same as mentioned above.

5. ELECTROPHORESIS

The submerged gel apparatus used was similar to that described by Fischer and Lerman (1979). Gels were 6.5% polyacrylamide containing a denaturant gradient either perpendicular or parallel to the direction of electrophoresis (100% denaturant=7M urea and 60% formamide). Parallel gradients had a short-stacking gel (3.8% acrylamide, no denaturant). The gels were run at 200V submerged in an aquarium containing TAE buffer (40 mM Tris acetate, [pH7.4], 20 mM sodium acetate, and 1 mM EDTA) that was maintained at a constant temperature of 65°C by a thermomix (B. Braun Melsungen, AG) heater and stirrer. After electrophoresis, the gels were fixed in a 7.8% acetic acid-26.5% methanol solution, dried, and exposed to Cronex 4 film (E.I. duPont De Nemours & Co., Inc., Wilmington, Del.) for 12 to 48 hours. Non-denaturing 6.5% polyacrylamide gels run at room temperature were also used to check the integrity of all heteroduplexes prior to loading on denaturing-gradient gels.

6. TRANSFECTION OF CEFs

CEFs were transfected by the method of Potter et al. (1984) with RSV Pr-C cDNA purified from a partial HindIII digestion of pATV-8. A bank of capacitors (effective capacity, 14 uF) charged to 1500 volts was discharged via an electronic switch (Model ZA 1000; PDS, Madison, Wis.) through the sample suspended in 0.5 ml of 1X ZAP buffer (140 mM NaCl, 0.75 mM Na_2HPO_4, 25 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid, [pH 7.2]) by using a Cell Chamber (PDS) 5 mm in length and with a cross-sectional area 1 sq. cm. at 0°C. Cells were then allowed to adhere to tissue culture dishes at 37°C. Medium was changed at 3-day intervals, and complete transformation was seen in 14 days. Individual transformed cells were grown at a limiting dilution of 0.2 transformed cells per 2.5-cm dish containing a feeder layer of cells genetically resistant to infection by RSV Pr-C. Growth of transformed cells was scored by light microscopy. These clones represented parental candidates (see Fig. 5).

7. PREPARATION OF RSV AND PURIFICATION OF VIRAL RNA

To generate sufficient RNA for analysis of the viral genotype, supernatant was collected from dishes containing a single colony, and the virus was amplified by addition to a 5-cm dish of susceptible CEFs followed by subsequent expansion. Supernatant containing virus liberated from these confluent transformed cells was then harvested until a total volume of 200 ml had been collected. The supernatant was cleared by centrifugation at 8,000 x g for 10 min and layered over a 3-ml 20% sucrose cushion. Virus was pelleted by centrifugation at 25,000 rpm for 2 hours in a SW27 rotor (Beckman Instruments, Inc., Fullerton, Calif.). The virus pellet was suspended in 4 ml of a solution of 100 mM NaCl, 10 mM Tris hydrochloride, pH 7.4, and 1 mM EDTA, and RNA was extracted according to the method of Palese and Schulman (1976).

8. FOCUS-FORMING UNIT (FFU) ASSAY

A 2.5 cm dish of confluent CEFs was incubated with serial dilutions of the supernatant for 40 minutes at 37°C. The cells were then overlaid with medium containing 0.75% agar, incubated at 37°C, and monitored for focus formation. After 7 days of incubation, no foci were seen at any dilution. To allow further time for foci to appear and to prevent cell death, the agar overlay was removed and replaced with liquid medium. On day 23, foci were seen in dishes of various dilutions. Because of the opportunity for viral spread across the monolayers during this procedure, the titer of FFU per milliliter thus obtained is only an estimate.

9. COLONY FORMATION IN SOFT AGAR

A culture of confluent CEFs growing in a 10-cm tissue culture dish (7×10^6 cells per dish) was prepared and infected with 2 ml of virus (approximately 400 FFU) in the presence of 0.3 ml of DEAE dextran (1:40) for 4 h at 37°C. Six plates, each with a bottom agar layer of 80% F-10 medium, 10% tryptose phosphate broth, 6% heat-inactivated calf serum, 2% heat-inactivated chick serum, 88 ug of folic acid per ml, 0.04% dimethyl sulfoxide, 0.18% sodium bicarbonate, 100 U of penicillin per ml, 100 ug

of streptomycin per ml, 0.25 ug of fungizone per ml, and 0.56% agar, were prepared during this time. Virus-infected cells were trypsinized and suspended in 6.5 ml of liquid F-10 medium. To each agar-containing plate, 1 ml of infected cell suspension was added. A top agar layer consisting of 50% F-10, 6% tryptose phosphate broth, 0.09% sodium bicarbonate, 4% heat-inactivated calf serum, 2% heat-inactivated chick serum, 0.25 ug of vitamins per ml (GIBCO Laboratories, Grand Island, N.Y.), 40% cultured supernatant from uninfected CEFs, 0.5% dimethyl sulfoxide, 100 U of penicillin per ml, 100 ug of streptomycin per ml, 0.13 ug of fungizone per ml, and 0.44% agar was added to these six plates and shaken immediately to evenly distribute cells. The entire culture was allowed to solidify for 15 min at room temperature and then incubated at 37°C. After 14 days, well-isolated colonies containing 100 to 500 cells were seen. With a 1-ml syringe containing less than 0.1 ml of 0.05% trypsin and 0.2 mM EDTA, colonies were picked and then added to a 2.5-cm dish of confluent susceptible CEFs containing 1 ml of liquid medium. Transformation of these dishes was monitored daily and expansion of these dishes was performed as required.

C. RESULTS

1. EXPERIMENTAL STRATEGY FOR DETERMINATION OF THE RSV MUTATION RATE

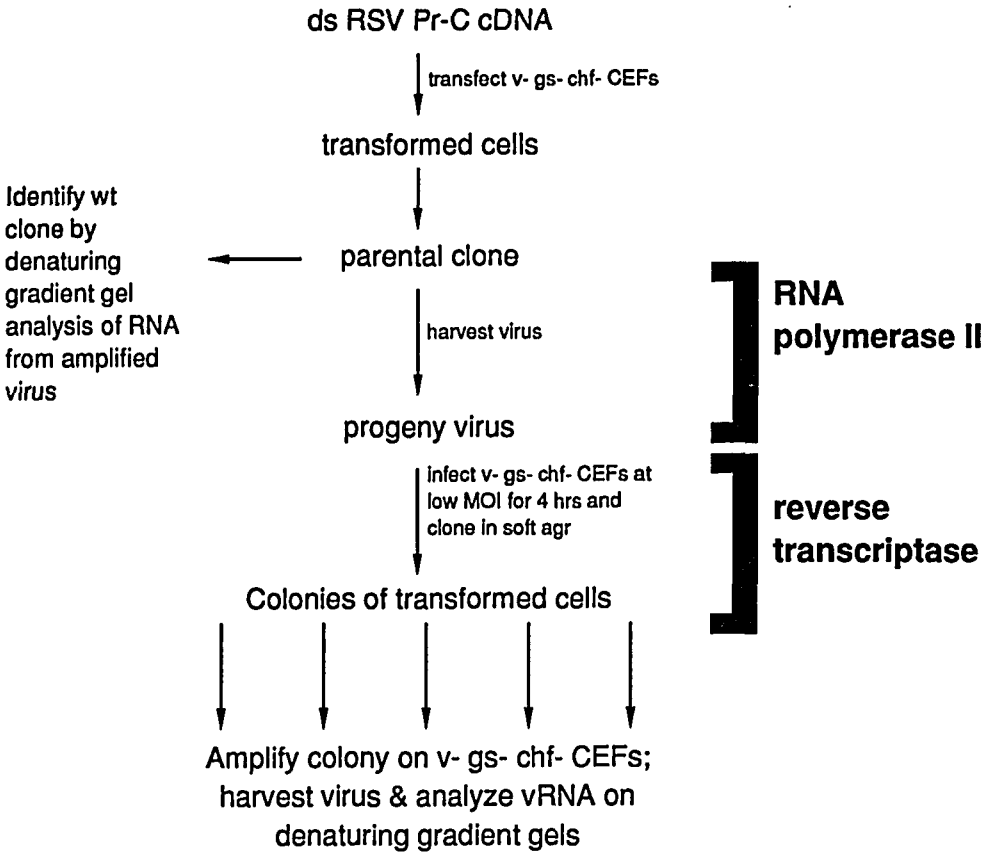
To determine the RSV mutation rate, variation within the progeny of a parental clone was examined after a single replication cycle. The replication cycle begins with a single colony of transformed cells (Fig. 5) and virus produced from that single colony was used to establish multiple transformed cell clones. In this single replication cycle, proviral DNA of the original parental colony was transcribed by RNA polymerase II to produce viral RNA for progeny virus. After infection of cells with this progeny virus, reverse transcription allowed synthesis and integration of proviral DNA. Mutations

Figure 5. EXPERIMENTAL DESIGN FOR MEASURING THE MUTATION RATE OF RSV DURING A SINGLE REPLICATION CYCLE.

The single cycle of replication involves transcription first of the parental provirus by RNA polymerase II to produce the RNA of the viral genome and then of the viral RNA by reverse transcriptase to make the provirus that is integrated in progeny colonies. Analysis by denaturing-gradient gel electrophoresis of RNA from the amplified virus detects the consensus sequence, which reflects that of the provirus in the progeny colonies. Individual mutations arising during the amplification would thus not be detected by this method.

FIGURE 5.

EXPERIMENTAL DESIGN FOR MEASURING THE MUTATION RATE OF RSV DURING A SINGLE REPLICATION CYCLE



occurring during this single cycle of replication (from a single colony of transformed cells to transformed progeny cells) may be the result of errors made by RNA polymerase II or reverse transcriptase or both. Since mutations rarely happen during provirus replication (Jolly et al., 1986), it is unlikely that mutations caused by cell DNA polymerase during growth of the original colony would be detected. Using denaturing-gradient gel analysis, mutations in the progeny were screened and a mutation rate for the retrovirus for a single replication cycle was calculated accordingly.

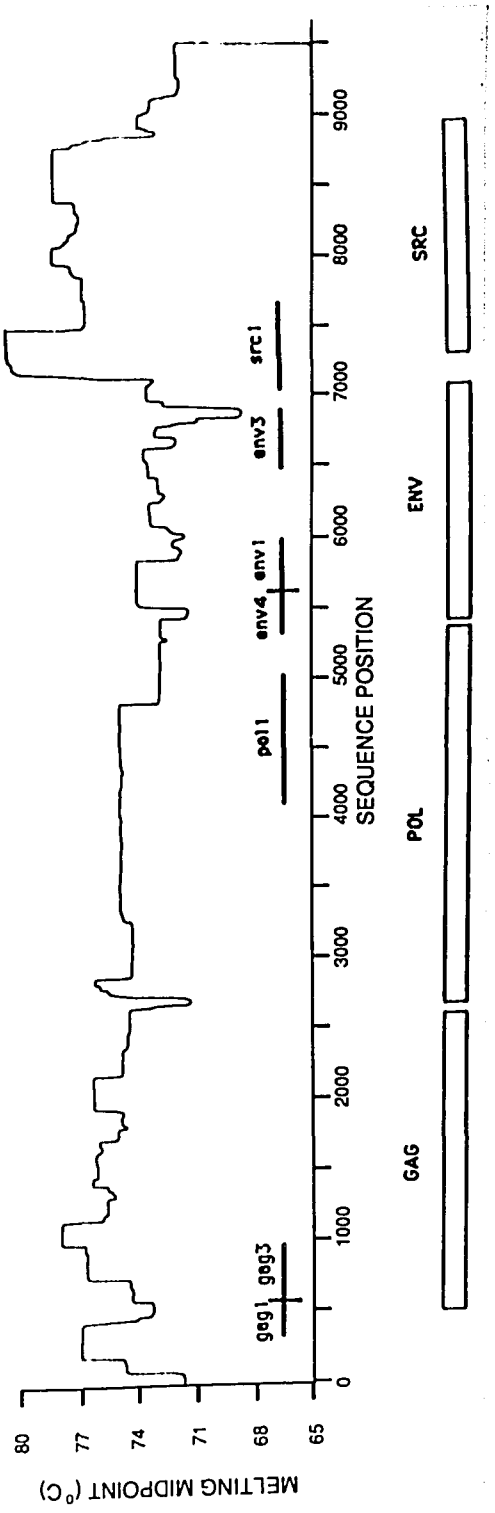
2. MELTING MAP OF THE ds DNA COPY OF THE RSV Pr-C GENOME

The first step in analysis of the RSV mutation rate by denaturing-gradient gel electrophoresis was the generation of a full-length melting map of the virus genome. Such a map is dependent on the availability of the nucleotide sequence. For this study the RSV Pr-C strain, which had been completely sequenced by Schwartz et al. (1983), was used. An expected melting progression along the ds DNA form of RSV Pr-C was calculated as described by Lerman et al. (1984). The abscissa in Fig. 6 represents nucleotide position within the RSV Pr-C genome, while the ordinate describes the temperature at which each base pair has an equal probability of helix or random chain configuration. Cooperativity among adjacent base pairs is strong enough to permit rather long domains of contiguous helix to undergo the transition from the helical to the random coil state within relatively narrow temperature intervals, giving a stepwise appearance to the melting map. Experimental results obtained with various dsRNA and DNA-RNA duplexes have shown that relative stabilities of different domains within the duplexes correlate well with those predicted by the theoretical melting behavior for the ds DNA duplexes (Smith et al., 1986). The sensitivity of denaturing gradient gel electrophoresis for the detection of single-point mutations in RNA molecules has been previously well characterized for both DNA (Lerman et al., 1984) and RNA molecules (Smith et al., 1986). Using a range of sequenced mutants, these workers showed that

Figure 6. MELTING MAP OF THE dsDNA FORM OF THE RSV PR-C GENOME.

The map indicates the temperature at which each base pair will be at equilibrium between helix or random chain configuration for the dsDNA molecule in an aqueous environment containing about 0.02M of sodium ions. Also indicated are the boundaries of the four retrovirus genes and the seven regions selected for mutational analysis. Numbering is according to that of Schwartz et al. (1983).

FIGURE 6.


MELTING MIDPOINT (°C)
80
77
74
71
68
65
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
SEQUENCE POSITION
gag1
gag3
pol1
env4
env1
env3
src1
GAG
POL
ENV
SRC

heteroduplexes between wild-type and mutant molecules are less thermodynamically stable than perfectly matched duplexes and undergo melting of mismatched domains at lower concentrations of denaturant. In both studies, all single-point mutations screened by heteroduplex analysis on denaturing-gradient gels were accurately and easily detected.

3. INITIAL CHARACTERIZATION OF VIRAL REGIONS CHOSEN FOR ANALYSIS

With the aid of the melting map, seven regions of the RSV Pr-C genome were chosen for further analysis (Fig. 6). Regions containing a relatively low-melting domain abutting a high-melting domain were selected. It is only in the low-melting domains of such heteroduplex regions that mutations can be detected reliably (Lerman et al., 1984). Another factor influencing our choice of regions for analysis was that we wanted to study at least one portion of each gene. Three regions of the env gene were chosen for analysis corresponding to regions that were previously reported to be conserved (*env4*) or hypervariable (*env1*) among different strains of avian sarcoma and leukosis viruses (ASLV) (Bova et al., 1988; Dorner and Coffin, 1986). These regions, therefore, were subcloned from pATV-8 into RNA polymerase transcription vectors pGEM-4. Control wild type duplexes were made by hybridizing ^{32}P-labeled transcripts of opposite polarity with one another (for specifics on construction of wild type duplexes, see Materials and Methods).

Since the melting map only predicts the location of the various domains, perpendicular denaturing-gradient gels were used to determine experimentally the percent denaturant at which melting of the domains of the RNA-RNA duplexes occurs (Fig. 7). An S1-treated wild type control heteroduplex for each region examined was loaded in a single well spanning the width of the cathode edge of the gel. Migration was from top to bottom, perpendicular to a denaturing gradient that increases from left to right. For example, Fig. 7A shows an analysis of the env1 region. RNA molecules

Figure 7. THE EXPERIMENTAL MELTING PATTERN OF dsRNA HYBRIDS.

The mobility of the radioactively labelled dsRNA hybrid was examined at 65°C on a 6.5% polyacrylamide gel containing a linearly increasing gradient of denaturant from 0 to 100%. Electrophoretic migration of the ds RNA hybrid is from top to bottom (perpendicular to the denaturing gradient). By using the first inflection point of this curve, it is possible to estimate the percent denaturant at which melting of the lowest melting domain, predicted from the theoretical melting map, occurs.

electrophoresis
⊕
⊖
A. env1
% denaturant
0 10 20 30 40 50 60 70 80 90 100
electrophoresis
⊕
⊖
B. env3
% denaturant
20 30 40 50 60 70 80 90
electrophoresis
⊕
⊖
C. env4
% denaturant
20 30 40 50 60 70 80 90 100

FIGURE 7.

on the left side of the curve (below 50% denaturant) are double stranded, while molecules on the right side of the curve (above 55% denaturant) are completely melted. Molecules between these two extremes have a completely melted low melting domain and an intact high melting domain, and migrated the most slowly. At the first inflection point of the experimental curve, melting of the low-melting domain predicted from the theoretical melting map for the env1 region, occurred. The perpendicular denaturing gels from the remaining regions showed patterns of varying sigmoidal shape (for example, the curves for the various env regions are shown in Fig. 7). The theoretical melting map was used to identify the location and extent of the lowest-melting domain for each region.

4. ISOLATION OF THE PARENTAL CLONE

The 9.3-kilobase fragment, representing the full-length RSV Pr-C genome, was isolated from a partial HindIII digestion of pATV-8. A large dish of confluent CEFs was transfected with this DNA by electroporation. In 14 days, complete transformation of the entire dish was observed. Transformed cells were then cloned at a limiting dilution (0.2 transformed cells per 2.5-cm dish) on cells genetically-resistant to infection by RSV Pr-C. Out of 98 2.5-cm dishes, 13 were found to contain a single focus. These clones represented parental candidates for the following experiments. Virus liberated into the supernatants of these dishes was collected and aliquoted. One aliquot of each candidate was amplified on susceptible CEFs. Within 3 days, complete transformation of these cells occurred. Virus was collected from the supernatants for analysis. Heteroduplexes were formed by hybridizing the amplified viral RNA of a candidate to ^{32}P-labeled transcripts derived from each of the subcloned regions. Nuclease S1 digestion followed by nondenaturing polyacrylamide gel electrophoresis did not detect any deletion-insertion mutations. Samples were then loaded onto a gel containing a gradient of denaturant parallel to the direction of electrophoresis. The presence of a

mutation would cause premature melting of the low-melting domain, leading to retardation of the sample in the gel. Our definition of a wild-type clone, therefore, was a candidate for which all seven regions being probed would exhibit patterns of migration identical to that of the control hybrid (Fig. 8). Additionally, although a candidate may have harbored multiple integrated proviruses (of which some may have contained mutations), we estimate that denaturing-gradient gel electrophoresis is capable of detecting mutant bands at up to one-tenth the intensity of the major band. Since the number of proviruses integrated per cell as a result of avian retrovirus infection is usually lower than 10 (Khoury and Hanafusa, 1976), it is unlikely that the chosen parental clone contained an undetected mutant provirus. Furthermore, if the mutations detected in the progeny clones were directly attributable to the presence of a mutated provirus in the parental clone, the mutations should all be identical. Because most mutations detected could be shown to be clearly different, due either to their locations in different regions of the genome or to the different mobilities of their heteroduplexes on denaturing-gradient gels (see below), it appears that the chosen parental clone did in fact contain only wild-type provirus. It should also be noted that of the five clones which emerged as possible parental candidates, the one with the highest virus yield (approximately 2.5 x 10^3 FFU/ml) was chosen as the parent for this study.

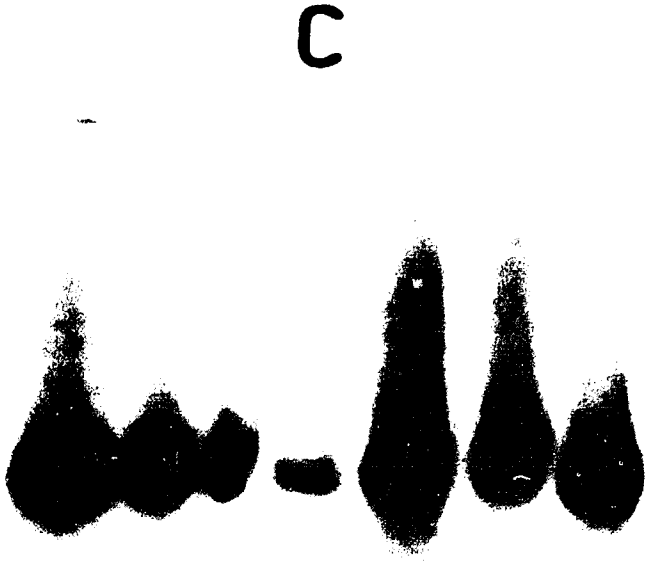
5. PROGENY GENERATION AND MUTATION RATE CALCULATION

A confluent dish of susceptible CEFs was infected at a multiplicity of infection of approximately 10^{-5} FFU per cell with an original aliquot of the parental virus. After 4 h, cells were split into six 10-cm dishes and grown in soft agar. After 14 days, approximately 60 colonies were seen in each dish. Well-isolated colonies were picked and amplified. Virus was harvested from the supernatant of each of 58 progeny colonies, and viral RNA was examined for mutation. Nuclease S1 analysis of RNA duplexes for the seven subcloned regions did not detect any insertion-deletion

Figure 8. SCREENING PARENTAL CANDIDATES IN THE *env4* REGION USING PARALLEL DENATURING-GRADIENT GEL ELECTROPHORESIS.

RNA-RNA hybrids were constructed using radioactively labeled minus-sense probes and viral RNA from parental candidates as described in Materials and Methods. All hybrids were initially examined for deletion-insertion mutations using 6.5% polyacrylamide gels. Hybrids for this region were then electrophoresed at 65°C on a gel containing a denaturant gradient (20 to 80%) parallel to the electric field. Patterns of migration for *env4* were identical to that of the control hybrid (lane **C**).

FIGURE 8.



mutations. Further analysis was then performed using denaturing-gradient gel electrophoresis (see Fig. 9). As previously discussed, only the low-melting domains of each region can be analyzed by this method, and from the theoretical melting maps, the target size for each of these domains is known. The data is summarized in Table 2. In all, 65,250 nucleotides of the RSV genome of 58 progeny RNAs were analyzed, with the detection of nine mutations, located within selective regions of the gag, env, and src genes. No mutations were detected in the 3' region of the pol gene and two regions of the env gene. Sequence analysis of a chosen mutant (that showed retarded migration of the env1 probe) confirmed that we were in fact detecting single-base mutations (data not shown).

The overall mutation rate for the virus was calculated by using the formula: mutation rate = 9 mutations per 65,250 nucleotides per replication cycle = 1.4×10^{-4} mutations per nucleotide per replication cycle.

D. DISCUSSION

We report here a method that has allowed the rapid and direct measurement of the mutation rate of a retrovirus. Progeny descended from a single virion after one replication cycle were collected and seven different regions of the genome were analyzed for mutations. CEFs were infected at a low multiplicity of infection, thus avoiding the possible amplification of defective particles which could drive virus evolution (Spindler et al., 1982) and lead to an erroneous calculation of the RSV mutation rate. In all, 65,250 nucleotides were screened, with the detection of nine mutations. The RSV mutation rate was calculated as 1.4×10^{-4} mutations per nucleotide per replication cycle. Since provirus replication does not appear to be error prone (Jolly et al., 1986), we attribute this rate to mutations incurred during the replication of the RSV genome by RNA polymerase II or reverse transcriptase or both. Our experimental

Figure 9. MUTATIONAL ANALYSIS USING PARALLEL DENATURING-GRADIENT GEL ELECTROPHORESIS FOR PROGENY RNA HYBRIDS.

Mutations (lanes **M**) in the *env1* (A), *src1* (B), and *gag3* (C) region are shown here. Mutant samples have a retarded migration through the gradient with respect to the control hybrid (lanes **C**). The denaturing gradient used for the regions shown here are as follows: (A), 30 - 70%; (B), 30 - 70%; (C), 50 - 100%.

FIGURE 9.

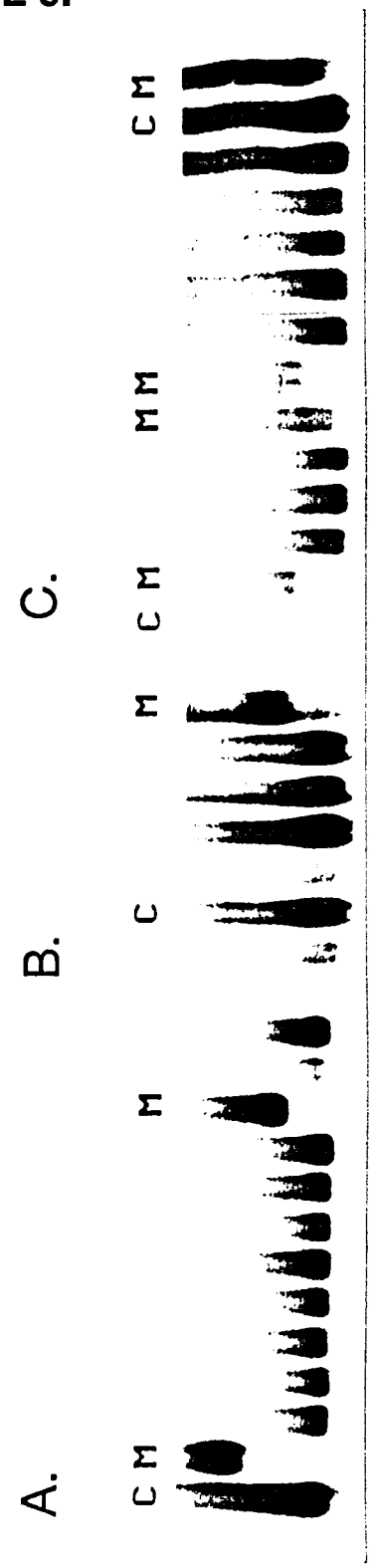


TABLE 2.

MUTATIONAL ANALYSIS OF SELECTED REGIONS OF THE RSV GENOME

Probe	Size of low-melting domain [a]	Nucleotides screened [b]	Mutations detected
gag1	104	6,032	1
gag3	156	9,048	4
pol1	249	14,442	0
env1	125	7,250	2
env4	159	9,222	0
env3	87	5,046	0
src1	245	14,210	2

[a] In nucleotides; as predicted by the theoretical melting map.

[b] For each region of the genome analyzed, the same 58 progeny RNAs were screened.

design, however, did not allow us to dissect the proportion of mutations contributed by each enzyme alone.

The use of denaturing-gradient gel analysis allowed several sites throughout the genome to be screened for mutations, overcoming any bias which may be due to the existence of regions of differential variability within viral genomes. An extreme example of such variability is the bacteriophage T4 rII locus, in which 10^4-fold differences in mutation frequencies were observed for single nucleotides at different sites in the gene (Salts and Rosen, 1971). Additionally, there are many reports demonstrating that different regions of viral genomes can exhibit different levels of variability in nature [for reviews, see Smith and Inglis (1987); Steinhauer and Holland (1987)]. Thus, studies analyzing only a small portion of a viral genome may give skewed estimates of the mutation rate of that virus. In this study, mutations were detected only within certain regions of the gag, src, and env genes. The region of the env gene in which we observed mutations corresponds to a region of the gp85 protein reported to be extremely variable among different strains of ASLV (Bova et al., 1988; Dorner and Coffin, 1986). In the pol gene, we analyzed a region corresponding to the zinc-binding site of the retroviral endonucleases (Johnson et al., 1986). Interestingly, although we analyzed the greatest number of nucleotides for this region, no mutations were detected. However, the numbers of mutations detected in each region are too small to conclude whether the distribution is nonrandom. We did not detect insertion-deletion mutation by nuclease S1 analysis in any region. This may be due to the deleterious effect of such mutations on proteins necessary for competent replication or transformation or both.

The design of these experiments enables us to calculate a minimum mutation rate for RSV under conditions in which the frequency of mutation due to recombination with endogenous viruses is very low. The v$^-$ gs$^-$ chf$^-$ chicken cells used express only very low levels of endogenous proviral message (approximately 0.5 to 1 copy of mRNA per

cell; Baker et al., 1981; Hayward et al., 1980; Wang et al., 1977). In contrast, the amount of viral RNA in a cell infected by an exogenous retrovirus is several orders of magnitude higher [3,000 to 20,000 copies per cell; (Hayward, 1977)]. Recombination between retroviruses is dependent on the expression of mRNA in a cell and most likely occurs during reverse transcription of RNA of heterozygous virions (that is, virions containing one endogenous and one exogenous viral RNA strand; for a review, see Linial and Blair, 1982). Because of the vast excess of exogenous viral RNA in infected v^- gs^- chf^- cells, the occurrence of copackaging heterozygous RNAs would be very low (most optimistically, 1 in 1,500 virions, assuming 1 copy of endogenous RNA per cell and 3,000 copies of exogenous RNA per cell). Even if one assumes that recombination would occur in 100% of such heterozygotes, the frequency of recombinants would be much lower than the number of mutants detected in this study (9 out of the 58 progeny clones examined). However, it is not possible to exclude the possibility that parental clone chosen for these experiments was aberrant in its expression of endogenous proviral message, and produced amounts of endogenous proviral message-specific RNA orders of magnitude higher than did previously examined cell lines. Nevertheless, because two of the nine mutations were located in the src gene, which is not present in endogenous viruses, we believe that we are in fact looking at polymerase error-induced mutations and not at those due to recombination.

Studies in which recombinants were selected for by phenotypic changes have shown that although the frequency of recombinants in chf^- cells was very low (less than 1 virion in 10^5), up to 5% of RSV progeny have been shown to be recombinant for an endogenous viral envelope marker after passage through chf^+ cells (Weiss et al., 1973). Therefore, although our calculation of the mutation rate for a retrovirus is already very high, it may in fact be much higher in cells expressing large amounts of endogenous viral RNA. Also, because the experiments described in this report selected for the

production of viable transforming virus, mutations that would produce nonviable progeny would not be detected. Therefore, the mutation rate presented here should be considered a minimum mutation rate.

It should also be noted that during the generation of progeny colonies, there can be a cycle or two of reinfection of the initially infected clone of cells, leading to more than one provirus in the clones so derived. However, if following insertion of a wild-type provirus, a mutation occurred in a second or subsequent round of infection and gave rise to a mutant provirus, a mixture of both wild-type and mutant virus would be produced. Therefore, a strong band on the denaturing-gradient gels representing the wild-type virus would be present as well as secondary retarded bands due to the mutant(s). However, as can be seen in Fig. 9, progeny characterized as mutants were those that gave a single dominant band of retarded mobility on a denaturing-gradient gel in the absence of a wild-type band. Therefore, the mutants detected in this study are most likely those that arose during a single cycle of replication.

Previous studies on variation in retroviruses have been undertaken. An in vitro error rate for reverse transcription of a DNA template from a mutant bacteriophage has been measured to be approximately 10^{-3} mutations per nucleotide polymerized (Gopinathan et al., 1979). This high error rate, however, may be able to be extrapolated to in vivo conditions. In another study, Coffin et al. (1980) repeatedly passaged the Prague B strain of RSV undiluted from one cell culture to another and used oligonucleotide fingerprint analysis to determine the variability of the virus. A frequency on the order of 3×10^{-4} bases per passage of virus was found, but because of the complexity of this system, this number cannot be used to develop a precise mutation rate (Coffin, 1986). Most recently, Dougherty and Temin (1986) employed a spleen necrosis virus-based vector that contained less than 1 kilobase of retrovirus sequences and two dominant selectable genes to estimate a mutation rate. By

monitoring for changes in expression of these selectable genes, they calculated that 0.5% mutant virions arose after a single replication cycle. This study, however, did not permit a precise calculation of the mutation rate in terms of mutations per nucleotide per replication cycle because the target size for mutation was not known. Nevertheless, the results from these studies do predict high variability in cloned retroviral populations, consistent with our results.

RNA viruses generally exhibit much higher variation in nature than do organisms that contain a DNA genome, and it has been postulated that this observation may be explained by the lack of fidelity of RNA polymerases (Holland et al., 1982). However, few direct determinations of mutation rates for RNA viruses have been done to test this hypothesis. Steinhauer and Holland (1986) used oligonucleotide mapping with T1 ribonuclease to calculate an error rate for the vesicular stomatitis virus RNA polymerase of $10^{-3.15}$ substitutions per base incorporated. Parvin et al. (1986) measured the mutation rates for influenza A virus and poliovirus type 1 as 1.5×10^{-5} and $<2 \times 10^{-5}$ mutations per nucleotide per cycle, respectively, by repeatedly dideoxy sequencing a single gene for each virus from numerous virus clones, each derived from the same plaque. Sedivy et al. (1987) also calculated a mutation rate for poliovirus (at 7.5×10^{-6} mutants per nucleotide per plaque generation) by using an inducible mammalian amber suppressor system to monitor the reversion frequency of a mutant of poliovirus. This poliovirus mutation rate is in close agreement with that of Parvin et al. (1986), as the latter authors estimated that poliovirus undergoes five replication cycles per plaque generation. However, the mutation rates calculated by Parvin et al. (1986) and Sedivy et al. (1987) are both much lower than the error rates observed by Ward et al. (1988) for the in vitro transcription of homopolymeric RNA templates by purified poliovirus RNA polymerase (which range from 7×10^{-4} to 5.4×10^{-3} noncomplementary nucleotides incorporated per total nucleotides incorporated, depending on the reaction conditions). This discrepancy

probably reflects the difference between in vivo and in vitro reaction conditions. Durbin and Stollar (1986) assessed virus variability by estimating the frequency for reversion of a particular Sindbis virus host-restricted mutant. Direct sequencing analysis of the E2 gene of revertants allowed them to estimate a mutation rate of $<10^{-6}$ errors per base incorporated, but they could not exclude the formal possibility that a second mutation outside the E2 gene might be involved in reversion to the ts^+ phenotype and consequently affect their reversion frequency estimate. Thus, the above results indicate that mutation rates are indeed high for RNA viruses (Holland et al., 1982), but that marked differences (by at least 2 orders of magnitude) exist between viruses. RSV, with a mutation rate of 1.4×10^{-4} mutations per nucleotide per cycle, is one of the fastest mutating organisms yet studied.

Previous results have suggested that mutation rates may correlate with the variability exhibited by viruses in nature and, therefore, with the effectiveness of vaccination against these viruses. There is epidemiological evidence indicating influenza A virus evolves more rapidly than poliovirus. Specifically, vaccine strains against influenza A viruses have to be changed frequently to protect against an evolving virus population. In contrast, vaccines against poliovirus are based on strains which have been used for the last several decades, ostensibly without loss of efficacy. Parvin et al. (1986) suggested that the higher mutation rate found for influenza A virus may be responsible for generating the diversity required for the rare event in which a mutant that could escape neutralization in a previously immunized host is produced. Our determination of the RSV mutation rate (which is approximately 9 and 65 times higher than that estimated for influenza A virus and poliovirus, respectively) indicates that RSV is an extremely mutable virus. This finding, in association with the reports of a rapid rates of evolution for human immunodeficiency virus (Hahn et al., 1986), Moloney murine sarcoma virus (Gojobori and Yokoyama, 1985), and RSV (Gojobori and Yokoyama,

1987), predicts that the high mutation rate of the retroviral genomes may be the source of their accelerated evolutionary rates. If mutation rates can indeed be correlated inversely with the effectiveness of vaccination against a given virus, these results suggest that prevention of retroviral-mediated disease via vaccination may be difficult.

IV.

THE E3 PROTEIN OF BOVINE CORONAVIRUS IS A RECEPTOR-DESTROYING ENZYME WITH ACETYLESTERASE ACTIVITY

Reinhard Vlasak,[1] Willem Luytjes,[2] Jason Leider[1] Willy Spaan,[2]

and Peter Palese[1]

Department of Microbiology, Mount Sinai School of Medicine,
New York, New York 10029-6574,[1] and
Department of Infectious Diseases and Immunology, Division of Virology,
State University of Utrecht, Utrecht, The Netherlands[2]

Originally published in J. Virol. **62:** 4686-4690 (1988). Reprinted with permission from the American Society for Microbiology.

A. INTRODUCTION

Members of at least three families of enveloped RNA viruses (i.e., *Orthomyxoviridae, Paramyxoviridae, and Coronaviridae*) bind to cell receptors containing sialic acid as receptor determinant (Rogers et al., 1986; Scheid and Choppin, 1974a; Scheid and Choppin, 1974b; Vlasak et al., 1988; Wiley and Skehel, 1987). In addition, virus-associated receptor-destroying activities have been described for members of these virus families (Gottschalk, 1957; Herrler et al., 1985; Nagai et al., 1976; Palese et al., 1974; Scheid et al., 1972; Vlasak et al., 1988). For parainfluenza and for influenza A and B viruses, the receptor-destroying enzyme is a neuraminidase, which removes sialic acids from cellular receptors. In parainfluenza viruses, the neuraminidase activity is located on the HN protein, which also possesses receptor binding/hemagglutinin activity (Scheid et al., 1972). In influenza A and B viruses, the neuraminidase is a glycoprotein which is distinct from the hemagglutinin protein (for reviews, see Air and Compans, 1983; Lamb, 1983). In contrast, for influenza C viruses and for bovine coronavirus (BCV), the receptor-destroying enzyme is not a neuraminidase, but an acetylesterase, which removes acetyl groups from O-acetylated sialic acids (Herrler et al., 1985; Vlasak et al., 1988). The esterase activity of influenza C virus is associated with the HE protein, which also possesses receptor-binding and fusion activity (Formanowski and Meier-Ewert, 1988; Herrler et al., 1988; Vlasak et al., 1987). However, no direct evidence was available that a specific BCV protein was associated with the esterase activity of viral preparations (Vlasak et al., 1988).

In the present report, we show that the BCV esterase activity resides on the E3 glycoprotein, which is one of three known surface proteins of the virus (Deregt and Babiuk, 1987; King and Brian, 1982). Enzymatic activity is inhibited by diisopropylfluorophosphate (DFP), indicating that the BCV receptor-destroying enzyme is a classical serine esterase, such as acetylcholinesterase (Cohen et al., 1967).

Furthermore, inhibition of the BCV acetylesterase by DFP inhibits viral replication, suggesting that the presence of an active viral esterase is essential for virus entry into host cells.

B. MATERIALS AND METHODS

1. VIRUSES AND CELLS

The BCV seed virus was obtained from Duphar B.B. Weesp (Amsterdam, The Netherlands). BCV was grown in Madin-Darby bovine kidney (MDBK) cells and purified as described previously (Vlasak et al., 1988). Influenza A/WSN/33 virus was grown in Madin-Darby canine kidney (MDCK cells as described previously (Brand and Palese, 1980). Erythrocytes derived from chickens (strain Rhode Island Red sex-linked chromosome X) were obtained from Pocono Rabbit Farm (Canadensis, Pa.).

2. ACETYLESTERASE ASSAY

Purified BCV preparations (5 ug) were incubated at room temperature in 1 ml of phosphate-buffered saline (PBS) containing 1 mM p-nitrophenylacetate. In order to prepare a 100 mM stock solution, p-nitropheynylacetate was dissolved in acetonitrile. The final acetonitrile concentration in the assay was 1%. Hydrolysis of the substrate was monitored at 400 nm with a chart recorder.

3. ENZYME INHIBITION ASSAY

DFP, phenylmethylsulfonyl fluoride (PMSF), and N-tosyl-L-phenylalanine chloromethyl ketone (TPCK) were dissolved in isopropanol in order to prepare 100X stock solutions. Purified BCV (5 ug) was preincubated with inhibitor for 10 min at room temperature in a volume of 50 ul. PBS was then added to a final volume of 990 ul. The reaction was started by the addition of 10 ul of substrate, and the activity was monitored at 400 nm.

4. HEMAGGLUTINATION ASSAY

A BCV suspension (22.5 ul containing 1,024 hemagglutination units) was mixed with 2.5 ul of 10 mM DFP (in 10% isopropanol), incubated at room temperature for 10 min, and used in the assay. Alternatively, DFP-treated BCV was purified over a 20 to 60% sucrose gradient. After the virus band was collected, the volume was adjusted to 500 ul with 20% sucrose. Controls were treated in the same say except that inhibitor was omitted. Hemagglutination assays were performed in V-shaped microtiter plates (Flow Laboratories, Inc., McLean, Va.) as described previously (Palese and Schulman, 1974).

5. PLAQUE ASSAYS

BCV and influenza A/WSN/33 virus were incubated with 1 mM DFP for 10 min at room temperature, purified over a 20 to 60% sucrose gradient, and titrated on MDBK cells by standard procedures (Parvin et al., 1986; Spaan et al., 1981).

6. PROTEIN LABELING WITH [^{3}H]DFP

Viral protein (140 ug) was labeled in PBS containing 0.1 mM [^{3}H]DFP (4.4 Ci/mmol; Dupont, NEN Research Products, Boston, Mass.) for 30 min at room temperature. Labeled virus was purified from unincorporated DFP on a 20 to 60% sucrose step gradient and 5-ug aliquots were analyzed on a sodium dodecyl sulfate-polyacrylamide gel (7% polyacrylamide) with or without beta-mercaptoethanol. After electrophoresis, gels were fixed, soaked in Amplify (Amersham Corp., Arlington Heights, Ill.), dried, and fluorographed for 2 to 6 days at -70^{0}C. Lanes of the gel containing molecular weight standards (Bio-Rad Laboratories, Richmond, Calif.) were silver stained.

7. PROTEIN LABELING WITH [^{35}S]METHIONINE

Confluent MDBK monolayers in 60-mm-diameter dishes were infected with BCV (multiplicity of infection of approximately 2) and incubated at 37^{0}C with 3T3 medium (Dulbecco modified Eagle medium supplemented with 10% heat-inactivated fetal calf serum, 50 U of penicillin G per ml, and 50 ug of streptomycin sulfate per ml, and 0.2% glucose) was added to the cells. After incubation at 37^{0}C for 30 min, cells were washed

with PBS and incubated at 37°C with labeling medium supplemented with 200 uCi of [^{35}S]methionine per ml (1,100 Ci/mmol; Dupont NEN). At 36 h postinfection, culture supernatants were collected, and labeled virus was purified on a 20 to 60% sucrose step gradient (see above).

C. RESULTS

1. ACETYLESTERASE ACTIVITY OF BCV

In a previous article, we showed that an acetylesterase activity is associated with BCV. This enzyme releases acetate from bovine submaxillary mucin at a rate comparable with that of influenza C virus esterase (Vlasak et al., 1988). To allow a more detailed characterization of the BCV enzyme, the synthetic low-molecular-weight substrate p-nitrophenylacetate was selected. Enzymatic measurements using this substrate involve determination of the cleavage product p-nitrophenol at 400 nm. This procedure is much less cumbersome than the cascade assay used for measuring acetate that is released from bovine submaxillary mucin. A purified BCV preparation effectively hydrolyzed this O-acetylester (Fig. 10). For further characterization of the BCV enzyme, different inhibitors were tested. DFP, a serine esterase and protease inhibitor, completely inhibited the BCV esterase, when preincubated with the virus at a 1 mM concentration (Fig. 10). Serine protease inhibitors PMSF and TPCK partially inhibited the BCV esterase (Table 3). EDTA had no effect, indicating that divalent cations are most likely not required for enzymatic activity, and dithiothreitol actually enhanced the activity.

2. SPECIFIC LABELING OF THE E3 PROTEIN OF BCV BY [^{3}H]DFP

Since DFP inhibits serine protease and serine esterases by binding covalently to the serine on the active site (Bender and Kezdy, 1965; Cohen et al., 1967), BCV preparations were radioactively labeled with [^{3}H]DFP by incubation at room temperature

Figure 10. HYDROLYSIS OF p-NITROPHENYLACETATE BY BCV.

Purified virus was incubated in a 1-ml cuvette containing 1mM p-nitrophenylacetate in PBS=1% acetonitrile. For DFP treatment, 5 ug of virus was preincubated with 1mM DFP for 10 min at room temperature (see Materials and Methods) and used in the assay. Hydrolysis of the chromogenic substrate was monitored at 400 nm using a chart recorder. OD_{400}, Optical density at 400 nm.

FIGURE 10.

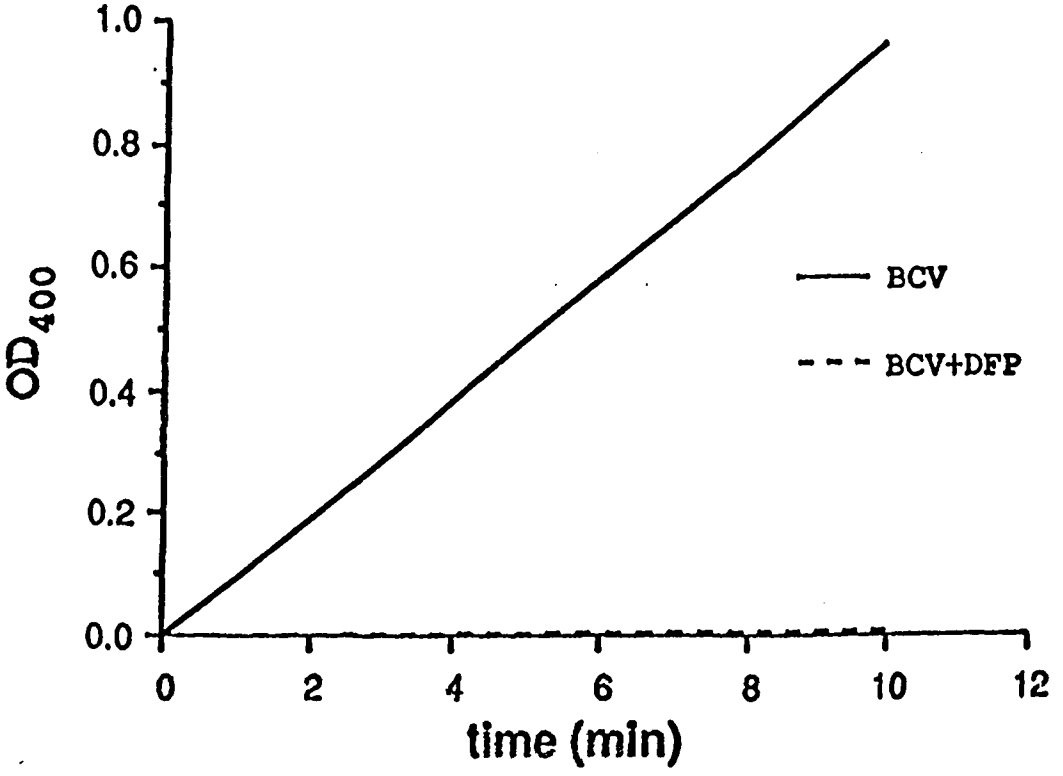


TABLE 3.

EFFECT OF DIFFERENT INHIBITORS ON BCV ESTERASE [a]

Inhibitor	Concn (mM)	% Activity
DFP	1	<1
PMSF	1	91
TPCK	0.1	79
EDTA	10	98
DTT	10	140

[a] Purified BCV was preincubated with DFP (diisopropyl fluorophosphate), PMSF (phenylmethysulfonyl fluoride), TPCK, (N-tosyl-L-phenylalanine chloromethyl ketone), EDTA (ethylenediaminetetracetic acid), or DTT (dithiothreitol) at the indicated concentration for 10 min at room temperature in a volume of 50 ul. After 20-fold dilution with PBS, acetylesterase activity was determined by using p-nitrophenylacetate (1 mM) as the substrate (see Materials and Methods).

for 30 min. Labeled virus was purified over a 20 to 60% sucrose step gradient, and viral proteins were analyzed on a sodium dodecyl sulfate-polyacrylamide (7% polyacrylamide) gel with and without beta-mercaptoethanol. After fluorography, a single labeled protein was detected, migrating with an apparent M_w of 62,000 (62K) under reducing conditions (Fig. 11, lane 2) and with an apparent M_w of approximately 125K under nonreducing conditions (Fig. 11, lane 3). In both instances, the [^{3}H]DFP-labeled protein migrated in the position of the [^{35}S]methionine-labeled E3 protein of BCV (Fig. 11, lanes 1 and 4). The E3 protein of BCV is a homodimer composed of two 62K subunits, apparently connected by disulfide bridges, and it is one of the viral proteins recognized as making up the virion structure (Deregt and Babiuk, 1987; King and Brian, 1982; King et al., 1985).

3. HEMAGGLUTINATION OF DFP-TREATED BCV

We then addressed the question of whether inactivation of the viral esterase affects the binding of BCV to erythrocyte receptors. BCV preparations were incubated with 1 mM DFP for 10 min at room temperature and purified over a 20 to 60% sucrose step gradient. DFP-treated and mock-treated BCV were collected, adjusted to a volume of 500 ul, and tested for receptor-binding activity by using a hemagglutination assay. No difference between treated and untreated BCV preparations was detected when hemagglutination assays were performed at 4^{0}C, indicating that there was no requirement of an active esterase for binding to cell receptors (Fig. 12). However, if the temperature was shifted up to 20^{0}C and then to 37^{0}C, untreated virus started to elute from erythrocytes as a result of the presence of the receptor-destroying activity. In contrast, the hemagglutination pattern of DFP-treated virus was stable under these conditions (Fig. 12). The same result was obtained when BCV was incubated with 1 mM DFP and used directly in hemagglutination assays without prior purification over a sucrose gradient. Clearly, inhibition of the viral esterase activity stabilized binding of

Figure 11. ANALYSIS OF [^{3}H]DFP-LABELED BCV PROTEINS.

[^{3}H]DFP-labeled and [^{35}S]methionine-labeled BCV was electrophoresed in a 7% polyacrylamide gel after solubilization in sample loading buffer with (lanes 1 and 2) or without (lanes 3 and 4) beta-mercaptoethanol. Lanes 2 and 3, [^{3}H]DFP-labeled BCV; lanes 1 and 4, [^{35}S]methionine-labeled BCV. Positions of viral proteins are indicated by arrows between lanes 2 and 3; positions of molecular weight markers (in thousands [K]) are shown on the right of both gels by arrows.

FIGURE 11.

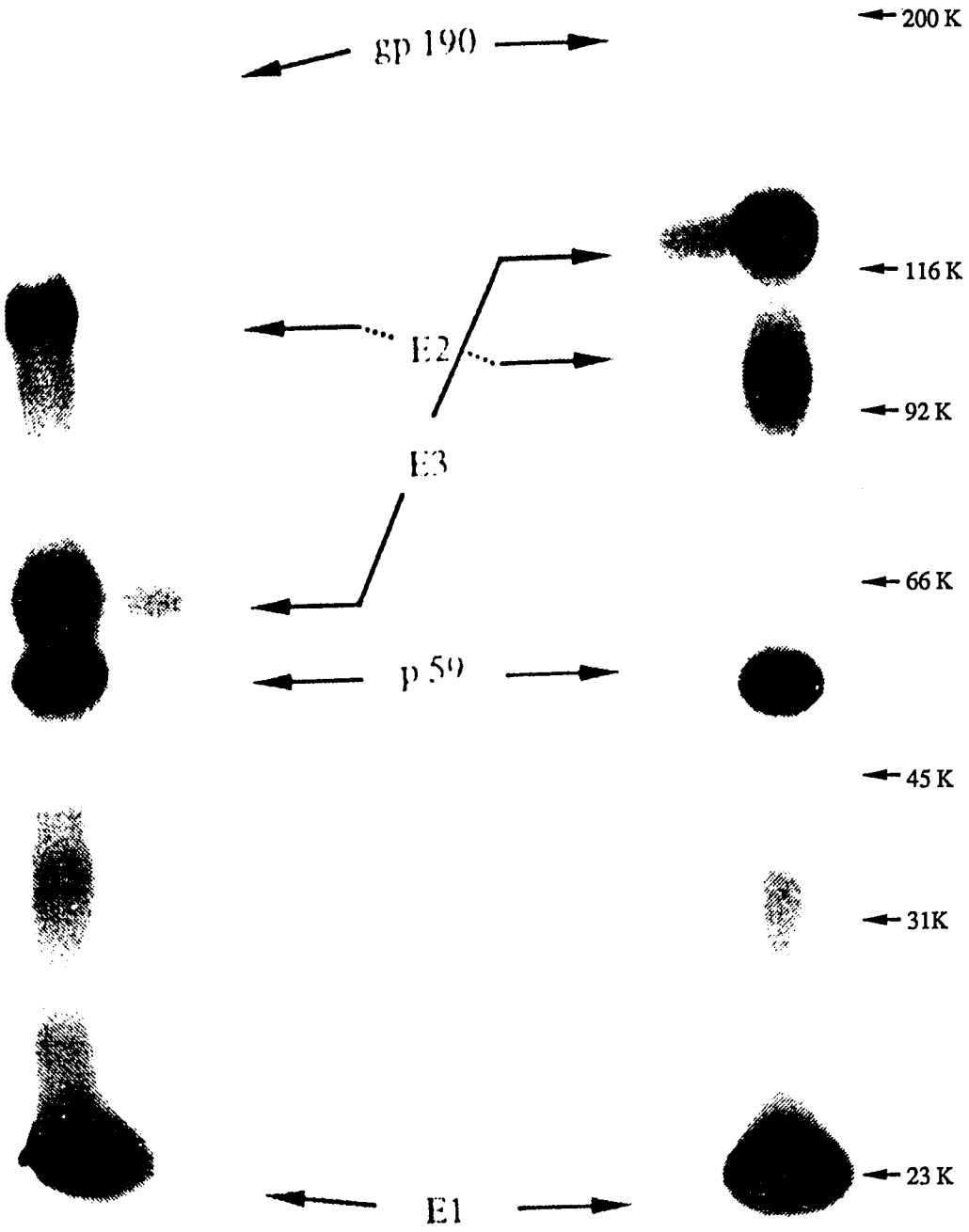
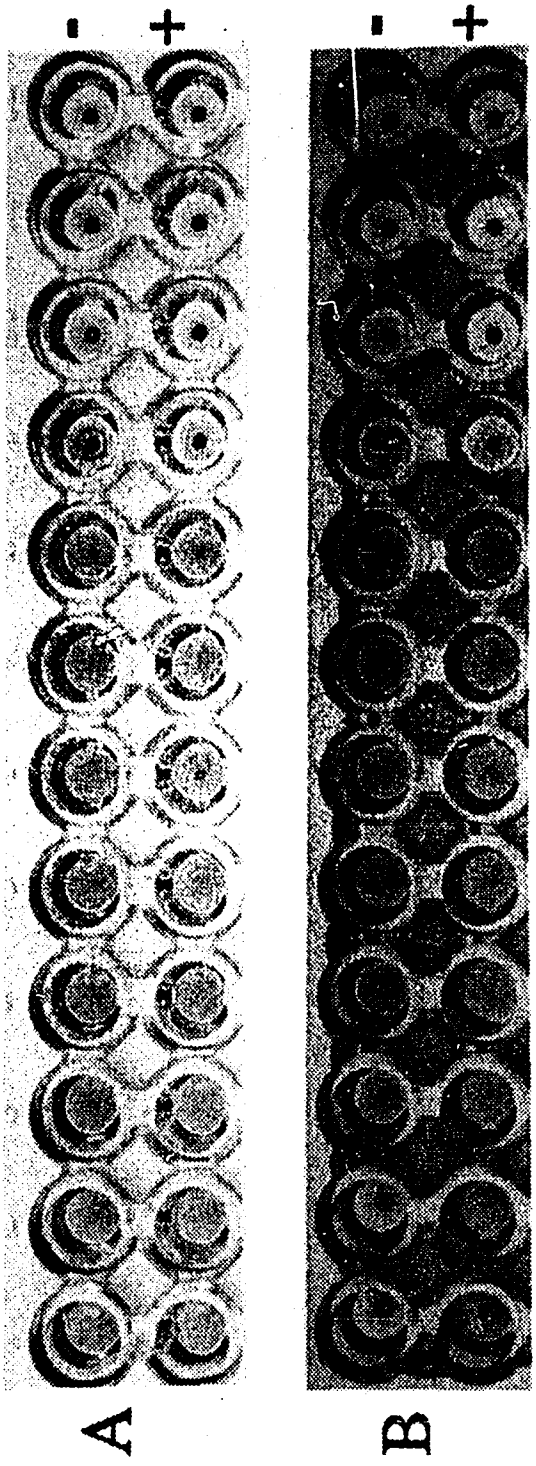


Figure 12. HEMAGGLUTINATION PATTERN OF DFP-TREATED BCV.

BCV (1,024 hemagglutination units) was incubated with 1 mM DFP for 10 min at room temperature. Hemagglutination was performed by using chicken erythrocytes and BCV preparations at serial 1:2 dilutions. (A) The microtiter plate was incubated at 4^0C for 1 h and photographed. (B) The same plate was incubated overnight at room temperature, followed by incubation at 37^0C for 30 min, and rephotographed. Symbols: +, BCV incubated with DFP; -, mock-treated BCV.

FIGURE 12.



BCV to erythrocyte receptors.

4. ESTERASE ACTIVITY REQUIRED FOR BCV REPLICATION

In order to explore the biological significance of the BCV esterase during virus replication, infectivity titrations were performed with DFP-treated BCV. Gradient-purified DFP-treated and mock-treated BCV were tested for acetylesterase activity and hemagglutination titers and then were used in 10-fold serial dilutions for plaque assays. DFP-treated BCV had approximately 100- to 400-fold-lower infectivity titers than mock-treated BCV (Table 4). To exclude the possibility of unspecific effects caused by DFP, assays with influenza A/WSN/33 virus were done in parallel. This virus does not possess an esterase activity, and DFP treatment does not result in phosphorylation of viral protein. No difference was found between the titer of DFP-treated and mock-treated influenza A/WSN/33 virus. These data strongly suggest that a functional BCV esterase is required for virus replication.

D. DISCUSSION

Previously we have shown that human coronavirus OC43 and BCV bind to sialic-acid containing cell receptors (Vlasak et al., 1988). Since the influenza C virus receptor-destroying enzyme, an O-acetylesterase (Herrler et al., 1985), removed coronavirus receptors, it was concluded that OC43 and BCV recognize receptors similar to those of influenza C virus. In addition, a receptor-destroying or acetylesterase activity was found to be associated with BCV (Vlasak et al., 1988).

In order to determine whether the latter activity could be attributed to a viral protein, experiments were directed at analyzing the specificity and catalytic mechanism of the enzyme. Since the influenza C virus acetylesterase activity was conveniently measured with the low-molecular-weight substrate p-nitrophenylacetate (Vlasak et al, 1987), we also used this assay for monitoring the BCV enzyme. Of several inhibitors tested, only DFP, a serine protease and esterase inhibitor, completely inhibited the BCV esterase.

TABLE 4.

PLAQUE FORMATION OF DFP-TREATED BCV PREPARATIONS [a]

Expt no. and virus	HA Titer [b]	% Esterase activity [c]	Titer (PFU/ml)
Expt 1			
BCV	2,048	100	2.6×10^7
BCV + DFP	2,048	<1	6.5×10^4
Expt 2			
BCV	128	100	2.5×10^6
BCV + DFP	128	<1	3.0×10^4
Expt 3			
BCV	8,192	100	5.4×10^7
BCV + DFP	8,192	<1	1.3×10^5
Expt1			
A/WSN/33	8	NA	1.5×10^7
A/WSN/33 + DFP	8	NA	1.3×10^7
Expt 3			
A/WSN/33	128	NA	6.2×10^7
A/WSN/33 + DFP	128	NA	5.6×10^7

[a] BCV and influenza A/WSN/33 virus were incubated with and without 1 mM DFP and purified over a 20 to 60% sucrose gradient (experiment 1 and 2). Alternatively, DFP-treated virus was used without further purification (experiment 3).

[b] Reciprocal of the highest diltuion of virus giving full hemagglutination after 60 min at 4^0C.

[c] Measured with 1mM p-nitrophenylacetate. NA, Not applicable.

Serine protease inhibitors PMSF and TPCK gave only partial inhibition, and no effect was detected by preincubation with EDTA, suggesting that divalent cations are not required for catalytic activity. Since the BCV esterase was quantitatively inhibited by DFP, it is suggested that the enzyme is a serine esterase. Similar findings were reported for other esterases (Cohen et al., 1967; Muchmore and Varki, 1987). By affinity labeling with the site-specific reagent DFP and analysis of the BCV protein on polyacrylamide gels, the E3 protein was identified as containing the viral esterase.

No difference in hemagglutination titers was observed between mock-treated and DFP-treated BCV in hemagglutination assays under standard conditions at 4°C. However, when the incubation temperature was raised, mock-treated BCV eluted form chicken erythrocytes, and only DFP-inactivated BCV retained its hemagglutination titer. These experiments suggest that inhibition of the receptor-destroying activity in BCV actually increases the stability of virus-receptor interactions. Similar results had been obtained earlier when the receptor-destroying activities of influenza viruses were inhibited (Bucher and Palese, 1975; Herrler et al., 1985).

To investigate the role of the BCV esterase during virus replication, plaque assay experiments were performed. Following inactivation of the esterase with DFP, infectious virus was titrated in tissue culture. Compared with mock-treated BCV, virus with an inactivated esterase had an approximately 100- to 400-fold reduced titer. In contrast, the infectivity of DFP-treated influenza A virus was not diminished. Thus, data obtained from these experiments could indicate a direct involvement of the acetylesterase in early phases of BCV replication. Since binding of BCV to sialic acid-containing receptors was not impaired by DFP inactivation of the esterase, it appears that the first step in viral replication, binding to cellular receptors, is independent of the presence or absence of a receptor-destroying activity. We thus speculate that an active esterase may be required for either endocytosis and/or uncoating of the virus with subsequent release of viral RNA

into the cytoplasm.

There are now three RNA virus families that include viruses which have been found to possess receptor-destroying activities (Fig. 13). Neuraminidases are associated with influenza A and B viruses (Air and Compans, 1983; Lamb, 1983) and with parainfluenza viruses (Scheid et al., 1972). Esterases have been shown to be part of influenza C virus (Formanowski and Meier-Ewert, 1988; Herrler et al., 1988; Herrler et al., 1985; Muchmore and Varki, 1987; Vlasak et al., 1987) and members of the coronavirus family, such as BCV (Vlasak et al., 1988). These receptor-destroying activities are either located in a separate protein (neuraminidases in influenza A and B viruses) or are part of a multifunctional protein (the HN and HE proteins of paramyxoviruses and influenza C virus). Finally, fusion activity may again be associated with a unique protein, as in the parainfluenza subgroup (Scheid and Choppin, 1974), or at the other extreme be part of the HE protein (Herrler et al., 1988; Ohuchi et al., 1982), which possesses several functions.

Taking into account that BCV and OC43 are antigenically closely related to mouse hepatitis virus (Hogue et al., 1984) yet appear to have an additional E3 surface glycoprotein, no biological function can be attributed to this protein by direct analogy. Since the receptor-binding and fusion activity in mouse hepatitis virus are located on the E2 protein (Luytjes et al., 1987; Sturman et al., 1985; Talbot et al., 1984; Wege et al., 1979), it might be hypothesized that the E2 proteins of BCV and OC43 have homologous activities. But what, then, is the precise function of the E3 protein? Earlier studies revealed that the E3 protein of BCV is the viral hemagglutinin (King et al., 1985). Thus, is the E3 protein an additional receptor-binding protein or is it the only receptor-binding protein, with the E2 protein alone possessing fusion activity? The present study suggests that the E3 protein of BCV has an acetylesterase activity necessary for virus replication. Thus, the interesting possibility arises that this protein has receptor-binding

Figure 13. SCHEMATIC REPRESENTATION OF THE FUNCTIONS OF THE SURFACE GLYCOPROTEINS OF ORTHOMYXOVIRUSES, PARAMYXOVIRUSES, AND CORONAVIRUSES.

Sendai virus and BCV are used as prototypes for paramyxoviruses and coronaviruses, respectively. HA, Hemagglutinin; NA, neuraminidase, HE, hemagglutinin-esterase; F, fusion protein; HN, hemagglutinin-neuraminidase; E2, peplomer protein; E3, coronavirus hemagglutinin-esterase. Boxes indicate proteins for which functions have been determined.

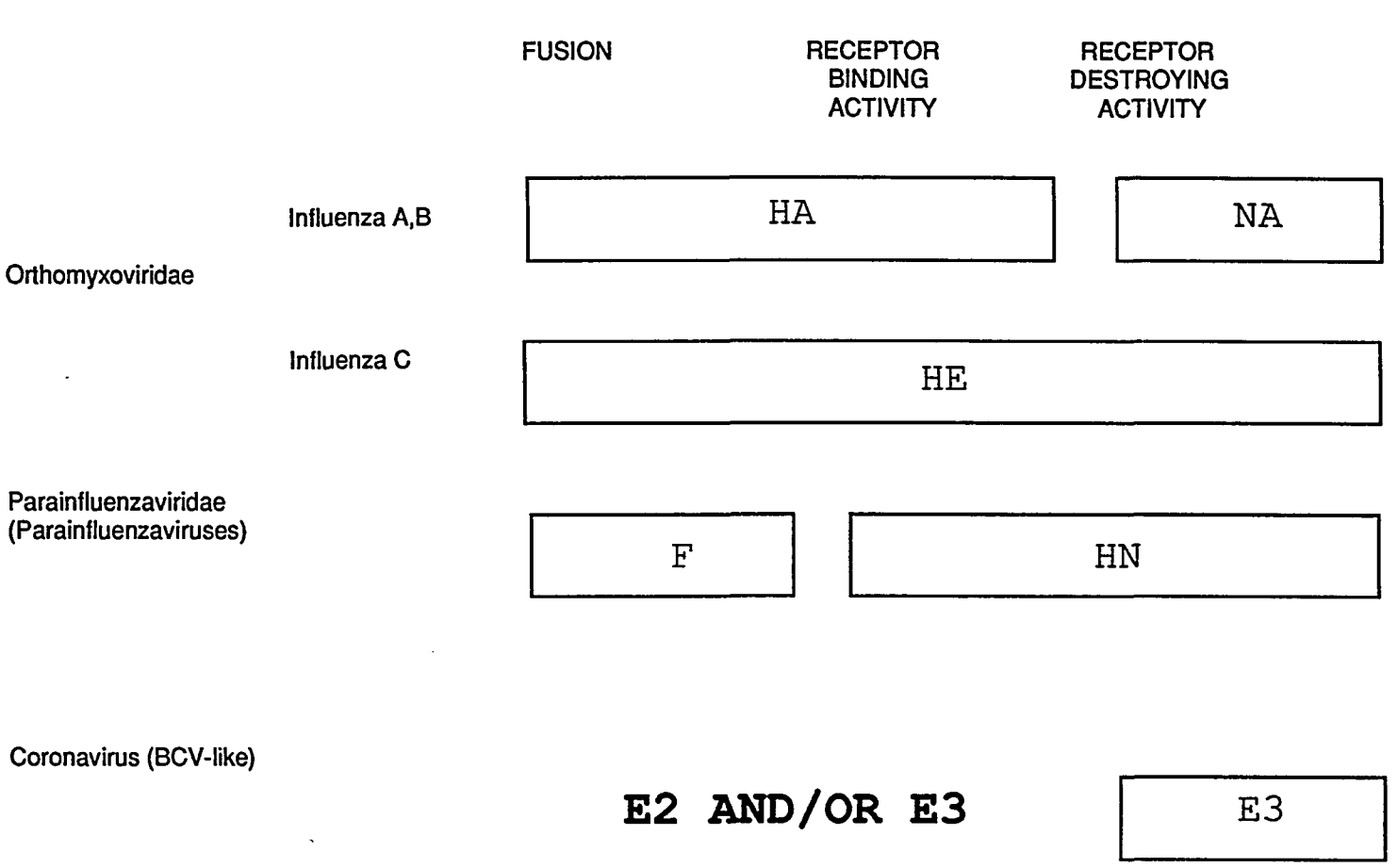

FUSION
RECEPTOR BINDING ACTIVITY
RECEPTOR DESTROYING ACTIVITY
Orthomyxoviridae
Influenza A,B
HA
NA
Influenza C
HE
Parainfluenzaviridae (Parainfluenzaviruses)
F
HN
Coronavirus (BCV-like)
E2 AND/OR E3
E3

FIGURE 13.

as well as receptor-destroying activities. Although a great deal has already been learned about the molecular biology of coronaviruses (Lai, 1986; Sturman and Holmes, 1983), identification of the precise roles of all proteins during coronavirus replication awaits further analysis. Additional experiments should also show whether all or only some coronaviruses possess receptor-binding/receptor-destroying activities and whether different strategies of attachment and uncoating prevail for different coronaviruses.

V. DISCUSSION

A. MUTATION RATES OF INFLUENZA A VIRUS AND POLIOVIRUS TYPE 1.

As described in chapter II, we used direct nucleotide sequence analysis to measure the mutation rates of the influenza A virus NS gene and of the VP1 gene of poliovirus type 1. Prior to our study, much evidence had accumulated suggesting that both of these viruses may have high mutation rates. The frequency of variants which had lost the recognition site for a neutralizing monoclonal antibody, as measured by the plaque reduction assay, was found to be comparable for influenza A virus and poliovirus type 1 (Emini et al., 1983; Lubeck et al., 1980). In the case of poliovirus, recent evidence suggesting a high mutation rate was a report that mutants of poliovirus resistant to guanidine occur at a frequency of between $10^{-6.7}$ and $10^{-7.4}$. Since a resistant phenotype requires two independent mutations for poliovirus, a nucleotide substitution frequency of between $10^{-3.1}$ and $10^{-3.4}$ was estimated for the gene encoding the nonstructural protein 2C (Pincus et al., 1986). Our study was the first to directly calculate the mutation rates of influenza A virus and poliovirus by using sequence analysis to measure the frequency of nucleotide substitutions in the genome of a population of virus derived from a single purified clone after only one plaque generation. We measured mutation rates of 1.5×10^{-5} and less than 2.1×10^{-6} mutations per nucleotide per infectious cycle for influenza A virus and poliovirus, respectively. The reasons for the poliovirus mutation rate being lower than that of influenza A virus and previous estimates may be due to greater restraints on VP1 than NS1 and/or differing polymerase error frequencies for the two viruses. Additionally, in reports using monoclonal antibodies, lack of knowledge about the genomic target size for the antibody-combining site, the number and location of mutations required for production of virus with an antibody-resistant phenotype, and/or complex passage histories of the virus during an unlimited number of virus replication

cycles may all have led to erroneous estimates of the poliovirus mutation rate. The estimate of poliovirus' nucleotide substitution frequency as 10^{-3} by Pincus et al. (1986) may have been misleading due to the multiple rounds of selection of the virus in the presence of guanidine that are required before the calculation of variant frequencies can be made.

Three noteworthy studies concerning the mutation rates of conventional (non-retrovirus) RNA viruses have been published since our work was done. Sedivy et al. (1987) calculated the mutation rate of poliovirus as 7.5×10^{-6} mutants per nucleotide per plaque generation by using an inducible mammalian amber suppressor system to monitor the reversion frequency of a mutant of poliovirus. This mutation rate is in close agreement with our result if we consider that poliovirus undergoes five replication cycles per plaque generation. Using purified poliovirus RNA polymerase, Ward et al. (1988) examined the fidelity of RNA replication by copying homopolymeric RNA templates in vitro. Depending on the reaction conditions, they estimated polymerase error frequencies of between 7×10^{-4} to 5.4×10^{-3} noncomplementary nucleotides incorporated. The reasons for the deviation by Ward et al. (1988) from the results obtained by us and Sedivy et al. (1987) may reflect the differences between in vitro and in vivo conditions such as the viral polymerase being more faithful during RNA replication in vivo than in vitro, the higher probability that most mutations arising are deleterious and therefore are selected against during viral replication, and/or the accuracy of the in vitro homopolymeric template they used to imitate polymerase-genome interactions in vivo. Recently, a nucleotide substitution frequency on the order of 10^{-3} to 10^{-4} was reported for vesicular stomatitis virus (VSV) by using oligonucleotide mapping with T1 ribonuclease to monitor for polymerase error frequencies at selected G residues in clonal pools of virus replicated in vivo (Steinhauer et al., 1989). However, there are several criticisms of their work, including the significant proportion (10 to 50%)

of supposedly error RNAs they detected which were actually due to incomplete RNase T1 digestion and the representativeness of the selected G residue they chose for examination as reflecting mutation frequencies at sites throughout the genome. Nonetheless, if we use their estimate of at least 8 to 10 replication cycles as having occurred between the replication of the original cloned virus and the production of the clonal pools they examined, a mutation rate estimate of approximately 10^{-5} mutations per nucleotide per cycle can be roughly estimated. In regards to these recent studies, our measurement of the mutation rate of influenza A virus indicates it is still among the most rapidly mutating conventional RNA viruses and that its mutation rate may play a critical role in driving its rate of evolution.

B. RETROVIRUS MUTATION RATE

As discussed in chapter III, the mutation rate of the retrovirus RSV was definitively determined as 1.4×10^{-4} mutations per nucleotide per cycle. Previous reports had indicated that the rates of evolution of retroviruses may be extremely high, on the order of 10^{-2} to 10^{-3} substitutions per site per year for the env gene of human immunodeficiency virus and 10-fold lower for the gag gene (Hahn et al. (1986)) and 1.31×10^{-3} per site per year for the v-mos gene of Moloney murine sarcoma virus (Gojobori and Yokoyama, 1985). However, a precise measurement for a single cycle of replication had not been done. Our results show that retroviruses are indeed among the most rapidly mutating viruses. By examining seven different regions of the genome, we avoided a possibly biased mutation rate estimate which could have been obtained by analyzing only a single genomic site. From studying several regions of the genome, we also gained preliminary evidence of relative mutation rate differences between genomic loci. These differences correlated with reports obtained by studying naturally occurring isolates. For example, we observed mutations in a region of the env gene known to

exhibit extensive variation among different strains of avian sarcoma and leukosis viruses (Bova et al., 1988). In the pol gene, however, we examined a region putatively coding for the highly conserved zinc-binding site of the retroviral endonuclease (Johnson et al., 1986) and did not observe any mutations.

Following our study, several other reports concerning retrovirus mutation rates have been published. Dougherty and Temin (1988) have used a spleen necrosis-based vector to calculate the mutation rate of a specific A-T to G-C substitution as 2×10^{-5} during a single round of replication, further evidence that retroviruses are extremely mutable. The fidelity of avian myeloblastosis virus and Moloney murine leukemia virus reverse transcriptases (RTs) during DNA synthesis in vitro has been examined by using the M13mplacZ-alpha gene as a mutational target and analyzing for substitutions, -1 frameshifts, and complex errors in the polymerization products. Both RTs are error prone, committing a misincorporation once every 30,000 nucleotides polymerized (Roberts et al., 1989a). The HIV-1 RT has also been examined in vitro and found to introduce base-substitution errors in DNA from the bacteriophage 0X174 amber-3 at estimated frequencies of 1/2000 to 1/4000 (Preston et al., 1989) and had an average error rate per detectable nucleotide incorporated of 1/1700 for a M13mp2-based fidelity assay (Roberts et al., 1989b). Certain template positions were also found to be mutational hotspots, with error rates as high as 1 per 70 polymerized nucleotides (Roberts et al., 1989b). Although caution must be exercised when analyzing in vitro error rates, the implications are that a considerable portion of the high mutation rates observed for retroviruses may be due to the error prone transcription by RT.

C. MECHANISM OF HIGH MUTATION RATES FOR RETROVIRUSES

Retroviruses are unique among RNA viruses in that they replicate their genome via DNA and RNA templates. During their life-cycle, retroviruses are found as both a

replicating, highly mutable RNA virus that uses enzymes of low fidelity for replication and a stably integrated provirus that is replicated infrequently by a faithful cellular DNA polymerase. Our results on the RSV mutation rate suggest that the source of retrovirus variation may be attributable to replication of the retrovirus genome by RNA polymerase II and/or RT. The in vitro error rate studies of RT provide additional evidence that this enzyme may be responsible for much of the observed retrovirus variation (Roberts et al., 1989a; Roberts et al., 1989b; Preston et al., 1989). A mechanism by which retroviruses readily mutate can be postulated. Enzymes of low fidelity replicate the retrovirus genome. The genome itself contains regions of high secondary structure and monotonous runs which lead to errors and slippage. Lax constraints for regions of encoded proteins permits enormous genomic plasticity. The result is production of a quasispecies of virus, a population in which there is a consensus sequence, but each infectious particle is genotypically subtly different (Eigen, 1971; Eigen and Schuster, 1977).

D. RECEPTOR-DESTROYING ENZYME OF BCV

As detailed in chapter IV, we identified the E3 protein of BCV as having receptor-destroying activity. This protein was shown to be an acetylesterase which inactivates O-acetylsialic acid containing receptors. Like influenza C virus, we found that the activity of the receptor-destroying enzyme is vital to viral replication as its inhibition leads to greatly reduced virus titers. We argue that an active receptor-destroying enzyme is needed in an early step(s) of virus infection, prior to transcription of the virus genome. Since hemagglutination of BCV was not impaired by inactivation of the esterase, we propose that the esterase functions at a step following binding of the virus to the cellular receptor, at the step(s) of endocytosis and/or uncoating of the virus.

Two recent reports have led us to speculate that the E3 of BCV is a serine

hydrolase in the same family of serine esterases as the hemagglutinin-esterase (HE) of the influenza C virus. Vlasak et al. (1989) recently identified the sequence of the catalytic active site of influenza C HE as Gly-Asp-Ser. The same consensus motif sequence is found in serine proteases, but differs from the Gly-Glu-Ser motif of other serine esterases, such as acetylcholinesterase. They proposed that the HE is a serine hydrolase constituting a new family of serine esterases. With the cloning and sequencing of the E3 protein of BCV, significant sequence similarity was found between the E3 polypeptide of BCV and the HE of influenza C virus (Parker et al., 1989). While the two proteins have significant similarity throughout, beginning with homologous amino acid sequences immediately after the putative signal sequence of E3 and ending just before the putative carboxy terminal anchoring domain of the E3, most striking is the presence of the Gly-Asp-Ser catalytic active sites in both. Following our characterization of the enzymatic activity of the E3, the indications suggest that the E3 protein of BCV is yet another serine hydrolase member of this new class of serine esterases.

SIGNIFICANCE

We have now determined the mutation rates of three RNA animal viruses. By sequence analysis, we demonstrated that the mutation rate of the NS gene of influenza A virus was significantly higher than that of the VP1 gene of poliovirus, 1.5×10^{-5} mutations per nucleotide per cycle versus $<2.1 \times 10^{-6}$ mutations per nucleotide per cycle, respectively. This difference in the mutation rates of the two viruses may reflect the success of vaccination against them; whereas vaccine strains against influenza A virus confer resistance to infection for only a few years, vaccines against poliovirus have been effective for several decades. For RSV, we were able to definitively determine the mutation rate of the virus as 1.4×10^{-4} mutations per nucleotide per single cycle of replication. Using the newly developed mutation detection technology of denaturing gradient gel analysis, we were able to rapidly screen 65,250 nucleotides distributed over seven different regions of the RSV genome. This result indicates that retroviruses are among the most highly mutable viruses yet observed. Based on our studies of the mutation rates of these three viruses, we now postulate that for some RNA viruses, the amount of variation seen in nature for a virus can be correlated with its mutation rate. For those viruses with relatively high mutation rates, such as influenza A viruses and retroviruses, this high variability could account for the diversity required for the generation of mutants that escape immune selection in a host previously immunized naturally or by prior exposure to vaccine strains. Thus, the hypervariability of some viruses may protect them from the immune surveillance system of their hosts.

Finally, we have identified the E3 protein of BCV as having receptor-destroying activity. We have found that the E3 is a serine hydrolase with activity similar to that of the HE of influenza C virus. Furthermore, as found for the HE of influenza C virus, inactivation of this enzyme with the serine protease/esterase inhibitor DFP leads to greatly reduced virus titers. Thus, our findings on the receptor-destroying enzyme of

BCV connotes that perhaps the most successful antiviral agents may be those developed against the less mutable regions of RNA viruses, such as the active sites of their enzymes.

BIBLIOGRAPHY

Air, G. M., and R. W. Compans. 1983. Influenza B and C viruses. In: Genetics of Influenza Viruses (P. Palese and D. W. Kingsbury, eds.) Springer-Verlag, Vienna, pp. 281-304.

Alexander, R., J. Leis, D. A. Soltis, R. M. Crowl, W. Danho, M. S. Poonian, Y.-C. E. Pan, and A. M. Skalka. 1987. Proteolytic processing of avian sarcoma and leukosis viruses pol-endo recombinant proteins reveals another pol gene domain. J. Virol. **61**:534-542.

Almoguera, C., D. Shibata, K. Forrester, J. Martin, N. Arnheim, and M. Perucho. 1988. Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. Cell **53**:549-554.

Anderson, K. B., and B. A. Nexo. 1983. Entry of murine retrovirus into mouse fibroblasts. Virology **125**:85-90.

Baker, B., H. L. Robinson, H. E. Varmus, and J. M. Bishop. 1981. Analysis of endogenous avian retrovirus DNA and RNA: viral and cellular determinants of retrovirus gene expression. Virology **144**:8-22.

Batschelet, E., E. Domingo, and C. Weissmann. 1976. The proportion of revertant and mutant phage in a growing population, as a function of mutation and growth rate. Gene **1**:27-32.

Beaton, A. R., and R. M. Krug. 1986. Transcription antitermination during influenza viral template RNA synthesis requires the nucleocapsid protein and the absence of a 5' capped end. Proc. Natl. Acad. Sci. U.S.A. **83**:6282-6286.

Bender, M. L., and F. J. Kezdy. 1965. Mechanism of action of proteolytic enzymes. Annu. Rev. Biochem. **34**:49-76.

Bernstein, H. D., N. Sonenberg, and D. Baltimore. 1985. Poliovirus mutant that does not selectively inhibit host cell protein synthesis. Mol. Cell. Biol. **5**:2913-2923.

Biggin, M. D., T. J. Gibson, and G. F. Hong. 1983. Buffer gradient gels and ^{35}S label as an aid to rapid DNA sequence determination. Proc. Natl. Acad. Sci. USA **80**:3963-3965.

Blondel, B., R. Crainic, O. Fichot, G. Dufraisse, A. Candrea, D. Diamond, M. Girard, and F. Horaud. 1986. Mutations conferring resistance to neutralization with monoclonal antibodies in type 1 poliovirus can be located outside or inside the antibody-binding site. J. Virol. **57**:81-90.

Bova, C. A., J. C. Olsen, and R. Swanstrom. 1988. The avian retrovirus env gene family: molecular analysis of host range and antigenic variants. J. Virol. **62**:75-83.

Braam, J., I. Ulmanen, and R. M. Krug. 1983. Molecular model of a eukaryotic transcription complex: Functions and movements of influenza P proteins during capped RNA-primed transcription. Cell **34**:609-618.

Brand, C., and P. Palese. 1980. Sequential passage of influenza virus in embryonated eggs or tissue culture: emergence of mutants. Virology **107**:424-433.

Bucher, D., and P. Palese. 1975. The biologically active proteins of influenza virus: neuraminidase. In: The Influenza Virus and Influenza (E. Kilbourne, ed.) Academic Press, Inc., New York, pp. 83-123.

Buonagurio, D. A., S. Nakada, J. D. Parvin, M. Krystal, P. Palese, and W. M. Fitch. 1986. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. Science **232**:980-982.

Callebant, P. E., and M. B. Pensaert. 1980. Characterization and isolation of structural polypeptides in hemagglutinating encephalomyelitis virus. J. Gen. Virol. **48**:193-204.

Chow, M., and D. Baltimore. 1982. Isolated poliovirus capsid protein VP1 induces a neutralizing response in rats. Proc. Natl. Acad. Sci. U.S.A. **79**:7518-7521.

Coffin, J. M. 1986. Genetic variation in AIDS viruses. Cell **40**:1-4.

Coffin, J. M., P. N. Tsichlis, C. S. Barker, and S. Voynow. 1980. Variation in avian retrovirus genomes. Ann. N.Y. Acad. Sci. **354**:410-425.

Cohen, J. A., R. A. Oosterbaan, and F. Berends. 1967. Organophosphorous compounds. Methods Enzymol. **11**:686-702.

Colman, P. M., W. G. Laver, J. N. Varghese, A. T. Baker, P. A. Tulloch, G. M. Air, and R. G. Webster. 1987. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. Nature **326**:358-363.

Compans, R. W., H.-D. Klenk, L. A. Caliguiri, and P. W. Choppin. 1970. Influenza virus proteins 1. Analysis of polypeptides of the virion and identification of spike glycoproteins. Virology **42**:880-889.

Cotton, R. G. H., N. D. Rodrigues, and R. D. Campbell. 1988. Reactivity of cytosine and thymine in single-base-pair mismatches with hydroxylamine and osmium tetroxide and its application to the study of mutants. Proc. Natl. Acad. Sci. U.S.A. **85**:4397-4401.

Dale, S., and H. Hanafusa. 1972. Penetration and intracellular release of the genomes of avian RNA tumor viruses. Virology **50**:440-458.

Darlix, J.-L., and P.-F. Spahr. 1983. High spontaneous mutation rate of Rous sarcoma virus demonstrated by direct sequencing of the RNA genome. Nucl. Acids Res. **11**:5953-5967.

Deregt, D., and L. A. Babiuk. 1987. Monoclonal antibodies to bovine coronavirus: characteristics and topographical mapping of neutralizing epitopes on the E2 and E3 glycoproteins. Virology **161**:410-420.

Deregt, D., M. Sabara, and L. A. Babiuk. 1987. Structural proteins of bovine coronavirus and their intracellular processing. J. Gen. Virol. **68**:2863-2877.

Domingo, E., D. Sabo, T. Taniguchi, and C. Weissmann. 1978. Nucleotide sequence heterogeneity of an RNA phage population. Cell **13**:735-744.

Dorner, A. J., and J. M. Coffin. 1986. Determinants for receptor interaction and cell killing on the avian retrovirus glycoprotein gp85. Cell **45**:365-374.

Dougherty, J. P. and H. M. Temin. 1986. High mutation rate of a spleen necrosis virus-based retrovirus vector. Mol. Cell. Biol. **6**:4387-4395.

Dougherty, J. P., and H. M. Temin. 1987. Determination of retroviral vector mutation rates and a promoterless retroviral vector. In: Gene Transfer Vectors for Mammalian Cells. (J. J. Miller and M. P. Calos, eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, pp. 18-23.

Dougherty, J. P., and H. M. Temin. 1988. Determination of the rate of base-pair substitution and insertion mutations in retrovirus replication. J. Virol. **62**:2817-2822.

Drake, J. W. 1969. Comparative rates of spontaneous mutation. Nature (London) **221**:1132.

Dubois-Dalcq, M. E., E. W. Doller, M. V. Haspel, and K. I. Holmes. 1982. Cell tropism and expression of mouse hepatitis virus (MHV) in mouse spinal cord cultures. Virology **119**:317-331.

Durbin, R. K., and V. Stollar. 1986. Sequence analysis of the E2 gene of a hyperglycosylated, host restricted mutant of Sindbis virus and estimation of mutation rate from frequency of revertants. Virology **154**:135-143.

Eigen, M. 1971. Selforganization of matter and the evolution of biological macromolecules. Naturwissenschaften **58**:465-523.

Eigen, M. and Schuster, P. 1977. The hypercycle, a principle of natural self-organization, Part A: Emergence of the hypercycle. Naturwissenschaften **64**:541-565.

Emini, E. A., S.-Y. Kao, A. J. Lewis, R. Crainic, and E. Wimmer. 1983. Functional basis of poliovirus neutralizing determined with monospecific neutralizing antibodies. J. Virol. **46**:466-474.

Fischer, S. G., and L. S. Lerman. 1979. Two-dimensional electrophoretic separation of restriction fragments of DNA. Methods Enzymol. **68**:183-191.

Fisher, A. G., B. Ensoli, D. Looney, A. Rose, R. C. Gallo, M. S. Saag, G. M. Shaw, B. H. Hahn, and F. Wong-Staal. 1988. Biologically diverse molecular variants within a single HIV-1 isolate. Science **334**:444-447.

Formanowski, F., and H. Meier-Ewert. 1988. Isolation of the influenza C virus glycoprotein in a soluble form by bromelain digestion. Virus Res. **10**:177-192.

Gojobori, T., and S. Yokoyama. 1985. Rates of evolution of the retroviral oncogene of Moloney murine sarcoma virus and its cellular homologues. Proc. Natl. Acad. Sci. USA **82**:4198-4201.

Gojobori, T., and S. Yokoyama. 1987. Molecular evolutionary rates of oncogenes. J. Mol. Evol. **26:**148-156.

Gopinathan, K. P., L. A. Weymouth, T. A. Kunkel, and L. A. Loeb. 1979. Mutagenesis in vitro by DNA polymerase from an RNA tumor virus. Nature (London) **278:**857-859.

Gottschalk, A. 1957. The specific enzyme of influenza virus and Vibrio cholerae. Biochim. Biophys. Acta **23:**645-646.

Hahn, B. H., G. M. Shaw, M. E. Taylor, R. R. Redfield, P. D. Markham, S. Z. Salahuddin, F. Wong-Staal, R. C. Gallo, E. S. Parks, and W. P. Parks. 1986. Genetic variation in HTLV-III/LAV over time in patients with AIDS or at risk for AIDS. Science **232:**1548-1553.

Hall, J. D., D. M. Coen, B. L. Fisher, M. Weisslitz, S. Randall, R. E. Almy, P. T. Gelep, and P. A. Schaffer. 1984. Generation of genetic diversity in herpes simplex virus: an antimutator phenotype maps to the DNA polymerase locus. Virology **132:**26-37.

Hanafusa, H. 1969. Rapid transformation of cells by Rous sarcoma virus. Proc. Natl. Acad. Sci. USA **63:**318-325.

Hay, A. J., B. Lomniczi, A. R. Bellamy, and J. J. Skehel. 1977. Transcription of the influenza virus genome. Virology **83:**337-355.

Hay, A. J., A. J. Wolstenholme, J. J. Skehel, and M. H. Smith. 1985. The molecular basis of the specific anti-influenza action of amantadine. EMBO J. **4:**3021-3024.

Hayashida, H., H. Toh, R. Kikuno, and T. Miyata. 1985. Evolution of influenza virus genes. Mol. Biol. Evol. **2:**289-303.

Hayward, W. S. 1977. Size and genetic content of viral RNAs in avian oncovirus-infected cells. J. Virol. **24:**47-63.

Hayward, W. S., S. B. Braverman, and S. M. Astrin. 1980. Transcriptional products and DNA structure of endogenous avian proviruses. Cold Spring Harbor Symp. Quant. Biol. **44:**1111-1121.

Herrler, G., G. Multhaup, K. Beyreuther, and H.-D. Klenk. 1989. Serine 71 of the glycoprotein HEF is located at the active site of the acetylesterase of influenza C virus. Arch. Virol. **102:**269-274.

Herrler, G., R. Rott, H.-D. Klenk, H.-P. Muller, A. K. Shukla, and R. Schauer. 1985. The receptor destroying enzyme of influenza C virus is neuraminate-O-acetylesterase. EMBO J. **4:**1503-1506.

Herrler, G., I. Durkip, H. Becht, and H.-D. Klenk. 1988. The glycoprotein of influenza C virus is the haemagglutinin, esterase and fusion factor. J. Gen. Virol. **69:**839-846.

Ho, D., R. J. Pomerantz, and J. C. Kaplan. 1987. Pathogenesis of infection with human immunodeficiency virus. N. Engl. J. Med. **317:**278-286.

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

Hogue, B. G., and D. A. Brian. 1986. Structural proteins of human respiratory coronavirus OC43. Virus Res. **5:** 131-144.

Hogue, B. G., B. King, and D. A. Brian. 1984. Antigenic relationship among proteins of bovine coronavirus, human respiratory coronavirus OC43, and mouse hepatitis coronavirus A59. J. Virol. **51:**384-388.

Holland, J., K. Spindler, F. Horodyski, E. Grabau, S. Nichol, and S. VandePol. 1982. Rapid evolution of RNA genomes. Science **215:**1577-1585.

Holland, T. C., S. D. Marlin, M. Levine, and J. Glorioso. 1983. Antigenic variants of herpes simplex virus selected with glycoprotein-specific monoclonal antibodies. J. Virol. **45:**672-682.

Homma, M., and M. Ohuchi. 1973. Trypsin action on the growth of Sendai virus in tissue culture cells. III. Structural difference of Sendai virus grown in eggs and tissue culture cells. J. Virol. **12:**1457-1465.

Johnson, M. S., M. A. McClure, C.-F. Feng, J. Gray, and R. F. Doolittle. 1986. Computer analysis of retroviral pol genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. Proc. Natl. Acad. Sci. USA **83:**7648-7652.

Jolly, D. J., R. C. Willis, and T. Friedmann. 1986. Variable stability of a selectable provirus after retroviral vector gene transfer into human cells. Mol. Cell. Biol. **6:**1141-1147.

Kato, A., K. Mixumoto, and A. Ishihama. 1985. Purification and enzymatic properties of an RNA-polymerase-RNA complex from influenza virus. Virus Res. **3:**115-127.

Katz, R. A., C. A. Omer, J. H. Weis, S. A. Mitsialis, A. J. Faras, and R. V. Guntaka. 1982. Restriction endonuclease and nucleotide and sequence analysis of molecularly cloned unintegrated avian tumor virus DNA: structure of large terminal repeats in circle junctions. J.Virol. **42:**346-351.

Kawakami, K., and A. Ishihama. 1983. RNA polymerase of influenza virus: III. Isolation of RNA polymerase-RNA complexes form influenza virus PR8. J. Biochem. **93:**989-996.

Khoury, A. T., and H. Hanafusa. 1976. Synthesis and integration of viral DNA in chicken cells at different times after infection with various multiplicities of avian oncornavirus. J. Virol. **18:**383-400.

King, B., and D. A. Brian. 1982. Bovine coronavirus structural proteins. J. Virol. **42:**700-707.

King, B., B. J. Potts, and D. A. Brian. 1985. Bovine coronavirus hemagglutinin protein. Virus Res. **2:**53-59.

Kitamura, N., B. L. Semler, P. G. Rothberg, G. R. Larsen, C. J. Adler, A. J. Dorner,

E. A. Emini, R. Hanecaf, J. J. Lee, S. van der Werf, C. W. Anderson, and E. Wimmer. 1981. Primary structure gene organization and polypeptide expression of poliovirus RNA. Nature (London) **291**:547-553.

Klenk, H.-D., and P. W. Choppin. 1970a. Plasma membrane lipids and parainfluenza virus assembly. Virology **40**:939-947.

Klenk, H.-D., and P. W. Choppin. 1970b. Glycosphingolipids of plasma membranes of cultured cells and an enveloped virus (SV5) grown in these cells. Proc. Natl. Acad. Sci. U.S.A. **66**:57-64.

Koch, R. E., and J. W. Drake. 1973. Ligase-defective bacteriophage T4. I. Effects on mutation rates. J. Virol. **11**:35-40.

Koenig, S., V. M. Hirsch, R. A. Olmsted, D. Powell, W. Maury, A. Rabson, A. S. Fauci, R. H. Purcell, and P. R. Johnson. 1989. Selective infection of human CD4$^+$ cells by simian immunodeficiency virus: Productive infection associated with envelope glycoprotein-induced fusion. Proc. Natl. Acad. Sci. U.S.A. **86**:2443-2447.

Krystal, M., D. Buonagurio, J. F. Young, and P. Palese. 1983. Sequential mutations in the NS genes of influenza virus field strains. J. Virol. **45**:547-554.

Lai, M. M. 1986. Replication of coronavirus RNA. In: RNA Genetics, vol. 1:RNA-Directed Virus Replication (J. J. Holland, P. Ahlquist, and E. Domingo eds.) CRC Press, Inc., Boca Raton, Fla., pp. 115-136.

Lai, M. M. C., C. D. Patton, and S. A. Stohlman. 1982. Replication of mouse hepatitis virus: negative-stranded RNA and replicative form RNA are of genome length. J. Virol. **44**:487-492.

Lamb, R. A. 1983. The influenza virus RNA segments and their encoded proteins. In: Genetics of Influenza Viruses (P. Palese and D. W. Kingsbury eds.), Springer-Verlag, Vienna, pp. 21-69.

Lamb, R. A., and P. W. Choppin. 1983. Gene structure and replication of influenza virus. Ann. Rev. Biochem. **52**:467-506.

Lamb, R. A., S. L. Zebedee, and C. D. Richardson. 1985. Influenza virus M_2 protein is an integral membrane protein expressed on the infected-cell surface. Cell **40**:627-633.

Latham, T., and F. I. Smith. 1989. Detection of single-base mutations in DNA molecules using the solution melting method. DNA **8**:223-231.

Lenard, J., and K. D. Miller. 1982. Uncoating of enveloped viruses. Cell **28**:5-6.

Lerman, L. S., S. G. Fischer, I. Hurley, K. Silverstein, and N. Lumelsky. 1984. Sequence-determined DNA separation. Annu. Rev. Biophys. Bioeng. **13**:339-423.

Linial, M., and D. Blair. 1982. Genetics of retroviruses. In: RNA Tumor Viruses (R. Weiss, N. Teich, H. Varmus, and J. Coffin eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y, pp. 649-783.

Loeb, L. A., and T. A. Kunkel. 1982. Fidelity of DNA synthesis. Annu. Rev. Biochem. **52:**429-457.

Lubeck, M. D., J. L. Schulman, and P. Palese. 1980. Antigenic variants of influenza viruses: marked differences in the frequencies of variants selected with different monoclonal antibodies. Virology **102:**458-462.

Luytjes, W., L. S. Sturman, P. J. Bredenbeek, J. Charite, B. A. M. van der Zeijst, M. C. Horzinek, and W. J. M. Spaan. 1987. Primary structure of the glycoprotein E2 of coronavirus MHV A59 and identification of the trypsin cleavage site. Virology **161:**479-487.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Melton, D.A., P.A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, M. R. Green. 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12:**7035-7056.

Miyamoto, J., and R. V. Gilden. 1971. Electron microscopic studies of tumor viruses. J. Virol. **7:**395-406.

Muchmore, E. A., and A. Varki. 1987. Selective inactivation of influenza C virus esterase: a probe for detecting 9-O-acetylated sialic acids. Science **236:**1293-1295.

Myers, R. M., Z. Larin, and T. Maniatis. 1985. Detection of single base substitutions by ribonuclease cleavage at mismatches in RNA:DNA duplexes. Science **230:**1242-1246.

Nagai, Y., H.-D. Klenk, and R. Rott. 1976. Proteolytic cleavage of the viral glycoproteins and its significance for the virulence of NDV by proteolytic cleavage. Virology **77:**125-134.

Nakada, S., R. S. Creager, M. Krystal, R. P. Aaronson, and P. Palese. 1984. Influenza C virus hemagglutinin: Comparison with influenza A and B virus hemagglutinins. J. Virol. **50:**118-124.

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.

Nottay, B. K., O. M. Kew, M. H. Hatch, J. T. Heyward, and J. F. Obijeski. 1981. Molecular variation of type 1 vaccine-related and wild polioviruses during replication in humans. Virology **108:**405-423.

Ohuchi, M., R. Ohuchi, and K. Mifune. 1982. Demonstration of hemolytic and fusion activities of influenza C virus. J. Virol. **42:**1076-1079.

Palese, P. 1986. Rapid evolution of human influenza viruses. In: Evolutionary processes and theory. (S. Karlin and E. Novo, eds.) Academic Press, Inc., New York, pp. 53-68.

Palese, P., and J. L. Schulman. 1974. Isolation and characterization of influenza virus recombinants with high and low neuraminidase activity: use of 2-(3'-methoxyphenyl)-N-

acetylneuraminic acid to identify cloned populations. Virology **57**:227-237.

Palese, P., and J. L. Schulman. 1976. Mapping of the influenza virus genome: identification of the hemagglutinin and the neuraminidase genes. Proc. Natl. Acad. Sci. USA **73**:2142-2146.

Palese, P., K. Tobita, M. Ueda, and R. W. Compans. 1974. Characterization of temperature sensitive influenza A virus mutants defective in neuraminidase. Virology **61**:397-410.

Panganiban, A., and H. M. Temin. 1984. Circles with two tandem LTRs are precursors to integrated retrovirus DNA. Cell **36**:673-679.

Parker, M. D., G. J. Cox, D. Deregt, D. R. Fitzpatrick, and L. A. Babiuk. 1989. Cloning and in vitro expression of the gene for the E3 hemagglutinin glycoprotein of bovine coronavirus. J. Gen. Virol. **70**:155-164.

Parvin, J. D., A. Moscona, W. T. Pan, J. M. Leider, and P. Palese. 1986. Measurement of the mutation rates of animal viruses: influenza A virus and poliovirus type 1. J. Virol. **59**:377-383.

Parvin, J. D., and L.-H. Wang. 1984. Mechanism for the generation of src-deletion mutants and recovered sarcoma viruses: identification of viral sequences involved in src deletions and in recombination with c-src sequences. Virology **138**:236-245.

Payne, H. R., and J. Storz. 1988. Analysis of cell fusion induced by bovine coronavirus infection. Arch. Virol. **103**:27-33.

Pfeifer, J. B., and R. W. Compans. 1984. Structure of the influenza C glycoprotein gene as determined from cloned DNA. Virus Res. **1**:281-296.

Pincus, S. E., Diamond, D. C., Emini, E. A., and Wimmer, E. 1986. Guanidine-selected mutants of poliovirus: mapping of point mutations to polypeptide 2C. J. Virol. **57**:638-646.

Plotch, S. J., M. Bouloy, I. Ulmanen, and R. M. Krug. 1981. A unique cap (m^7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate primer that initiate viral RNA transcription. Cell **23**:847-858.

Portner, A., R. G. Webster, and W. J. Bean. 1980. Similar frequencies of antigenic variants in Sendai, vesicular stomatitis and influenza A viruses. Virology **104**:235-238.

Potter, H., L. Wier, and P. Leder. 1984. Enhancer-dependent expression of human k immunoglobulin genes introduced into mouse pre-B lymphocytes by electroporation. Proc. Natl. Acad. Sci. USA **81**:7161-7165.

Preston, B. D., B. J. Poiesz, and L. A. Loeb. 1989. Fidelity of HIV-1 reverse transcriptase. Science **242**:1168-1171.

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

Raymond, F. L., A. J. Caton, N. J. Cox, A. P. Kendal, and G. G. Brownlee. 1986. The antigenicity and evolution of influenza H1 haemagglutinin, from 1950-57 and 1977-1983: Two pathways from one gene. Virology **148:**275-287.

Roberts, J. D., K. Bebenek, and T. A. Kunkel. 1989a. The accuracy of reverse transcriptase form HIV-1. Science **242:**1171-1173.

Roberts, J. D., B. D. Preston, L. A. Johnston, A. Soni, L. A. Loeb, and T. A. Kunkel. 1989b. Fidelity of two retroviral reverse transcriptases during DNA-dependent DNA synthesis in vitro. Mol. Cell. Biol. **6:**469-476.

Robertson, J. S., M. Schubert, and R. A. Lazzarini. 1981. Polyadenylation sites for influenza virus mRNA. J. Virol. **38:**157-163.

Rogers, G. N., G. Herrler, J. C. Paulson, and H.-D. Klenk. 1986. Influenza C virus uses 9-O-acetyl-N-neuraminic acid as high affinity receptor determinant for attachment to cells. J. Biol. Chem. **261:**5947-5951.

Rossmann, M. 1988. Viral receptors and drug design. Nature (London) **333:**392-393.

Saag, M. S., B. H. Hahn, J. Gibbons, Y. Li, E. S. Parks, W. P. Parks, and G. M. Shaw. 1988. Extensive variation of human immunodeficiency virus type-1 in vivo. Nature **334:**440-444.

Saitou, N., and M. Nei. 1986. Polymorphism and evolution of influenza A virus genes. Mol. Biol. Evol. **3:**57-74.

Salts, Y. and A. Rosen. 1971. Neighbor effects in the mutation of ochre triplets in the T4 rII gene. Mutat. Res. **13:**109-113.

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

Sawicki, S. G., and D. L. Sawicki. 1986. Coronavirus minus-strand RNA synthesis and effect of cycloheximide on coronavirus RNA synthesis. J. Virol. **57:**328-334.

Scheid, A., L. A. Caliguiri, R. W. Compans, and P. W. Choppin. 1972. Isolation of paramyxovirus glycoproteins. Association of both hemagglutinating and neuraminidase activities with larger SV5 glycoprotein. Virology **50:**640-651.

Scheid, A., M. C. Graves, S. M. Silver, and P. W. Choppin. 1978. Studies on the structure and function of paramyxovirus glycoproteins. In: Negative Strand Viruses and the Host Cell (B. W. J. Mahy and R. D. Barry, eds.) Academic Press, New York, pp. 181-193.

Scheid, A., and P. W. Choppin. 1974a. Identification of biological activities of paramyxoviruses glycoproteins: activation of cell fusion, hemolysis, and infectivity by proteolytic cleavage of an inactive precursor protein of Sendai virus. Virology **57:**475-490.

Scheid, A., and P. W. Choppin. 1974b. The hemagglutinating and neuraminidase

protein of a paramyxovirus: interaction with neuraminic acid in affinity chromatography. Virology **62:**125-133.

Schwartz, D. E., R. Tizard, and W. Gilbert. 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32:**853-869.

Sedivy, J. M., J. P. Capone, U. L RajBhandary, and P. A. Sharp. 1987. An inducible mammalian amber suppressor: propagation of a poliovirus mutant. Cell **50:**379-389.

Sheffield, V. C., D. R. Cox, L. S. Lerman, and R. M. Myers. 1989. Attachment of a 40-base-pair G+C-rich sequence (GC-clamp) to genomic DNA fragments by the polymerase chain reaction results in improved detection of single-base changes. Proc. Natl. Acad. Sci. U.S.A. **86:**232-236.

Smith, D. B., and S. C. Inglis. 1987. The mutation rate and variability of eukaryotic viruses: an analytical review. J. Gen. Virol. **68:**2729-2740.

Smith, F. I., T. E. Latham, J. A. Ferrier, and P. Palese. 1988. Novel method of detecting single base substitutions in RNA molecules by differential melting behaviour in solution. Genomics **3:**2-7-223.

Smith, F. I., J. D. Parvin, and P. Palese. 1986. Detection of single base substitutions in influenza virus RNA molecules by denaturing gradient gel electrophoresis of RNA-RNA or DNA-RNA heteroduplexes. Virology **150:**55-64.

Sobrino, F., M. Davila, J. Ortin, and E. Domingo. 1983. Multiple genetic variants arise in the course of replication of foot-and-mouth disease virus in cell culture. Virology **128:**310-318.

Spaan, W. J. M., P. J. M. Rottier, M. C. Horzinek, and B. A. M. van der Zeijst. 1981. Isolation and identification of virus specific mRNAs in cells infected with mouse hepatitis virus (MHV-A59). Virology **108:**424-434.

Spindler, K. R., F. M. Horodyski, and J. J. Holland. 1982. High multiplicities of infection favor rapid and random evolution of vesicular stomatitis virus. Virology **119:**96-108.

Steinhauer, D. A., and J. J. Holland. 1986. Direct method for quantitation of extreme polymerase error frequencies at selected base sites in viral RNA. J. Virol. **57:**219-228.

Steinhauer, D. A., and J. J. Holland. 1987. Rapid evolution of RNA viruses. Ann. Rev. Microbiol. **41:**409-433.

Steinhauer, D. A., J. C. de la Torre, and J. J. Holland. 1989. High nucleotide substitution error frequencies in clonal pool of vesicular stomatitis virus. J. Virol. **63:**2063-2071.

Stoltzfus, M. C. 1988. Synthesis and processing of avian sarcoma retrovirus RNA. Adv. Virus Res. **35:**1-38.

Sturman, L. S., and K. V. Holmes. 1983. The molecular biology of coronaviruses. Adv.

Virus Res. **28**:35-112.

Sturman, L. S., C. S. Ricard, and K. V. Holmes. 1985. Proteolytic cleavage of the E2 glycoprotein of murine coronavirus: activation of cell fusing activity of virions by trypsin and separation of two different 90K cleavage fragments. J. Virol. **56**:904-911.

Szewczyk, B. W., G. Laver, and D. F. Summers. 1988. Purification, thioredoxin renaturation, and reconstituted activity of the three subunits of the influenza A virus RNA polymerase. Proc. Natl. Acad. Sci. U.S.A. **85**:7907-7911.

Talbot, P. J., A. A. Salmi, R. L. Knobler, and M. J. Buchmeier. 1984. Topographical mapping of epitopes on the glycoproteins of murine hepatitis virus-4 (strain JHM): correlation with biological activities. Virology **132**:250-260.

Ulmanen, I., B. A. Broni, and R. M. Krug. 1981. The role of two of the influenza virus core P proteins in recognizing cap structures (m^7GpppNm) on RNAs and in initiating viral RNA transcription. Proc. Natl. Acad. Sci. USA **78**:7355-7359.

Ulmanen, I., B. A. Broni, and R. M. Krug. 1983. Influenza virus temperature-sensitive cap (m^7GpppNm)-dependent endonuclease. J. Virol. **45**:27-35.

Varghese, J. N., W. G. Laver, and P. M. Colman. 1983. Structure of the influenza virus glycoprotein antigen: neuraminidase at 2.9A resolution. Nature **303**:35-40.

Vlasak, R., M. Krystal, M. Nacht, and P. Palese. 1987. The influenza C virus glycoprotein (HE) exhibits receptor binding (hemagglutinin) and receptor destroying (esterase) activities. Virology **160**:419-425.

Vlasak, R., W. Luytjes, W. Spaan and P. Palese. 1988. Human and bovine coronaviruses recognize sialic acid containing receptors similar to those of influenza C viruses. Proc. Natl. Acad. Sci. USA **85**:4526-4529.

Vlasak, R., T. Muster, A. M. Lauro, J. C. Powers, and P. Palese. 1989. Influenza C virus esterase: Analysis of catalytic site, inhibition, and possible function. J. Virol. **63**:2056-2062.

Wang, S. Y., W. S. Hayward, and H. Hanafusa. 1977. Genetic variation in the RNA transcripts of endogenous virus genes in uninfected chicken cells. J. Virol. **24**:64-73.

Ward, C. D., M. A. M. Stokes, and J. B. Flanegan 1988. Direct measurement of the poliovirus RNA polymerase error frequency in vitro. J. Virol. **62**:558-562.

Webster, R. G., W. G. Laver, G. M. Air, and G. C. Schild. 1982. Molecular mechanisms of variation in influenza viruses. Nature (London) **296**:115-121.

Wege, H., H. Wege, K. Nagashima, and V. ter Meulen. 1979. Structural polypeptides of the murine coronavirus JHM. J. Gen. Virol. **42**:37-47.

Weis, W., J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel, and D. C. Wiley. 1988. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature **333**:426-431.

Weiss, R. A., W. S. Mason, and P. K. Vogt. 1973. Genetic recombinants and heterozygotes derived from endogenous and exogenous avian RNA tumor viruses. Virology **52**:535-552.

White, J., J. Kartenbeck, and A. Helenius. 1982. Membrane fusion activity of influenza virus. EMBO J. **1**:217-222.

Wiley, D. C., and J. J. Skehel. 1987. The structure and function of hemagglutinin membrane of influenza virus. Annu. Rev. Biochem. **56**:365-394.

Wiley, D. C., I. A. Wilson, and J. J. Skehel. 1981. Structural identification of the antibody-binding sites of Hong Kong influenza haemagglutinin and their involvement in antigenic variation. Nature **289**:373-378.

Wilson, I. A., J. J. Skehel, and D. C. Wiley. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3A resolution. Nature (London) **289**:366-373.

Winters, E. F. Yamamoto, C. Almoguera, and M. Perucho. 1985. A method to detect and characterize point mutations in transcribed genes: amplification and overexpression of the mutant c-Ki-ras allele in human tumor cells. Proc. Natl. Acad. Sci. U.S.A. **82**:7575-7579.

World Health Organization. 1985. New approaches to vaccine development: memorandum from a WHO meeting. Bull. WHO **63**:479-484.

Zarling, D. A., and H. M. Temin. 1976. High spontaneous mutation rate of an avian sarcoma virus. J. Virol. **17**:74-84.

Zebedee, S. L., and R. A. Lamb. 1988. Influenza A virus M_2 protein: Monoclonal antibody restriction of virus growth and detection of M_2 in virions. J. Virol. **62**:2762-2722.

Zvonarjev, A. Y., and Y. Z. Ghendon. 1980. Influence of a membrane (M) protein on influenza A virion transcriptase in vitro and its susceptibility to rimantadine. J. Virol. **33**:583-586.