CHAPTER 1

VIRUSES: AN OVERVIEW

MILTON W. TAYLOR and HOWARD V. HERSHEY
Program in Microbiology and Genetics, Department of Biology
Indiana University, Bloomington, IN 47405, U.S.A

INTRODUCTION

It is very difficult to be the opening author of a treatise of this type. It is even more difficult to write a chapter entitled "Viruses — An Overview." It is like a lecture on *"Animals — A Brief Summary"*.

The diversity in the animal kingdom — from sponges to insects to the primates, we, perhaps erroneously, call "sapiens" — make a brief summary impossible. The degree of diversity and variability within the viral "kingdom" which co-evolved with and infects all the other kingdoms and which is itself highly variable in morphology, basic chemistry, and molecular biology make it an extremely difficult topic to discuss in a short summary paper. So we will look for some commonalities among viruses by asking a few simple questions about virus structure, function, and evolution.

Perhaps the best way to start is to ask "what is a virus?" Early attempts to define a virus have been fraught with contradictions and confusion. This, of course, reflects the history of infectious diseases and their relationship to viruses. As "new" diseases were diagnosed, clinicians attempted to relate them to infectious agents, the majority of which were thought to be of bacterial origin. Microbiologists in the 1940-50's tried to find a common evolutionary pathway from the then recognized viruses to bacteria, hoping that viruses would prove to be the missing link in evolution at the procaryotic end of the scale. There have also been historical controversies as to whether viruses are "inert" molecules, living organisms, or autocatalytic proteins. None of these controversies seem relevant today. To repeat Lwoff's famous saying (1) "Viruses should be considered as viruses because viruses are viruses".

A more functional definition is the one presented by Luria and Darnell in their textbook, General Virology, in 1967 (2). They define viruses as "entities whose genome is an element of nucleic acid either DNA or RNA, which reproduce inside

R. P. Bercoff (ed.), *The Molecular Basis of Viral Replication*
© Plenum Press, New York 1987

living cells and use their synthetic machinery to direct the synthesis of specialized particles, the virion, which contain the viral genome and transfer it to other cells". Thus, viruses are here defined by virtue of their obligate intracellular parasitism at the genetic level. Although this definition is now 20 years old, it is still generally acceptable. More recently, S. Harrison described a virus particle as "a structure for transferring nucleic acid from one cell to another", adding that "the nucleic acid may be either RNA or DNA and, in both cases particles of varying complexity are found. Observed structures reflect requirements for efficient and accurate assembly, for exit and re-entry, and for correctly localized disassembly" (55).

The concept of the virus as discussed by Lwoff in his 1957 paper (1), and the definitions above emphasize three characteristics of the virus particle: i) its infectivity, i.e., the ability to be transferred from cell to cell, ii) ability to exist in a non-cellular state, and iii) the obligate parasitism at the genetic level.

1. HISTORICAL BACKGROUND

Although the published history of virology begins with Jenner's experimental reports on vaccinia virus, (3) we know that the observation that transfer of pus from a lesion of an smallpox infected individual to a non-infected individual could result in immunity (variolation) was recognized among people in the Far and Middle East who suffered from periodic outbreaks of the disease centuries before Jenner's time. Of course neither they nor Jenner knew the nature of the causative agent of smallpox although Jenner does refer to it as a virus. In fact, the transfer of infectious smallpox material from one individual to another was introduced into England from Turkey a long time before Jenner's experiments. However, Jenner noted an inverse correlation between the severity of smallpox and exposure to cowpox (3). Jenner's paper makes interesting reading and is well documented. It is also noteworthy that Jenner's first attempts at publication were rejected ! The first vaccine used on a worldwide scale to eradicate a human viral diease — smallpox — contained live vaccinia virus developed by Jenner almost 200 years ago.

Continued research on immunization against disease has resulted in vaccines against a large number of other human viral pathogens, including such major diseases as polio, yellow fever, measles, mumps, and rubella, as well as many economically important viral pathogens which infect domestic animals. This list will surely become larger as more viral antigens are isolated by recombinant DNA technology, e.g., rabies and hepatitis B vaccines.

By the late 1800's, Koch and Pasteur had established the germ theory, which attributed disease to bacterial-like organisms. Such organisms were retained by the porcelain filters used at that time. However in 1892, Ivanovski comunicated to the Imperial Academy of Sciences of St. Petersbourg that the causative agent of the tobacco mosaic disease was filterable (Fig. 1). He proposed that the agent of the disease was some type of filterable toxin or a small microbe, and it was not until a few years later that Beijerinck proposed a living (reproducing) organism as the causative agent of tobacco mosaic disease, an organism smaller than all known bacteria. Beijerinck (4), who was unaware of Ivanovski's work, proposed that the

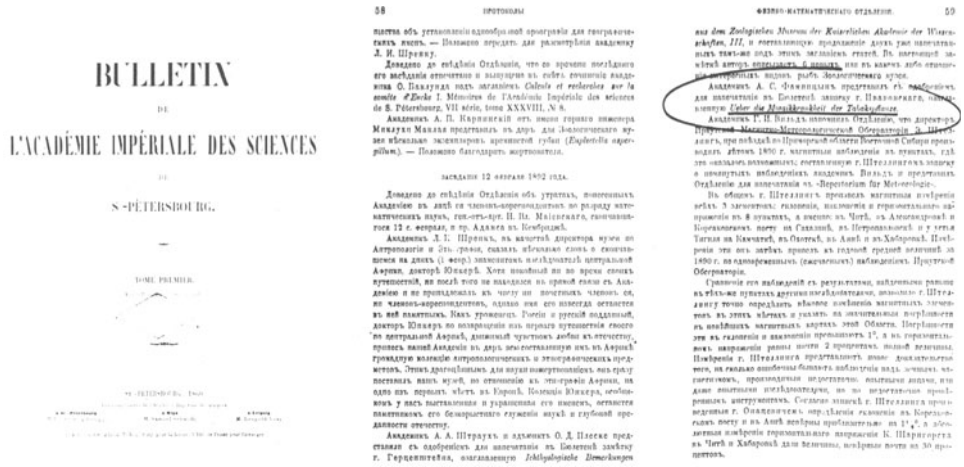

Figure 1. Ivanowki's communication to the Imperial Academy of Sciences of St. Petersburg.

tobacco mosaic disease was caused by a novel type of organism that existed in fluid or soluble form, and he called this *contagium vivum fluidum*. In essence, he rejected the idea that all infectious agents must be cellular in origin and proposed a non-cellular form (liquid!). This was a speculative jump that many feel opened up the field of virology. Beijerinck also recognized the obligatory parasitic nature of viruses by showing that there was no independent reproduction in the test tube. In order to reproduce "they must become incorporated into the living protoplasm of the cell", he writes (4). He also quantitated the amount of material necessary for infection, and showed a relationship between dilution and severity of the disease. Perhaps, Beijerinck should be regarded as the real founder of modern virology.

In 1898, Freidrich Loeffler and Paul Frosch investigated the outbreaks of Foot-and-Mouth (FAM) disease in German cattle. Like Ivanovski they discovered an infectious, filterable agent, and like Ivanovski they discussed the possibility that this substance might be a soluble toxin. However, they later rejected this conclusion and suggested that this was "an agent capable of reproducing... so small that the pores of a filter which will hold back the smallest bacterium will still allow it to pass" (5).

These pioneer virologists established working criteria, all negative, for identifying what we now know as viruses. Viruses i) unlike bacteria could not be seen through a light microscope, ii) could not be cultivated in cell-free medium, and iii) are not retained by filters known to prevent passage of bacteria. However, the concept that these were an entirely new class of biological entities was not yet considered. They were assumed to be "small" microbes, although called viruses by all the scientists at this time.

Perhaps one of the most important discoveries of modern virology was made at the turn of the century, and long ignored for 40 or so years. This was the discovery of the transmissability of avian leukemia by Ellerman and Bang (6) in

Denmark in 1908 and of a sarcoma of chickens by Peyton Rous (7) in the U.S. in 1911. Unfortunately these discoveries were relegated to the rank of avian curiosities, and their importance to virology and medicine was not recognized for many decades.

In 1915 and independently in 1917, the host range of viruses was expanded by the discovery of d'Herelle and Twort (8, 9) of bacterial viruses. The bacteriophage has since become one of the best studied organisms on earth. Modern molecular biology would not have developed without the work of the Cold Spring Harbor group of Hershey, Luria, and Delbruck in the 1940's who laid the groundwork for the quantitative aspects of virology (10). Much of this work was stimulated by the speculations of the physicist Schrodinger in his book "What Is Life?" which directed many people trained in the physical sciences to explore these small replicative "minimal" organisms (10).

Around the 1930's, two major discoveries were made that helped characterize the virus further. William Elford, of the National Institute for Medical Research, London, used a material called "collodion" to construct a range of membrane filters with different pore sizes (11). Using these filters, Elford estimated the size of several viruses. Two important results derive from his experiments, i) viruses were shown to be particulate entities with a definite size, and ii) viruses causing different diseases had different sizes, although viruses causing any specific diseases were identical in size. He estimated, for example, that the size of the Foot-and-Mouth disease virus was 10 nm. Thus, Elford's work gave some indication of how small viruses really were (12).

In 1935, Wendell Stanley, an organic chemist, reported the crystallization of tobacco mosaic virus (13). Although this led to controversy as to whether viruses were living organisms or auto-replicating proteins, it demonstrated the proteinous nature of viruses. However because of lack of knowledge of the nucleic acid component, it was difficult to explain the mechanism of viral replication. Stanley proposed that TMV was an "autocatalytic" protein which required the living cell for multiplication (13). More important — although not clearly understood at the time — was the demonstration a few years later that bacteriophage contain a nucleic acid (14). The concept that viruses were quite different from bacteria was beginning to be understood.

The importance of the bacteriophage research of the 1950's and 1960's by Luria, Hershey, Lwoff, and many others will never be too much stressed (10): Their research made virology into a quantitative science, gave birth to modern molecular biology, and led to the basic discoveries that opened up nucleic acid research and genetic engineering.

In parallel with the advances in virology, major advances were being made in the field of cell-culture. The art (for at first it was more art than science) of cell culture, began with the work of Alexis Carrel, who, in 1910, showed that it was possible to maintain chick tissues in culture by growing them in plasma clots supplemented with extracts from living chick embryos (15).

William Earle (a former student of Carrel's) established the first truly immortal cell line in the early 1940's (16). These cell lines were established by treating primary mouse fibroblasts with the chemical carcinogen methylcholanthrene. One

of these cell lines, the L-cell, was established from mouse embryo fibroblasts in 1943. This cell line is widely used today and has proved invaluable in virology. A mutant derivative of this cell line has become a major tool in gene isolation experiments. In addition to the work of Earle, one must mention the work of George Gey, who established many human and rodent cell lines at about the same time (17).

Enders, in the late 1940's (18) showed that it was possible to culture poliomyelitis virus in various human embryonic tissues of non-neural origin. This led to the era of well-funded polio research and the development of methods for quantitating animal viruses, different tissue-culture media, and animal virus plaque assays (19). Dulbecco in 1963, demonstrated that viral transformation could be quantitated in a similar manner (20).

Closer to our own time, the work of Saul Spiegelman should be mentioned and in particular the *in vitro* replication of bacteriophage RNA (21). The discovery of the reverse transcriptase by Baltimore and Temin (22, 23) was a landmark in tumor virology and has profoundly altered all thinking in the area of cell biology and eukaryotic development. More recently the characterization of oncogenes by Bishop and Weinberg (23, 24), and the isolation of viruses of the HTLV/LAV series (Human Immune-deficiency Virus, HIV) by the groups of Gallo and Montagnier (26,27) have resulted in major insights into virus organization and replication.

2. VIRAL STRUCTURE

Since other chapters of this volume will describe in detail the molecular biology of individual viral species and virus-host interaction, we shall give here a simplified overview of viral structure and viral classification. Basically, virus are placed into three structural groupings based on electron microscopy. They are either

(a) *spherical* ("isometric"),

(b) *rod* shaped or filamentous (rigid or flexible) (Fig. 2) or

(c) *complex* (implying either a combination or neither of the above) (Fig. 3).

Many viruses posses lipid bilayer membranes, in part derived from the host cell, but usually with viral proteins inserted into the host lipid bilayer. As originally hypothesized by Crick and Watson (28), based on the limited coding potential of viral nucleic acids, the viral capsids are in most cases made up of repeating subunits. These capsid proteins protect the internalized nucleic acid from degradation, and may also act as means of cell attachment.

Viruses as we see them are symmetrical objects. It is important to remember that proteins themselves are *not* symmetrical and are irregular in shape. If a symmetrical arrangement did not occur, the same set of amino acids would have different patterns of noncovalent bonding in different places. Thus, because of the physical constraints of forming a symmetrical structure from asymmetrical proteins, spherical and rod-shaped structures fit the optimum energy requirements (Fig. 4a,

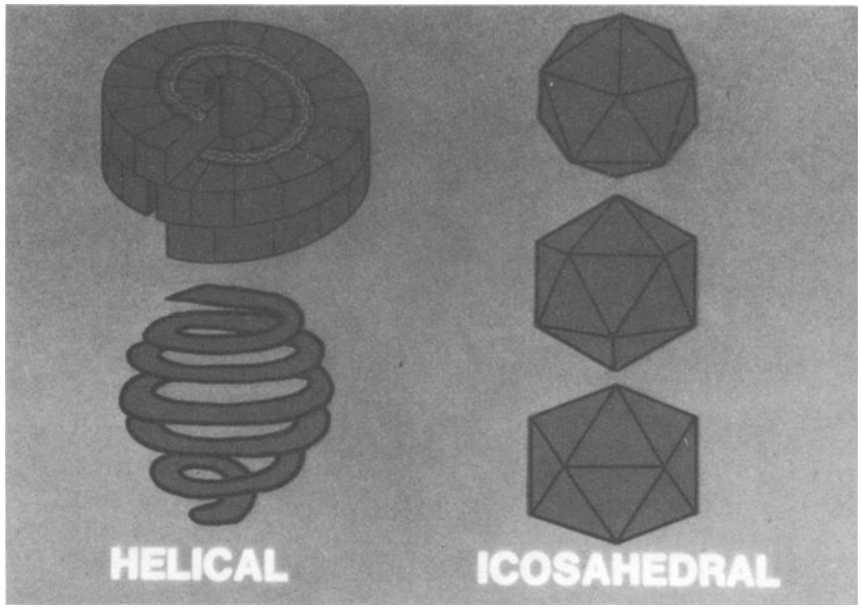

Figure 2. Hypothetical structure of a virus particle

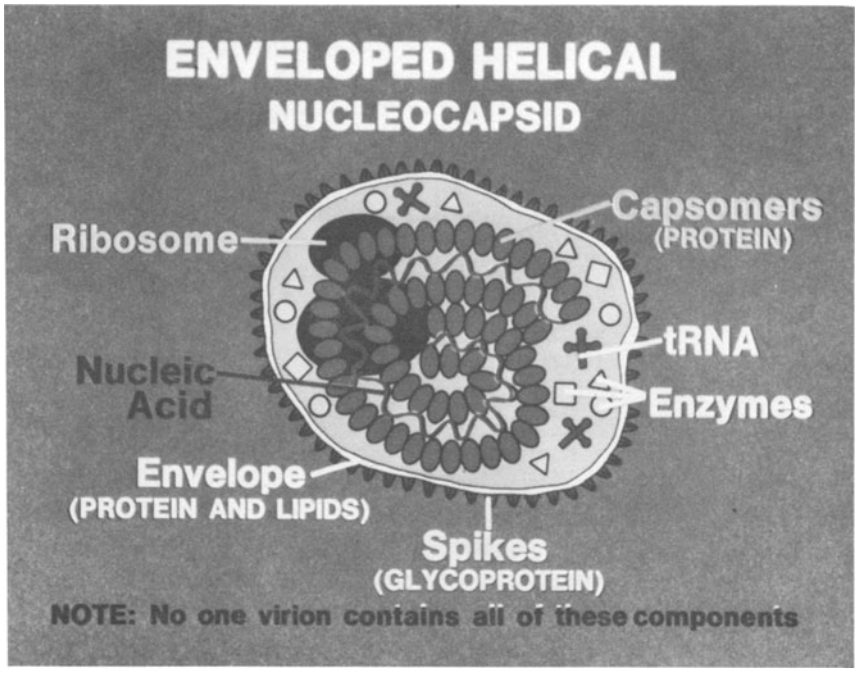

Figure 3. Hypothetical structure of a complex virus

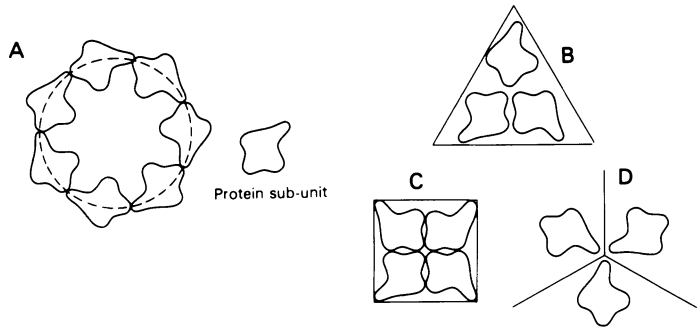

Figure 4. (A): Arrangement of identical asymmetrical components around the circumference of a circle to yield an asymmetrical structure. (B): Asymmetrical subunits located at the vertices of each triangular facet. (C): Asymmetrical subunits at each corner of a square with face represented in (D).
(Adapted from *Introduction to Modern Virology* (1974) S. B. Primrose, Halsted Press, with permission)

b, c, d). Although there may be slight deviations from symmetry this bonding of identical proteins is essential to self-assembly. Since viruses are fairly stable structures, the maximum number of bonds must be formed between the subunits and there are only limited ways in which this can be done.

Helical Structure

It turns out that one of the simplest ways of arranging non-symmetrical protein units is to place them around the circumference of a circle (Fig. 4a) to obtain a disc-like structure. If we examine the assembly process of TMV, we find disk-like structures as intermediates during self-assembly depending on the pH of the incubation buffer (Fig. 5) (29). However, because of the interaction between these discs and viral RNA the disc-like structures form a helical structure (29, 30). All filamentous viruses have helical protein structures, which probably reflects the constraints of wrapping disk structures around a long nucleic acid.

Icosahedral Viruses

Most spherical animal viruses have icosahedral symmetry. Multiplying the number of subunits per face by the number of faces gives the number of subunits that can be arranged around such a closed shell. For the icosahedron it turns out that 60 subunits or multiples of 60 are the number required (Fig. 6a, b). Since all spherical viruses are icosahedral, there must be some constraints on building other structures. Many viruses have more than 60 subunits (60N) but are still icosahedrons. The number of subunits does not have to equal the number of structural proteins (30).

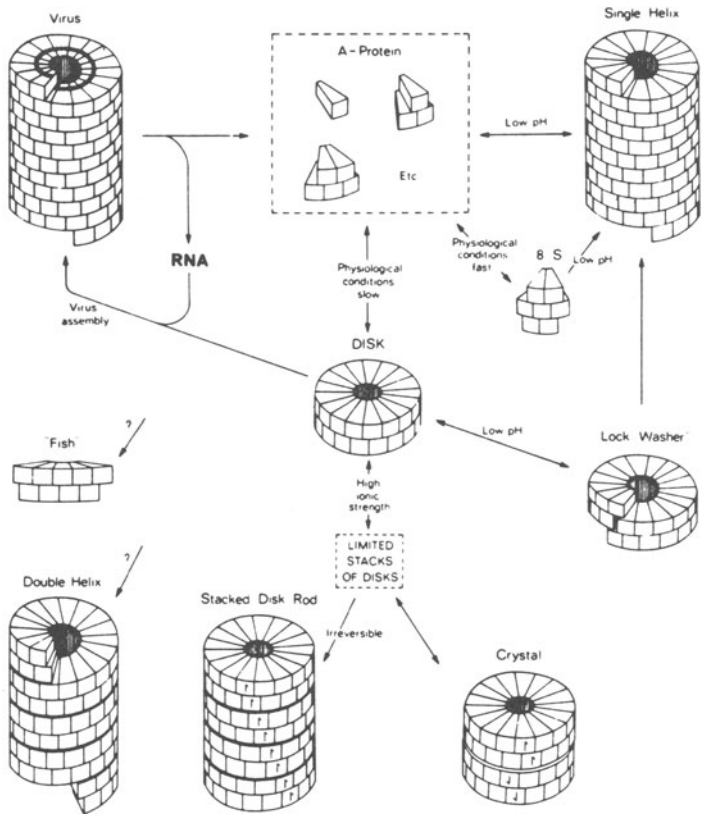

Figure 5. Diagram of interconversions that have been observed between some of the better aggregates of TMV proteins. (Reproduced from ref. 29, with permission).

Complex Viruses

Not all viruses are obviously helical or spherical. Viruses such as pox, herpes, rhabdo, T-phage, and λ-phage have complex morphology. In some cases this is due to the presence of a lipid membrane, and a helical/spherical basic structure (nucleocapsid) is found within the lipid membrane. Some of the plant viruses are flexible rods. These are basically helical but they have no straight axis of symmetry and so the subunits are quasi-equivalently related. Complex viruses such as the bacteriophage, are assembled independently from distinct sub-assemblies of icosahedral heads, rod-shaped tails, and tail fiber assemblies, and are then put together in the presence of a scaffolding protein.

Another important aspect of structure is the relationship between the viral nucleic acid and the capsid protein. In the case of helical viruses, such as TMV, there is a specific interaction between the viral nucleic acid and the protein

subunits. In the case of isometric viruses, however, the condensation of nucleic acid is often independent of the protein structure and other viral and non-viral nucleic acids can be packaged into the capsid protein. In these cases it appears that the only restriction is that the viral RNA fit into the shell structure.

3. CLASSIFICATION OF VIRUSES

Because viruses contain either RNA or DNA, double-stranded or single-stranded, circular or linear, and these features can change quickly upon entry into the host, different viruses often have little in common with each other than their parasitic nature. The taxonomic scheme proposed by the ICTV (International Committee on Taxonomy of Viruses) (31) uses these structure and biochemical differences as the basis of its classification scheme. The hierarchy of this scheme subdivides viruses on the basis of their nucleic acid (RNA or DNA), viral structure (e.g., helical, isocohedral), whether they are enveloped, and genome structure (e.g., linear, d-s) (Fig. 7a,b).

Family names end in *viridae*, subfamily names in *virinae*, and genera, like species, in... *virus*. This taxonomic scheme is a mixture of the old and the new, since the names of some groups — such as adenoviridae and herpesviridae — refer to the original source of isolation or pathology of the virus, whereas the actual classification scheme is based on structure, type and character of nucleic acid, and in the case of retroviridae, on the presence of an enzyme, the reverse transcriptase.

Baltimore (32) has modified this scheme to use the mode of gene replication and expression to classify viruses (Fig. 8). In his classification scheme, mRNA (or + strand RNA) plays a pivotal role since protein synthesis occurs by the same mechanism for all viruses. All viruses are assigned to a numbered class based on the mode of synthesis of mRNA. All mRNA is designated (+) RNA. RNA which is complementary to the mRNA is designated as (-), and those which are non-

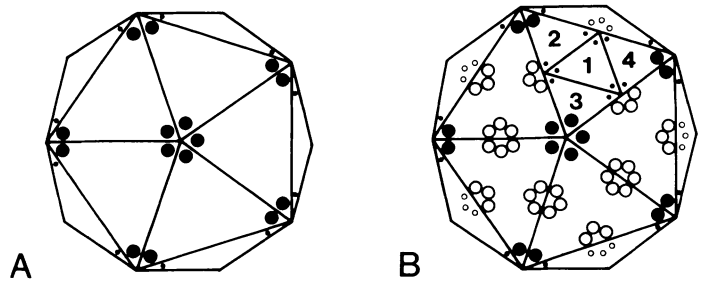

Figure 6. Arrangement of 60n identical subunits on the surface of an icosahedron. (A): n = 1, and the 60 subunits are distributed such that there is one subunit at the vertices of each triangular face. (B): n = 4, each triangular facet is divided into smaller (but identical) equilateral triangles.
(Reproduced from *Introduction to Modern Virology*, S. B. Primrose (1974), Halsted Press, with permission)

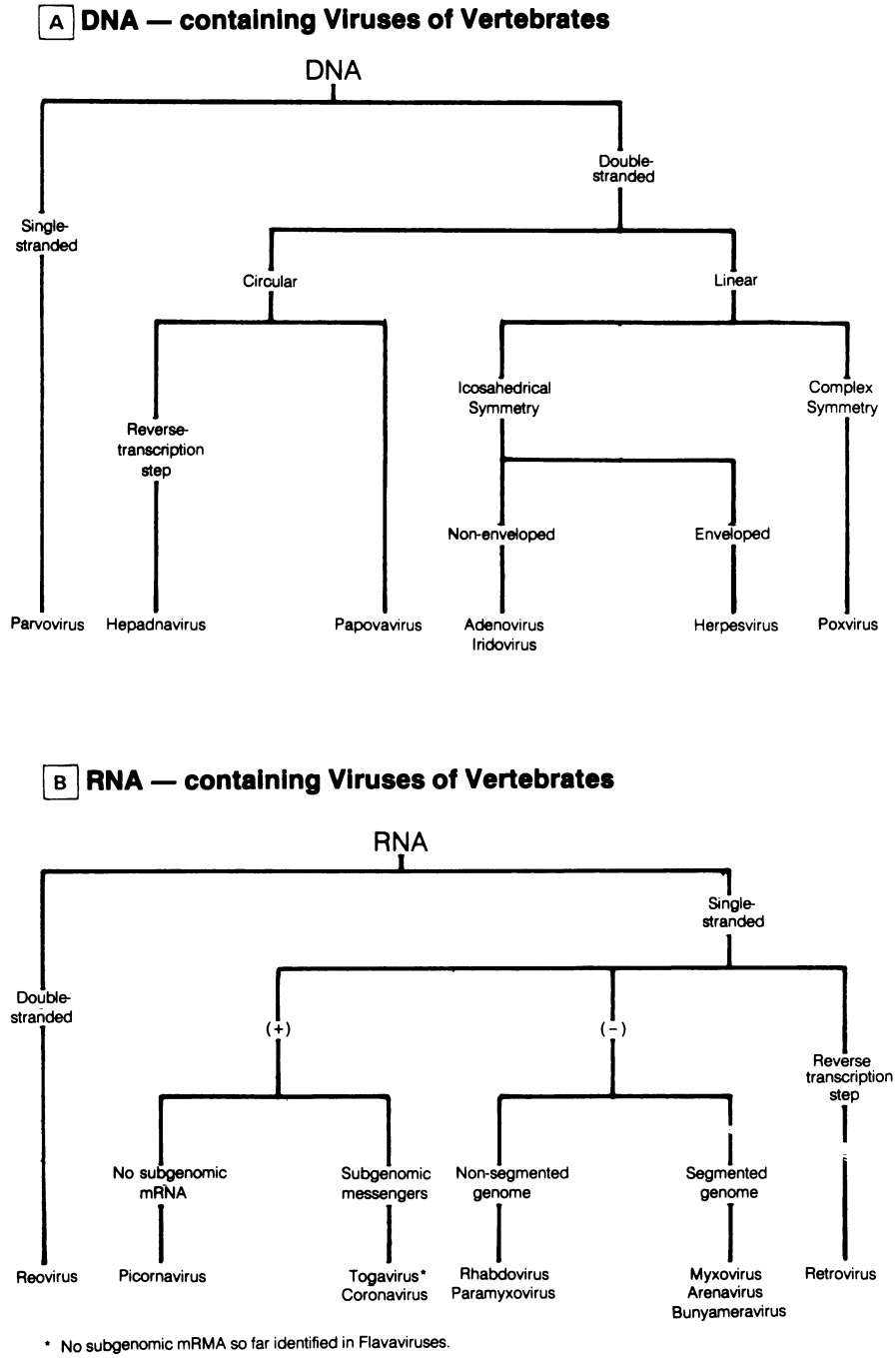

Figure 7. Classification of Animal Viruses

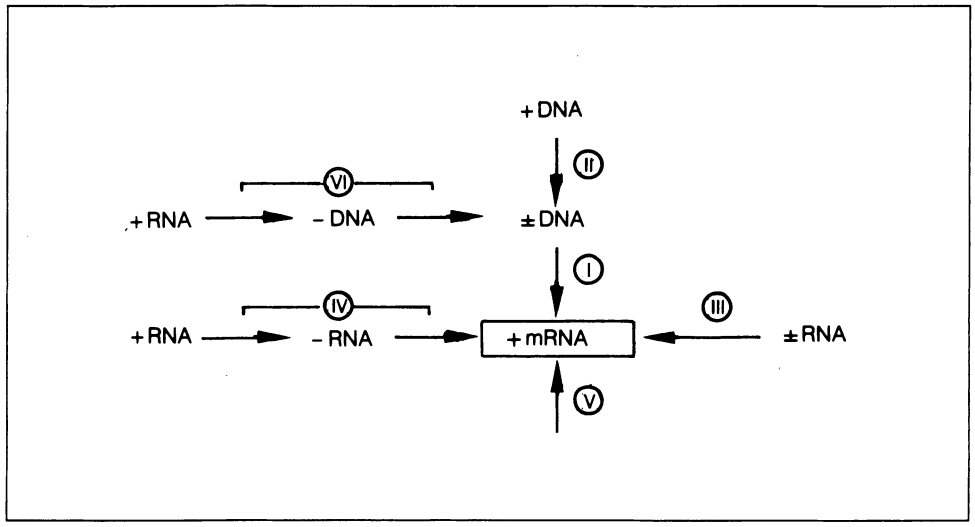

Figure 8. Functional Classification of Viruses according to D. Baltimore. Representative examples: I = T4 phage, vaccinia virus; II = ΦX174; III = reovirus; IV = RNA phages, poliovirus; V = vesicular stomatitis virus, Newcastle disease virus; VI = RNA tumor viruses.
(Adapted from ref. 32)

complementary as (+). Using this terminology, six classes of virus can be distinguished. We can subdivide the classes using other characteristics, such as enveloped, non-enveloped, segmented genome, etc.

Neither of these classification schemes do imply any phylogenetic relationship beyond those which can be imputed through nucleic acid hybridization, genetic recombination, and nucleic acid sequences of common regions. The presumption is that viruses which differ at few nucleotides have a more recent common ancestor than those that differ at a hundred. These techniques, however, are mute on the phylogenetic relationship between viral groups too disparate to have common features.

Since there is no fossil record of viruses, we cannot even consider viruses as a monophyletic group, since it is entirely possible (in fact, probable) that viruses do not always share a common ancestry, but arose independently more than once from different sources and by different mechanisms. We will discuss possible mechanisms later. Moreover, because viruses are parasitic, their further evolution is closely co-ordinated to the evolution of their hosts. Thus, the differences between for example, single-stranded circular DNA animal and plant viruses may reflect *either* common ancestry with subsequent divergence due to separate evolution of the host or may reflect two separate evolutionary events (by similar or different mechanisms) resulting in two independently-evolved viruses, one specific to plants and the other to animals. One must be cautious, moreover, in ascribing total linkage of viral evolution to a single host species since many plant and animal

viruses are spread through the intervening vehicle of fungus or insect vectors, and these viruses must be coadapted for existence in both hosts.

4. VIRAL ONCOGENESIS

Probably one of the most exciting areas of virology is the role of oncogenic viruses. Again, as in so many other areas of virology, our model system is really derived from the bacteriophage work and the concept of provirus. The existence of bacteria that had the ability to generate phage *de novo* — i.e., were lysogenic — goes back to the 1920's with the work of Bordet (33) and Bail (34). However, the theory of lysogeny was really crystallized by the research of Lwoff in the 1950's (1). In temperate phage we have a system in which the phage DNA can be incorporated into the host genome, express unique repressor functions, prevent super-infection by the same phage, can be excised out, and on occasion carry (transduce) nearby genes.

The analogy of the prophage concept to RNA tumor viruses has proven very fruitful. The oncogene theory of Todaro and Huebner (35) is basically a restatement of the prophage theory with some modifications for animal viruses. The reader is referred to the proper section of chapter 16 of this volume for a detailed discussion of the oncogene theory. For the purpose of this introductory review enough to remind that the theory stated that many forms of neoplasia arise by the action of carcinogens on the expression of retrovirus oncogenes resident in the cell. Although this model is incorrect, the analogy between lysogeny and retrovirus integration proved to be a good one. The RNA tumor viruses integrate their DNA into the host chromosome, place nearby genes under the control of their Long Terminal Repeats (LTR's), and "transduce" oncogenes (Fig. 9). We can detect these rare, and often defective, transducing viruses by the characteristic of the oncogenes. Not only has the study of retroviruses given insight into the process of oncogenesis, but it has also contributed greatly to our understanding of cellular biology and cell differentiation (see chapter 16 of this volume).

5. ORIGINS OF VIRUSES

It is obvious that many of the early theories proposing viruses as very primitive organisms, as precursors of bacteria, or as precursors of lower eucaryotes, are difficult to accept in light of modern molecular biology. Viral evolution is ongoing, and the discovery of new viral strains affecting man or other animals, probably reflects ongoing recombination and evolution. Of course, one can not give a definitive answer to the question of where viruses came from; rather, one can only present a number of different hypotheses.

Three main theories have been advanced to explain the origin of viruses: i) viruses originated very early in evolution before the development of cellular life, that is, viral nucleic acid were among the first molecules replicating in the "primordial soup"; ii) viruses are the result of the degeneration of more complex parasitic organisms that have lost many of their key components, thus utilizing the protein synthetic and genetic apparatus of the cell; iii) viruses are derived from genetic

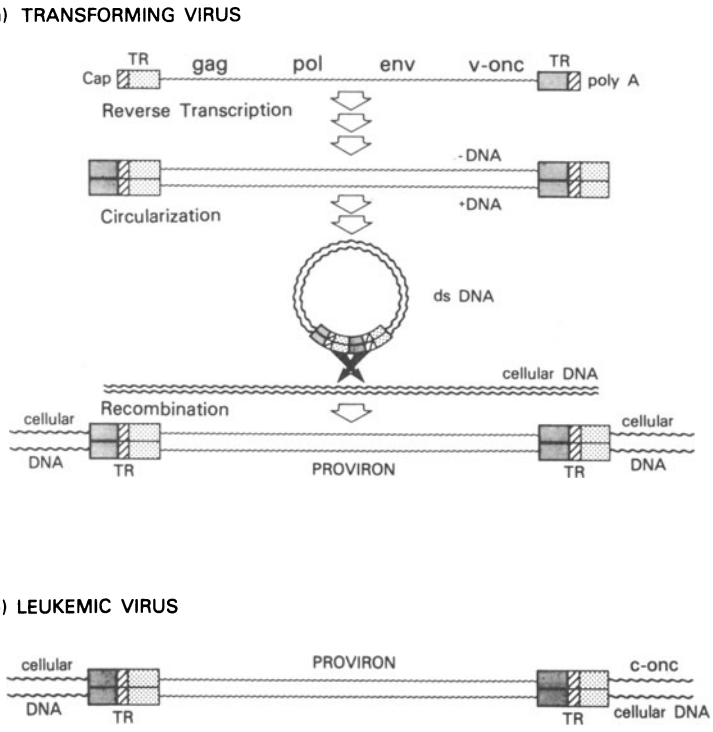

Figure 9. Model to explain the formation and differences between (A) transforming viruses, (B) leukemic viruses, and expression of oncogenes.

elements of the cell which have evolved a semi-independent existence, presumably by "modular recombination" (36).

The more classical theory of parasitism (38, 39) is that viruses have originated by loss of essential functions from some type of free-living procaryotic organism. If we accept this basic premise, our candidate organisms for the original parasite would be eubacteria, archaeobacteria, rickettsia, mycoplasma, chlamydia or possibly a fungal eucaryote. Matthews (39) has argued that it is chlamydia-like organisms that have given rise to viruses such as pox virus. The chlamydia are obligate parasites, and lack an energy-generating system. They have two phases in their life cycle, outside the cells they exist as infectious *elementary* bodies about 300 nm in diameter. They have an outer cell wall, inner wall layer, and a plasma membrane. The genome DNA has a molecular weight of 4×10^8, and the chlamydia also contain RNA.

On entering the cell the elementary body is converted into the noninfectious *reticulate body*. This body is surrounded by a bilayer membrane derived from the host and divides by binary fission, giving rise to thousands of progeny within a few hours. Sometimes infection can lead to an immune state, and prevent superinfection.

What do chlamydia and poxviruses share in common, and what do they not?

a) Size of cell/viral particle — Approximately the same.

b) Genomic size — Approximately the same.

c) Cell Wall: Present in chlamydia, absent in pox virus.

d) Growth: Neither chlamydia or pox virus grow outside of cell. Both are obligate *intracellular* parasites.

e) Energy-yielding system: Absent in both.

f) Inability to synthesize amino-acids. Dependent on host cell.

g) Presence of two nucleic acids: Pox viruses do not contain RNA.

This theory is quite tenable for larger viruses such as T2, T4, herpes, pox viruses, etc., although it does not explain the formation of specialized organelles such as tail fibers, etc. Smaller viruses could be generated as defective interfering derivatives of a larger virus.

An alternate possibility, is that complex viruses may have resulted not by subtractive processes but from recombination type events, or physical joining/reordering of DNA segments of "primitive" viruses and host genes, an additive process. That herpes, T2 and T4, have genes for functions normally found in the host cell, such as thymidine kinase, tRNAs, and other enzymes of nucleotide biosynthesis may be explained as resulting from recombination, or transduction type events between virus and host cells or even by recombination between different viruses.

At the present time it is often speculated (and accepted) that certain groups of viruses arose from genetic elements of the cell or from chromosomal DNA (RNA?) which have somehow evolved an independent existence.

The late Herman Muller, the Nobel Prize winner who worked in our own department, suggested in 1922 (40) that bacteriophages may be derived from genes, "if these d'Herelle bodies were really genes, fundamentally like our chromosome genes, they would give us an utterly new angle from which to attack the gene problem. They are filterable, to some extent isolable, can be handled in test-tubes, and their properties, as shown by their effects on the bacteria, can then be studied after treatment. It would be very rash to call these bodies genes, and yet at present we must confess that there is no distinction known between the genes and them. Hence we cannot categorically deny that perhaps we may be able to grind genes in a mortar and cook them in a beaker after all." This statement was obviously farsighted and very modern.

By genetic elements we imply either segments of chromosomal DNA, transposons, insertion elements, or plasmid-like elements. As will be discussed below, we should also consider the reverse possibility: that such elements may be degenerate viruses. Let us now consider some possible examples.

During RNA processing, introns of varying length are spliced out of the heterogeneous nuclear RNA (HnRNA). It has been speculated that these RNAs might occasionally circularize and replicate autonomously. This has been suggested as a model for the generation of autonomously replicating RNAs. Viroids are small (250-400 base) closed circular single stranded RNAs, that are infectious to many

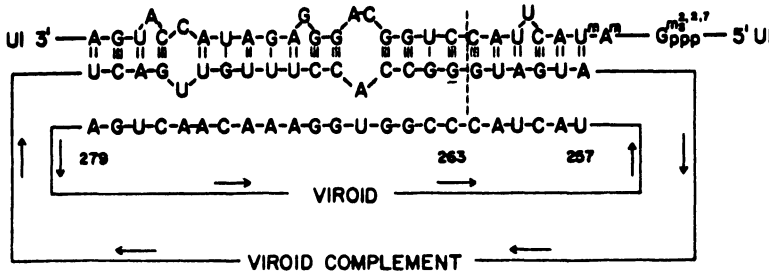

Figure 10. Possible relationship between viroid RNA and UIRNA (Reproduced from ref. 41, with permission).

plant species. Viroids can replicate in isolated nuclei and seem to have concatemeric precursors.

Small nuclear RNAs (snRNAs) associated with RNP particles are believed to be involved in the processing of primary transcription products (see chapter 4 of this volume). The 5' end of one such RNA, U1, has been shown to exhibit complementarity with the ends of many eucaryotic introns. Although no homologous sequences have been found between viroid and U1 sequences, some homologous sequences have been found between viroid RNA complements and U1 RNA. Figure 10 illustrates the possible base pairing interaction between the Potato Spindle Tuber viroid (PSTV) RNA complement and the 5' end of U1 RNA (41), (42). It should be pointed out that this is the only case in which such homologous regions have been detected.

Zimmern (43) has proposed a model in which intron-like RNA's (termed signal RNAs in the model), interact with "antenna RNAs", derived from structural or other genes. The product of signal RNA integrated into an antenna RNA is a fuson. Such fusons could code for replicases involved in subsequent independent replication of the fuson. Antenna RNA could be activated in different ways by signal RNA. Integration of the fuson or signal sequence into the genome might lead to amplification. Zimmern (43) argues that RNA viruses might have evolved from the attachment of a signal RNA carrying an origin of replication, to the antenna mRNA for a polymerase recognizing that origin, thus perpetuating a self-replicating RNA molecule. Recombination between this molecule and other mRNAs might lead to the formation of a complete RNA virus.

Another candidate for the origin of viruses is the transposon, or transposable element. Transposable elements are generally integrated into the host DNA with short reiterations of cellular DNA at either ends, have inverted terminal repeats at either end which can be similar to retrovirus LTR's in the case of complex transposons, have a long open reading frame, and many other characteristics similar to a retrovirus.

Although we can argue that viruses evolved from transposons the reverse may also be true, i.e., transposons and Ty-elements may be derived from retroviruses, or all may have a common ancestor. The recent finding that yeast carrying Ty

elements have reverse transcriptase and produce virus-like particles similar to the intracisternal A-like particles of the mouse suggests that this transposable element may be a defective retrovirus (44). Likewise the copia element of *Drosophila* produces a virus-like particle resembling a retrovirus (45). Thus, before concluding that transposons are precursors to viruses, one should consider the opposite direction of evolution, i.e., transposons may have evolved from defective retroviruses.

The best argument against the "transposon" origin, but in favor of transposons being relics of retroviruses, is the recent finding (46) that Hepatitis B virus, a DNA virus, is possibly derived from an "ancient" retrovirus. It is now a well established fact that Hepatitis B virus (Hepadnaviruses) replicates through an RNA intermediate by reverse transcription (see chapter 17 and ref.47), a process similar to retrovirus replication. Miller and Robinson (46) have compared the nucleotide sequences of thirteen viral genes by means of a computer search program, and have recently reported that extensive homology exists over a 100 nucleotide segment of the conserved region of hepatitis DNA and retroviral U5 sequences. An examination of retrovirus-like sequences in human and simian chromosomes shows a similar pattern of conserved sequences.

Although hepadnaviruses do not contain an integrase, they contain two 11 or 12 base-pair direct repeats in the region homologous to the retrovirus U5 sequence. Thus, hepadnaviruses may be capable of transposition. When the nucleocapsids of Hepatitis B viruses and retroviruses are compared, similar regions of homology are found.

The copia and 17.6 transposable element of *Drosophila*, cauliflower mosaic viruses of plants, the Ty element of yeast, endogenous retrovirus-like elements of mammals, retroviruses, and hepatitis B virus share homology over several regions of their genome. The data indicate that all may have evolved from a common ancestor. Was this ancestor a large RNA virus, or a small transposon-like element that later incorporated host genes and become autonomous and infectious?

This last possibility can be explored by comparing the codon usage of viral genes and cellular genes. If one examines the codon usage of oncogenes, we find that there is a strong bias at the third position for cytidine rather than uridine, and guanine rather than adenine as in genes of eukaryotic cells. This bias is reversed in eukaryotic viruses. UUC is favored 3 to 1 over UUU for phenylalanine in the *src* gene, but UUU is favored for phenylalanine in the virus proteins. By examining the sequence in hepatitis viruses, it can be seen that in the X gene the codon usage preference is similar to that of eukaryotic genes, whereas the other genes definitely show a different pattern. It would thus seem that all of these viruses were derived from an ancestral retrovirus capable of undergoing recombination with its host cell.

It would be useful to do a similar computer-assisted analysis for herpes genes such as thymidine kinase.

6. VIRAL EVOLUTION

It is widely believed (48, 49) that RNA preceded DNA as the genetic material. The argument for this is that the 2' hydroxyl group in RNA tends to make the RNA more labile than DNA and thus selection would tend to favor DNA. RNA viruses

are unique, in that they are the only self-reproducing organism in the biosphere that utilize an RNA genome.

DNA replication has associated proofreading mechanisms and repair systems. DNA replication in *E. coli* has an error rate of one false nucleotide incorporated for every 10^6-10^7 bases polymerized. Post-replicative repair systems may decrease this another thousand-fold.

The estimated error rate for RNA replicase, however, lies between 10^{-3}-10^{-6}. There are no known RNA repair enzymes. Thus RNA is intrinsically prone to error (referred to as noisiness by Reanney (38)). That this can lead to changes in the viral population emerged from experiments that used the small RNA-containing phage Qβ (50); 15% of the clones arising from a multiply-passaged population of Qβ had fingerprint patterns that deviated from those of the RNA of the population as a whole. Almost all progeny virus had at least one base sequence differing from wild type. This phenomenon is obviously the same as that referred to below as antigenic drift in foot-and-mouth-disease virus and influenza virus.

The noisiness of RNA replication has probably had an effect on the evolution of *RNA genomic size*. The largest known RNA genome is about 8×10^6 daltons. This may reflect the error rate, as the larger the genome the greater the possible error. Some RNA viruses may have better their performance by segmentation. Segmentation may allow for an escape from some of the deleterious effects of this high mutation rate by the ability to shuffle viral segments, and the greater likelihood of selection of the "correct" combination. Also segmentation substitutes for the type of recombination that might occur in DNA viruses. This is best illustrated by recombination (or antigenic shift) in influenza viruses (see chapter 13 of this volume).

There seems to be pressure toward smaller genomes in RNA viruses. *In vitro* experiments of Spiegelman and colleagues using Qβ replicase and a Qβ template (21) appeared to confirm that small RNA templates reproduce more progeny than large RNA templates. This also appears to be true, in general, of DI particles, where the small DI RNA out-replicates the intact viral RNA. Thus the evolutionary pressure would be for smaller RNA genomes.

Another aspect of evolution which one must consider is speciation; how do new species of virus arise, and how (or why?) do viruses change their host range? Many viruses can multiply in both insects and vertebrates, or insects and plants (but not in all three). Obviously this reflects ecological opportunity, the ability of insects to feed on vertebrates, or plants. Thus the spread of certain viruses will follow insect feeding mechanisms, and may have little to do with evolutionary pressures. This may explain why a particular rhabdoviruses can multiply in vertebrates or plants but not in both.

Studies with wound tumor viruses (51, 52) showed that when WTV was maintained in sweet clover plants for up to two years without passage through an insect vector, mutants arose that could no longer colonize the insect. Some of these mutants lacked segments of the genome, but retained their ability to replicate effectively in plants. Thus changing the ecology of the virus resulted in changing its host range. We know from studies of influenza viruses that antigenic shift results in large changes in the character of the virus (chapter 13).

Botstein (36) has proposed a generalized modular theory of viral evolution: He suggests that viral evolution is not so much the result of mutational events, but rather the result of recombination between individual blocks of genes, (functional blocks), here termed modules. Each module would specify a specific function. According to this theory, each viral type is the result of a combination of genes modules, which successfully occupy a niche in the environment. It is obvious that such a system does work in the generation of new lambdoid phage, where we can have the immunity of one phage, and host range of another (e.g., recombination between λ and 434, λ and Φ80).

Similar mechanisms may be at work in animal virus systems. Segmented genomes do undergo reassortment. Recombinants can be found between apparently non-related viruses, such as SV-40 and adenoviruses and, at the nucleic acid level, although possibly not at the functional level, between SV-40 and ΦX174 (53).

Similar mechanisms are also at work between defective viruses and infectious virus, endogenous retroviruses and exogenous retroviruses, and similar pathways

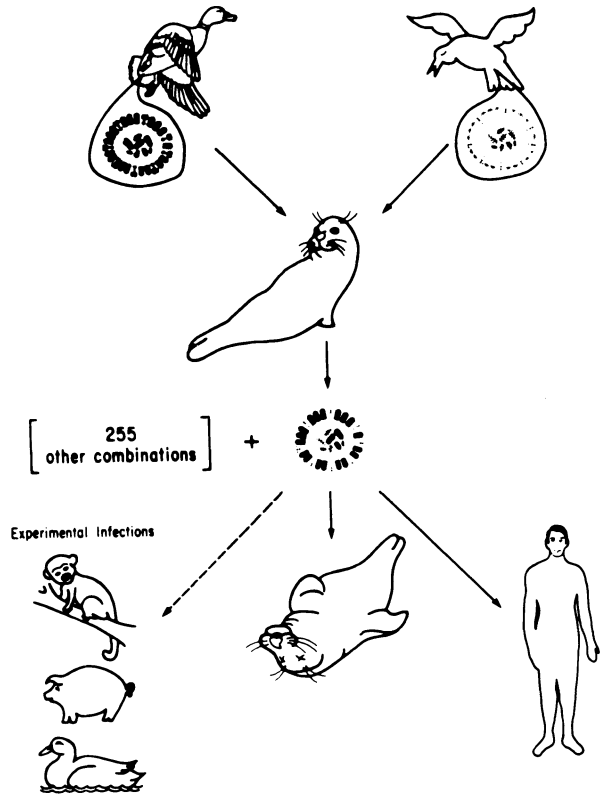

Figure 11. Diagramatic representation of the origin of seal influenza virus A/Seal/Mass/1/80 (H7N7). (Reproduced from ref. 54, with permission).

can be expanded to include the incorporation of host genetic material, such as oncogenes, into retroviruses.

At the microevolutionary scale, two other mechanisms leading to viral variation are antigenic drift, and antigenic shift. The best example of these microevolutionary mechanisms are presented by influenza viruses, although other examples could be mentioned — such as HTLV-3 and rhinoviruses. Antigenic drift involves minor changes in coat protein, or, in the case of influenza, in hemagglutinin and neuraminidase (54). Antigenic shift involves major antigenic changes.

Antigenic drift can be mimicked in the laboratory by growing influenza virus in the presence of monoclonal antibodies to the specific hemagglutinin (HA). Antigenic variants occur at a frequency of 1×10^{-5}/ml (54). Such variants have single amino acid changes in the HA polypeptide chain. In nature antigenic drift occurs by the accumulation of point mutations. Sequence analysis has shown that two mutations or more are necessary for a new strain to occur that escapes neutralization by the antibody to the parental virus. Thus, the accumulation of point mutation results in the evolution of new substrains of virus.

In *antigenic shift* large changes occur in the character of influenza virus. Sequence data indicates that the new subtypes have occurred by genetic reassortment. Genetic reassortment has been shown to occur between influenza A viruses of humans and lower animals *in vivo* (54). This type of evidence substantiates in part the Botstein (36) theory of viral evolution. This reassortment is diagrammed in Fig. 11 (55).

From the discussion above it is obvious the sources of viral evolution could be recombination between the viral genomes, recombination between viruses and the host genome; or segment reshuffling in segmented viruses.

7. REFERENCES

1) Lwoff, A. (1957) J. Gen. Microbiol. **17**,239-253.

2) Luria, S.E. and Darnell, J.E., Jr. (1967). In: *Virology*, John Wiley and Sons, Inc., New York.

3) Jenner, E. (1961) In: *Milestones in Microbiology*. T.D.Brock,ed., Prentice-Hall Inc., New Jersey.

4) Beijerinck, M.W. (1961) In: *Milestones in Microbiology*, T.D. Brock, ed., Prentice-Hall Inc., New Jersey.

5) Loeffler, F. and Frosch, P. (1961) In: *Milestones in Microbiology*. T.D. Brock, ed., Prentice-Hall Inc., New Jersey.

6) Ellerman, V. and O. Bang. (1908) Zentrablatt fur Bakteriologie, Parasitenkunde, Infektionskrank-heiten und Hygienes **46**, 595-609.

7) Rous, P. (1911) J. Exper. Med. **13**,397-411.

8) d'Herelle, F. (1961) In: *Milestones in Microbiology*, T.D. Brock, ed., Prentice-Hall Inc., New Jersey.

9) Twort, F. (1915) Lancet *11*,1241-1243.

10) Cairns, J., Stent, G.S., and Watson, J.D. (1966) In: *Phage and the Origins of Molecular Biology*, Cold Spring Harbor Laboratory, Cold Spring Harbor.

11) Elford, W.J. (1931) J. Pathol. Bacteriol. **34**, 505-521.

12) Elford, W.J. (1938) In: *Handbuch der Virusforschung*, R.Doerr and C. Hallaue eds., Springer-Verlag, Vienna.

13) Stanley, W.M. (1935) Science **81**, 644-645.

14) Schlessinger, M. (1935) Nature (London) **138**, 508-509.

15) Carrel, A. (1912) J. Exp. Med. **15**, 516-528.

16) Earle, W.R, Schilling, E.L., Staele, T.H., Straus, N.P., Brown, M.F., and Shelton, E. (1943). J. Nat. Cancer Inst. **4**, 165-212.

17) Gey, G., Coffman, W., and Kubiceck, M. (1952) Cancer Research **12**, 364-365.

18) Enders, J.F., Weller, T.H., and Robbins, F.C. (1949) Science **109**, 85-87.

19) Dulbecco, R. (1952) Proc. Natl. Acad. Sci. USA. **38**, 747-752.

20) Dulbecco, R. (1963) Science **142**, 932-36.

21) Spiegelman, S., Haruna, I., Holland, I.B., Beaudreau, G. and Mills, D. (1965) Proc. Natl. Acad. Sci. U.S.A. **54**, 919-927.

22) Baltimore, D. (1970) Nature (London) 13226, 1209-1211.

23) Temin, H.M. and Mizutani, S. (1970) Nature (London) **226**, 1211-1213.

24) Bishop, J.M. (1984). In: *The Microbe 1984. Part I Viruses*, Mahy, B.W.J. and J.R. Pattison, eds. Cambridge University Press, Cambridge, pp. 121-147.

25) Weinberg, R.A. (1982) Adv. in Cancer Res. **36**,149-63.

26) Gallo, R.C., Sarin, P.S., Gelmann, E.P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V.S., Mann, D., Sidhu, G.D., Stahl, R.E., Zolla-Pazner, S., Leibowitch, J., and Popovic, M. (1983) Science **220**, 865-867.

27) Barre-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Daguet, C., Axler-Blin, E., Vizenet-Grum, F., Rouzioux, C., Rozenbaum, W. and Montagnier, L. (1983) Science **220**, 868-871.

28) Crick, F.H.C. and Watson, J.D. (1956). Nature (London) **177**, 473-475.

29) Butler, P.J.G. and Durham, A.C.H. (1977) Adv. Protein Chem. **31**,187-251.

30) Klug, A. (1979) Harvey Lecture **74**, 141-172.

31) Wildy, P. (1971) In: *Monographs of Virology* , J.L.Melnick, ed., **5** Karger, Basel.

32) Baltimore, D. (1971) Bacteriol. Rev. **35**, 234-241.

33) Bordet, J. (1925) Annales de l'Institut Pasteur, **39**, 711-763.

34) Bail, O. (1925) Medizinische Klink (Munchen) **21**, 1271-1273.

35) Todaro, G.J. and Huebner, R.H. (1972) Proc. Natl. Acad. Sci. USA.**69**, 1009-1015.

36) Botstein, D. (1980) In: *Animal Virus Genetics*, Fields,B.W., Jaenisch, R., and Fox, C.C. , eds, Academic Press. New York pp. 363-384.

37) Hoyle, F. and Wickramasinghe, C. (1978) New Scientist, pp. 946-948.

38) Reanny, D.C. (1982) Ann. Rev. Microbiol. **36**, 47-73.

39) Matthews, R.E.F. (1983) Int. Review of Cytology, Suppl. 15. J.F. Danielli, ed. Academic Press, New York, pp. 245-280.

40) Muller, H.J. (1922) Am. Naturalist **56**, 32-50.

41) Diener, T.O. (1981) Proc. Natl. Acad. Sci. U.S.A. **78**, 5014-5015.

42) Kiefer, M.C., Owens, B.A., and Diener, T.O. (1983) Proc. Natl. Acad. Sci. USA. **80**, 6234-6238.

43) Zimmern, D. (1982) Trends in Biochemical Sciences. **7**, 205-207.

44) Garfinkel, D.J., Boebe, J.D., and Fink, G.R. (1985) Cell **42**,507-517.

45) Shiba, T. and Saigo, R. (1983) Nature (London) **302**, 119-24.

46) Miller, R.H. and Robinson, W.S. (1986) Proc. Natl. Acad. Sci. USA. **83**, 2531-2535.

47) Summers, J. and Mason, W.S. (1982) Cell **29**, 403-415.

48) Eigen, M., Gardiner, W., Schuster, P. Winkler-Oswatisch, R. (1981) Sci. Am. **244**, 88-118.

49) Reanney, D.C. (1979) Nature (London) **288**, 598-600.

50) Domingo, E., Sabo, D., Taniguchi, T., & Weissman, C. (1978) Cell **13**, 735-744.

51) Reddy, D.V.R., Black, L.M. (1974) Virology **61**, 458-73.

52) Reddy, D.V.R., Black, L.M. (1977) Virology **80**, 336-46.

53) Dorsett, D.L., Keshet, I., and Winocour, E. (1983) J. Virol. **45**, 218-228.

54) Murphy, B.R. and Webster, R.G. (1985). In: *Virology*, B.N.Fields, ed. Raven Press, New York. pp. 1179-1240.

55) Harrison, S. C. (1984) In: *The Microbe 1984*, B.W.J. Mahy & J.R. Pattison, eds., Cambridge University Press, Cambridge, **1**, pp.29-73