A Search for Simplicity

Biophysicists search for the smallest known virus that can give a clue to the hereditary pattern of the gene

by Robert L. Sinsheimer

On most campuses the gulf between physics and biology is so vast and deep that an onlooker might assume that one group was composed of matter and the other of anti-matter, so that annihilation would be the fate of anyone who attempted to bridge the gulf. Having resided in a physics department for eight years, I know physicists are not as different as all that, and this lack of contact between these areas is, to my view, a pity, for biologists have certainly much to learn from physics that will be of the greatest value in the solution of their problems.

I am not so sure that the converse is valid; I am not sure that biology can contribute much to physics in a fundamental sense, although it can certainly offer the stimulation of an entrancing field, with wide scope and opportunity for the ingenious application of physical principles and concepts and techniques.

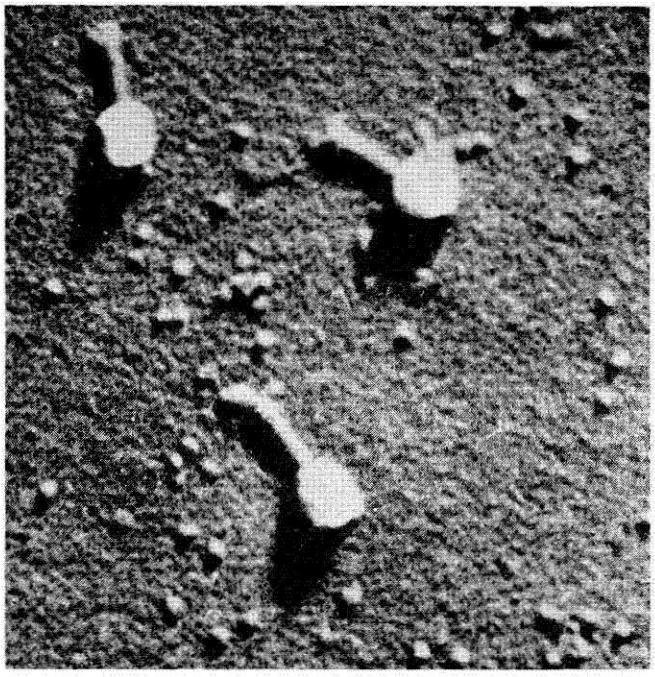
Actually, of course, an increasing number of people do work in the tenuous area between biology and physics, and some of these people even go so far as to call themselves biophysicists. Physics is the study of the properties of matter and energy. Biology is the science of life in its myriad manifestations. One of the most important areas of modern biology can be loosely called molecular biology. Molecular biology might be defined as the attempt to achieve an understanding of some of the remarkable phenomena of living organisms in terms of the structures and the physical and chemical interactions of their components—that is, in terms of the structures and reactions of the molecules, the macromolecules, the particles, which are found in living matter in an orderly and often highly organized pattern.

If the people who are engaged in this activity tend to derive their concepts, their techniques, and their mode of approach from physics, they are likely to consider themselves as falling into a loose category called biophysicists. If, conversely, they tend to draw their concepts and techniques and mode of approach from chemistry, they are likely to consider themselves biochemists.

Actually, very few of the people in this field are so narrow as to restrict themselves to the use of a purely physical or a purely chemical approach. In general, the biological phenomenon—be it photosynthesis or gene duplication, muscle contraction or nerve conduction—is the source of inspiration, and the scientist seeks to attack the problem with whatever means he can devise—be they biological, chemical or physical in nature. But I think it is clear that if a man's background is in physics and mathematics he is more likely to choose a quantitative physical approach than a qualitative chemical one, at least if he can devise one to answer the problems at hand.

Operationally, biophysics can be best defined in terms of the activities of those people who consider themselves to be working in the field. And so perhaps I can best give you a glimpse of a small area of the field by briefly describing the more recent activities of a self-designated biophysics laboratory at Caltech.

The broad biological problem with which I have been vitally intrigued for some time is represented by the word gene. More specifically, I am interested in the problems of the structure of the gene, the manner of its action, and the mode of its replication. Now genes are, formally speaking, simply units of heredity, and initially the idea of a gene was a purely abstract conception. It is a biological observation that a plan—a pattern of heredity—is passed from each generation to the next generation. Each organism, each of us, develops from a single cell which, at the time, possesses within itself a plan that leads, in due course, to the development of the mature individual. And each individual retains many



Tadpole-shaped T2 virus particles. Magnification is about 100,000 diameters. Smaller objects are extraneous material. The T2 virus is a representative of the most extensively studied class of bacterial viruses.

copies of that plan which he may pass on to his descendants.

One of the first triumphs of genetics was to show that this plan is unitary in character. It is made up of discrete entities or factors called genes. When stocks are crossed equal numbers of genes from each stock are pooled—in a precise way, to be sure. When a mutation—a change in the hereditary plan—is observed, it is found most often to be the result of a modification of one of these units.

To a physical scientist it usually seems almost self-evident that there must be some physical object, some structural entity, corresponding to each gene. To a biologist, more accustomed to thinking in terms of dynamic interacting systems, this has not been so obvious. Lest we be smug, let us remember that the discreteness of nature, the existence of atoms, was debated for many decades. In any case, it has only been relatively recently that techniques have been available to make the issue of the nature of the gene anything other than an ideological debate.

For this discussion I am going to assume that there is a discrete structure corresponding to the unit of heredity. To learn, then, the physical basis of these units of heredity, to learn how they are copied at every cell division, to learn how they direct the activities of the cells, of the embryo, of the whole organism as it develops and the hereditary plan unfolds—this is the broad problem. It is clearly a central problem in biology and it is one that has attracted many workers from varied disciplines.

As stated, the problem seems deceptively simple. In truth, it is extraordinarily complex and, with our present limited vision, difficult of access. There are many conceivable approaches to the problem, and how any individual will approach it is certainly a function of his background, his knowledge, and his intuition.

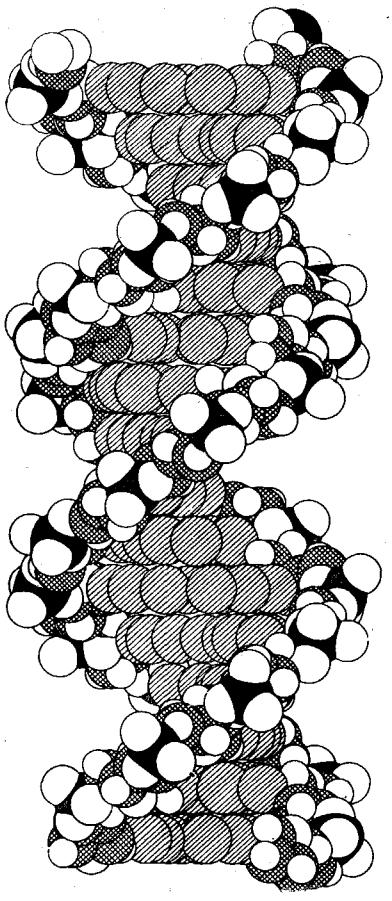
Physicists and people of like mind prefer to work with simple isolated systems, with arrangements in which as many variables as possible are controlled, or else are at least surely ignorable. They also, and with good reason, prefer systems that are amenable to quantitative analysis. Now, one may question whether even in principle (technical difficulties aside) a gene can ever function in an isolated system. For instance, one may ask whether its function might require a sort of feedback from the result of its function, but we are not ready for such questions yet. Surely, at present, a gene can only find expression in a living organism. But in a search for simplicity one can take certain plausible steps. One can first of all look for an organism which is both qualitatively and quantitatively of a minimum complexity—an organism for which the hereditary plan might therefore be of a minimum complexity. If one takes this approach and looks about for the simplest (or more accurately, the smallest) object that can cause, in an accurate way, its own duplication—that carries a plan of heredity which is at least believed to be unitary in character and which is mutable in the usual way—one's attention is soon drawn to the viruses.

Parasitic genes

It must be said right away that viruses lack many of the attributes of life. They must, in a sense, sponge on their host for preformed parts, for catalysts, for energy. But viruses do have the remarkable power to reorganize the synthetic machinery of the host and use it to bring about their own specific replication. They seem to represent in a stripped-down form just that portion of the living organism which is central to our problem of the gene. A virus, in this sense, represents an organized set of independent—and parasitic—genes.

Viruses are known to prey upon nearly all living species, and here again, in a search for simplicity one is led to concentrate upon the bacterial viruses, the bacteriophages. The simplicity here may be a little deceptive, but at least the host, the bacterium, is a single-celled, undifferentiated creature that may be grown in a precisely defined environment under quantitatively reproducible conditions.

Further, the bacterial virus particles may be quantitatively assayed by methods of great delicacy. It is easily possible to detect single virus particles. Thus one has means to detect in the range of 10^{-15} to 10^{-17} grams of virus, dependent upon the particular virus. Yet it is quite feasible to prepare 10^{15} particles. These features have been, and remain, a basic advantage that has permitted bacterial virus work to far out-distance work on other viruses. Bacterial virus work is quantitative and the conditions of bacterial virus growth can be readily and quantitatively varied.



Molecular model of the two-chain helical structure for DNA proposed by Watson and Crick. DNA is believed to carry hereditary information in the sequences of its four or five component building blocks.

Now, bacterial viruses are of many kinds and qualities. All of them are particles with structure and with some degree of functional organization. A bacterial virus that has been easily the most intensively studied goes by the name of T2.

This virus (shown on p. 22) is about 600 angstroms across and 800 A long, in the head; it has a tail which is about 1000 A long, and has a particle weight of about 250 million. It has, thus, about 1/2000 of the mass of its host bacterium. Structurally, this particle is composed of an outer sheath which is protein in nature, and an inner core which is composed of a small number (of the order of 10) of large macromolecules of a substance called nucleic acid. Each of these nucleic acid macromolecules is of the order of 13 or 14 million in molecular weight. The sheath is differentiated into the head, the tail, and an organ of attachment at the tip of the tail which is the means by which the particle attaches to its host bacterium.

If a suspension of these particles is mixed with a suspension of susceptible bacteria, the particles, in Brownian motion, collide with and attach to the outer membrane of the bacteria by the tail.

At this moment the fate of the bacterium is sealed. About 20 minutes later the infected bacterium will burst, and as it does, some 200 virus particles just like the one that initiated infection will be spewed out. This is a really dramatic result. It was first discovered here at Caltech 20 years ago by Emory L. Ellis and Max Delbruck. In the ensuing two decades, a great many ingenious experiments have been performed to find out what goes on inside the host cell in those 20 minutes.

The more salient results of these investigations have told us that in the first minute or so, in some manner, a hole is produced in the bacterial wall at the site of attachment and the nucleic acid core of the virus is, so to speak, injected into the bacterium. The protein sheath has done its work; it can now actually be dispensed with. The injected nucleic acid brings about, first, its own replication (the production of more virus nucleic acid); then the development of new protein coats; and, finally, the assembly of the mature virus particles that are released when the cell bursts, or lyses.

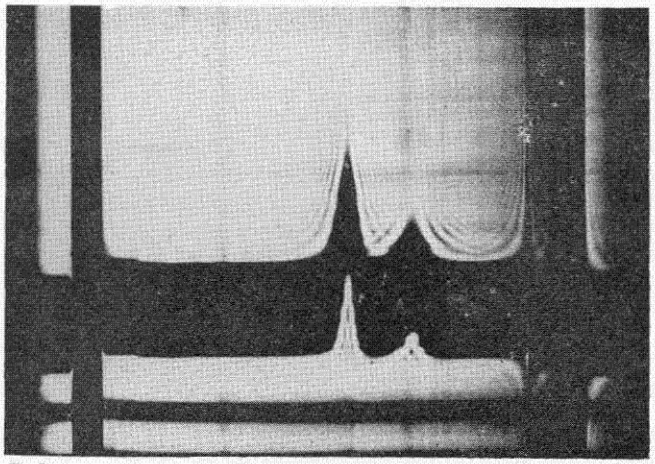
The hereditary plan

Now, these virus particles have definite hereditary traits that are passed on to their progeny. They may or may not attack certain host bacteria; the time of lysis may vary from strain to strain; they may differ in temperature resistance, in resistance to radiation, in their chemical composition, in their antigenic nature. Viruses mutate to new characters; viruses can also be mated by infecting the same cell with several particles; if these have different characters, recombinant types with new assortments of the parental characters are found, indicating the unitary nature of the hereditary pattern. So, these viruses have a hereditary plan that determines both the nucleic acid and the protein component of the progeny. Yet, effectively, only the nucleic acid enters the cell; so, logically, it must carry the hereditary plan. It seemingly must, here at least, be the bearer of genes.

There is not time here to buttress this conclusion, to cite all of the other evidence, direct and indirect, that is available from other organisms to confirm the view that these macromolecules we call nucleic acids are the hereditary factors. For viruses, the evidence seems conclusive.

If the nucleic acid carries the hereditary pattern, in what symbols, in what code, is it written? Here we enter what is at present a realm of pure speculation. Seemingly, the pattern must be expressed in the structure of the nucleic acid molecules. Fortunately, we believe we know the general form of the structure of these macromolecules.

The nucleic acid found in bacteriophages such as T2 consists of two helical chains (shown above) wound about each other and linked to each other by weak but specific bonds. Each chain is a linear sequence of monomeric units called nucleotides, of which there are usually only four or five kinds. For want of a better thought, it is presumed that the information, the hereditary pattern, must reside in the sequence of these four or five kinds of nucleotides along the chain.



Sedimentation velocity pattern of a partially purified ØX174 virus preparation. Sedimentation is from right to left in the analytical ultracentrifuge. The faster-moving peak on the left is the virus. The slower-moving peak on the right represents incomplete virus particles.

While very plausible, there is no experimental evidence for this hypothesis. To go further—to test this hypothesis, to decipher this code—it is clear that we will have to learn the structure of these macromolecules in much greater detail.

Now, in the T2 virus there are some 8 or 10 of these macromolecules, each composed of a double strand containing some 20,000 nucleotides per strand. There is good reason to believe that the 8 or 10 strands are each different and each important. As an hereditary pattern this is still a rather complex object to hope to analyse in detail. A mutation, for instance, *might* (I emphasize *might*) involve a change in only one nucleotide. This would be one part in 400,000.

We have sought, therefore, to find a virus with an even simpler (again I really mean smaller) hereditary pattern. We have looked among the bacterial viruses, for the reasons already mentioned, for the smallest we could find, and we have sought to isolate and study it.

The exotic ϕX174

The virus to which we have in this way been drawn has the somewhat exotic name of ØX174, and we have been devoting the better part of a year to learning how to grow it, how to isolate and purify it, and to establishing its size and some features of its structure.

Many of the technical problems encountered in this work are not very pertinent here. The biological problems associated with growing large quantities of the virus had to be solved, and the chemical problems of isolating relatively purified virus from the crude lysate while preserving the viability of the particles, caused us some pause. But, in time, we produced a highly purified preparation which we then examined in the analytical ultracentrifuge—a device which enables us to measure the rates of sedimentation in a strong centrifugal field of the component or components of a suspension.

Our preparation at this stage had two major macromolecular components, one of which had about 1½ times the sedimentation rate of the other in the centrifugal field (as shown at the left). By the use of a partition cell—that is, a centrifugal cell with a porous plate across its middle—and then by making assays of the virus titer remaining above the partition after various times of centrifugation, it was possible to correlate the virus infectivity with the more rapidly sedimenting component. The slow component was not infective.

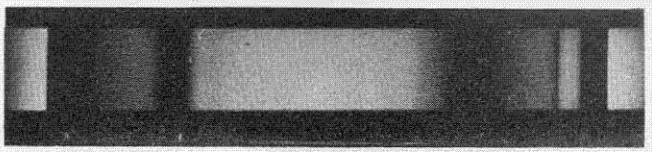
Seeking pure virus

To obtain pure virus, it was now necessary to separate these two components. This proved extraordinarily difficult to do. The slower component displayed essentially identical properties to those of the faster moving component in regard to all of the readily available chemical methods of fractionation. Indeed, this is not surprising, as we are now quite sure that the slow component is related to the virus; it is, in fact, an incomplete virus—the protein shell, lacking most of the internal nucleic acid.

Small amounts of pure virus could be obtained by straight differential centrifugation, but this was a Pyrrhic method involving great losses of material. Fortunately, at this stage, the newly developed method of density gradient centrifugation—developed here at Caltech by Jerome Vinograd, research associate in chemistry, and Matthew Meselson, research fellow in chemistry—came to our rescue. In this technique a dense salt solution is spun at high speed until a stable density gradient is established in the solution. Because of the centrifugal field, the salt tends to move centrifugally, but the resultant concentration gradient sets up a centripetal diffusion flow and at equilibrium a density gradient will exist from the inner to the outer radius of the cell. This is quite analogous to the density gradient in the atmosphere.

If particles are suspended in this stratified salt solution they will, of course, seek their own density. Because of diffusion, however, the particles will not all come to exactly that level corresponding to their density, but will form a band about this level. It may be shown that the width of this band is inversely related to the molecular weight of the particle.

Fortunately—as is shown below—it turned out that the



A B

Density gradient sedimentation of a partially purified ØX174 virus preparation in the analytical ultracentrifuge. Density increases from right to left. The dark band A indicates, by absorption of the ultraviolet light, the position of the virus particle. The dark band B indicates the position of the less-dense incomplete virus particles.

two centrifugal components of our preparation had different densities. The virus has a density of 1.40 and forms a band near the outer edge of the cell (the band may be detected by the ultraviolet absorption of the particles), while the slower sedimenting component has a density of 1.32 and forms a band near the *inner* edge of the cell. These bands are now spacially separated and the two components can be obtained physically separate.

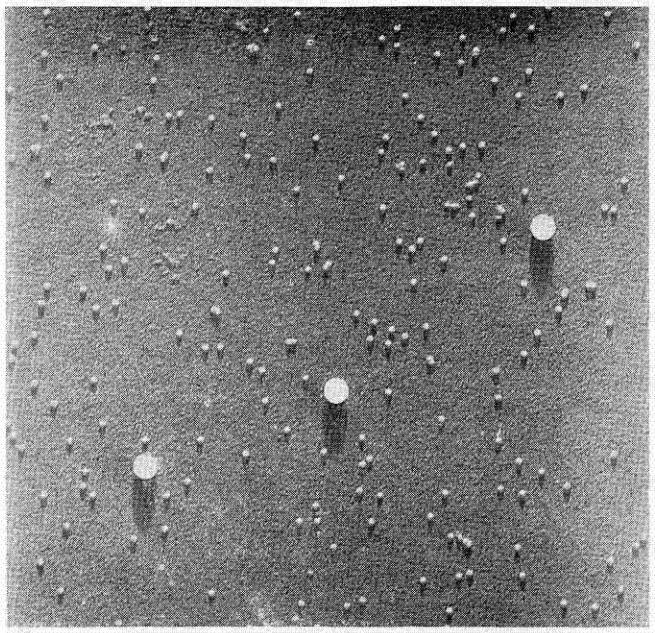
Incomplete virus

We may note that the lighter band is broader than the virus band, indicating a lower particle weight. This is in accord with the idea that it is an incomplete virus.

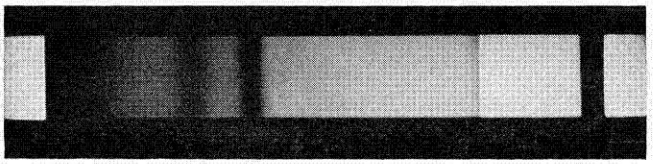
Having obtained a pure virus preparation, the particle weight of the virus could now be determined by another physical technique—that of light scattering. If a suspension of particles is illuminated by a parallel, monochromatic beam of light, of wave-length 10 or more times greater than the dimensions of the particle, the amount of light scattered at 90° to the beam will depend only upon the number of particles and their effective refractive index. The latter can be measured in a separate experiment, so a measure of the number of particles present in a given suspension can be obtained from a scattering experiment. If the mass of virus present in the suspension is known, then the mass per unit particle is readily calculated.

The particle weight of this virus (which is pictured below), calculated from the light scattering data, is only 6.2×10^6. This is about 1/40th of the weight of the T2 virus.

As was expected, this virus has nucleic acid; chemically it can be shown to be about 25 percent by weight



Electron micrograph of ∅X174 virus particles. The three large spheres are polystyrene latex of known diameter. Magnification is about 23,000X.



A B

Density gradient sedimentation of a mixture of DNA from ∅X174 and T2. Density increases from right to left. The dark band A indicates the position of the DNA of ∅X174, while the dark band B indicates the position of the less-dense DNA from the T2 virus.

of nucleic acid. The important possibility then arises that this nucleic acid is present as a single molecule of weight 25 percent of 6.2×10^6 or 1.6×10^6. To test this, the nucleic acid was extracted from the virus and its molecular weight determined by light scattering. The molecular weight of the nucleic acid is found to be 1.8×10^6. Thus the virus has but one molecule of nucleic acid, only 5500 nucleotides. The nucleic acid content is thus 1.5 percent of that of the T2 virus previously discussed.

A surprising feature

It often happens in physics that when the scale of a particular phenomenon is changed by an order of magnitude or two, new features appear and, somewhat to our surprise, this may have happened in this instance. In the density gradient sedimentation picture above, two bands (A and B) appear. One is produced by DNA from the ∅X virus. The other is produced by the addition of some nucleic acid of the T2 virus. The latter band is clearly narrower than the band produced by the ∅X nucleic acid, indicating it is bigger in molecular weight, but also this band is displaced. The nucleic acid from ∅X is significantly denser than that from the T2 virus. Since all the nucleic acids previously studied—from bacteria, from salmon sperm, from calf thymus glands—have been almost exactly the same density as that of the nucleic acid from T2 virus, there seems to be something unique about the nucleic acid from the small ∅X virus.

It must be said that we have never had a nucleic acid of this low a molecular weight, and it is possible that this altered density is an artifact resulting from the action of our isolation procedure upon such a singularly small nucleic acid. There is, however, supporting radiobiological evidence for the view that this nucleic acid is in some way unusual even when inside the virus.

Here the prologue ends. We have established the size of this virus and the unimolecular nature of its nucleic acid. The virus is almost two orders of magnitude less complex than the previously studied T2. I feel our search for simplicity has been successful and can stop here. We can, and have, isolated mutants of this virus; we can almost surely create other mutants when needed. It is time to begin to examine the structure and the functions of these particles in the most intensive way.