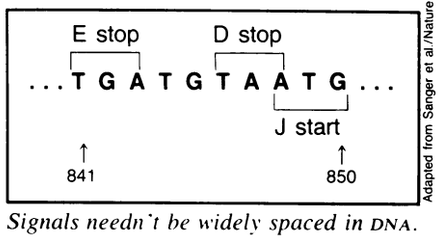
Full Gene Sequence of DNA Virus Solved

Scientists have just solved a new puzzle, the complete genetic information of a DNA-containing virus. Fitting together analyses of more than a hundred DNA fragments, Fred Sanger and his co-workers have determined the exact sequence of the 5,375 nucleotides making up the DNA strand of bacterial virus phiX174. The work, done in Cambridge, England, at the Medical Research Council, is published in the Feb. 24 NATURE. The entire sequence takes up two and a half printed pages in the journal.

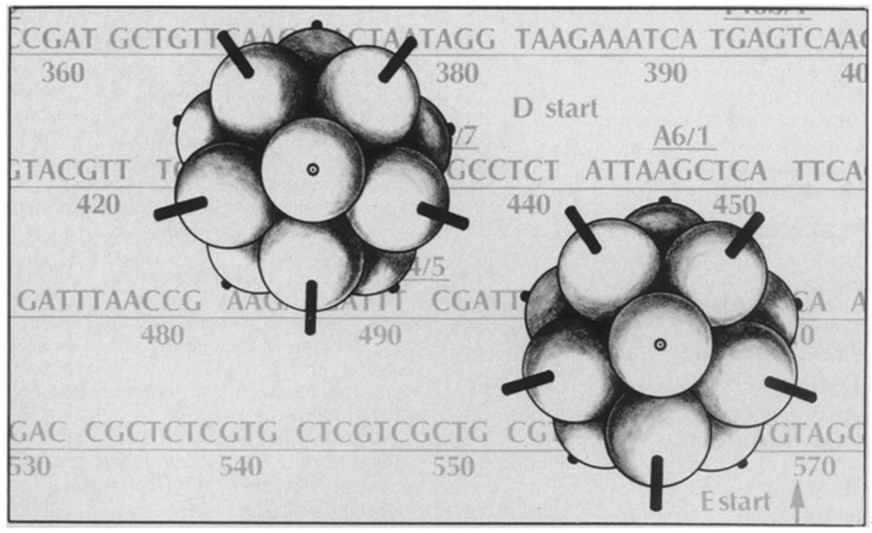
PhiX174 is the second living organism for which the entire genetic structure has been determined. Last April, Walter Fiers and co-workers at the University of Ghent in Belgium reported the 3,569-nucleotide RNA sequence of bacterial virus MS2. That virus has only three genes. The phiX174 has nine.

Sequencing of the entire DNA took less than two years, J. C. Fiddes, one of Sanger's collaborators, said at a symposium on genetics in Park City, Utah, this week. The work was possible because new techniques allowed rapid analysis of DNA. Sensitive gel electrophoresis can separate DNA fragments, up to 150 nucleotides long, that differ in length by only one nucleotide. The British researchers could immediately determine the nucleotide sequence of a piece of phiX174 DNA by examining patterns of bands in a set of gels containing fragments prepared under different conditions.

The nucleotide sequence of phiX174 provides both confirmations and surprises for those who study the structure of genetic information. Basically, the detailed description conforms to the canons of molecular biology. All known proteins were represented in that circular, single-stranded DNA by series of three nucleotide groups (called codons), and the key for relating the codons and amino acids in the proteins was the same as that described in other organisms.



Signals needn't be widely spaced in DNA.



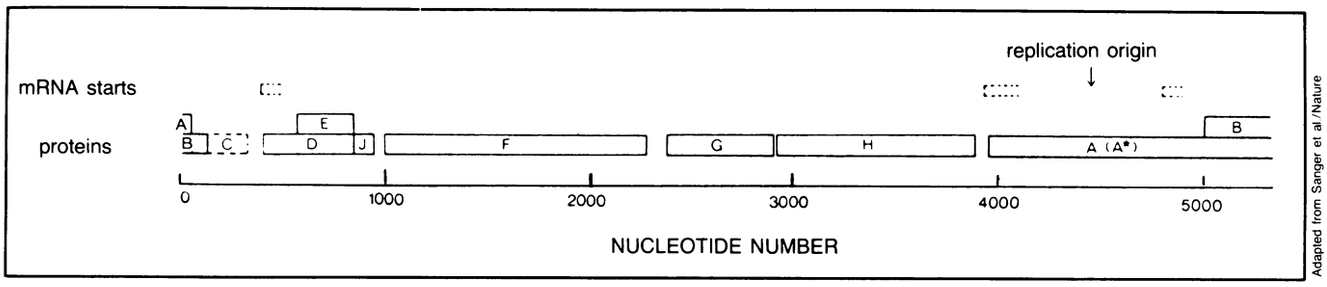
PhiX174 virus and a portion of its just-elucidated sequence of 5,375 nucleotides.

The biggest surprise from the sequence was first reported a few months ago. The phiX174 virus, with its nine genes, can use the same stretch of DNA to code for two different proteins—and it does so at least twice (SN: 11/13/76, p. 310). Sanger's detailed analysis revealed that the gene for protein E is entirely inside the gene for protein D, and the gene for protein B is within that for protein A. Different amino acids result because the DNA is translated in different reading frames. Beginning at different start signals, protein synthesis groups the nucleic acid letters into different codon words. The nucleotide sequences of mutant viruses defective in the A, B, D or E proteins further confirmed that the genes overlap.

The nucleotide sequence also hints at how the overlapping genes may have evolved. In phiX174, the four nucleotides of DNA are not used randomly, Sanger points out. Thymidylic acid appears more often than any of the others and is far more likely to be the third nucleotide in a codon. In gene E, however, only 14.3 percent of the codons end in thymidylic acid, and only 15.8 percent do so in the overlapping region of A (compared to 42.1 percent in D and 34.2 percent in B). This difference indicates that D and B genes arose first, and that later mutations resulted in E and in an extension of A.

"The most striking feature of the phiX DNA sequence is the way in which the various functions of the genome are compressed within the 5,375 nucleotides," the researchers find. This compactness provided yet another surprise. The three-nucleotide signals for the end of one protein and for the start of the next need not be separated by long stretches of nongene DNA. The signal terminating the D protein actually shares its final nucleotide with the start signal for protein J, and the stop signal of gene A probably overlaps one nucleotide with the start signal for gene C. Because a similar overlap has been observed between two bacterial genes, these ways of concentrating information into DNA may be relevant to organisms beyond tiny viruses.

Biologists have also looked to the DNA sequence for clues to the mechanism of protein synthesis. Ribosomes, the cell organelles that link the appropriate amino



Map of the phiX174 virus: Genes and overlapping genes fill length of DNA molecule.

acids into protein molecules, contain special RNA. Researchers have suggested that before protein synthesis, nucleotides of ribosomal RNA must bind temporarily to nucleotides of messenger RNA, the intermediary between DNA and protein formation. When Sanger and colleagues examined the nucleotide sequences before the start signal of each phiX174 protein, they found that messenger RNA would have a stretch of at least four nucleotides that could bind with the nucleotides at one end of the ribosomal RNA. The sites where ribosomes bind have also been isolated from pieces of messenger RNA both in Sanger's laboratory by Nigel L. Brown and Michael Smith and by Rockefeller University biologists Jeffrey V. Ravetch, Peter Model and Hugh D. Robertson. Two articles also in the Feb. 24 NATURE describe the sequences of these sites and locate the sites on the overall phiX174 map. In all these ribosome binding sites there is a sequence of nucleotides that should be able to form stable base pairs with the end of the ribosomal RNA. The number of these possible base pairs, however, does not reflect how frequently ribosomes bind to each site and make the appropriate protein, so other factors must be involved in control of protein synthesis.

The sequence has certainly not answered all questions. For example, the researchers learned little about what is special about the position on the DNA molecule, the replication origin, where copies of the viral DNA always begin. Sanger found that this origin falls within the gene for protein A, but that area showed no obvious symmetry or special shape.

Furthermore, the analysis opened the possibility of there being even more phiX174 proteins. "With the presence of two pairs of overlapping genes, the genome has more coding capacity than had been originally supposed on the assumption that each gene was physically separate," Sanger says. He points out that there are sites in the A, F, G and H genes where sequences for other proteins could possibly start.

Nobody knows whether the mechanisms uncovered by solving the phiX174 sequence will also appear in organisms that are less limited in total DNA. But molecular biologists now have new phenomena to watch for and plenty to think about. Everything the virus phiX174 is or does should be explicable beginning with that DNA sequence. □

One man's Mars: No Martians

For months, the Viking project's biologists have been struggling to find a theory that would explain the seemingly conflicting results of their experiments on the Martian surface. Now the first such theory has been offered, and, says Vance Oyama of the National Aeronautics and Space Administration's Ames Research Center, it shows "no need to invoke biological processes."

It is far from universally accepted, and it requires invoking several specialized chemical processes such as catalysis and polymerization. But, says Oyama, it fits all the facts, and elegantly so. Furthermore, if Oyama's theory is correct, it bears not only on the biology-instrument results, but also on the nature of the magnetic material that clings to Viking's test magnets, as well as on atmosphere-surface interactions and even on the early evolution of the planet.

The theory begins with the simplest of photochemical effects in the atmosphere—the dissociation of the dominant carbon dioxide by solar ultraviolet radiation into "activated" carbon monoxide and single atoms of oxygen. With more UV, some of the CO is further dissociated into carbon and oxygen. The carbon then combines with CO to form carbene (C_2O), which in turn combines with still more CO to form the first key element in Oyama's proposal, carbon suboxide, or C_3O_2. The carbon suboxide molecules, Oyama says, can then unite into a carbon suboxide polymer, $(C_3O_2)_N$. All of these reactions, he adds, are known from laboratory studies, and the resulting polymer has a reddish cast very much like that of the Martian surface.

The next step is to plug this result into the data from the three different kinds of biology experiments aboard the landers. Oyama's explanation begins with the pyrolytic-release instrument—the one of the three that has suggested reduction processes.

The PR instrument works by exposing a soil sample to an atmosphere that includes CO and CO_2 whose carbon atoms are radioactive carbon 14. After an incubation period, the sample is incinerated, or pyrolyzed, and the resultant gases are monitored for radioactivity as an indication of how much of the CO and/or CO_2 was incorporated into the soil.

During the design of the Viking experiments, it had been suggested by some observers that the radioactive carbon isotopes might kill off the very microorganisms that they were trying to measure. Viking's biologists have long felt that the radioactivity levels are too low to do any damage, and Oyama sees no evidence of life to kill anyway, but the isotopes do play an important role in his explanation of the PR results.

The decay of the carbon 14 (into nitrogen 14) releases a beta particle, he points out, and the energy of this decay product (0.156 MeV, Oyama says), multiplied by the 22 microcuries of radioactivity in each PR "cycle," is more than sufficient to fracture carbon-carbon, carbon-hydrogen or carbon-oxygen bonds. This breakage has the effect of "activating" the red carbon suboxide polymer, making it "receptive" to incorporating the provided carbon monoxide. When the polymer is heated to 625°C, it yields about 4 percent of itself as the original, but now labeled, monomer form, which sticks to the experiment's organic-vapor trap and appears, upon further heating, as the critical "second peak" in the experiment's data.

If water vapor is present in the PR experiment when the sample is exposed to the labeled atmosphere, says Oyama, this second peak is lower. In Oyama's own "gas-exchange" experiment, which exposes a sample to a *non*labeled nutrient solution and simply monitors changes in the composition of the test-cell atmosphere, the prominent release of oxygen is also reduced. But the reason, Oyama says, is very different.

In the Martian atmosphere, the same photodissociative reactions which lead to the formation of carbon suboxide also lead, by another path, to activated oxygen

The closest look yet at Phobos

Viking orbiter 1 has provided the closest look ever taken at the Martian moon Phobos (see cover), revealing a wealth of detail and enabling the first accurate calculations of the tiny satelliate's density. A very preliminary look, says Joseph Veverka of Cornell University, suggests about 2 grams per cubic centimeter, consistent with earlier speculation that Phobos resembles carbonaceous chondritic material. Veverka and others feel that this is evidence that Phobos (and Deimos) may be captured objects originally formed far out in the asteroid belt between Mars and Jupiter. The photos have shown many small, bowl-shaped craters which, combined with the flatter floors of larger craters, should yield an estimate of the depth of the planetoid's regolith, or fragmented surface, with substantial implications for its evolution and history. One of Phobos's major mysteries—a series of strange, parallel striations crossing part of the surface—shows up clearly in the closest photos as grooves, about 100 to 200 meters wide and tens of kilometers long, rather than as "chains" of small craters. (Such chains have also been seen, suggesting secondary ejecta from larger craters.) The grooves do not appear on every side. This implies that they are symptoms neither of Mars-caused tidal distortions nor of strata seen edge-on. They could be related to the formation of Phobos's largest crater, Stickney. This will seem even more likely if the other moon, Deimos, has none. □