ΦX174 phage

φX174 (phiX174) is a virus that infects the bacterium E. coli. Hence φX174 is a **bacteriophage**.

The **phi X 174** or **ΦX174** bacteriophage was the first **DNA-based genome to be sequenced** by F. Sanger and his team in 1977.

Fiers and Sinsheimer had demonstrated the physical, covalently closed circularity of phi X 174 DNA in 1962.

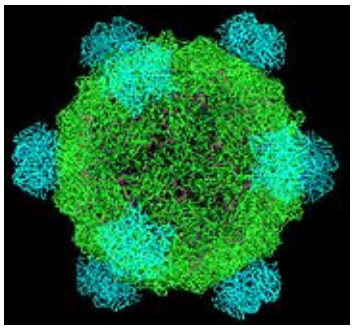
Nobel prize winner Arthur Kornberg used phi X 174 as a model to first prove that DNA synthesized in a test tube by purified enzymes could produce all the features of a natural virus, ushering in the age of synthetic biology.

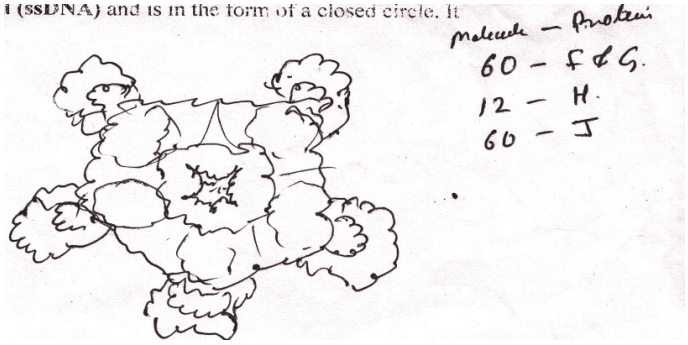
In 2003, it was reported by Craig Venter's group that the genome of ΦX174 was the first to be completely assembled *in vitro* from synthesized oligonucleotides.

Structure:

This bacteriophage has a [+] circular single-stranded DNA genome of 5386 nucleotides encoding 11 proteins. Of these 11 genes, only 8 are essential to viral morphogenesis. The GC-content is 44% and 95% of nucleotides belong to coding genes.

Each complete infectious particle (virion) of φX174 consists of a protein coat which envelopes a core that contains both protein and DNA. The coat of the virus contains 60 molecules each of two proteins (F and G) and 12 molecules of another protein (H). The core of the virion contains one molecule of DNA and 60 copies of a fourth protein, the J protein. **The DNA molecule is single-stranded (ssDNA)** and is in the form of a closed circle. It contains **5386 nucleotides**.



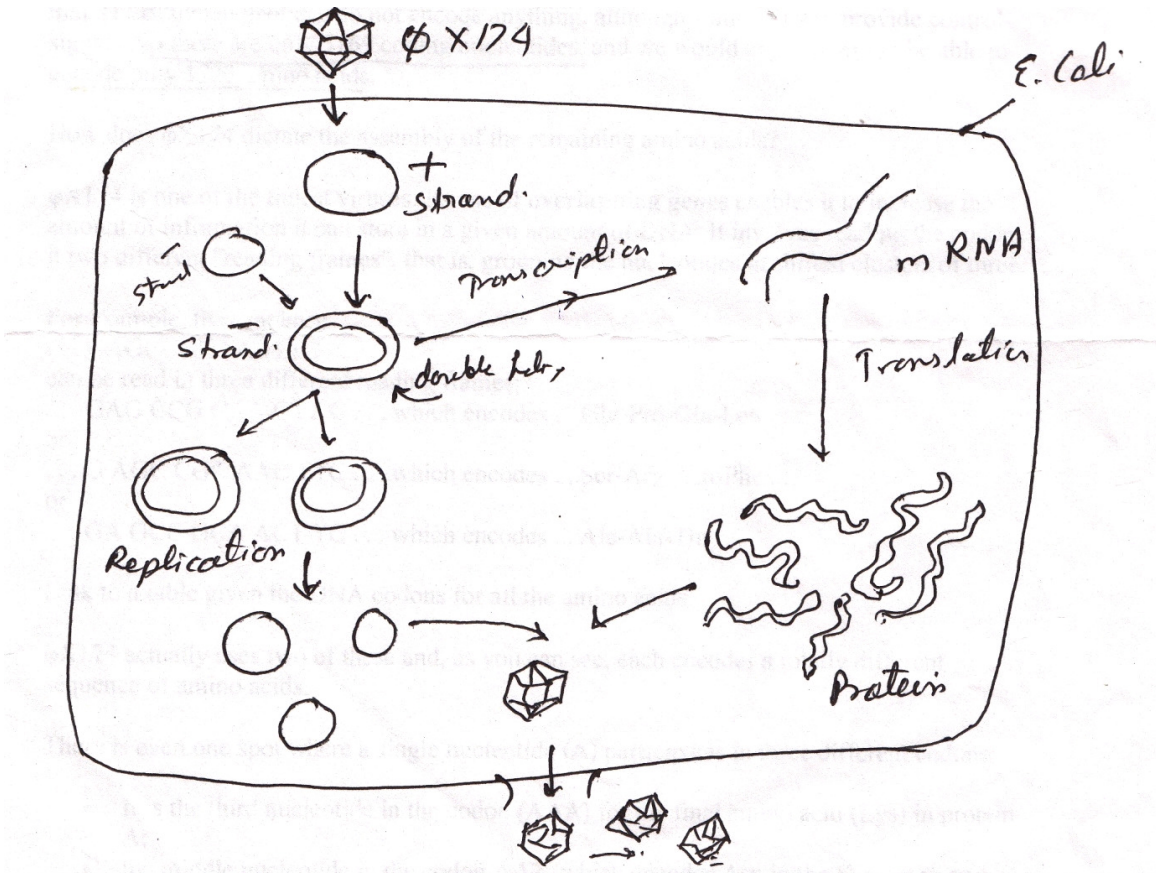


Protein	Number of Amino Acids	**Function**
A	513	Stage II and stage III DNA replication

A*	341	An unessential protein for viral propagation. It may play a role in the inhibition of host cell DNA replication and superinfection exclusion
B	120	Internal scaffolding protein, required for capsid morphogenesis and the assembly of early morphogenetic intermediates. Sixty copies present in the procapsid
C	86	Facilitates the switch from stage II to stage III DNA replication. Required for stage III DNA synthesis
D	152	External scaffolding protein, required for procapsid morphogenesis. Two hundred and forty copies present in the procapsid.
E	91	Host cell lysis
F	427	Major coat protein. Sixty copies present in the virion and procapsid
G	175	Major spike protein. Sixty copies present in the virion and procapsid
H	328	DNA pilot protein need for DNA injection, also called the minor spike protein. Twelve copies in the procapsid and virion
J	38	DNA binding protein, needed for DNA packaging. Sixty copies present in the virion
K	56	An unessential protein for viral propagation. It may play a role optimizing burst sizes in various Hosts
Total	1986	

Multiplication:

When φX174 attaches to its host, its ssDNA molecule is inserted into the cell. Here the DNA strand (+) serves as the template for the synthesis of a complementary (−) strand. The two strands form a double helix, which then replicates itself several times.



The minus strands of these DNA molecules then serve as templates for the synthesis of

- mRNA molecules.
- some 200 complementary (+) strands of DNA, each of which will later be packaged into the core of a new virion.

  The protein-synthesizing machinery of the host cell translates the viral mRNA molecules into 11 different kinds of proteins. Four of these are the four (F, G, H, and J) that will be incorporated into new virions. As for the other 7 proteins

- A, A*, and C play roles in the replication of viral DNA;
- B and D assist in the assembly of the virion proteins into new virions;
- E lyses the host cell so the newly-synthesized virions can escape;
- K boosts virion production;

but none of these proteins become part of the virion.

The 11 proteins encoded by φX174 DNA range in size from the A protein, which contains 513 amino acids, to the J protein, which contains only 38. The 11 proteins

together contain a total of 1986 amino acids (the A* protein is simply a shortened version of the A protein). This raises a question. With 3 nucleotides needed to specify one amino acid, φX174 would need 5958 nucleotides to encode 1986 amino acids (5958/3 = 1986). But its DNA molecule contains 5386 nucleotides, only enough to encode 1795 amino acids. Furthermore, it turns out that 217 of the nucleotides do not encode anything, although some of them provide control signals. So there are only 5169 coding nucleotides, and we would expect them to be able to encode only 1723 amino acids.

How does φX174 dictate the assembly of the remaining amino acids?

φX174 is one of the tiniest viruses. Its use of **overlapping genes** enables it to increase the amount of information it can store in a given amount of DNA. It involves reading the codons it two different "reading frames", that is, grouping the nucleotides in shifted clusters of three.

For example, the sequence
. . . GAGCCGCAACTTC . . .
can be read in three different reading frames:
. . . GAG CCG CAA CTT C . . . which encodes . . Glu-Pro-Gln-Leu . .
or
. . . G AGC CGC AAC TTC . . . which encodes . . Ser-Arg-Asn-Phe . .
or
. . . GA GCC GCA ACT TC . . . which encodes . . Ala-Ala-Thr . .

Link to a table given the DNA codons for all the amino acids.

φX174 actually uses two of these and, as you can see, each encodes a totally different sequence of amino acids.

There is even one spot where a single nucleotide (A) participates in three different codons:

- It is the third nucleotide in the codon (AAA) for the final amino acid (Lys) in protein A;
- the middle nucleotide in the codon AAT, which encodes Asn in the K protein; and
- the first nucleotide in ATG, the codon that places methionine (Met) at the start position of protein C.