Self-Assembly of Biological Structures

D. J. KUSHNER

Department of Biology, University of Ottawa, Ottawa, Ontario, Canada

INTRODUCTION

Much of contemporary biological thought is concerned with information transfer: with the "central dogma" that links DNA, RNA, and protein; with the universality of the genetic code and colinearity between nucleotide and amino acid sequences; with determining how polypeptide chain synthesis starts and stops. A polypeptide chain may contain an impressive amount of information. Once the primary structure is established, the chain can usually take on higher configurations without further directions from the genome. That is, genetic control of primary structure determines secondary, tertiary, and quaternary structure.
As regards secondary and tertiary structure, ribonuclease and other proteins can be unfolded by breaking disulfide bridges or by other treatments and can then resume in vitro their native, active form (52). The molecular weights (MW) of individual polypeptide chains usually do not exceed 50,000, and those of MW greater than 70,000 are rare (114, 218). Proteins of higher MW have quaternary structure; that is, they contain more than one folded polypeptide chain, usually linked by noncovalent bonds, hydrogen, hydrophobic, or ionic (218). [For a discussion of hydrophobic bonds, consult references 126 and 218 and current texts of biochemistry. It has often been pointed out that the so-called “hydrophobic bonds” between nonpolar groups depend less on the attraction between such groups (van der Waal’s forces) than on their mutual avoidance of water to make a thermodynamically more stable configuration.] The association and dissociation of these chains is greatly influenced by the environment. To cite but two examples among many (reviewed in references 85 and 218), hemoglobin is made of two pairs of polypeptide chains; oxygen binding increases the attraction between these chains, with important physiological consequences (95, 177). The glutamic dehydrogenase of beef liver, of MW 2,000,000, undergoes reversible stepwise dissociation into eight enzymically active subunits, and these can be dissociated by urea into identical polypeptide chains, of MW 50,000, that are enzymically inactive but become active on reassociation. Association-dissociation equilibria (and enzymatic activity) are influenced by steroids, nucleotides, and inorganic ions (218).

At the same time that the quaternary structure of proteins was becoming understood, the following questions were asked for many biological systems. Once information is given for specifying structural proteins and enzymes (which can catalyze the formation of other compounds), what more is needed to build recognizable cell structures? To make a cell organelle, a flagellum, or a virus, is information required beyond that which specifies the component molecules, or can such molecules spontaneously assemble into larger structures under the proper environmental conditions?

The idea that large structures can most efficiently be built up by repetition of smaller ones is teleologically attractive. First, it reduces the amount of information required. This conservation may be necessary: the nucleic acids of some small viruses contain only enough information to specify two or three protein molecules (35, 73). Second, in making anything that contains a large number of small subunits, whether we consider these as the amino acids that go into the protein coat of a virus or the individual nuts, bolts, wires, and so on that make up an automobile, there is much less chance of error if a subassembly process is used and the parts thus made are placed in the final product. This assumes that any defective subunit, whether protein or taillight, will be rejected before final assembly. Without such a check for accuracy, there are very few chances of making a perfect final product. Caspar (32) calculated that if 1,000 identical proteins of 100 amino acid residues each were combined into a virus shell and if there were a 1-in-10,000 chance of error in assembly at each step, there would be about a 90% chance of making a correct virus. However, with the same probability of error at each step the chance of building up a single protein molecule of this size (108 amino acid units) would only be 0.005%.

Such a calculation implies the same probability of error in fitting a protein monomer into a larger structure as in fitting an amino acid into a polypeptide chain. It is difficult to predict whether this would be so. However, as will be seen below, most polymerizations that form large structures are equilibrium processes, tending to the lowest free energy, and do not involve covalent bond formation. It would be relatively much easier to correct a wrong protein subunit than a wrong amino acid. The first would simply require dissociation of part of the structure and subsequent reassociation with a proper subunit. The second would require the enzymatic breakdown and resynthesis of the whole chain.

The automobile analogy and some of the above calculations were suggested by Crane (39) in 1950 in a remarkable paper on principles of biological growth, which also forecast from quite simple structural considerations that most biological rod-shaped structures would turn out to be helical in construction (see Fig. 1, 5). Crane pointed out that subassembly stages are more efficient in permitting more hands, or machines, to work on the automobile at once, and the same argument must apply to biological structures. It is easy to imagine how slowly, say, the tobacco mosaic virus could be formed if all the protein in its coat had to be synthesized at one place in the infected cell.

There is now very little doubt that some structures in living cells are formed by self-assembly of subunits. This has been most clearly shown for certain viruses, with which many of the ground rules and basic definitions of self-assembly were first worked out. The structural regularity of some viruses and the relative simplicity of their genetic material provide the clearest examples of how gene products are assembled into biological structures. However, self-assembly has also been proposed to account for the formation of such
diverse structures as membranes, microtubules (which may be the essential structures in flagella of eucaryotic cells, mitotic apparatus, and other structures), ribosomes, strands of muscle and connective tissue, bacterial flagella, and pili.

The most detailed recent treatments of the formation of biological structures by self-assembly are found in the 1966 Ciba Foundation Symposium (241), in the 1966 Symposium of the Society for General Physiology (cited in reference 38), and in the 1967 Symposium of the International Society for Cell Biology (234). These books and some of the reviews cited should be consulted for detailed treatment of the topology of viruses and other structures and for the thermodynamics of self-assembly. This review, intended as a broad introduction to current research and thought patterns on the role of self-assembly in forming biological structures, will attempt to present an overall picture of this complex, rapidly expanding, profoundly important subject.

STRUCTURE AND ASSEMBLY OF THE SIMPLER VIRUSES

Principles of virus construction are discussed in detail in several reviews (see 33, 35, 87, 115).

In 1956, Crick and Watson made the basic suggestion that small viruses, almost all of which appear to be rods or bumpy spheres in the electron microscope, are formed of identical protein subunits packed together in a regular way to provide a protective shell for the nucleic acid (40). After Caspar's (31) demonstration that the bushy stunt virus had cubic symmetry, Crick and Watson (40) suggested that many small spherical viruses would prove to have such symmetry. If a polyhedron (or a structure made of subunits packed regularly around a central point) has cubic symmetry, its vertices or subunits lie on at least one set of threefold axes arranged as the diagonals of a cube. If an arrangement of subunits possesses cubic symmetry, each subunit can be said to lie on a sphere (115), and the arrangement may be roughly spherical in appearance. Viruses with such symmetry are often called "spherical" and the term will be thus employed here.

Solids with cubic symmetry include the regular tetrahedron (3:2 symmetry), the cube (4:3:2 symmetry), the octahedron (4:3:2 symmetry), the dodecahedron (5:3:2 symmetry), and the icosahedron (5:3:2 symmetry). For example, a tetrahedron has axes of threefold and twofold rotational symmetry, axes around which the body can be rotated to three or two positions without change of appearance—similarly for the other symmetries given. The same symmetry patterns are found in derivatives of these bodies. Of these solids, the icosahedron and its derivatives seem to be used exclusively for the spherical viruses; the octahedron may be used for certain bacteriophage heads (23).

During the next few years, electron microscope and X-ray diffraction studies showed that spherical viruses from such different sources as plants, mammals, insects, and bacteria had icosahedral (5:3:2) symmetry (87). Caspar and Klug (35) showed that the icosahedral design led to the minimal energy arrangement of a closed shell built of regularly bonded, identical subunits. This design would be expected if subunits assembled spontaneously to build the shell of a spherical virus. The idea that such shells are built by self-assembly of the subunits is now central to studies of virus structure and synthesis, and it guided these studies even before it was demonstrated directly.

Subunits, Capsomeres, and Capsids

The word subunit has been given different meanings by students of proteins and larger structures. The subunit of a large protein molecule is often considered smaller than a protein but larger than a single polypeptide chain (218). Membrane subunits, if they exist, may be very large and complex indeed (see below). Crick and Watson (40) considered subunits, in their proposed building scheme, not necessarily as single protein molecules but rather as asymmetric units which could be used to build the shell of a spherical virus. They could consist of several identical protein molecules joined in some asymmetric manner (see Fig. 8) or of an aggregate of different protein molecules.

The individual protein monomers that make up a virus coat are now usually spoken of as subunits, and the term will be so used in the following discussion.
Detailed discussions of the problems involved in arranging subunits on an icosahedral surface are contained in references 14, 32-33a, 35, 87, and 115. Most commonly, subunits are arranged in pentagons (pentamers) and hexagons [hexamers (Fig. 3)]. The number of pentamers is always 12. The number of hexamers depends on the type of virus. Thus, the spherical bacteriophage ϕX174 has only 12 pentamers; its surface is covered with 60 subunits, each in exactly equivalent position (Fig. 2, 3). The turnip yellow mosaic virus (TYMV) and other small plant viruses have, in addition to 12 pentamers, 20 hexamers (see Fig. 3 and 8); the herpes virus has 150 hexamers and the adenovirus has 240 hexamers (35).

A structure built along similar principles, by arranging pentagons and hexagons on an icosahedral surface, is the geodesic dome, developed by Buckminster Fuller (cited in reference 35). Radar domes are familiar examples of this, as was the U.S. Pavilion at EXPO-67 (Fig. 4).

In viruses, the subunit clusters (pentamers and hexamers), which are the morphological units seen in the electron microscope, are usually called capsomeres; the closed virus shell or tubing (see below) built of a number of structural units is called the capsid. This terminology, originally proposed in 1959 (142) and slightly redefined in 1962 (34), will be followed here. When T-even bacteriophages are considered, the term "head capsid" will be used to describe the empty head.
However, as more has been learned about morphopoiesis, especially that of bacteriophages, the nomenclature has become more detailed and sophisticated, involving not only subunits but also those units that hold together the component parts of large viruses. A detailed system of “Terminology in Morphopoiesis” is soon to appear in Science (E. Kellner, personal communication).

Fig. 3. Models constructed from pentagonal and hexagonal members arranged in 5:3:2 symmetry. All contain 12 pentagons. The total number of members are: A, 92; B, 42; C, 12; D, 32 (87). The arrangement of pentagons and hexagons in C and D correspond, respectively, to those of the phage φX174 (60 subunits arranged in 12 pentagons) and the CCMV and other small RNA-containing viruses (180 subunits arranged in 12 pentagons and 20 hexagons). Reprinted by permission of the copyright owner, Academic Press.

Equivalent and Quasi-Equivalent Positions of Subunits

The analysis of Caspar and Klug (35) showed that if a rod-shaped virus [tobacco mosaic virus (TMV) or tobacco rattle virus] were built by the assembly of a number of identical subunits, so that the design of the coat was determined by the specific binding properties of the subunits, the subunits would assemble in a helical pattern. Helical designs are also characteristic of parts of the tails of certain bacteriophages and probably, as will be seen below, of other rod-shaped structures (e.g., flagella and microtubules). All subunits except those on the ends of straight helices are in equivalent positions. The 60 subunits that make up the capsid of the phage φX174 are also all in equivalent positions. However, the subunits in all viruses with icosahedral symmetry that contain hexamers as well as pentamers are held in a state of bonding termed “quasi-equivalent” (33, 35). The concept of quasi-equivalent bonding arises because it is often impossible to represent minimal energy designs for ordered structures built of equal units with specific binding properties with all these units in exactly equivalent positions. Rather, “it is often possible to form a stable ordered structure in which the bonds between equal units are systematically deformed in a number of slightly different ways. If each unit forms the same type of bonds with its neighbors, then all the units will be quasi-equivalently related. Quasi-equivalence in ordered structures can be defined as any small, nonrandom variation in a regular bonding pattern which leads to a more stable structure than strictly regular bonding” (33). In most icosahedral shell designs, quasi-equivalent bonding is a geometrical necessity (33). Such a design will have the greatest possible number of most stable bonds formed and will,
therefore, be the structure of the lowest free energy.

**Tobacco Mosaic Virus**

The self-assembly of the component parts of the TMV has been studied in much more detail than that of other virus. (References 32 and 128 contain detailed reviews and theoretical treatment of this assembly.) The common strain of TMV is a rod about 300 nm long and 18 nm in diameter. (See Fig. 2, which illustrates the TMV and the bacteriophages φX174 and T4, as well as Fig. 5D.) It consists of an outer helix of pitch 2.3 nm composed of between 2,100 and 2,600 identical protein subunits, each of MW 17,530, and a ribonucleic acid (RNA) chain of MW 2.0 to 2.5 × 10^6. [The range in proteins and RNA does not reflect heterogeneity within virus preparations, but rather the difference between different analytical methods and interpretations (128).] The RNA chain is coiled in a helix 8.0 nm in diameter along the inner ends of the protein subunits, with three nucleotides per subunit (Fig. 5D). Each subunit makes the same bonds with its neighbors and, hence, all (except those on the ends of the virus particle) are in equivalent positions. In certain variant virus strains, the packing of protein subunits is slightly different and they may be in quasi-equivalent positions (32, 115).

TMV can be dissociated into protein and RNA by many agents, including concentrated acetic acid, urea, sodium dodecyl sulfate, phenol, heat, and pressure. The protein obtained by mild alkali degradation (e.g., at pH of about 10), called the A protein, has been used in most reconstitution studies. This protein is a cyclic trimer of the 17,530 MW subunit (see Fig. 7). It dissociates reversibly into these subunits on high dilution, at pH 13, or in 67% acetic acid.

The effects of salt on protein-RNA and protein-protein dissociation, together with other evidence, suggest that TMV protein is bound to its RNA by salt links and that protein molecules are bound to each other by hydrophobic bonds (32, 33, 128).

The A protein and RNA, under proper conditions of temperature and pH, can spontaneously reassemble into virus particles that are indistinguishable from the original TMV both morphologically and in biological activity (59, 60). Under the best conditions, up to 80% of the original virus activity can be recovered after degradation and reconstitution (59). In the absence of RNA, the dissociated protein units may polymerize to give a helical structure, morphologically indistinguishable from the intact virus although, of course, noninfectious. The proteins can also polymerize as double discs containing 32 subunits, and these discs can further polymerize in stacks to form long rods similar to but distinguishable from the helical structure (32, 128).

**Polymerization of TMV subunits and other proteins.** Polymerization of TMV protein should be considered in the context of what is known of the formation of other helical structures from protein subunits. Detailed physico-chemical treatments of the process of building up long protein aggregates by condensation of smaller subunits in a manner involving the same basic binding pattern between subunits are found in reviews by Caspar (32, 33), Lauffer and Stevens (128), and by Oosawa's group (170-172).

The simplest examples of polymerization studied are the globular-fibrous (G-F) transformation of actin and the formation of flagella from flagellin molecules. (See below and also Fig. 5, which shows proposed helical structures of F actin and of flagella.) These phenomena somewhat resemble crystal formation in solution, or gas condensation. Simply stated, no polymerization takes place until a certain critical subunit concentration is reached. Then polymerization may occur very quickly indeed: with actin, the initial rate of polymerization is proportional to the third or fourth power of actin monomer concentration. Adding small fragments (seed fragments) of actin or of flagella greatly increases the rate of polymerization. At equilibrium, a mixture of monomers and long polymers exists.

Even if each subunit in a polymer has only a slightly lower free energy in the bound state than in the free, polymerization can occur. Conversely, very small changes in monomer free energy, which might be brought about by altering the tempera-
ture, pH, pressure, or by allosteric effects, can reverse polymerization (33a).

The polymerization of TMV protein appears to be more complex than that of actin or flagellin. It may take place in discrete steps and is influenced by the presence of RNA.

Ionic strength, pH, and other environmental conditions interact in determining whether TMV protein polymerization occurs. Briefly, raising ionic strength increases polymerization (32, 128). At pH 5 and below, the A protein is polymerized at all temperatures studied; raising the pH leads to depolymerization. In neutral or slightly alkaline pH values, the protein is depolymerized, but it can polymerize in the presence of TMV RNA (32, 128).

Lauffer et al. (127) observed that raising the temperature favors aggregation. In going from 5 to 30°C, in a solution of 0.1 ionic strength (pH 6.5), subunits were changed in stages from the triplet form (A protein) to complete rods. Both enthalpy and entropy increase when TMV protein polymerizes, even though the order in the polymer increases. The thermodynamic changes are thought to result from a decrease in the ordered structure of water caused by transferring the surface of the TMV protein subunits from an aqueous (free) to a more organic (polymerized) environment (128).

As will be seen, the theme of increased entropy accompanying polymerization runs through many self-assembly studies.

Models of TMV assembly. Two proposed models of TMV assembly (32, 128), based mainly on kinetic and thermodynamic observations, will be described briefly. Both consider the subunit as an ellipsoid 7.0 nm long and about 2.5 nm in its greatest diameter (see Fig. 5D, 7).

Caspar's (32) treatment assumed, to simplify the calculation, that each protein subunit is bonded equally strongly to its six nearest neighbors (see Fig. 5D). The analysis that followed suggested that TMV protein aggregates by passing through stable intermediates of increasing size, the first being the A protein. It was postulated that subsequent aggregates would contain 7 subunits, 32 subunits (a two-turn disc, apparently corresponding to the double discs actually observed) and 48 to 49 subunits (three-turn helix), and that linear polymerization of these segments would lead to a completed rod.

Caspar suggested that the sides of the elliptical protein subunit contain nonpolar residues that attract each other in stereospecific fashion. About two protons per subunit are bound on polymerization of the TMV A protein and released on disaggregation. [Recently it was found, however, that protons are not bound when the protein polymerized in the cold (128).] Caspar suggested that these protons are hydrogen-bonded to carboxyl-carboxylate pairs and that they prevent the normal electrostatic repulsion such groups would have for each other (Fig. 6). If the two carboxyl groups were ionized, this repulsion could strain the subunit structure and lead to dissolution. This may account for the great sensitivity of TMV protein polymerization to slight changes in pH (32, 33).

This model suggests that carboxyl-carboxylate binding may control polymerization and prevent condensation of the protein alone at a pH above 6.5. If RNA is present, the additional stabilization of nucleotide binding should overcome the electrostatic repulsion within the ordered structure.

Caspar suggested that such a stabilization would have a definite biological advantage for the virus since in the infected cell coat proteins could not be lost by forming uninfected virus-like aggregates lacking nucleic acid (32, 33). There seems to be little direct evidence, however, that such aggregates are actually absent from infected cells. Even if they are present, they are not necessarily lost for future virus manufacture. If the RNA does confer added stability to the protein helix, proteins in RNA-free aggregates should eventually, through association-dissociation equilibria, be shifted into infective viruses. In this case, such uninfected aggregates might represent convenient storage forms of TMV protein. The same argument may hold for the aberrant viral capsids observed in cells infected with other viruses.

More recently, Lauffer and his co-workers proposed a theory of TMV protein aggregation in which water plays a controlling role (128). According to this, the subunit ellipsoids cannot form salt linkage, hydrogen bonds, or ionic bonds with neighboring ellipsoids, but they contain zones that can interact with water. Each ellipsoid is thought to contain four kinds of such zones, each with a characteristic "melting" temperature.

![Fig. 6. Proton association equilibria of carboxyl-carboxylate pairs, believed to be involved in the sensitivity of TMV protein subunits to slight changes in pH. These pairs are also thought to be rather specific lead-binding sites; X-ray studies of TMV that has bound lead ions suggest that one of these sites is about 2.5 nm and another about 8.4 nm, from the center of the virus helix, that is, near the inside and outside surfaces (32; Fig. 5D). Reprinted by permission of the copyright owner, Academic Press.](http://mmbr.asm.org/)
above which thermodynamic conditions are no longer favorable for contact with water but are favorable for these zones on neighboring subunits to turn towards each other and avoid contact with water (in other words, to form hydrophobic bonds). At the lowest temperatures, only subunits with the lowest melting temperatures interact to form the A protein trimer; other structural changes occur as the temperature is increased and more zones turn from an aqueous to an organic environment (see Fig. 7). The effects of pH and ionic strength are interpreted in terms of their effects on the interaction between subunits and water.

These two models have been characterized by Lauffer and Stevens (128) as “approximately equivalent in degree of arbitrariness.” The amino acid sequences in the TMV subunit are now known (cited in 128); presumably, the exact conformation of the subunits will soon be determined and can be compared to that proposed in these models.

**Structure, synthesis, and role of RNA.** The configuration of viral RNA in the TMV is determined by packing of the protein subunits. Isolated, the RNA has no definite secondary structure of its own, and other RNA molecules can easily be fitted among the same protein subunits. The coat protein, whose subunits contain 158 amino acids, is coded for by an RNA molecule containing at least 6,300 nucleotides (128). This virus, then, presumably contains much more information than is needed for its subunits and for any specific RNA polymerase. Dissociation studies showed that some parts of the RNA bind protein more strongly than others, and Caspar suggested that the excess nucleotides include sequences that bind especially strongly to protein. Such sequences might help stabilize the virus while it is being formed (32).

In infected plants, infectious nucleic acid synthesis precedes that of coat protein. In later stages, coat protein synthesis may take place independently of virus RNA (32). Recent work on the intracellular assembly of TMV has been discussed (54, 82, 199). The RNA of this virus seems to be synthesized in the nucleus of plant cells but to combine with TMV protein in many parts of the cytoplasm. Although the process of assembly has not been observed in vivo, the spontaneous combination of RNA and protein in vitro makes it unnecessary to postulate any further informational requirements for this assembly.

**SPHERICAL VIRUSES**

After showing that the icosahedral pattern of subunit arrangement represents the lowest free-energy form of building a closed shell of such subunits, Caspar and Klug (35) stated that a convincing demonstration that this design principle is correct would be the spontaneous polymerization of subunits into icosahedral shells. In the past few years, evidence has accumulated that both this and the in vitro self-assembly of infectious spherical virus from protein and nucleic acid can happen. Before describing these results, earlier studies on spherical structure and growth in infected cells will be discussed. Such work suggests that the ease with which protein shells can

---

**Fig. 7. Lauffer's 1966 model for TMV polymerization, in which affinity to water plays a controlling role. Distribution of centers that react with water is shown in Fig. 7A. As the temperature rises, more and more of these centers interact with each other to form trimers (stage II, Fig. 7A, and upper left, Fig. 7B). At a higher temperature, more polymerization occurs (stage III, Fig. 7A, and A'-A2, Fig. 7B); if the temperature is raised further, the small ends of the ellipses will come together and the polymers will collapse (stage IV, Fig. 7A, and the lower part of Fig. 7B). This process could continue to form stable double discs. (See text and reference 128 for further details.) A later model has also been proposed with slightly different distribution of centers that react with water (128). Reprinted from Biochemistry 5:2440–2446. Copyright (1966) by the American Chemical Society. Reprinted by permission of the copyright owner.**
form may make the details of internal virus structure difficult to understand.

**Questions of Virus Structure and Virus Precursors**

Klug et al. (116) discuss both construction principles and structure of several spherical viruses. The turnip yellow mosaic virus (TYMV) is a small RNA virus, of about $5 \times 10^4$ MW and 30 nm in diameter, possessing icosahedral symmetry. Thirty-two distinct morphological units or "knobs" appear on its surface (Fig. 8a, 8b). Its protein shell contains 180 identical protein subunits of MW 20,100 ± 100. These are about 3.5 nm apart on the periphery but more tightly packed at the inner radius so that the surface is grooved. The RNA is not simply wrapped as a ball in the center of the virus but penetrates deeply into the protein coat. Possibly it makes a zig-zag path through the protein shell, and it may contribute to the 32 morphological units observed by the electron microscope (Fig. 8a, 8b). This hypothesis is based both on X-ray data and on the observation that mild alkali treatment of the intact virus breaks about 30 evenly spaced points on the RNA chain. The implication is that these points may be closer to the surface, as they would be if they contributed to the surface knobs (116).

The turnip crinkle virus, another RNA virus of about $9 \times 10^4$ MW, may have a more complex structure. Upon treatment with alkali, this virus yielded a particle 50% its diameter. The soluble protein produced by alkali treatment could reaggregate to form such smaller particles, and it is not certain from this or from later work (129) whether the latter represent true inner shells or reaggregation products of outer shell protein. (Note recent work on phage fr capsid reconstitution discussed below.)

Some success was reported in reconstituting whole virus particles by adding RNA to an excess of soluble virus protein (116), and more details of this work are found in Leberman’s review on virus disaggregation and assembly (129). Apparently, studies of infectivity were not included in this work. The structure of the tomato bushy stunt virus appears similar to that of the turnip crinkle virus (116).

The DNA-containing papilloma viruses have a diameter of 56 nm and particle weights of $40 \times 10^4$ to $50 \times 10^4$ (116). When viruses were purified by centrifuging on cesium chloride gradients, both complete viruses and incomplete particles, lacking nucleic acid, were found. Klug et al. (116) found in the upper layer, which contained the incomplete particles (now often called "top components"), empty shells of the same diameter as the virus and protein shells of the same diameter but containing a protein core. In other preparations, these workers found particles about 38 nm in diameter which were thought to represent the core itself, although the possibility was admitted that they might have arisen by abnormal reassembly of normal protein subunits. It was suggested that DNA lies between the inner and outer shell and that it might penetrate the inner core.

More recent work on viruses of the papilloma-polyoma family illustrates how difficult it may be to draw conclusions concerning inner viral struc-
ture from the protein shells found when viruses are broken up. Thus, Mattern et al. (155, 156) found in CsCl gradients of small-plaque polyoma virus a 48-nm shell containing a 38-nm shell; they also found separate shells of 48, 38, and 20 nm. They suggested that the complete virus is composed of a 22-nm shell of capsomeres, surrounded by a shell of DNA, and then a 38-nm shell of capsomeres and a 48-nm shell. In addition, some particles seemed to have a phospholipid layer, bringing the diameter to 60 nm. Similar observations were made on SV₄₀ (simian virus 40).

However, Koch et al. (120) and Anderer et al. (6) carried out similar studies on preparations of SV₄₀ and found shells with diameters of about 41.1, 31.5, 27.4, and 76.8 nm. The largest of these shells seemed to be constructed of 420 structural units with an icosahedral surface lattice. These workers argued from geometrical considerations that such shells could not all be fitted into one whose symmetry was that of the largest, and they concluded that the smaller shells were derived from reassembled morphological units of the disintegrated normal shells. Three different polypeptide chains of about 16,000 MW were found in these shells. Two contained 45% each of the total protein and were thought to be structural units of the shell. A third, more basic protein was thought, on the basis of evidence provided by partial tryptic digestion of the virus, to be inside the particle and to play a role in orienting DNA.

Tubular structures of different diameters may be found in preparations of viruses of the papilloma-polyoma family (116). It has been suggested (35) that such tubes could result from faulty assembly of protein subunits in the absence of nucleic acid and, indeed, that one of the most probable mistakes would be tubular structures based on a hexagonal mesh. Tubes contain a greater ratio of hexamers to pentamers than do icosahedral forms, and conditions leading to the preferential formation of hexamers might favor tube formation (14; Fig. 9). Tubular structures called polyheads also appear in cells infected by mutant T-even bacteriophages and can be produced in vitro from the cowpea chlorotic motte virus [CCMV (14)].

Mattern et al. (155, 156) suggested that filamentous forms which appear early in some cells infected with polyoma virus might be virus precursors. However, if these filaments are caused by improper subunit assembly, they could also be envisaged as serving as a storage form for virus protein subunits before conditions were suitable for intact viruses to be made from subunits. In this sense, they might be considered indirect viral precursors.

Polyoma viral proteins are able to enclose host instead of viral DNA. In infected mouse cells, in addition to infectious capsids containing circular viral DNA, some contain fragments of mouse DNA, apparently excised at random. These "pseudoviruses" are noninfective and their biological significance is not known (159).

The adenovirus is an interesting variant of the icosahedral form, with 12 fibers extending from it, one on each vertex, and with fivefold symmetry. It contains only three antigens (167, 231); despite this relative simplicity, no attempts at in vitro self-assembly have been reported so far.

Even before in vitro self-assembly of spherical viruses was observed, the idea that it occurred in infected cells was supported by demonstrations that virus nucleic acid and protein synthesis take place independently (as with TMV, discussed above) or that they can be uncoupled from each other. In cells infected with pseudorabies virus, 5-fluorouracil, which inhibits DNA synthesis, causes noninfectious particles to be formed without the DNA core (186). In HeLa cells infected with poliovirus, newly synthesized virus RNA can be coated with existing virus protein in less than 5 min, even if further protein synthesis has been inhibited with cycloheximide (12). In monkey kidney cells infected with echovirus 12, viral protein synthesis, or RNA synthesis, can be inhibited by p-fluorophenylalanine or 2-(α-hydroxybenzyl)-benzimidazole, respectively, without inhibiting the synthesis of the other component (75). Cytosine arabinoside inhibits DNA formation in cells infected with herpes simplex virus, but if this drug is added after some viral DNA is already formed, infective virus may appear. This was interpreted to show that "existing DNA genomes are coated by viral capsids in the absence of DNA synthesis" (134).

A central question in virus formation, which cannot be discussed in detail here, is whether the protein shell is formed first and the nucleic acid packaged in it or the shell is built up around a

Fig. 9. Geodestix models composed of pentamers and hexamers. (A) Small icosahedron (cf. Fig. 3C). (B) Ellipse. (C) CCMV (cf. Fig. 3D). Increasing the proportion of hexamers to pentamers can cause elongation (14). Reprinted by permission of the copyright owner, Academic Press.
nucleic acid core. Most virus preparations contain empty shells. These could represent viral precursors, viruses that have lost their contents during fixation, or defective forms made by self-assembly in the absence of nucleic acid. Maizel et al. (146) suggested that empty capsids of type 1 poliovirus, found in considerable numbers in infected cells, had probably not arisen by virus breakdown but were products of faulty assembly. This conclusion was based partly on the fact that the protein composition of the empty capsids was different from that of the intact virus. However, Jacobson and Baltimore (99) have since provided evidence, by labeling studies, of a precursor-product relationship between the empty capsids (which they called procapsids) and the virus. They suggested that the difference in chemical composition might have arisen at the time the viral RNA joined the procapsids.

**Reconstitution of Simple SphericalViruses**

During the last two years, several spherical viruses have, with varying degrees of success, been reconstituted in vitro from their component parts. The most active research is being carried out on small spherical plant viruses and on RNA-containing bacteriophages.

**Plant viruses.** The CCMV is a small RNA-containing virus 25 nm in diameter (Fig. 10A). The MW of its RNA has been estimated as $1.1 \times 10^6$ (15). Its surface structure is consistent with an arrangement of 180 subunits of 19,500 MW in the form of a truncated icosahedron made of 32 capsomers (20 hexamers and 12 pentamers). Brome mosaic virus (BMV) is physically similar to CCMV, although serologically unrelated. Broad bean mottle virus (BBMV) is slightly larger and also serologically unrelated. Each of these viruses contains 180 protein subunits of MW 20,300 and 20,500 to 20,900, respectively (81).

At neutral pH in the absence of Mg$^{+}$, CCMV particles swell and become susceptible to attack by nucleases. After nuclease action, a variety of forms appear: ellipsoids, small icosahedra, double shells, and tubes with no apparent relation to the original virus. These were probably formed by the release of subunits and by their repolymerization (14).

Bancroft and his colleagues (13, 15, 81) have carried out extensive experiments on the self-assembly of these viruses. RNA was isolated by phenol extraction, and protein was solubilized by overnight dialysis against 1 M NaCl or other salts. This treatment breaks down the virus protein to lower molecular-weight forms whose physical properties suggest that they are dimers of the coat subunits. For either CCMV or BMV, combining protein and nucleic acid and dialyzing against lower salt concentrations leads to the formation of a nucleoprotein with the same sedimentation value as the intact virus and with almost the same infectivity (Fig. 10B; 13, 81). A snake-venom phosphodiesterase had little or no effect on the RNA of reconstituted CCMV or BMV (13, 81) although it completely inactivated the isolated RNA. Reassembly of the BBMV was also possible; however, because it was difficult to measure infectivity of this virus quantitatively, it was less intensively studied.

Hybrid viruses could be made by reassembling the RNA from either CCMV or BMV with the protein from the other of these viruses or from BBMV. All hybrid viruses were highly infectious; the host range and the progeny produced depended on the RNA rather than on the protein. The protein from a virus that does not attack a given host does not block the infectivity of a viral RNA that does attack the host. RNA in the hybrid viruses was resistant to snake-venom phosphodiesterase (81). Infectious particles, similar to reconstituted viruses, could also be made by using the RNA from one virus plus mixtures of proteins from different viruses; immunoelectrophoretic measurements suggested that particles with mixed protein coats had been formed (232). Possible reasons that the protein components of hybrid or mixed-protein viruses had relatively little effect on infectivity or on host range will be discussed below.

Under the conditions used in the above experiments [0.01 M tris(hydroxymethyl)aminomethane (pH 7.4), 0.01 M KCl, and $5 \times 10^{-3}$ M MgCl$_2$], none of these virus proteins would form a virus-like shell (top component) in the absence of nucleic acid. However, in 0.2 M NaCl and 0.01 M acetate buffer (pH 5.0), the disassembled protein of either CCMV or BMV formed particles which had the same diameter as the original virus but which, unlike the top component of other plant viruses, differed in charge from the original virus (15). These structures (called pseudo-top components; Fig. 10C) appear to have the same icosahedral arrangement of subunits on their surfaces as does the original virus (55).

**RNA bacteriophages.** A few RNA-containing bacteriophages attack male strains of *Escherichia coli*. These spherical viruses, some of which are so closely related that their coat-protein subunits may differ by only one amino acid (51, 135), are of special interest because of their small size (diameter of 20 to 26 nm and MW of about $3.6 \times 10^6$) and their genetic simplicity, which has permitted the closest correlation yet possible between genetic information and morphopoiesis.
Another attractive feature of these phages is that their RNA can direct the synthesis of coat protein and RNA synthetase in cell-free systems, so that one can envisage the future possibility of following all stages of morphogenesis in vitro (30, 74, 217).

Complementation tests with several amber mutants of the phage R17 suggest that only three cistrons exist (73). One of these codes for the coat protein, which forms a shell containing about 180 identical subunits of 13,800 MW (235). A second protein produced on phage infection is a phage-specific RNA synthetase (136, 137). Without a protein formed by the third cistron (A protein), phages yield uninfected progeny that do not adsorb on the host. This protein (MW of about 35,000 to 40,000) is probably present in only one copy per phage particle (211, 212).

This genetic simplicity correlates well with the low MW of the RNA of R17 and related phages [1.1 x 10^6 (64)], an amount that would be expected to code for only about 10^3 MW of protein (164)]. Steitz (211) calculated that, after coding for the coat protein, the A protein, and the RNA synthetase, some nucleotides in the R17 genome would remain unused. She thought that these might be needed for recognition of RNA by the RNA synthetase and the coat protein.

In reconstitution studies of the RNA phages, different workers have isolated RNA by phenol extraction and protein by treatment with 67% acetic acid, a procedure that yields individual subunits (79). The A protein was isolated by treating viruses with high concentrations of guanidinium chloride and urea (191).

Upon mixing protein and RNA in neutral or alkaline solutions containing 0.1 to 0.15 M NaCl and 10^{-4} M Ca^{++} or Mg^{++}, particles resembling intact viruses were formed; however, their infectivity was very low. Sugiyama et al. (216) found that the infectivity of reconstituted phage MS2 was less than 10^{-4} that of the native phage, compared on the basis of RNA content. Hohn (84) found that the reconstitution of phage fr probably did not significantly increase infectivity.

Roberts and Steitz (191) found that adding the A protein to coat protein subunits and RNA of phage R17 increased infectivity of mixtures of the two other components several hundredfold. However, even with A protein added, the efficiency of conversion of RNA strands to infective particles was low. Only about one in every 10^4 particles formed was infective. The reconstituted infective particles were more sensitive than was the native virus to ribonuclease.

Although Hohn first found that subunits of phage fr would not assemble in the absence of RNA (84), Herman et al. (79) observed that at higher concentrations the protein subunits alone assembled to form particles resembling the intact virus. These were thought to be reconstituted icosahedral shells, even though detailed electron microscopic examinations of their surface structure were not reported.

Most recently, Hohn (personal communication) observed that the particles formed by phage fr subunits in the absence of RNA were heavier than the empty shells obtained upon phage degradation. The particle weights calculated indicated that the former could contain about 60 additional subunits. If these were assembled as a small icosahedron, its diameter would be 15.2 nm (6); in fact, a few particles of about this size were found after reconstitution of the RNA-free protein. This work, which is of obvious significance for interpreting the cores found in spherical viruses, suggests that the 180-subunit shell condenses on the surface of the 60-subunit shell.

**Comparison of plant and bacterial virus reconstitution.** The small plant viruses and the RNA phages resemble each other in size, RNA content, and number and arrangement of subunits. It seems reasonable to think that the genome of each group carries the same amounts of information, even though detailed genetic analysis of the plant viruses is not possible. Why, then, are the reconstituted plant viruses so much more infectious than the reconstituted phages? One difference in experimental approach is that workers with plant viruses were content to break the protein shells down to dimers rather than monomers. This could have led to a closer fit in the new protein shell, and it might account for the fact that RNA is better protected from enzymatic degradation in the reconstituted plant viruses than in the RNA phages.

The successful reconstitution of infectivity in plant viruses, however, tells us little about the accuracy of self-assembly because the isolated RNA of these viruses is also highly infectious. That of the CCMV seems to be as noninfectious as that in the intact virus (13), and that of the BMV about half as infectious (81). [In contrast, TMV RNA is only about 3% as infectious as that in the intact virus (81).] The high infectivity of naked RNA and the fact that this RNA is infectious when enclosed in the protein from a different virus (see above) indicate that specific attachment of the virus protein on the plant cell surface is not essential for infectivity. Dissociation on the surface to yield free RNA would be sufficient, and if the RNA enters the cell any dissociation would tend toward completion.

The bonds holding subunits together on these spherical viruses are still poorly understood. Even
though RNA is not essential for assembly, it seems to stabilize it; protein polymerization without RNA requires higher protein concentration (79) or lower pH (14, 15). It has been suggested (14) that paired hydrogen bonds forming between unionized carboxylic acid side chains control attraction between units. Thus, the structural determinants of the spherical viruses may be similar to those of TMV (see above). It is only now that extensive kinetic and thermodynamic studies on spherical virus polymerization can be carried out.

Reconstitution of Small DNA Phages

Takai (219) showed that disaggregated coat protein and DNA in the spherical phage ϕX174 could recombine to give a structure that protected DNA from deoxyribonuclease. The infectivity of this structure was not studied.

The male-specific E. coli phage fd is a thin rod about 800 nm long and 5 nm in diameter. It contains a single strand of DNA (1.7 × 10⁶ MW) that runs from one end of the virus to the other and back again "like a circle of string pulled taut from opposite ends of its circumference" (153, 154). Electron microscope, X-ray diffraction, and optical rotary dispersion measurements suggest that the phage coat is constructed of helically arranged subunits of high α-helical content (117-119). The coat protein can be broken down by phenol and sodium dodecyl sulfate treatment to monomers of about 10,000 MW. The subunits alone can combine to give irregular rods of the same diameter as the intact virus, but shorter. The subunits also combine with fd DNA and with other single-stranded DNA molecules, such as that of the phage ϕX174, but not with double-stranded DNA. Subunits can recombine with fd DNA in vitro to yield infective particles, provided that at least 3% of the coat protein is still attached to the DNA. With nucleic acid alone, the proteins form irregular, noninfective rods. The yield of the infectious particles is low, corresponding even under the best conditions to only one completion in each 10⁷ DNA protein complexes originally present (117-119).

In Vitro Formation of Hybrid Viruses

The reassembly of hybrids between protein and RNA of different small spherical plant viruses (81) was described above. Hybrids can also be made with nonviral nucleic acids or even with polyanions other than nucleic acids.

Protein of CCMV formed complexes with RNA from TMV, bacteriophage f, ribosomal RNA, and s-RNA, as well as with the single-stranded DNA from phage S18. These complexes resembled the native CCMV in the electron microscope. The hybrid with TMV RNA was more infective to tobacco plants than was the RNA alone, although much less so than native TMV. The phage DNA was much less infective as a hybrid than in the naked form (81).

TMV proteins can combine with TMV RNA (157), with yeast RNA, and with synthetic polynucleotides (215). Sugiyama (215) found that RNA of the spherical phage MSV combined with TMV protein to give rods shorter than, but otherwise indistinguishable from, TMV. The length of the rods was determined by the length of the RNA, which was apparently arranged as in the native TMV. These rods were infective neither to E. coli nor to tobacco leaves, although infectious MSV RNA could be recovered by phenol extraction.

As described above, in high concentrations proteins of the RNA phage fr can form empty capsids by self-assembly (79), but in low-protein concentrations such particles are formed only in the presence of RNA or certain polyanions. T. Hohn (J. Mol. Biol., 1969, in press) found that besides fr RNA, polyuridylic acid and polyvinylsulfate could induce assembly in low-protein concentrations. The appearance of the particles formed varied with the kind of polyanion used. After particles were formed the RNA could, in some cases, dissociate from it and could be used again. Thus, RNA seemed to be playing a catalytic role in morphogenesis. (If this occurs in virus-infected cells, empty capsids could be easily formed and mistakenly considered as virus precursors. This possibility adds yet another difficulty to the task of deciding which comes first, the full or the empty capsid.)

Hohn's results should not be interpreted to mean that there is no specificity of interaction between fr RNA and its protein. His more recent results (Fed. Eur. Biochem. Soc. J., 1969, in press) show that these form a complex ("complex I") containing six protein subunits per RNA molecule. Such a complex is not formed with RNA from phage Qβ. Complex I does not dissociate easily and possibly could be considered a precursor of completed phage particles in infected cells.

COMPLEX PHAGES

T-Even Bacteriophages

By far the most complex viruses in which morphopoiesis has been studied, and those in which analysis has reached the greatest degree of sophistication, are the T-even phages of E. coli, especially T₄ (Fig. 2 and 13). This phage may conveniently be thought of as composed of a head (including a neck), a tail (made of two
coaxial hollow tubes, the outer being the contractile sheath and the inner known as the "core," "tube," or "needle," plus the termination of these tubes in a hexagonal baseplate, and tail fibers. (For reviews of the structure of this and other phages, see references 23, 24, 47, 107, 160, 202.) The complexity of structure and assembly is reflected in the genetic analysis of phage T4 morphopoiesis.

Many of the phage genes involved in T4 formation have now been mapped. Studies of such genes have relied heavily on conditional lethal mutants, those in which the expression of the mutation occurs only in certain hosts or at certain temperatures. When the mutation is expressed, incomplete or abnormal phages may be formed. The various genetic markers now known for this phage include those for DNA synthesis (early functions) and more than 40 genes involved in the formation or assembly of various phage structures or substructures. These late-functioning genes act independently of each other after one of the early-functioning genes has acted (47, 53, 105, 107). (For detailed reviews of phage morphopoiesis, see references 47, 105, 106, 113, and 242.)

The phage contains at least 15 different proteins and because of the number of genes involved in phage formation, it has been suggested that it may contain many more (242). One of the central problems in phage morphopoiesis is whether different gene products form part of the final structure or assist catalytically in its assembly. The different precursors of phage structures are present at the same time and, thus, their assembly is not regulated by sequential synthesis (105, 242).

The shape of the head capsid and the surface arrangements of its capsomere suggest that it is an elongated version of an icosahedral virus (105). Most of the head-capsid protein is of one kind ("majority protein," of 80,000 MW), although small amounts of other proteins are present and appear to be of great importance in determining form (105, 106). One interesting effect of mutations in the late-functioning genes is the appearance of aberrant phagelike structures with small heads (polyheads) that may be as long as the infected bacterium and several layers thick or of phages with altered head shape or with two tails (105, 107, 125; Fig. 11). Polyheads are sometimes observed also in cells infected with normal phages, if lysis is inhibited (107). Polyheads are serologically related, but not identical to normal phage head capsids. The appearance of such altered phages has been especially helpful in genetic analysis of head formation.

To list briefly the effects of some specific genes, gene 23 produces the "majority protein" (protein 23); if this gene does not function, no heads of any kind can be produced. Gene 31, an early-functioning gene, governs in some way the solubility of protein 23. In lysates of cells infected with phage deficient in gene 31, protein 23 is bound to the cell envelope instead of being free in the supernatant fluid (105, 107). Mutations in gene 20 lead to polyhead production. A mutation in gene 66 causes a short-headed variant of T4. Gene 22 seems to contribute a minor component to the phage head; mutations in this gene and in gene 40 produce multilayered polyheads (105, 106).

Mutants in gene 21 produce headlike particles containing little or no nucleic acid. These bodies (called "e" particles") are usually found along the inner envelope of the infected cell (53, 107; Fig. 12).

The contractile sheath of phage T4 is built up of 144 subunits of 50,000 MW (105), specified by gene 18 (113). Very recent genetic and in vitro studies (113) show that in tail formation the baseplate is completed first, and then the core (tube) forms on the baseplate and the subunit product of gene 18 polymerizes on the core-baseplate structure. Genes 3 and 15 appear to be required for the stabilization of the sheath (see Fig. 13). Occasionally, abnormal "poly-sheaths" are formed, but the genes involved in their formation are still not understood. Eisering et al. (49) studied the genes responsible for tail-fiber formation and assembly in phage T4D (genes 34-38 and 57). Of these, gene 37 seems to control the major structural element.

Polglazov et al. (179) claim to have produced contracted sheaths by reaggregation of disrupted sheath subunits of phage T4, although it has been pointed out that the physiological significance of this is still uncertain (105). Polglazov's group (180, 181) also reported the in vitro reaggregation of T4 capsid subunits, prepared by alkali disintegration, into sheets one subunit thick, with some parts displaying hexagonal symmetry similar to that found on polyheads.

Speculations on the control of morphopoiesis: orders of morphopoiesis and morphopoietic cores. In contrast to the small spherical viruses whose subunits are in equivalent or quasi-equivalent positions, the head capsids of T-even phages contain 1,000 to 2,000 subunits, some of which must occupy positions of quite different curvature than others (105). The shape does not appear to conform to any minimal-energy design, and this suggests that more information is required to specify it than that contained in its subunits alone. This suggestion is supported by the genetic complexity or head-capsid formation; some
FIG. 10. Electron micrographs of native CCMV (A), reassembled virus (B), and "pseudo-top component" (C). The viruses are 25 nm in diameter. Photographs were supplied by J. B. Bancroft.
eight genes are required to specify head shape (105, 242; Fig. 13). In Kellenberger's terminology, 
the normal prolate head is of a high "morphopoietic order" (105).

The concept of orders of morphopoiesis relates to the amount of information 
needed to make a given structure. In first-order morphopoiesis, 
under the proper conditions of pH, temperature, 
and ionic environment, identical subunits contain 
all the information needed to build the final product in its characteristic shape. This occurs when TMV protein subunits form helical or 
stacked-disc structures in the absence of nucleic acid (32) and, probably, in the formation of capsids of small spherical viruses.

Kellenberger considers that, in addition to subunits, morphopoietic factors may be involved 
in determining structure. A morphopoietic factor is defined as "any substance which interacts specifically with the subunit (or a polymerized product of it) in such a way that this interaction is a necessary step in achieving the final shape. A morphopoietic factor may or may not be found integrated into the final product" (105).
By stipulating specificity, this definition requires that the morphopoietic factor be a large molecule, since ions or low-molecular-weight compounds (such as nucleotides, which certainly affect protein interactions) are not specific. It would be a gene product (protein) or, possibly, synthesized by gene products (e.g., nucleic acid). The order of morphopoiesis is determined by the sum of the number of different kinds of subunits and of morphopoietic factors.

Since more than one gene could be involved in making a morphopoietic factor, the order of morphopoiesis is not necessarily the same as the number of controlling genes. It might be difficult to distinguish a morphopoietic factor incorporated in the final product from a subunit; presumably such a distinction would be easy enough in the case of the A protein of the R17 phage, which seems to be present in only one copy per virus particle.

Kellenberger and his colleagues have suggested that morphopoietic cores might play an essential role in the assembly of both heads and tails (105, 107, 125). These cores are envisaged as three-dimensional bodies, presumably produced under the direction of the viral genome, that determine the shape of the capsid by directing assembly of subunits on their surface. Such a core in a complete virus would have to share the inside of the head capsid with the nucleic acid.

The existence of such cores was postulated because certain variant head-related structures, which lack nucleic acid, enclose material different from the surrounding medium. These structures include the phage lambda, polyheads of T-even phages, and “tau particles,” produced by phages mutant in gene 21 (53, 103, 105, 107). Electron micrographs suggest that the cores of tau particles and polyheads may have an organization, usually a ring-shaped structure (Fig. 12). This was thought to represent a highly hydrated structure. The possibility that a core could be made by self-assembly of hydrated protein subunits is discussed by E. Kellenberger (personal communication) in the 1968 Nobel Symposium.

More detailed arguments for and against morphopoietic cores of phage heads have been presented (105, 107). Kellenberger (105) points out that the concept does not as yet provide for the control of core size. The most unsatisfactory aspect of the idea, in my opinion, is that it shifts the problem of determining shape from protein subunits, about whose ability to assemble in defined patterns something is known and more can be learned, to structures about which much less is known. If the core is made of a highly hydrated protein structure, one could envisage a mechanism for controlling the amount of protein produced but not for controlling shape.

This objection vanishes, of course, if cores are essential for head capsid assembly. If they are, it should be possible to demonstrate subunit assembly on cores in vitro. With the present state of the art of phage assembly, such experiments seem possible.

For tubular structures, the concept of a morphopoietic core is more satisfactory. It is well known that tubes can form by helical polymerization of identical subunits, and it is easy to envisage a second helix polymerizing on the surface of such a tube. Kellenberger suggested (105) that the central tube of the phage tail might act as a core for sheath assembly, and more recent studies (113) have shown that this happens.
Putting Larger Parts of Bacteriophages Together

T₄ phage. Edgar and Wood (48) showed that extracts of E. coli infected with phages containing mutants of gene 23 (involved in producing head proteins) have no heads but contain components necessary for tail assembly. Other phage mutants could produce tailless heads; when these were added to the first extracts, infective particles were formed, the number almost corresponding to the number of heads. This reconstitution seems to entail the construction of a complete phage particle. Early steps in tail assembly were demonstrated by mixing extracts made with mutants deficient in different tail genes. At least two steps, association of tail proteins and their attachment to phage, could be distinguished. In general, mutants that made heads but no tails complemented mutants that made tails but no heads in producing active phages when extracts of infected cells were mixed. There was no such complementation within the group of headless or tailless mutants. Even though not all steps in phage synthesis could be demonstrated in vitro, the assembly of tails, their attachment to virus particles, and the union of head and tail could be demonstrated.

After further combined genetic and in vitro studies, Wood, Edgar and co-workers (47, 242), proposed a pathway for the entire assembly of T₄ (Fig. 13). This is thought to occur in stepwise fashion with most, although not all, of the steps under gene control. According to this scheme, the head, tail, and tail fibers are assembled independently of each other and then unite, first the head with the tail, and the fibers are added last. The rate of fiber attachment is stimulated by a heat-labile factor (presumably a protein), a product of gene 63; however, fibers can attach without the catalytic action of this factor.

According to this scheme, the empty head capsid is made first by the combined action of eight different genes; then DNA is introduced under the action of three more genes. After several further steps, the tail joins the head. Mutants defective in genes 16, 17, or 49 accumulate empty capsids. Tracer experiments (not yet described in detail) also suggest that these capsids are precursors of complete heads (113, 242). This possibility seems especially interesting in view of the uncertainty of whether empty capsids in preparations of spherical viruses are virus precursors or result from aberrant self-assembly.

In mutant-infected cells that accumulate empty heads, DNA strands several times longer than those of the phage itself also accumulate (113). DNA strands two or more times the length of T₄ DNA, apparently made of linearly linked T₄ DNA molecules, also appear in normal phage infection (61). These results could imply that certain genes are involved in cutting and packaging DNA and that empty head capsids have an active role in such packaging. Why DNA strands many times longer than their template should be formed is not known.

Even if the complex phage-head capsids are made empty and then filled, it seems dangerous to conclude that the spherical viruses are made in this way. So far, T-even phage head capsids have not been formed in vitro, but in vitro studies of spherical virus capsids show that, in general, these are more easily formed in the presence of nucleic acid than in its absence. It is very difficult to imagine that empty capsids will form in the infected cell if nucleic acid is present. The empty capsids observed in such cells could be due to loss of nucleic acid on fixation, or they could be caused by dissociation of nucleic acid after full capsid formation. [See Hohn's speculations on the catalytic role of nucleic acid (J. Mol. Biol., in press).] With this cautionary note, I leave a subject better settled by experimentation than by further discussion.

The way in which most of the morphopoietic genes of the T-even phages act is still not known. Most of the work so far reported has used crude cell extracts although experiments with better-defined systems are under way (242). One enzyme (specified by gene 63) seems to be involved in attaching tail fibers. Enzymatic functions for other gene products may be found. It is tempting to think that small organic molecules, which can affect the conformation and interaction of many proteins (i.e., "allosteric effects"; 63, 209, 218) and influence the assembly of other structures (see below), are involved in phage morphopoiesis. Possible sites of allosteric action are discussed by Kellenberger (105), but there is still no direct proof that such action occurs. Dialysis experiments showed that cations, but no other dialyzable factors, were required for tail-fiber attachments (242).

Lambda phage. Phage lambda of E. coli is smaller than T₄ and contains about 25% as much DNA (238). Its MW is 37 × 10⁶; its head, hexagonal in projection, is 54 to 57 nm in diameter and its tail is 140 to 150 nm long. Two protein subunits have been isolated from this phage. Their molecular weights are 55,000 and 110,000 and their amino acid contents are different. It is thought that these might represent head and tail, respectively (46). Weigle (238) found several phage genes that could be characterized as head donors or tail donors. Testing combinations of lysates containing defective phages showed that reassembly of heads and tails could occur, with less than 1% efficiency of correct assembly. A kinetic study of assembly of heads and tails...
(239) showed that it depends on kinetic encounter of the particles and that only about 1 in 100 of these encounters produces an active phage. Tails alone can adsorb to the host bacterium. Here, they can attach heads and become infective.

**Phage P22.** The DNA phage P22 of *Salmonella typhimurium* has a polyhedral head about 55 nm in diameter to which is attached a baseplate consisting of a central core with six spikes arranged around it. Both head synthesis and tail synthesis are especially temperature-sensitive in different phage strains (93, 94), and this was useful in obtaining separate heads with attached cores and tail parts. When these were mixed in vitro under the proper environmental conditions, active phage was obtained, with a high efficiency of reactivation.

**BACTERIAL FLAGELLA**

Bacterial flagella are especially attractive for studying protein synthesis and self-assembly. Although some details of their structure are still uncertain, they are smaller and simpler than flagellia or cilia of eucaryotic cells. Many are about 20 nm in diameter and may be several micrometers long. More details on the range of flagellar structures in the microbial world are given by Rhodes (188). Flagella can be mechanically removed from bacteria and disintegrated by heat, acid, alkali, or other agents into identical subunits, the flagellins, globular proteins of MW varying from 20,000 to 40,000 in different species. (See 2, 5, 111, 145, and 149-152 for references to MW and to amino acid composition.) The flagella of several species consist only of protein, although some may also contain small amounts of carbohydrate (2, 150). Different electron microscope and X-ray diffraction studies suggest that certain flagella are constructed of several helices (from three to five; see 111, 138) made of flagellin molecules, about 5 nm in diameter, wound around a common center, with about eight subunits per turn (111, 138; Fig. 5A). Some flagella appear to be composed of vertical rods instead of helical arrangements of flagellin molecules (138). Lowy and Hanson (138) compared flagellar structures to that of F-actin, a linear polymer of G-actin molecules 5.5 nm in diameter and assembled in a double helix (Fig. 5B). As shown below, actin self-assembly in many ways resembles that of flagellin.

In the last 5 years, several laboratories have demonstrated the in vitro self-assembly of flagellins into structures that look very much like bacterial flagella (2, 4, 9, 140, 172; S. Asakura, 7th Int. Congr. Biochem. Tokyo, 1967, p. 229-230). Abram and Koffler (2) disintegrated the flagella of *Bacillus pumilus* by acid treatment into subunits of 30,000 to 40,000 MW and found that when the pH value was raised these slowly reaggregated to give both straight filaments and wavy, flagella-like filaments. By adjusting the pH, the straight filaments could be transformed to wavy filaments, but not the reverse. The flagella-like filaments seemed to represent the most stable organized state.

Martinez et al. (152) recently isolated nonmotile mutants of *B. subtilis* that produced straight instead of normal, wavy flagella. Flagellins of the mutant and parent strains (which differed only in one amino acid) could be reassembled in vitro to yield, respectively, straight or wavy flagella (see Fig. 15.)

Oosawa, Asakura, and co-workers (170, 172) studied the theoretical physical chemistry of flagellin (and actin) polymerization. Their analysis deals mainly with the formation of helical or tubular polymers in which each monomer is bound to four or more neighboring monomers. According to this, a nucleus must first be formed which can then grow into a polymer by the addition of monomers (Fig. 14). Nuclear formation is expected to be rate-limiting, and in some cases it can be shown that addition of a preformed nucleus or "seed" is essential to polymer formation. The theoretical analysis of Oosawa and Higashi (170), which is limited to helical and tubular polymers but which may be pertinent to other kinds, also examines how small molecules that can bind to the protein monomers can influence the process of polymerization and the formation of one type of polymer to another.

This group (78, 170, 172) found that polymerization of flagellin from different *Salmonella* strains seems to require a nucleus made of so many monomers that it cannot form spontaneously. Flagellin prepared by heat treatment (MW, 38,000) will not polymerize on cooling unless small seed fragments of flagellin (0.2 to 0.3 μm long), prepared by sonic oscillation, are added. In studying the determinants of flagellar form, two types of flagella were used as sources

![Fig. 14. Model of polymerization as nucleation and growth of nuclei (slightly modified from reference 172 to show that processes are reversible (cf. 170)).](image-url)
of both flagellin and seeds: normal, wavy flagella and "curly" flagella, with wavelengths 50% that of the normal ones (Fig. 15). These were mixed in all possible combinations. From the types of reconstituted flagella formed (Table 1), it was concluded that seeds may specify the type of flagella that grow on them, but that flagellin molecules and flagellar fragments of the curly type have greater ability to determine structure than do those of the normal type. However, the two structures have almost the same free energy. Upon low-temperature storage or upon lowering the pH, normal, reconstituted flagella become wavy. Upon raising the pH or upon addition of adenosine triphosphate (ATP) or pyrophosphate, the transformation is reversed (170, 172).

Flagellar fragments are not needed for polymerization of flagellin from other bacteria, or even from all Salmonella species (2, 4, 152). Possibly, in such cases, very small nuclei are needed to initiate polymerization; however, this has not been studied directly.

Flagellar Growth, Flagellin Synthesis, and Self-Assembly

A good deal is known about the effects of the environment on flagellar formation (109, 121, 190), and in certain experiments the in vivo formation of flagellin molecules can be differentiated from their assembly.

Flagella of several species originate in basal granules, discs, or hooks attached to the cytoplasmic membrane (3, 68, 108; Fig. 16). Flagellar growth can be dissociated from cell division; commonly, this is done by raising the tempera-

<table>
<thead>
<tr>
<th>Seed</th>
<th>Flagellin</th>
<th>Type of reconstituted flagella</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Normal</td>
<td>Normal</td>
</tr>
<tr>
<td>Curly</td>
<td>Normal</td>
<td>Curly</td>
</tr>
<tr>
<td>Curly</td>
<td>Curly</td>
<td>Curly</td>
</tr>
</tbody>
</table>

* From reference 172.

Fig. 15. Reconstituted Salmonella flagella filaments. (a) Normal strain (SJ 25). (b) Curly strain (SJ 30). (c) Straight strain (SJ 814). This nonmotile mutant of Salmonella typhimurium is described in reference 91. Negative staining with phosphotungstic acid. Figures were supplied by S. Asakura. X 8,800.
Fig. 16. Attachment of flagella to cytoplasmic membrane (cm) of Bacillus stearothermophilus. Most of the flagella attached to the membrane originate in spherical (s), disc-shaped (d), or mushroom-shaped structures (arrows). Phage and phage tails are represented by ph and t, respectively. Negative staining with phosphotungstic acid (3). × 71,000.

A similar argument holds for G-actin polymerization (see below).

Kerridge observed that S. typhimurium could regenerate flagella in the absence of net protein synthesis or when protein and RNA syntheses were about 75% inhibited by 8-azaguanine (108). One possible explanation was that the flagella were assembled from an intracellular...
pool of flagellin already present. Kerridge (110) could find no evidence for any substantial amounts of such a pool and suggested that this might indicate that the site of flagellin synthesis is closely associated with the basal granule (7th Int. Congr. Biochem. Tokyo, 1967, p. 231–232).

More recently, Aamodt and Eisenstadt (1) found that actinomycin D, which inhibits RNA and protein synthesis by 99% in S. typhimurium, prevents flagellar regeneration. These combined results imply that synthesis of flagellar protein is less sensitive than that of other cell proteins to certain inhibitors.

Martinez and Gordee (151) observed that chloramphenicol inhibited protein synthesis but not flagellar regeneration in Spirillum serpens and B. subtilis from which flagella had been removed mechanically or by growth at high temperatures. Immunological and tracer experiments gave direct evidence that these bacteria have pools of preformed flagellin which serve as precursors for flagella. B. subtilis does not always have such pools: Dimmitt et al. (44) found that chloramphenicol and puromycin inhibited flagellar regeneration in this organism.

In Proteus vulgaris, during the logarithmic growth phase few flagella are produced and intracellular flagellin accumulates; in later growth, flagella are formed and the intracellular flagellin decreases (cited in 2). Variations in flagellin concentration during the growth cycle might explain why inhibition of protein synthesis does not always have the same effect on flagellar regeneration. So far, flagella formation seems most simply interpreted in terms of the self-assembly of flagellin molecules that are synthesized within the cell.

Asakura et al. (9) pointed out that the in vitro polymerization of some flagella was much slower than their de novo formation by living bacteria. When S. serpens divides, each daughter cell inherits a tuft of flagella at one end and grows a tuft at the other. The inherited tuft stays the same size as at division (151). However, Martinez and Gordee observed that if the cell is deflagellated, both ends grow fresh tufts. It is clear from these examples that some regulating devices for flagellar synthesis and self-assembly must exist.

At which end do flagella grow? In theory, flagella could grow by adding flagellins to the flagellar tip or to the end inside the cytoplasmic membrane, followed by extrusion of the newly synthesized part, or they could grow by adding flagellins to both ends. Although workers with flagella seem to favor implicitly the first mode of growth, only recently have attempts been made to distinguish between these possibilities. Asakura et al. (10) showed, by labeling Salmonella flagella with antibodies, that flagellin molecules are added to flagellar fragments in only one direction. D. Kerridge (personal communication) used the fact that flagella formed in the presence of p-fluorophenylalanine have abnormal morphology to obtain evidence that growth of Salmonella flagella in vivo does take place at the distal end.

If flagella do indeed grow at the tips, by means of an intracellular flagellin pool, it would seem almost essential for any efficient utilization that the flagellin molecules travel down the center of hollow flagella. Although many flagella appear to be hollow in cross section, there is still uncertainty as to whether they really are or possess a core (2, 111, 138). The fact that a core is not needed for in vitro assembly of flagellins suggests that one may not be required by the living cell, but this question and the intimate mechanism of flagellar growth are still matters about which it is only possible to speculate.

Theories of Flagellar Motion

Although microbiologists have long been fascinated by the great activity of bacterial flagella, how these organelles move remains a mystery. Lowy and Hanson (138) suggested that interconversion between a straight and helical arrangement of subunits was involved in contractility. The transformation between normal and “curly” flagella of S. typhimurium, catalyzed by ATP, was mentioned above, and the possibility was considered that such transformations may play a part in cell locomotion (9, 170). However, the ATP causing these transformations was not dephosphorylated and, hence, did not supply energy (9).

Klug (115) proposed a mechanism of flagellar movement which assumes that the movement originates in the basal body and explains how a wave can travel down the flagellum without requiring that energy be supplied throughout its length. The theory assumes that the flagellum of the living cell has a helical shape (which appears as a sine wave after drying for electron microscopic examination). Now, although in a rod made of a helical arrangement of subunits all are in equivalent positions (see above and Fig. 5), this changes once the rod is bent. Each subunit in a helically shaped flagellum can be considered as part of a coiled coil, a helix bent into another helix. Some subunits will be under compressive strain and others, on the opposite side of the flagellum, under tensile strain. The arrangement of subunits could be said to be quasi-equivalent, i.e., that corresponding to the minimal free energy state of the whole flagellum. The deforma-
tion of position of each flagellin molecule needed
produce such an arrangement could be very
small, of the order of only 0.1 nm. If such an
arrangement exists, an input of energy leading to
a slight perturbation of subunits at the base
could send a helical wave down the whole flag-
ellum.

This interesting theory predicts that bacteria
with straight flagella would not be motile; so far,
this seems to be true (91, 152). It also predicts
that the self-assembly of such a structure would
require high activation energy. This is supported
by the fact that flagella of certain bacteria need
seeds for reconstitution. On the other hand, not
do all. Furthermore, in reconstituted flagella of
B. pumilus the wavy form is more stable than the
straight (2). Although it is certainly reasonable to
think that a flagellum in water has a helical
rather than a flat sine-curve shape, this should
be tested further; after drying, either form would
appear as sine curves. Individual flagella of
living bacteria cannot be seen, but the problem
of their geometry might be solved by studying
sections of fixed flagella. To my knowledge, this
has not yet been done.

The much more complex flagella and cilia of
eucaryotic cells are discussed below in the section
on microtubules.

PILI
These hairlike fibers project from the surfaces
of several species of gram-negative bacteria. They
are smaller than flagella (diameter about 7 nm
and from 0.5 to 2 μm long); commonly, a few
hundred pili are found per cell (25). The presence
of the F pilus of F+ (male) E. coli is controlled by
the F-factor episome, which genetically controls
fertility. There is good correlation between the
presence of F pilus and the ability to serve as
donors of the bacterial chromosome; the presence
of such pili is usually also accompanied by sensi-
tivity to male-specific bacteriophages (25, 201).

Pili contain more than 99% protein. After
heating at low pH values or after certain other
treatments they break down into subunit “pilin”
molecules of about 17,000 MW, containing a
preponderance of nonpolar side chains. Pili are
probably composed of a single helix of such
subunits, containing 3\(\frac{1}{4}\) per turn (Fig. 5). When
pilin solutions are dialyzed against 0.2 M KCl,
they spontaneously repolymerize to form strands
indistinguishable in the electron microscope from
native pili (25).

There is an obvious resemblance between
pilin and flagellin self-assembly, and the regenera-
tion of pili by the living cell also seems to be
similar to that of flagella. Bacteria from which
pili have been removed by rapid stirring in a
blender can regenerate fresh pili. Regeneration
is inhibited by streptomycin but not by chlor-
amphenicol, a result which, together with labeling
experiments with radioactive amino acids, led
Brinton to conclude that pili grow by an assembly
of previously synthesized subunits. The pool of
such subunits might be a large one. In the con-
tinuing presence of chloramphenicol, at least
four cycles of regeneration were observed after
removal of pili (25).

The holes in pili are probably too small for
pilin molecules to pass through (26; Fig. 5C).
If pili are formed by self-assembly in vitro, it
seems likely that this assembly takes place within
the cell, that is, that pili grow at the proximal
end. (See discussion of bacterial flagella.)

MICROTUBULES
Slatterback (207) introduced the term micro-
tubule to denote fine tubular components of the
cytoplasm, of which examples had previously
been described in a great variety of eucaryotic
cells. Microtubules can be up to several microm-
eters long and their diameter in different cells
varies from about 14 to 27 nm. The thicker ones
may contribute especially to cellular elasticity.
(For reviews, see 112, 182 and 207.)

Microtubules are found in eucaryotic cilia and
flagella (see below) and in the caudal sheaths of
spermatozoa. They form the elements that
apparently give structure to tentacles and orga-
nelles of motion of protozoa. Microtubules are
found in cells of higher plants, where their
distribution suggests that they may play a part
in wall deposition, and in cells of developing
organisms, where they may guide asymmetric
development. They make up most of the spindle
fibers in the mitotic apparatus. Very long micro-
tubules are found in nerve cell processes, and
microtubules are found in melanophores and
in many other specialized cells. They are called
tubules because their center is less dense than
their periphery, but there is no evidence that
material moves through this center (182, 207).
Arrays of microtubules may guide the movement
of material through the cytoplasm. Even though
all their functions are not understood, it is clear
that microtubules play very important architec-
tural roles in eucaryotic cells.

They are probably confined to eucaryotic
cells. Structures resembling microtubules ob-
served in Proteus species (96) are now thought
to be tubular variants of bacteriophages (36a;
J. F. M. Hoeniger, personal communication).

Microtubules appear to have a subunit struc-
ture. Electron micrographs could be interpreted
Fig. 17. Temperature effects on a system of microtubules in the protozoan, Actinosphaerium nucleofilum. (a) Reappearance of axopodia after warming. The needle-like pseudopodia (axopodia), each of which contains a system of microtubules (the axoneme), had withdrawn into the cortex after 2.25 hr at 4°C. The figures show the development of axopodia after 3 (top), 15, and 45 min at room temperature. Arrow indicates axonemes beginning to form in the cell cortex. Reforming axopodia appear in the same numbers and approximately in the same places as before cold treatment. The highly birefringent spot in the center of the cell is due to an ingested rotifer. X 100. (b) Cross section of a retracting axopodium after 2.25 hr in the cold. Symbols: M, mitochondrion; G, granule; E, excretion body. X 38,000. (c) Cross section of a recovering axopodium. After warming the cold-treated protozoan, the axoneme (a double coiled arrangement of 22-nm microtubules) appears. This arrangement is the normal one for protozoa not exposed to cold (221). X 76,000. Reprinted by permission of the copyright owner, The Rockefeller University Press.

to show that they are made of a helical arrangement of subunits (about 5 nm in diameter) of low pitch, containing 12 to 13 per turn, or to show that 12 or 13 linear columns of subunits are grouped around a center (71, 86, 130, 182). It is tempting to think that the first arrangement (which might provide for a more orderly self-assembly) holds, but there seems to be no strong
reason for favoring it. The difficulties of interpreting electron micrographs of microtubules, difficulties which also apply to interpreting electron micrographs of bacterial flagella, are discussed by Hookes et al. (86), who outline a "computer approach" to ultrastructural studies.

The formation and disappearance of microtubules probably involve processes of polymerization and depolymerization similar to those of actin, flagellin, or TMV protein subunits. At low temperature or high hydrostatic pressure, microtubules quickly disappear, and they reappear quickly when the temperature is raised again or the pressure is lowered (19, 182, 194, 221; Fig. 17). The conditions that cause polymerization and depolymerization and the speed of these processes have suggested to the authors cited and to others (175, 240) that microtubules are formed by self-assembly. Wilt et al. (240) found that the structural protein of the mitotic apparatus of sea urchin eggs is synthesized before activation of the egg by fertilization, that is, before the apparatus itself is formed; this suggests that it is formed by assembly of proteins already present.

Structure and Formation of Eucaryotic Cilia and Flagella

These structures are larger and much more complex than bacterial flagella. In cross section they present, within an enclosing membrane, the familiar pattern of nine pairs of tubules surrounding two central tubules (Fig. 18a). Reviews of the structure and morphopoiesis of cilia and flagella have been presented (65, 66, 86). Despite the obvious difficulties involved in studying the formation of such complex structures, some hopeful progress has been made both in genetic and in chemical analysis. Randall and his colleagues (86) found a number of different mutations that affect the appearance, length, and motility of flagella of Chlamydomonas reinhardii. At least four genes are involved in forming the central pair of tubules. So far, the specific gene products have not been identified.

Gibbons (65, 66) succeeded in reconstituting parts of the cilia of Tetrahymena pyriformis, although not the microtubules themselves. After removing the outer membrane by digitonin extraction, he isolated the "axoneme," the complex fibrous part of the cilium. This contains inner and outer tubules, fibers running between them, and characteristic projections (arms) on the outer tubules (Fig. 18a, 18b). The axoneme contained most of the adenosine triphosphatase of the cilia.

Extraction with ethylenediaminetetraacetic acid (EDTA) removed the two inner tubules, the arms from the outer tubules (Fig. 18b), and almost all of the adenosine triphosphatase activity of the axoneme. Two proteins, 14S and 30S "dyneins," were isolated from the EDTA extract, the second (MW of about 5,400,000) being a polymerized form of the first (MW of about 600,000). Preparations of 30S dynein contained rods 9 nm in diameter and up to 500 nm in length; these seemed to be composed of globular subunits of MW 55,000.

When 30S dynein was added to EDTA-extracted axonemes, the arms of the outer tubules were restored (Fig. 18b). Adding ATP to axoneme suspensions lowered their light-scattering, and this was also true of the reconstituted material. However, motility, (which can be observed when ATP is added to certain preparations of isolated cilia) was not restored (65, 66).

More recently, certain proteins from the microtubules of cilia and sperm flagella, as well as from mitotic apparatus, were isolated (187, 213). Stephens (213) found that protein subunits of about 60,000 MW could be obtained by mild detergent treatment of microtubules from sea.
Fig. 18b. Control: axonemes extracted with EDTA, lacking the "arms" usually found on the outer tubules (cf. Fig. 18a). Recombined: recombination and reformation of arms take place after adding MgSO$_4$. The same kind of reconstitution of arms occurs if 30S "dynein" is added (65). $\times$ 120,000. Reprinted by courtesy of I. R. Gibbons.
urchin sperm tails and that upon lowering the detergent concentration the subunits formed fibers that resembled the original microtubules. Polymerization was increased by 0.1 mM salt. Fragments of intact microtubules (seeds or nuclei) were needed for this polymerization, a necessity that may be related to the occurrence of orienting centers for microtubule formation in the cell [See (182) for a discussion of such centers which include the centriole and centromeres involved in spindle fiber formation.] The in vitro polymerization of this preparation of microtubular protein was not inhibited by cold or by colchicine; in these respects, its behavior differed from the in vivo behavior of the protein of mitotic apparatus.

**RIBOSOMES**

The overall composition and gross structure of ribosomes are well known. In bacterial ribosomes, the 70S form readily dissociates into 50S and 30S subunits of about 1.8 x 10^6 and 0.7 x 10^6 MW, respectively, each containing about 60% RNA and 40% protein (222). Below this level the fine structure is not known, but it is obviously much more complex than that of most structures considered above, which are made of a few kinds of units many times repeated. *E. coli* ribosomes contain about 23 different proteins in the 30S subunit and 43 in the 50S subunit. There is chemical evidence that some 13 of the 30S proteins are present in one copy per ribosome and reason to think that this is also true for the rest of the proteins (162, 228). The spatial arrangement in which these proteins are packed into the ribosome is not known, and it is only now becoming possible to obtain evidence on the biological function of some of the individual proteins.

Active work on ribosome synthesis is going on in several laboratories. (See reference 158 for a review to 1965, 147, 197, and other papers quoted for more recent reviews.) These studies are based partly on labeling experiments with RNA precursors in growing cells and partly on experiments with inhibitors, such as chloramphenicol or 5-fluorouracil, that lead to the formation either of incomplete ribosomes or of related particles (98, 122, 131, 143, 147, 148, 173, 174, 197). The picture emerging from these studies is (for bacterial ribosomes) that 16S and 23S ribosomal RNA are probably associated with a small amount of protein when they first appear in the cytoplasm. Then, ribosomal protein from a cytoplasmic pool that can, in some but not all cases, represent 3 to 4% of the soluble protein (67, 197) is added on in steps to form, finally, the 30S and 50S subunits.

Mangiorotti and co-workers (147, 148) observed that in exponentially growing *E. coli* nearly all newly formed messenger RNA (mRNA) appears in polysomes, but that all newly formed ribosomal RNA (rRNA) is free of polysomes until it is part of a complete ribosomal subunit. Their work (148, 198) also indicates that 30S and 50S subunits do not combine spontaneously but need mRNA and transfer RNA (tRNA), as well as certain ions, to couple into the 70S monomer. That is, they are thought to join together at the beginning of protein synthesis and to come apart when a polypeptide chain is completed (147, 148).

Work by Kaempfer et al. (102) also showed, by a variant of the classic experiment by Meselson and Stahl (158), that ribosomes exchange subunits during normal growth of *E. coli*, presumably by dissociation into and formation from a pool of free 30S and 50S subunits. However, the subunits themselves are stable and do not exchange their protein with that of the cytoplasmic pool. This observation strengthens the conclusion of Santer et al. (197), based on immunological evidence, that the ribosomal proteins in the cytoplasm are a source of those in the ribosomes, not their breakdown products.

Ribosome formation in animal cells can continue if protein synthesis but not RNA synthesis is inhibited by cycloheximide (233). This indicates that pools of ribosomal proteins are present.

The picture of proteins being added on in steps (cf. 67) to an rRNA molecule can be profitably thought of as a case of self-assembly, although a very complex one. Recent in vitro studies have supported this idea. Several groups of workers have shown that chemically dissociated *E. coli* ribosomes can reassemble (166, 210). When 70S ribosomes or 30S and 50S subunits are centrifuged in CsCl gradients, they dissociate into free protein and smaller ribonucleoprotein cores containing about 60% of the original protein (166). The 50S subunit yields a 40S core, and the 30S subunit a 23S core. The degree of breakdown depends on the concentration of Mg++, which seems to antagonize the dissociative effect of CsCl (132, 210). If the protein and the core are mixed and CsCl is removed by dialysis, reassociation occurs to yield ribonucleoprotein particles that resemble the original ribosomal subunits in physical properties and, as far as has been studied, in biological properties. The capabilities of different ribosomal subunits tested are listed in Table 2.

The results can be summarized simply. The native subunits were active in the functions indicated in the table. The cores were inactive, as were the proteins released from the ribosomes (split proteins, 166). Cores could be reactivated if recombined with their own protein but not if
TABLE 2. Reactions carried out by native and reconstituted ribosomal subunits*

<table>
<thead>
<tr>
<th>Subunit</th>
<th>Reaction tested</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amino acid incorporation$^b$</td>
</tr>
<tr>
<td>30S</td>
<td>+</td>
</tr>
<tr>
<td>50S</td>
<td></td>
</tr>
</tbody>
</table>

* From references 166 and 210.

$^b$ Polyuridylic acid-dependent polypehylala-nine synthesis. The reconstituted 30S or 50S subunit was mixed with a native 50S or 30S subunit, respectively, to measure this activity.

$^c$ Polyuridylic acid.

combined with proteins from the other subunits. A mixture of reconstituted 50S and 30S subunits was also active in amino acid incorporation, provided that the individual cores were mixed first with their own split proteins.

Nomura's group (88, 166, 223, 224, 227) separated the split protein from 30S and 50S subunits into acidic and basic fractions and analyzed the role of these fractions and some of their individual proteins in various activities of the subunits. All proteins are not needed for every function. For example, acidic split proteins from 50S particles are essential for amino acid incorporation, whereas basic proteins are stimulatory but not essential. Basic, but not acidic, proteins are indispensable for the tRNA binding of these subunits. Basic split proteins are needed for amino acid incorporation of 30S subunits and for tRNA and mRNA binding. The acidic split proteins are dispensable, but they stimulate amino acid incorporation and tRNA binding (166).

Five separate proteins in the basic fraction of 30S split proteins have been isolated and tested individually in reconstitution experiments for their effect on ribosomal functions. They affect these functions differently (see 166, 223, 224, 227), and this points to the functional heterogeneity of ribosomal proteins, already known to be chemically heterogeneous.

Traub and Nomura (226) were also able to dissect 23S core particles (arising from 30S ribosomes) into "core proteins" and free (16S) tRNA by treatment with urea and LiCl; they then obtained efficient reconstitution of active 30S ribosomal subunits from RNA and protein. Both protein and RNA requirements for this reconstitution were quite specific.

This reconstitution takes place at 37 to 40°C, but not in the cold. More recent results indicate that at low temperatures some proteins combine with RNA; then, after incubation at 40°C, a rearrangement of the bound proteins may occur, enabling them to bind other proteins and build an active 30S particle (P. Traub, personal communication). The temperature dependence of ribosome reconstitution suggests to this reviewer that hydrophobic bond formation between proteins may be needed to make an active ribosomal subunit.

The first reconstitution studies showed that the chemical determinants of streptomycin sensitivity or resistance reside in the core of the 30S subunit rather than in its split proteins (166, 210). The inhibition of protein synthesis caused in E. coli by the colicin E3 is also due to changes within the 30S core (210). Further dissection of the 30S core has now shown that streptomycin resistance is determined by one of the core proteins instead of by RNA (226).

The reconstitution of cores or RNA with split proteins takes place very rapidly (in less than 1 min under physiological conditions) without any other cellular components being present (210, 224, 226). Such a self-assembly process could well function in the cell. According to Lerman et al. (132), the reversible dissociation and reassociation of ribosomes, which proceed not gradually but in very rapid, discrete steps, suggest that there is an organized quaternary structure within ribosomal particles and that the experimental formation of ribosomal subunits from cores and split proteins was a model of natural ribosome biogenesis. This is also supported by the more recent work of Itoh et al. (97).

To sum up, both experiments on living cells and reconstruction experiments with isolated ribosomes create a picture of ribosome formation by stepwise self-directed addition of proteins to RNA. We still cannot be certain that all protein in reconstituted ribosomes is in the right place, but to the extent that in vitro tests reflect ribosomal function, these experiments provide some of the most impressive examples of self-assembly from a lower to a higher level of organization.

The ribosomes of E. coli are dissociated by high salt concentrations, and they reassociate when the salt is removed. A curious reverse case of ribosomal dissociation and reassociation occurs in the extremely halophilic bacterium Halobacterium cutirubrum, which needs nearly saturated NaCl for growth and survival. The ribosomes of this organism also require very high salt concentrations for stability (3.0 to 4.0 M KCl + 0.1 M MgCl₂, to remain in the 70S
form), and in the absence of salt they dissociate into acidic proteins and particles richer in nucleic acids than the original ribosomes (18). (This probably reflects the fact that such organisms growing in media containing high NaCl have a high internal KCl concentration.) Bayley and Griffiths (17) showed that intact ribosomes function in amino acid incorporation with a cell-free extract of these bacteria in the presence of high salt concentrations. Bayley (16) showed that ribosomal subunits could be partly reassociated by raising the salt concentration. So far, only the physical characteristics of such reassociated structures have been studied.

**FIBERS OF MUSCLE AND CONNECTIVE TISSUE**

**Collagen**

Two reviews by Cohen (37, 38) discuss the structure of several fibrous proteins (collagen, the keratins, and myosin) and the possible role of self-assembly in the formation of these structures. Her reviews also deal with the structural role played by α-helices and superhelices made of more than one α-helix (the coiled-coil form). This form is subject to little conformational change, and it confers stability on charged polypeptide chains in an aqueous environment. Thus, it is suitable for making stiff rods such as are found in collagen, one of the most abundant of animal proteins (37, 38, 42, 43).

It is well known that collagen can be dissolved in dilute acids and reconstituted from these solutions as fibrous material resembling, although not the same as, native collagen. Tropocollagen, the basic associating unit, consists of a three-chain coiled coil, about 280 nm long and 1.4 nm wide, of about 300,000 MW; the two ends are different, so that tropocollagen may be considered to have a head and a tail. In native collagen and reconstituted aggregates, subunits are arranged head to tail. The tropocollagen molecules in aggregates are not perfectly matched laterally, but one lap by a quarter of their length in successive layers. This seems to be responsible for the distinctive banding pattern of native and reaggregated collagens. (See Fig. 19 and references 37, 38, 42, 43, 168, and 169 for review.)

The three chains of tropocollagen are themselves composed of subunits held together by ester-like bonds. Petruska and Hodge (178) suggested that formation of chains by subunits of different lengths would provide a vernier mechanism which would determine the length of the assembled tropocollagen. By this scheme, two of the chains in tropocollagen would each contain five subunits, and the third would contain seven slightly smaller subunits; subunits would be added until all chains had come to the same length (Fig. 20).

The idea that self-assembly occurs seems to be a guiding one in current research on collagen structure. However, collagen is obviously not formed by simple quaternary association of monomers. Covalent bonds hold together the subunits of tropocollagen. Furthermore, in the animal body native collagen has covalent cross links, essential for its tensile strength (37).

**Actin**

Students of muscle biochemistry have known for years that raising the ionic strength of solutions of G actin makes these molecules polymerize into F actin (42). Both reconstituted F-actin and actin filaments isolated from muscle appear to consist of two helically wound strands of roughly spherical, identical G-actin subunits of about 5.5 nm in diameter (Fig. 5). Both electron microscopic and X-ray diffraction data agree very well with this picture of actin fibers (76). Similarities between the structure of polymerized

![Tropocollagen Molecule](image1)

![Collagen Fibril](image2)

**Fig. 19.** Collagen organization, showing the "head and tail" structure of tropocollagen (above) and the overlapping of tropocollagen subunits that leads to the banding pattern of collagen (71a). Reprinted by courtesy of J. Gross.
Actin and myosin polymerization, as well as the association of actin and myosin, are favored by raising the temperature (11). Entropy increases as actin polymerizes, a situation similar to that found in the TMV protein. Again, this apparently paradoxical entropy increase upon formation of an ordered structure seems caused by the change in bound water from an ice-like state (on an organic surface) to a liquid state.

Flagellin polymerization also increases with increasing temperature, up to about 26°C. Above this temperature, conformational changes in the flagellin molecules inhibit their ability to polymerize. The thermodynamic aspects of flagellin polymerization below 26°C do not seem to have been investigated to the point that would permit us to decide whether this process also involves an increase in entropy (62), although it probably does.

**Myosin**

Myosin of skeletal muscle is a large, complex molecule whose structure is still imperfectly understood (229, 243). It can be degraded by guanidine or at alkaline pH to subunits: two large ones about 200,000 MW and a few smaller ones of 20,000 MW. Despite their size (see Introduction), the large subunits are probably single polypeptide chains. So far, they have been dissociated only by breaking covalent bonds. More impressive is the fact that they are synthesized on polysomes containing 50 to 60 ribosomes. A comparison of the size of polysomes involved in the synthesis of several other proteins shows that each ribosome corresponds to a polypeptide MW of 3,000 to 4,000. Thus, the large number of ribosomes involved in myosin synthesis seem to be carried on a monocistronic messenger able to specify an MW of about 200,000 (80).

In the electron microscope, myosin consists of straight rods about 130 nm long and 2 nm in diameter with a thicker (globular) region on one end 25 nm long and 4 nm in diameter (37, 38). The thicker portion has adenosine triphosphatase activity and is the site to which actin is bound in muscle fibers. If the KCl concentration of a solution of such molecules is lowered from 0.6 to 0.15 M, protein aggregates are formed, spindle-shaped as are the myosin fibers in muscle, with surface projections like those on such fibers. These projections seem to be formed of the thickened region of myosin. On either longitudinal half of the aggregates, the myosin molecules are arranged in opposite polarities; that is, the thicker parts all point one way on one side of the center and the other way on the other side (89, 90; Fig. 21).
membrane" concept proposed by Dr. [192]. Reprinted by permission of the copyright owner, Academic Press.

Recently, A. Stracher (personal communication) dissociated the globular part of myosin by treatment with high salts; adenosine triphosphatase activity was lost but could be reconstituted by mixing the two protein fragments.

Detailed discussions of muscle structure may be found in references 37, 38, 90 and 139. Although it would seem reasonable that actin and myosin molecules contain in their own structure the determinants for assembly into the filaments found in muscle fibers, we are still at some distance from understanding all factors that contribute to the molecular architecture of muscle.

MEMBRANES

The complexity of membrane function is reflected in membrane structure. Among other functions, membranes are responsible for maintaining ionic differences between the inside and outside of the cell or between the cell compartments they enclose. They may carry out the transport of ions and of many specific low-molecular-weight substances, as well as the transport of macromolecules by pinocytosis and phagocytosis. The membranes of mitochondria and the cytoplasmic membrane of bacteria are sites of oxidative phosphorylation, a function that may be intimately related to transport. This variety of functions argues for a complex structure, or structures, since it is by no means certain that all membranes have the same structure (see 123, 205). No generally accepted membrane model exists, and some of the proposed models of membrane structure differ fundamentally from each other. A detailed discussion of current theories of membrane structure and function cannot be given here (see 21, 22, 123, 183, 192, 214). Some models (Fig. 22a, 22b, 22c) will be briefly described, keeping in mind the questions: is there a subunit structure in membranes and are membranes made by the self-assembly of such subunits?

The classical Danielli-Davson model (41; Fig. 22a) pictures membranes as being constructed internally of a lipid bilayer, the nonpolar ends facing each other and the polar ends facing outward towards layers of protein or carbohydrate, or both. This model is the basis of the "unit membrane" concept proposed by Robertson (192). Under the electron microscope, structures called unit membranes appear as two electron-dense lines with a lighter central zone between them, the whole structure being about 6 to 7 nm wide in membranous organelles of eucaryotic cells.

**Fig. 21. Possible arrangement of myosin molecules in myosin filaments, showing reversed polarity on either side of the center (89). Reprinted by permission of the copyright owner, Academic Press.**

**Fig. 22a. Danielli-Davson model of lipid bilayer membrane structure (41, 192).**

**Fig. 22b. Benson's model of membrane structure (22). Picture supplied by A. A. Benson.**

**Fig. 22c. Subunit model of Green et al. (69), showing single subunits (A) and subunits united to form a two-dimensional sheet (B). Reprinted by permission of the copyright owner, Academic Press.**
cells and 7.5 to 10 nm wide in surface membranes. However, depending on how one looks at them, electron micrographs of thin sections of both fixed and freeze-dried membranes often show regions of transverse density as well (204, 206). Sjöstrand (204) suggested that the transverse densities were caused by cross layers of proteins blocking off small lipid micelles. According to other workers, however, such apparent cross structures are artifacts caused during preparation of material for study in thin sections (192, 214).

The Danielli-Davson model predicts that hydrophilic bonds exist between lipids and proteins. However, there is evidence that in membranes of chloroplasts (100, 101, 236), erythrocytes (36) and H. halobium (144) lipids and proteins are bound hydrophobically. In line with this, Benson, Weier, and co-workers (21, 22, 236, 237) proposed a membrane model in which the protein chains, instead of being spread on the outer surfaces, weave back and forth across the membrane, making hydrophobic contacts at the interior with the hydrocarbon chains of the lipids and hydrophilic contact at the exterior with their polar groups (Fig. 22b). This concept seems more consistent than the unit membrane concept with a subunit model of membrane design. (Membrane "subunits" are thought of as containing several different proteins, lipids, and other molecules.)

Korn (123) criticized the arguments for the unit-membrane concept. He pointed out the difficulties of extending X-ray diffraction data on myelin sheaths, on which this concept is partly based, to membranes with greatly different physiological functions and chemical composition (see also 205), and he also discussed the difficulties of exact chemical interpretations of electron micrographs, difficulties apparent to anyone who has attended any recent conference on membrane structure.

Green's group proposed that membranes are made of repeating units (Fig. 22c). This model arises from studies of mitochondrial membranes, but is thought to be applicable to all (69, 70, 77, 176). Conformational changes in such subunits are believed to be involved in energy transformations carried out by the membrane, that is, in ATP production and in mechanical or other work (77, 176). It is suggested (70) that membranes are composed of cuboidal-repeating units that have phospholipids on two surfaces and that the four surfaces that do not contain lipids can interact, hydrophobically, to form a two-dimensional sheet (Fig. 22c), the open ends of which can finally fuse into a spherical or tubular system. In vivo formation of membranes is thought to go on through self-assembly of such repeating units.

Membrane Synthesis in Cells

Membrane movement has for some time been thought to involve the formation of new membrane in one region and membrane breakdown at another, causing a flow from the first place to the second (20). Campbell (29) suggested that such formation involves condensation of membrane components. Understandably, however, with such uncertainties as exist on membrane structure, knowledge of how membranes are synthesized has not progressed far.

Siekevitz et al. (200) studied the development of membranes in two systems where a net synthesis occurs: the rapid accumulation of smooth endoplasmic reticulum in parenchymal cells of livers of newly born rats and the appearance of chloroplast lamellae that develop in the etiolated cells of a Chlamydomonas strain when it is exposed to light. Different electron transport enzymes appear as the membranes of the endoplasmic reticulum develop, but not all of the enzymes are found in the membrane at the same time. There seems to be a basic membrane structure on which newly synthesized enzymes are inserted. In the Chlamydomonas, membrane synthesis, but not that of electron-transport enzymes, is dependent on chlorophyll formation, and the enzymes are synthesized as soon as the new membrane is formed. Labeling experiments showed that the turnover of total lipid is faster than that of total protein, and some proteins turn over faster than others (200). These admittedly tentative results do not, so far, support the idea of discrete subunits coming together to form a membrane. This work emphasizes how very complex membranes are and how much farther it will be necessary to go to get any clear picture of their formation.

Membrane Reconstitution

Limited success has been obtained in reconstitution experiments. Extracting lipids from mitochondria greatly lowers their rate of respiration. The various parts of the electron-transport chain can be reactivated by adding back the phospholipid extracted or, indeed, quite different phospholipids (200, 200). Although a critical review of current knowledge and theories of mitochondrial structure and function is not possible here, the reader is referred to papers from the laboratories of Green and Racker (25a, 25b, 122a) for recent discussions of disaggregation and reconstitution both of mitochondrial membranes and of the electron transport chain and adenosine
membranes associated with smaller particles away with sodium detergents and that, original such reconstitution in with bile treatment though adenosine lysoseikticus 183, 185, 193, the reconstituted structure containing less protein than the original, for those of Micrococcus lysodeikticus (28), and for the L phase of Streptobacillus moniliformis (184). By treatment with bile salts, Green and co-workers disaggregated many different membranes: the inner and outer membranes of mitochondria, chloroplast membranes, and membranes from bovine photoreceptors, microsomes, and erythrocytes. After removal of the detergent, well-defined vesicular membranes were formed (69).

Terry et al. (220) observed that the degree of disaggregation and reaggregation of Mycoplasma laidlawii membranes depended on the SDS concentration used in the first process and the Mg++ concentration present during the second. Several different kinds of reaggregates were obtained, although all contained some triple-layered membrane-like pieces. However, the reaggregates obtained from membranes that had been disaggregated with the lowest possible SDS concentration resembled the original membranes, not only in appearance but also in equilibrium density.

Are such reconstruction experiments really models for membrane formation in the cell? We do not know. One difficulty in interpretation arises from the lack of a biological test for activity of reconstituted membrane-like structures. To my knowledge, none of these structures has been shown to carry out active transport or to alter the ionic constitution on either side of it. Considering how small reconstituted membranes are, demonstration of these properties would be quite difficult. Some of the enzymes characteristic of membranes, notably the reduced nicotinamide adenine dinucleotide oxidase of Mycoplasma and the respiratory chain enzymes, are found in reconstituted membranes, but the adenosine triphosphatase that seems to be associated with active transport, the Mg++-dependent, Na++ + K+ activated adenosine triphosphatase, has not been found (183).

It is by no means certain that treatment with detergents does in fact degrade membranes to lipoprotein subunits. It may also separate lipid from protein. Subunits of M. laidlawii contain both lipid and protein and appear as a single peak in the analytical ultracentrifuge. However, by careful density-gradient centrifugation, the protein and lipid can be shown to be separate from each other in the presence of detergent. Presumably, they recombine when the detergent is dialyzed away (50).

Finally, we must realize that a great many different “structures,” some quite attractive, can be obtained by adding lipids to water. These include the layered structures of the well-known “myelin figures” obtained by mixing phospholipids with water and a variety of shapes obtained if cholesterol and saponin are also present (123, 141, 214; Fig. 23). Could not some of the trilaminar structures, appearing after detergent is dialyzed away from disaggregated membranes, be due to the lipids alone? The sympathetic reviewer can only recommend caution in interpretation and boldness in devising experiments to test the physiological significance of reaggregated membranes.

REGULAR MACROMOLECULAR LAYERS OUTSIDE OF BACTERIAL CELL WALLS

The outside of the cell walls of several bacterial species, Spirillum serpens, Micrococcus radiodurans, and several Halobacterium species, as well as the outer enveloping membrane of tablets of Lampropedia hyalina, are composed of lipoprotein globules of about 10 nm in diameter, with centers spaced about 14 nm apart, packed in a hexagonal array on the cell surface (165, 195; Fig. 24, 25). Murray (165) studied the fine structure of the external layers of S. serpens and suggested, from electron microscope observations, that each of the 10-nm subunits was itself composed of six hexagonally arranged subunits, about 6 by 2 by 2 nm in size. He also pointed out that hexagonal packing would do well for the cylindrical part of S. serpens, although topology demanded that each hemispherical end of this organism should have six pentagonal arrangements of subunits, just as six pentagons are required for each half of a spherical virus. A few such pentagons were found.

F. Buckmire and R. G. E. Murray found that the external layer of S. serpens is not essential for the life of the cell (personal communication). It is absent from certain mutants and also from cells grown without Ca+++, but when the proper salts are added it appears very quickly, even in the presence of chloramphenicol.

The outer layer may be removed from walls of S. serpens by treatment with guanidine hydrochloride. When such material is added to wall fragments from S. serpens mutants in the pres-
ence of Ca++, a hexagonal array of particles appears on the surface of the wall fragments. These interesting results (which are still in the preliminary stage; F. Buckmire and R. G. E. Murray, Bacteriol. Proc., p. 47, 1969, and personal communication) are consistent with a formation of the orderly array of particles by self-assembly of lipoprotein from an intracellular pool. Presumably, for such a mechanism to operate the bacteria would have to excrete the proteins or lipoproteins.

Preliminary experiments (C. C. Brinton, Jr., J. C. McNary, and J. Carnahan, Bacteriol. Proc., p. 48, 1969) have shown that protein subunits of the tetragonal surface layer of *Bacillus brevis* can also assemble in vitro.

**FIBRIN CLOTS**

The formation of fibrin polymers when blood clots can also be considered an example of self-assembly, although the structure formed is not limited in shape or size. Thrombin, a specific proteolytic enzyme, splits two pairs of highly charged polypeptides from the fibrinogen monomers (probably from their ends), leaving fibrin molecules which then interact to form an insoluble aggregate, the fibrin clot. When such clots are formed in vitro, the subunits are not held together by covalent bonds and can be disaggregated by urea or low pH. In vivo, however, a fibrin stabilizing factor carries out transpeptidization within the aggregate to make a much more stable clot. So far, covalent bond formation after self-assembly has been observed only in such clots and in the extracellular collagen fibers. Present knowledge of the size and shape of fibrinogen and fibrin molecules and of the possible arrangement of the latter in the fibrin clot was reviewed by Cohen (37, 38).
Fig. 24. Regular macromolecules outside bacterial cell walls. (a) A folded cell wall fragment from Spirillum serpens showing structured outer surface and smooth inner surface. (b) Phosphotungstic acid-negative stain of S. serpens wall fragment showing regular packing of polygonal structures. Markers in a and b indicate 100 nm (165). Reprinted by permission of the National Research Council of Canada.
FIG. 25. Surface of the extreme halophile, Halobacterium cutirubrum. Replica was prepared, as described in reference 124 by J. Y. D’Aoust. × 51,000.

CONCLUDING REMARKS

Accuracy of Self-Assembly

Considering the range and importance of self-assembly processes and their present interest, we may expect that other examples will soon be described and that our understanding of those already known will become deeper. Obviously, we cannot make a qualitative distinction between formation of such biological structures as viruses and bacterial flagella and the association of polypeptide chains to form larger protein molecules. Some of the latter, indeed, can be structurally as complex as small viruses. For one example, the iron storage protein ferritin consists of an iron hydroxide complex surrounded by a protein shell of 20 identical, roughly spherical subunits (MW, about 25,000) which seem to be situated on the vertices of a pentagonal dodecahedron (78, 83).

The erythrocrorurins, heme-containing respiratory proteins of certain invertebrates (133), and the hemocyanins, copper-containing respiratory proteins of arthropods and molluscs (26, 27, 85), have molecular weights of three million or more. Changes in pH and ionic strength make them dissociate and reassociate and, even though some architectural details are still uncertain, they appear to have very ordered structures. That proposed for the erythrocrorurins consists of two six-membered rings laid atop each other (133). Models proposed for the hemocyanins are made of stacked rings of up to twenty subunits each (26, 85).

Van Holde (85) pointed out that the formation of hydrophobic bonds is very important in the polymerization of many multi-subunit proteins and that, in general, multichain proteins have higher proportions of nonpolar residues than do single-chain proteins. Increased entropy accompanies the polymerization of the subunits of such
proteins, just as it accompanies the polymerization of subunits of more complex structures. The same kind of thermodynamic driving force seems to be involved.

Large proteins could serve as valuable model systems for studying in vitro self-assembly, for they offer some of the best ways of testing its accuracy. Very small changes in the configuration of enzymes or oxygen-carrying proteins change their kinetic behavior towards substrates, inhibitors, and allosteric effectors. Thus, a detailed kinetic study should permit one to determine any small differences between a native protein and a disaggregated, reconstituted one.

Unfortunately (since to be certain that a structure is formed by self-assembly in the cell one must demonstrate such formation in vitro), most biological structures lack such stringent criteria for architectural exactness. The infectivity of a reconstituted virus does not necessarily indicate how well it is made. Isolated bacterial flagella do not move, and the accuracy of reconstruction can be measured only by physical tests. As far as they have been tested, reconstituted ribosomal particles do appear to be accurately made. A still more stringent proof would be the demonstration that they can participate in making a specific protein in vitro.

Predictions on the Scope of Self-Assembly

“The important consideration is, in what order are we to proceed and when, if ever, are we to stop”? (From “Kai Lung’s Golden Hours” by Ernest Bramah.)

To rephrase the question asked at the beginning of this essay, once information is given to specify the necessary structural and enzymatic proteins, what more is needed to make cell structures or, for that matter, to make a complete cell? In his interesting calculations of theoretical minimal cell size, Morowitz (163, 164) assumed that no more information was necessary. With this assumption, and others, he arrived at a figure smaller (though not ridiculously so) than the size of any known cell.

From present biochemical knowledge Morowitz postulated a cell in which the only structural elements were deoxyribonucleic acid (DNA), ribosomes, and a semipermeable membrane and which would contain essential energy-yielding and synthetic systems. It would have some 45 enzymes, including 20 for amino acid activation, 8 for energy metabolism, 9 for nucleotide triphosphate production, and others. Assuming an average MW of 40,000 for each enzyme and making a few reasonable calculations on the other apparatus needed by a cell lead to an estimate of one with a total MW of about $62 \times 10^6$, of which more than 75% is DNA. If such a cell contains 75% water, its diameter can be calculated to be about 100 nm (164).

Separate calculations based on the magnitude of random thermal motion to which cells of different sizes are subject and on the probability that all essential cell components would be partitioned on division showed that it is very unlikely that biological structures smaller than 40 nm in diameter would be both complex and stable enough to be self-replicating (164).

The smallest known cell, Mycoplasma laidlawii strain H39, is at least ten times as large as the "biochemical" minimal cell. It contains DNA of $500 \times 10^6$ MW and, by using the calculation that a segment of double-stranded DNA of 800,000 MW is needed to code for a protein of 40,000 MW, can be estimated to code for about 625 proteins (163, 164).

Even picayune adjustments of Morowitz's biochemical model would probably not so much as double the theoretical minimum. (I think his estimate of 500,000 for total molecular weights of ribosomal and membrane proteins is much too low and that other enzymes of intermediary metabolism would be needed, since it is unlikely that even the richest medium would supply the proper amounts of every amino acid and base.) Smaller cells than M. laidlawii H39 may well await discovery, but this, the smallest known, does seem to have substantially more information than it needs to specify only the basic enzymatic and structural subunits. I think it likely that some of the 625 genes act in ways still unknown to determine the simple structure of these cells.

As cell complexity increases, so does the superabundance of genetic material. The DNA content of E. coli suggests that its genome contains enough information to code for about 2,800 different protein monomers, probably many more than are needed for structural proteins and for the most complex scheme of intermediary metabolism (164). The DNA content of most eucaryotic cells [including both highly differentiated ones, where all genes are probably not expressed (45), and single-celled organisms] is from 100 to 1,000 times as great as that of E. coli and other bacteria (208). What is all that genetic information doing?

This is hardly a question to answer quickly! It is unlikely that eucaryotic cells need many extra metabolic enzymes. On the contrary, the ability of so many procaryotic cells to grow on simple media is usually taken to mean that their synthetic capabilities are at least as great as those of more complex cells. The latter are distinguished rather by their size and structural complexity, their elaborate processes of mitosis and meiosis, their
possession of specialized organelles, and their division into many interconnecting compartments through intricate membrane systems (43).

It would seem naïve to suggest that such structural complexity arose only by self-directed assembly of basic subunits, especially when quite enough genetic material is present to determine structure in other ways. This, apparently, is what a large proportion of the genome of the complex T4 phage does. More genes seem to be involved in making eucaryotic cilia than those that specify the structural proteins. Although all of the genes involved in determining structures as complex as mitochondria, chloroplasts, and chromosomes have certainly not been analyzed, I predict that when and if this is done we will find that gene products direct form in many subtle ways.

Studies on the control of cellular and subcellular form appear to be at a crossroads. Studies of the simplest systems have told us a great deal about what might be considered the rough determinants of form: the energetic considerations which lead to the polymerization of large molecules, the kind of designs which, being minimal energy states, are readily formed by self-assembly. Some simple mechanisms, dependent only on the subunits themselves, for controlling size have been suggested. These include the control of rod-shaped virus length by their length of nucleic acid and the "vernier" mechanisms evoked for phage tails (7) and collagen subunits.

Judging from the present energy of research, much of that done in the near future will doubtless serve to widen and deepen our knowledge of how self-assembly of subunits leads to the formation of simpler structures. Hopefully, a still wider range of studies, in addition to those on phage morphopoiesis, will be aimed at finding other ways in which gene products direct biological form. However indispensable self-assembly of smaller subunits is to the formation of complex viruses, cell organelles, and all of the networks of membranes of the cell, I believe that as our understanding grows many other gene products will also prove to be at work, guiding the processes of polymerization or, as it were, trimming and fitting together the simple polymers that self-assembly supplies.

ACKNOWLEDGMENTS

Some of this review was written at the Institute of Molecular Biology, University of Geneva, during a visit made possible by a travel grant from the National Research Council of Canada.

I am indebted to A. Tissières for making room for me in his laboratory and to R. Truitt for discussions on ribosome structure. G. Mangiarotti, M. Yanagida, U. Laemmli, M. Kates, and J. G. Kaplan made valuable suggestions on an earlier version of this manuscript. I am especially indebted to E. Kellenberger for valuable discussions on various aspects of self-assembly problems. The following very kindly supplied some of the pictures used: S. Asakura, A. A. Benson, J. Bancroft, F. Buckmire, D. Caspar, E. Kellenberger, R. Edgar, F. Eiserling, A. Glauert, R. Murray, A. Klug, E. Boy de la Tour and J. Y. D'Anust. I also thank others who gave me permission to use published figures and who informed me of unpublished work. I am greatly indebted to H. March and J. Hélie for the drawings and to G. Ben-Tchavtchaud for photographing the figures.

LITERATURE CITED

SELF-ASSEMBLY OF BIOLOGICAL STRUCTURES


63. Gerhart, J. C., and A. B. Pardee. 1964. Aspartate trans-


