Journal of Supramolecular Structure 2:512–514 (1974)

UNDELIVERED SUMMARY REMARKS FOR THE 1974 SQUAW VALLEY MEETING ON ASSEMBLY MECHANISMS

W. B. Wood

Division of Biology, California Institute of Technology, Pasadena, California 91109

The status of research on macromolecular assembly is similar in several respects to that of research on macromolecular synthesis in the late 1950's. The work of that era can teach us some lessons, but it also has left us with some preconceptions that may be misleading us in our attempts to understand assembly mechanisms.

The problem of macromolecular assembly is becoming a central theme of research in cellular and developmental biology, and I believe that its importance will continue to increase during the next several years. In thinking about how to summarize this meeting, I was struck by the similarities between our present situation and the status of research on macromolecular biosynthesis in the middle and late 1950's. I think it may be instructive to look at these similarities in order to gain some perspective on where the emerging field of macromolecular assembly stands now and where it may be going.

By the mid-1950's, the general structures of glycogen, nucleic acids, and proteins were completely understood, but this knowledge contributed little to explaining their biosynthesis. The structures hinted at the energetic requirements for synthesis, and nucleic acid structures provided the clue to the general nature of replication, but two more developments were required before the mechanisms of biosynthesis could be worked out. These were the in vitro systems that allowed biochemical definition of the synthetic reactions, and genetic systems that allowed the number and nature of catalytic elements to be defined mutationally. The first of these developments led, by the early 1960's, to a general understanding of how glycogen, nucleic acids, and proteins are synthesized, and the second development has subsequently contributed to unraveling the details, particularly of DNA synthesis.

Looking at macromolecular assembly as a higher level of biosynthesis, we are perhaps just now beginning to emerge from the structural stage. We know the detailed structures of some oligomeric enzymes and simple viruses and the general structures of muscle filaments, collagen fibrils, flagella, microtubules, and complex viruses. Such information alone tells us only a very limited amount about how these structures are assembled. We are getting information about assembly in a few systems, but only those in which we have the possibility of genetic analysis or in vitro assembly or both.

I find it useful to think about supramolecular structures as the end products of assembly pathways, just as glycogen and nucleic acids are the end products of biosynthetic pathways. As we begin trying to understand assembly pathways, it seems to me that a few

preconceptions, left over from the earlier research into biosynthesis, may be misleading us. Let me enumerate four of these preconceptions.

First, we may be relying too heavily on the notion of self-assembly. The intermediates in our pathways are primarily proteins. We are used to thinking of proteins not as substrates but rather as specific catalysts and combining elements, hence the notion that all of the information for assembling supramolecular structures is contained in their subunits. But imagine trying to design the parts of, say, a sewing machine so that they could be assembled reliably by putting one of each in a large box and shaking it for two weeks. In our experience with mechanical devices, it is far easier to design assembly lines that employ tools, and I believe we will find that "tools" in the form of accessory proteins and other subcellular elements are important in assembling biological structures as well. If so, we can learn what they are and how they work only by studying assembly processes as well as their finished products.

A second possibly misleading preconception is that enzymic catalysis is involved only in breakage and formation of covalent bonds. A chapter subtitle in Watson's *Molecular Biology of the Gene* states: "Enzymes not involved in making (breaking) of weak bonds" [sic]. I believe that this statement will turn out to be wrong, and that we will have to broaden our concept of catalysis in order to understand the roles of accessory proteins in macromolecular assembly. Some of these proteins we know to be traditional enzymes that catalyze posttranslational covalent modifications, such as the hydroxylation of procollagens and the cleavage of viral capsid proteins that were discussed at this meeting. Others, however, are nontraditional; we have heard direct evidence for a recycled scaffolding protein in viral capsid assembly, indirect evidence for catalysis of conformational change in viral tail fiber assembly and attachment, and hints of possible accessory factors in ribosome and microtubule assembly. Conceivably we are entering a new area of enzymology, where we will find classes of specific morphogenetic catalysts that operate in many assembly pathways, just as classes of metabolic enzymes – dehydrogenases, isomerases, carboxylases, and so on - are found in many biosynthetic pathways.

A third possibly misleading preconception is that we adequately understand the genetic control of protein structure. To elucidate protein synthesis, the biochemists of the 1950's and early 1960's had to confront problems of information transmission as well as enzymatic mechanism, a confrontation that resulted in the solution of the genetic code. The present situation in macromolecular assembly is again similar in that we need to think about informational as well as chemical mechanisms. Our problems are more difficult, however, since much of the information we have to deal with is not linearly coded. We tend to be overly impressed with the elegant simplicity of the coding relationship between nucleotide and amino acid sequences, and we pretend that it explains the transmission of information in gene expression. But in fact, it is only the simplest part of the explanation. Even for single protein molecules there is a great deal of ignorance concealed in the statement that amino acid sequence determines three-dimensional structure. We understand this relationship in a qualitative way, but not yet quantitatively enough to predict protein conformations. The problem takes us into the realm of molecular ecology, as Paul Weiss has called it (1), since the environment as well as the protein contributes information to the folding process. Similar but more complex problems confront us as we move up from single protein molecules in solution to higher levels of organization in cells and tissues, where the environment includes preexisting structural elements and nonaqueous phases as well as ions and molecules in aqueous solution. What kind of information directs microtubules or collagen filaments to form specific higher-order arrays, and how is this information translated?

This introduces what may be a fourth misleading preconception, that almost all of the heritable structural information in cells is carried in DNA. In theory, structural information can be maintained in and transmitted by any cell component that is continuous from one generation to the next. The ciliate cortex of Paramecium is such a component; it carries the information for its gross organization and transmits this organization to progeny independently of the genes, as Sonneborn showed several years ago (2, 3). Thus we, like the biochemists of the 1950's, must not only work out assembly pathways for biological structures, but we also must learn where the information for assembly resides and how it is utilized.

All of this may seem fairly obvious. However, one consequence of this line of thought is that if we are in about the equivalent of 1957, then in four or five years things should be really popping, and we should plan to get together again to see what progress has been made.

REFERENCES

1. Weiss, Paul, in "The Neurosciences, a Study Program," F. O. Schmitt (Ed.), p. 801. Rockefeller University Press, New York (1967).
2. Sonneborn, T. M., in "The Nature of Biological Diversity," J. M. Allen (Ed.), pp. 165–221. McGraw-Hill, New York (1963).
3. Beisson, J., and Sonneborn, T. M., Proc. Nat. Acad. Sci. U.S. 53:275–282 (1965).