

solvent which causes the breakdown of base pairing. When RNA synthesis is stopped by actinomycin D, the poly A is also largely suppressed. Evidently therefore the poly A is transcribed from the DNA as part of the RNA chain, and is not added afterwards.

These results then provide support for the messenger precursor hypothesis, and agree with yet another study with a similar aim by Darnell, Wall and Tushinski (*ibid.*, 1321). Working with HeLa cells they too find poly A in the messenger, and in small quantity in the nuclear RNA. They have also hybridized both species of RNA with the endogenous DNA under conditions that lead to the detection only of molecules with sequences multiply "reiterated" in the DNA. The hybridized RNA is then isolated by digesting away the unbound part with ribonuclease. The hybridized messenger is enriched in A, compared with total messenger. It seems then the poly A is a part of, or at least adjacent to, the rapidly hybridizing RNA sequence, otherwise it would be liberated during the nuclease step.

The function of the poly A tract in the messenger is a matter of conjecture. Because the messengers are monocistronic and the poly A is not translated, it is likely that it occupies one end of the RNA chain. From the length of the haemoglobin message and that of the poly A tract, it seems clear that there is no room for the elaborate "leader" sequences that are found in viral RNA messengers, and are of uncertain function. The poly A may be required for translation, or for attachment to the messenger, involving perhaps the group of proteins that appear always to be attached to the messenger, without evidently getting in the way of the translation machinery. This RNA-protein package, in fact, seems to be the form in which the messenger is expelled from the nucleus, and the possibility that the function of the poly A relates to the transport of the messenger into the cytoplasm may also be entertained.

BACTERIOPHAGES

Pathway to a Phage

from our Cell Biology Correspondent
ANYBODY who has looked at high resolution electron micrographs of bacteriophage T4 particles and marvelled at their intricate architecture can readily understand why more than one molecular biologist has been content to devote years fathoming out the genes involved and the steps in their morphogenesis. And the story they have to tell is becoming as fascinating as the T4 phage is intricate. The tally of genes required for T4 morphogenesis has now reached 46, no less, and that

is a minimal estimate. In the current issue of the *Journal of Molecular Biology* (58, 685 and 693; 1971) two groups, Meezan and Wood at CalTech and King at the MRC Laboratory at Cambridge, describe in detail the functions of just four (genes 48, 54, 19 and 3) of the twenty-one genes involved in the assembly of the phage tail.

Both groups find that in *Escherichia coli* infected with strains of T4 carrying amber mutations in any one of genes 48, 54 or 19 phage baseplates—hexagonal structures with six projecting spikes which in the mature phages are attached to the tail core and its surrounding sheath—accumulate, but complete tails do not form. Because extracts of cells infected with one or other of these mutants complement and yield some infectious phage the proteins specified by these genes are able to catalyse tail formation *in vitro*. Moreover such complementation experiments indicate that the products of these three genes are required in a unique order (gene 48 product first, followed by gene 54 product and finally gene 19 product) for the polymerization of the tail core on the baseplate. Exploiting temperature sensitive mutants, King has also shown that gene 19 specifies the major structural protein of the tail core and the proteins specified by genes 48 and 54, not themselves part of the core, somehow act on the baseplates rendering them a suitable substrate for the polymerization of the gene 19 structural protein.

King has also determined the role of another gene, 3; its product is not involved in the polymerization of the tail

core, neither is it the structural protein of the surrounding tail sheath (that is specified by gene 18) but the gene 3 product does react with the terminal annuli of the tail core in the absence of the sheath. It seems that the gene 3 product acts on the free end of the core and converts it to a substrate for the attachment and stabilization of the sheath protein. And once these four genes have acted in the correct order and the baseplate, core and sheath have assembled, the product of yet another gene, 15, can form the actual joint linking the sheath to the tip of the core.

As King takes pains to point out, the assembly of the tail of T4 depends on a sequence of interactions between particular gene products and a macromolecular precursor structure; apparently the various structural proteins, unable to interact strongly when in solution, strongly interact with the macromolecular substrate produced by the preceding step in the morphogenetic pathway. The gene 15 protein, for example, which joins the terminal subunits of the tail core and sheath does not react in solution with the proteins specified by genes 18, 3 and 19, but strongly interacts with them once they are assembled into a sheath and core attached to the baseplate. Presumably during assembly these proteins undergo either allosteric or chemical changes which make them suitable substrates for, and able to interact with, the next protein to be polymerized. Such a highly regulated pathway, which may well be the evolutionary consequence of the fact that the phage is assembled at sites distant from those where the

Sequence and Synthesis of Substance P

IT has been known for forty years that a substance, identified only recently as a peptide, and present in a variety of tissues, possesses hypotensive and a variety of other physiological activities. Appreciable amounts of this material are hard to come by, but in view of its broad neural activity, which is quite different from those of the hypothalamic peptide hormones, the determination of its structure is of considerable interest. The complete sequence is reported by Chang, Leeman and Niall in next Wednesday's *Nature New Biology*. The sequence is arg-pro-lys-pro-gln-gln-phe-phe-gly-leu - met - NH₂. Preparations derived from three different bovine tissues were found to have the same sequence. Substance P is inferred to occur throughout the central nervous system.

The sequence is also compared with those of two other peptides of similar activity, one, physalaemin, from the skin of an amphibian, the others edoisin, from the salivary glands of

a cephalopod. These also have eleven residues, of which four, including the last three at the N-terminus, are identical in all three molecules. In most other positions there are conservative substitutions. Physalaemin and edoisin have pyrrolidone carboxylic acid as the C-terminal residue.

In a second article Tregear *et al.* describe the synthesis of substance P by the solid-phase Merrifield method. The product is active in respect of the various functions associated with natural substance P. This work provides a good example of the value of modern peptide synthesis technique, for it makes available in relatively large quantity a biological material which can otherwise only be isolated in very small amounts. The authors point out the possibilities now open for its biological characterization. The synthetic material will be used to prepare antibodies, with the aid of which it should be possible to locate and assay it in nervous tissue and elsewhere.