another recently found control factor. The lactose and other inducible operons of *E. coli* cannot be expressed completely unless cyclic AMP is present in the cell. The effect of the nucleotide is mediated by a protein factor (termed CAP), which seems to act on the initiation of transcription by RNA polymerase at the promoter of the operon. This factor too seems to act on the complete enzyme.

The similarity between psi and CAP also extends to their interaction with nucleotides. Although no one has suggested that psi interacts with the all pervasive cyclic AMP, Travers throws out a tantalizing hint that it may interact with ppGpp, a guanosine tetraphosphate. The implications of this interaction are exciting. In normal (stringent) *E. coli* cells, starvation for essential amino-acids blocks not only protein synthesis but also the synthesis of ribosomal and transfer RNA, although messenger RNA synthesis seems to continue. The effect is abolished in a (relaxed) mutant strain, in which all RNA synthesis continues irrespective of the state of protein synthesis. Comparison of normal *E. coli* with the mutant suggests that the effect is mediated by ppGpp, which accumulates in stringent cells starved for essential amino-acids.

Psi is a good candidate for the controlling protein which must mediate the effect of ppGpp on stable RNA synthesis. Just as CAP is a protein which when activated by cyclic AMP acts as a general control element which is required for the complete expression of many bacterial operons, psi may be a general control element which, unless inactivated by ppGpp, keeps RNA polymerase transcribing ribosomal RNA, and perhaps also transfer RNA genes. (The action of psi may also explain the greater overall rate of expression of ribosomal RNA genes.) This response to small molecules would allow bacteria independently to adjust the synthesis of different classes of RNA to any changes in the environment.

PROTEINS

Sequence Sequel

from our Molecular Biology Correspondent

FOLLOWING on Jornvall's sequence of the 374 residues of liver alcohol dehydrogenase, about which I wrote a few weeks ago, comes another *tour de force* of protein chemistry. With five of his colleagues, E. L. Smith reports an all but complete sequence for bovine glutamate dehydrogenase (Smith *et al.*, *Proc. US Nat. Acad. Sci.*, **67**, 724; 1970). In its native state, at low concentration, this enzyme is a hexamer, which undergoes an indefinite concentration-dependent association to form long rods. The monomer has a considerably greater molecular weight than those of the other known NAD-dependent dehydrogenases: according to the new sequence there are 506 residues in the chain, and the corresponding weight of 56,000 agrees with physically measured values.

The sequence is regarded as tentative, because at two points overlap has not yet been achieved, and here and there a few residues have not been fully

ordered. The first conclusion from this work is that, because the sequence is unique, the subunits are identical. Taken with the other known sequences for dehydrogenases of different quaternary structure, this seems to establish a clear pattern. There is some polarization of composition, with clusters of charged and hydrophobic residues at opposite ends of the chain. Since there is little perceptible homology of sequence between liver alcohol dehydrogenase and glyceraldehyde 3-phosphate dehydrogenase, which are much closer together in size, it is not surprising that a comparison, using a computer search, of the glutamate dehydrogenase sequence with that of the latter also shows scant similarity. A strange exception occurs at one point in the chain, however: a run of twelve residues—94 to 105 in glutamate dehydrogenase, and 209–220 in glyceraldehyde 3-phosphate dehydrogenase—are identical in six positions, and have conservative substitutions, with single-base mutations in the rest. Moreover, both proteins contain enzymatically essential lysines, and Smith *et al.* have identified the one in their enzyme by way of its interaction with pyridoxal phosphate. This reagent can be used as a covalent label, for reduction of the labile Schiff base adduct leaves it irreversibly bound. The vital residue turns out to be lys-97, which is in the conserved peptide sequence. Glyceraldehyde 3-phosphate dehydrogenase also contains a lysine in the corresponding position of this segment, but the only lysine known to be functionally vital was identified by Harris and Perham as residue 183. It is not obvious what one is to make of this fugitive hint of common gene ancestry.

Smith *et al.* have in addition identified what seems to be an activity-controlling ligand binding site. It is known that modification of one tyrosine residue leads to loss of response towards the non-competitive inhibitor, GTP. The tyrosine, labelled with a 3-nitro group, has been identified as residue 412.

It is worth noting that subunit enzymes with identical chains now seem to be very common. An interesting example is the enzyme, thiolase, in which thiols operate as acyl acceptors in catalysis. The molecular weight is 170,000, breaking down to 42,000 in a denaturing medium. Gehring and Harris (*Europ. J. Biochem.*, **16**, 487; 1970) have ascertained that there is only one N-terminal tripeptide sequence, and that the number of tryptic peptides corresponds to the known content of lysine and arginine, if there are four chains. It is reasonable to infer that the subunits are identical. The protein is inactivated by reaction with iodoacetamide at the level of one molecule per chain. In their second article (*ibid.*, 492) Gehring and Harris report how they isolated peptides from the protein so modified, and showed that the substituent is an o-thiol, which is in a unique sequence at least to the extent of twenty-six residues. It is the same thiol groups which are acetylated when fragments are prepared from the enzyme-substrate complex, and are therefore marked as the catalytic residues.

VIRUSES

Doing without Cells

from our Cell Biology Correspondent

FOR the past several years the ribonucleoprotein particles which contain vaccinia virus DNA, and

which can be found in the cytoplasm of cells infected with this virus, have usually been referred to as "DNA factories". They have now been rechristened in the Latin by Dahl and Kates (*Virology*, **42**, 453; 1970), who apparently believe that virosomes is a better name because it does not "create a misleading impression concerning their composition and function". Be that as it may, Dahl and Kates have succeeded in isolating these complexes from infected HeLa cells. The virosomes are heterogeneous in size and density and the various classes of virosomes are related by some complex precursor product relationship which has yet to be fully resolved. What is more interesting is that the dense virosomes isolated three hours after infection contain nearly all the newly replicated viral DNA and an endogenous RNA polymerase activity (Dahl and Kates, *ibid.*, 463). *In vitro* the isolated virosomes can synthesize RNA which competes in hybridization experiments with both "early" and "late" viral RNAs made *in vivo*.

To convince themselves of the reality of "early" and "late" viral RNAs in infected cells, Dahl and Kates transcribed *in vitro* purified viral DNA with RNA polymerase from *E. coli*. They then analysed the sequences present in RNA extracted from infected cells at early and late times by competition hybridization experiments with the RNA made *in vitro* by the bacterial enzyme. The RNA isolated from cells six hours after infection competed with about 97 per cent of the RNA made *in vitro*. By contrast the RNA extracted from cells 1.5 hours after infection only competed with about 40 per cent of the *in vitro* RNA. Apart from proving that sequences are present in "late" RNA which are not present in "early" RNA, this experiment indicates that *E. coli* RNA polymerase transcribes one and the same strand of vaccinia DNA as is transcribed in cells. Green and Westphal have respectively shown that *E. coli* polymerase asymmetrically transcribes the correct strand of adenovirus 2 and SV40 DNAs. Presumably the bacterial enzyme readily recognizes the promoter sequences in these viral DNAs, which hints that such sequences may be universally similar.

In the same issue of *Virology* (page 508), Bukrinskaya, Bykovsky and Zhdanov report that a ribonucleoprotein complex isolated from Sendai virus particles associates with ribosomes from Ehrlich ascites or chick embryo cells; and, in a suitable cell-free system, it stimulates a five-fold increase in the incorporation of leucine and phenylalanine. This stimulation is inhibited by puromycin, the omission of ATP or the S-100 fraction, and the extent of the incorporation depends on the amount of ribonucleoprotein complex added. Bukrinskaya *et al.* suggest, therefore, that they are detecting protein synthesis directed by Sendai RNA. They further claim that because the incubation mixture after thirty minutes contains an RNA polymerase activity, which is insensitive to actinomycin D, one of the proteins being made is viral RNA polymerase. A full report of these experiments will be awaited with interest.

Jockusch, Bull and Kaesberg (*ibid.*, 401) report that an *E. coli* cell-free system programmed with bacteriophage Q β RNA makes three of the four proteins which are found in extracts of infected *E. coli*. This result, therefore, leaves open the question of whether the Q β genome contains four rather than three cistrons, as does the characterization of Q β mutants reported by Horiuchi and Matsushashi (*ibid.*, 49). They find that the

twenty-two amber mutants which they have tested fit into only three complementation groups.

BACTERIOPHAGE

Heads and Tails

from our Cell Biology Correspondent

THE assembly *in vitro* of such macromolecular complexes as bacteriophage R17, tobacco mosaic virus and ribosomes of *E. coli* from purified preparations of nucleic acids and proteins leaves little doubt about the importance of self assembly mechanisms in the genesis of viruses and cell organelles. But such observations can belie the amount of genetic information which may be required for the synthesis of comparatively simple macromolecular structures, for work with the DNA bacteriophages, such as the T phages and lambda phage, indicates that the number of genes implicated in the morphogenesis of these phage particles far exceeds the number of genes which specify molecules actually found in the particles.

Bacteriophage lambda, for example, is by any standards a simple structure; each particle consists of a head, the phage DNA and associated internal proteins enclosed in a roughly spherical protein coat, and a tail, an aggregate of protein. It is perhaps surprising therefore that genetic analyses of the lambda chromosome have implicated at least eighteen genes in the production of heads and tails, especially since Siminovitch's group (*Virology*, **42**, 375, 390; 1970) have extended earlier work and proved that only three species of proteins account for at least 95 per cent of the total protein in each phage particle.

Buchwald, Steed-Glaister and Siminovitch, using two amber mutants, have isolated pure, empty heads which sediment at 114S and have a molecular weight of 18×10^6 and tails which sediment at 43S and have a molecular weight of 6.1×10^6 . These two components account for about 75 per cent of the protein in the phages. Buchwald, Murialdo and Siminovitch have analysed on SDS polyacrylamide gels the proteins of dissociated heads, tails and whole phage. The head protein has a molecular weight of 37,500, the tail protein has a molecular weight of 32,500; and a protein present in phage, but not in heads or tails, has a molecular weight of 11,000. These three proteins account for 57, 19 and 19 per cent respectively of the total phage protein. Minor protein components no doubt account for the other 4-5 per cent—three to six minor proteins can be detected by electrophoresing dissociated whole phage.

Genetic analyses reported by others and now confirmed by Buchwald *et al.* show that the head protein is coded by gene *E* while gene *V* specifies the tail protein. The gene which specifies the third major protein, that with a molecular weight of 11,000, has yet to be identified and the same is true of the genes which specify the minor protein components. If there are as many as six minor proteins specified by the phage, the roles of nine genes will be accounted for; but what do the other nine genes implicated in the morphogenesis of the phage do? Some at least may well specify proteins required for the assembly of the phage but not components of the particles. If that is the case it will be extremely difficult to define precisely the function of these putative proteins, but their existence should caution against over-emphasis of the role of molecular self assembly.