The Arabinose Operon *in vitro*

ALTHOUGH the operon model of gene organization and regulation can account for both the negative control of gene expression by repressor molecules which prevent transcription and the positive control of expression by activators which promote transcription, it is far easier to envisage how a repressor might work than how an activator might work. And in the event, negatively controlled operons have proved to be far more amenable to experimentation than their positively controlled counterparts. Genetic and biochemical analyses of the *lac* operon, for example, have reached the point where the complete switch gear required for the regulated expression of this small segment of the *Escherichia coli* chromosome can be reconstructed *in vitro*. Pure repressor protein can be obtained in quantities sufficient to allow protein chemists to analyse the structure and active sites of functional repressor and determine the amino-acid sequence of its subunit polypeptide chains, and surely soon, if it has not already, *lac* repressor will attract the attention of X-ray crystallographers.

By contrast, analysis of positively controlled operons—the arabinose operon of *E. coli* is the classic example—is in its infancy. Biochemical analysis in particular has been so slow for one outstanding reason ; the regulator protein, the so called araC protein specified by the *araC* gene, has defied assay and without an assay there is no hope of purifying this molecule This gloomy picture may change rapidly, however, now that two groups, Greenblatt and Schleif and Zubay, Gielow and Englesberg, have, as they described last week in *Nature New Biology* (**233**, 164 and 166 ; 1971), shown that it is possible to assay the araC protein, albeit in a rather cumbersome way, because it is required for the expression of the arabinose operon in a cell-free system which supports coupled transcription translation.

What comprises the arabinose operon ? It is a sequence of three structural genes, which specify enzymes involved in arabinose metabolism, a regulator region and the *araC* gene which specifies the regulatory protein. Experiments with wild type *E. coli* and a host of variously mutated strains have led to an intriguing model of the control of this operon which envisages that a protein specified by the *araC* gene can act as both repressor and activator. The idea is that the araC protein can exist in one or other of two conformations which are in equilibrium. In its repressor state the araC protein is assumed to bind to an operator site and prevent transcription of the three structural genes. In its activator state, assumed by the araC protein when arabinose is present (presumably as a result of an allosteric change following the binding of arabinose), the araC protein permits rather than represses transcription. In short, it is believed that arabinose shifts the equilibrium between repressor and activator in favour of the latter. But the proof of the model is in the test tube and a complete reconstruction of this regulatory system depends on obtaining pure araC protein.

As Greenblatt and Schleif comment, in a few words which belie the years of frustrating research which has failed in its chief objective, attempts to detect the araC protein by virtue of its putative ability to bind either to arabinose or to arabinose operon DNA have been totally unsuccessful. For this reason these two workers, and independently Zubay and his colleagues, turned to the cell free system previously developed by Zubay's group to study the regulated expression *in vitro* of the well characterized *lac* operon. Both groups have taken the line that if the araC protein together with arabinose really are absolutely required for the expression of the arabinose operon *in vivo*, as the model predicts, the same should also be true of a cell free system programmed with arabinose operon DNA, which in practice means the DNA of the transducing phage $\phi80dara$ that carries the arabinose operon. Synthesis of ribulokinase, one of the three enzymes specified by the operon, should in the presence of arabinose depend on the addition to the cell free system of extracts from cells known to exhibit regulated inducible expression of the operon and therefore presumed to contain araC protein.

As both groups report, this is precisely the case ; by exploiting various strains carrying mutations in the arabinose operon they have shown that an active *araC* gene product is essential for ribulokinase synthesis. AraC protein acts as an activator *in vitro*. But what about its other putative function, acting as a repressor ? Greenblatt and Schleif set up the cell free system and added regulatory protein extracts from a mutant strain which is constitutive (araC^c) ; this strain expresses the arabinose operon in the absence of arabinose and this expression cannot be inhibited by adding the sugar D-fucose, which is the case with wild-type cells. *In vitro* they found, surprisingly, that, with extracts containing araC^c protein, arabinose is still required for ribulokinase synthesis but expression is not inhibited with D-fucose. When, however, more and more of an extract containing wild type araC protein is added to the cell free system the synthesis of ribulokinase, while remaining dependent on arabinose, becomes increasingly sensitive to inhibition by D-fucose and repression of the operon is observed as a result of the presence of the wild type araC protein.

Greenblatt and Schleif interpret in an interesting way the curious discrepancy that in constitutive mutant cells the expression of the operon occurs in the absence of arabinose whereas *in vitro* when a cell free system is supplemented with araC^c protein arabinose is required for ribulokinase synthesis. They suggest that the mutation to constitutivity does not freeze the araC protein in the activator state and render it functionally independent of arabinose but rather that the mutation increases the spectrum of co-activator molecules such that in addition to arabinose some other (unspecified) molecule, which must be present in all growing *E. coli* cells, can cause the araC^c protein to assume the activator configuration.

All this is heady stuff, not least because by exploiting this assay for the araC protein it should be possible to monitor its purification. And what tales the protein chemists and crystallographers will have in store once enough of the araC protein is available for them to get to work and define at a few Ångstroms resolution the conformational gyrations of a protein which can function as both gene specific repressor and activator.