Yeast genetics

Genetic control mechanisms: transcriptional twisting

Ian W. Dawes

YEAST geneticists and molecular biologists have long been advocating yeast as a model eukaryotic organism as well as an organism for exploitation by the biotechnologist. Reservations about the usefulness of such a model are based largely on the relatively uncomplicated nature of the yeast life cycle and partly on differences from higher eukaryotes in the mitotic and meiotic processes. At a recent meeting*, however, it became clear that much progress has been made in the study of those cellular processes for which yeast is ideally suited. One example of such a process is the regulation of transcription.

Now that important regulatory sequences have been identified upstream of several genes in yeast, the search is on in earnest to find proteins binding to these sequences and to determine how this binding leads to specific control as well as initiation of transcription. With regard to the nature of the upstream sequences concerned with positive control, the seven genes for allantoin degradation (*DAL1* to *DAL5, DAL7, DUR1,2*) which are all under the control of the *DAL81* gene, contain two types of consensus region (T.G. Cooper, University of Tennessee, Memphis). For efficient constitutive expression, Cooper proposed an upstream expression sequence (UES) which is found in front of all the above genes at least once. For the highly regulated *DAL4, DAL7* and *DUR1,2* genes there is a second sequence, an upstream modulation sequence (UMS) present in multiple copies, which is involved directly in the inducibility of these genes. For the less inducible genes (*DAL1* and *DAL2*) there are fewer UMS elements, and these exhibit only limited homology with the consensus region. For the constitutive genes (*DAL3* and *DAL5*) there is no UMS.

The distinction between expression and regulation sequences is not radically different from what has been reported for other genes that have been examined (and maybe soon the terminology will be agreed), but the nice point of this work is the extent to which the correlation between control sequences and inducibility holds for a group of genes under one type of control.

It also illustrates the extent to which UMS and UES sites can be fairly scattered around in the 5′ upstream region of a naturally regulated system. For negative control of the *CAR1* gene, Cooper's group has found a 13-base pair sequence (an upstream repression sequence, URS) that is necessary for repression, and which can control *CYC1* expression if cloned upstream or downstream of the upstream activating sequence (UAS) elements. This sequence is present in the 5′ flanking regions of more than a dozen genes of widely diverse functions, including nitrogen catabolism, protein synthesis and glycolysis, and it may be part of a more global regulatory system, or possibly a generic sequence associated with a repressor binding site as it does show some homology with the phage λ repressor binding site.

Several points have emerged about proteins binding at the UAS elements. The DNA-binding proteins that have been sequenced have short domains for DNA binding relative to the overall length of the molecules (for example, the amino-terminal 74 residues from 881 for the product of the regulatory *GAL4* gene: (G. Gill, Biochemistry and Molecular Biology, Harvard University) and 60 carboxy-terminal residues from 282 for the *GCN4* gene product conferring general amino-acid regulation (K. Struhl, Harvard Medical School). These DNA-binding regions show the helix-turn-helix motif of prokaryotic repressors. Other regions of these control proteins are necessary for the activation of transcription, thus fusions of the carboxy-terminal end of the *GAL4* gene product and the *Escherichia coli lexA* protein will bind to *lexA* promoters and initiate transcription in a heterologous yeast system (Gill). The elegant work from Struhl's laboratory with the *GCN4* gene product shows that in addition to the short DNA binding site at the carboxy terminus, this protein has a central and even shorter (19-residue) acidic domain needed for the activation function. There are similar acidic regions in other yeast regulatory proteins and in the HMG proteins thought to be important in regulation of higher cell development (Struhl). Deletions between the DNA-binding and acidic domains do not abolish the activity of the protein *in vivo*.

The presence of a single DNA-binding regulatory protein at the UAS is not sufficient for control — other proteins are involved. This has been found for the upstream control of the two cytochrome *c* genes *CYC1* and *CYC7* (L. Guarente, Massachusetts Institute of Technology). *CYC1* expression is regulated by haem and glucose from two elements, *UAS1* and *UAS2*. Dissection of the 73-base pair *UAS1* region reveals one protein (plus) binding the one region, and two others (HAP1 and RC2) to another. How do the proteins binding at these UAS sites affect the initiation of transcription which can be located from 250 to 1,000 or more base pairs downstream from the UAS site? Rudi Planta (Vrije Universiteit, Amsterdam) described the proponents of various hypotheses as being either "twisters, sliders, oozers or loopers" (see Ptashne, M. *Nature* **322,** 697–701; 1986). Gill, a confirmed looper, suggested for the *GAL1* gene that up to four molecules of the *GAL4* product bind at UAS_G and interact with other proteins binding nearer the start of transcription, thereby introducing a loop in the DNA and bringing the UAS and the transcription start site into juxtaposition.

With all the current discussion of whether the human genome should be completely mapped and sequenced, and who will pay for the exercise, it is interesting to see how far similar exercises have progressed in an organism with only a fraction of the haploid DNA content. M. Olson (Washington University School of Medicine, St Louis) reported on progress in the task of restriction-mapping the entire yeast genome at the 2-kilobase level. By using 'bottom up' screening of about 5,000 λ inserts for overlaps, about 600 local maps have been identified, which should reduce to about 200 on a second pass through the restriction data. The other approach of 'top down' mapping using OFAGE electrophoresis to separate the 16 yeast chromosomes and map them individually with the rare cutting restriction enzymes *Not*I and *Sfi*I, generating about 200-kilobase fragments, is also well under way, with more than half of the 55 *Sfi*I fragments mapped, and complete maps for eight of the chromosomes. One of the early goals of this work should soon be realized: the production of an ordered set of λ clones to allow precise location of a newly cloned gene in a single hybridization experiment. The organization of chromosome 1 (the smallest chromosome, in which 250 kilobases can encode only about 100 genes) is the furthest advanced. Kaback and colleagues (New Jersey Medical School, Newark) have mapped 160 kilobases of almost contiguous DNA representing 70 per cent of the chromosome, and have also studied transcription from genes previously undetected by mutation.

Any report of a meeting of the size and interest of this one is bound to be fragmentary and selective. Fortunately, for the first time the complete abstracts of a meeting in this series are available in published form as a supplement to the journal *Yeast*. □

*13th conference on yeast genetics and molecular biology, held at Banff, Alberta, 1–5 September 1986.

Ian W. Dawes is in the Department of Microbiology, University of Edinburgh, West Mains Road, Edinburgh EH9 3JG, UK.