Arabinose C Protein: Regulation of the Arabinose Operon in vitro

JACK GREENBLATT
Biological Laboratories, Harvard University, Cambridge, Massachusetts 02138

ROBERT SCHLEIF
Graduate Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02154

The detection of the gene ara C protein using a DNA-dependent in vitro protein synthesizing system represents the first isolation of an operon specific positive regulator with an in vivo role that has been genetically defined.

The protein product of the gene ara C is required for normal expression of the l-arabinose operon in Escherichia coli. If the C protein is absent, either because of deletion of the C gene or a nonsense mutation within it, the arabinose operon cannot be induced. The C protein is therefore a positive regulator of the operon. C gene mutants known as C- have also been isolated which have high constitutive levels of arabinose enzymes in the absence of L-arabinose. Cells containing a C+ gene in addition to a C- gene are not constitutive. From this and other observations it was concluded that, in addition to its positive control properties, C protein also has repressing properties. The biochemical basis for these control functions remains unknown.

Here we report the detection of the gene ara C protein using a DNA-dependent in vitro protein-synthesizing system in which ribulokinase, an enzyme of the arabinose operon, is synthesized subject to both the positive and the negative controlling influences of exogenously added C protein. Furthermore, the ribulokinase synthesis is L-arabinose and 3',5'-cyclic AMP dependent and is inhibited by D-fucose, as predicted from the in vivo behaviour of the operon. Further purification and a more detailed study of the arabinose C protein are now possible.

The C Protein Assay

The arabinose C protein is a regulatory protein possessing no known enzymatic activity to guide its purification. The in vivo behaviour of the arabinose operon makes it likely that C protein binds L-arabinose and also suggests that it may bind to DNA or to RNA polymerase. Nonetheless, our attempts to detect C protein using these properties have not been successful. We therefore turned to the ability of C protein to induce the synthesis of arabinose operon enzymes. The cell-free DNA-dependent in vitro protein-synthesizing system developed by Zubay and his collaborators synthesizes enzymatically active β-galactosidase. This system requires the addition of lac operon DNA and responds appropriately to lac repressor, galactoside inducers and 3',5'-cyclic AMP. We have used the system they have described to identify the action of C protein.

DNA containing the arabinose operon and a sample to be tested for C protein activity were added to a cell-free system containing no arabinose C protein and no arabinose operon enzymes. After 1 h of protein synthesis, we assayed for the production of ribulokinase, an enzyme coded by the operator proximal gene of the arabinose operon (ara B).

To assay the small quantities of active enzyme synthesized in an in vitro system it is necessary to have a sensitive assay and to begin with a system devoid of the enzyme. Ribulokinase transfers a phosphate from ATP to ribulose. If radioactive ribulose is used as a substrate, the amount phosphorylated can be measured by precipitating and counting the ribulose phosphate product. In our assay conditions the rate of synthesis of product is constant for 4 h, after which it slowly decreases. When the incubation time is less than 15 h, the assay response is linear with concentrations of enzyme up to 0.04 μg. The assay has a sensitivity of 0.002 μg corresponding to about 5 × 10⁶ dimeric ribulokinase molecules. The requirement for a system devoid of ribulokinase was met by using as a source of cell-free extract in the protein synthesis system cells deleted for the arabinose operon so that no arabinose enzymes whatever could be present.

The activity of the protein-synthesizing system was monitored by measuring its ability to synthesize β-galactosidase. This required that the protein-synthesizing extract be prepared from a strain with deletions in both the arabinose and lactose genes. The lac operon DNA used was that located on the transducing phage λphac₃ (ref. 7).

The transducing phage λdara was used as a source of arabinose operon DNA because its genome is a hundred times more concentrated in the operon than is E. coli DNA. We had already shown that in λdara the arabinose operon and the gene ara C have replaced some of the phage late genes and that they are transcribed and translated under the control of the system which controls the late genes. Thus extracts prepared from cells in which this phage is growing should be a rich source of C protein. These extracts must not, however, contain any ribulokinase. A point mutation was therefore put into the ribulokinase gene of the λdara phage and it was grown in cells deleted of their chromosomal arabinose genes. By these means we obtained the essential ingredients for an in vitro system.
assay of C protein: a rich source of C protein uncontaminated by ribulokinase, a source of arabinose operon DNA, and a system with which to attempt its transcription and translation into enzymatically active ribulokinase.

Induction of Ribulokinase Synthesis

Ribulokinase is synthesized when C protein-containing extract, l-arabinose, and 3',5'-cyclic AMP are added to the cell-free protein synthesizing system described above and defined explicitly in the legend to Fig. 1. The background level of the ribulokinase assay was measured by including chloramphenicol or omitting λdara DNA during the in vitro synthesis. All data shown are the actual counts measured without subtraction of the very reproducible background. Detectable ribulokinase synthesis requires the presence of C protein (closed circles in Fig. 1), and is completely inhibited by the addition of chloramphenicol (open circles in Fig. 1). The synthesis is dependent upon the addition of λdara DNA, the response being linear with added DNA up to at least 50 μg/ml (data not shown). Ribulokinase synthesis is routinely obtained and the amount of ribulokinase synthesized varies from 0.1 to 0.3 μg/ml of in vitro synthesis reaction.

The preparation of the C protein containing extract used to stimulate ribulokinase synthesis was different from that of the arabinose deletion crude extract used in the in vitro system (see legend to Fig. 1). The C protein extract was dialysed for a much shorter time at 4°C and it was not incubated for 1 h at 37°C. It was therefore possible that the stimulation shown in Fig. 1 could be caused by some small molecule or by some macromolecule other than C protein. In addition, the stimulation could be effected by some regulator of phage lambda since the C protein-containing extract was prepared from cells growing λdara and the template arabinose operon DNA is located within λdara DNA. All these possibilities were largely excluded, however, by testing an extract prepared from cells growing λhy80, the phage most closely related to λdara. That this extract does not stimulate ribulokinase synthesis, even though one prepared exactly in parallel from cells growing λdara does (Fig. 2), suggests that the product of a gene of the λdara phage is required for ribulokinase synthesis and that the implicated gene is not a phage gene.

Proof that C Protein is Involved

If the C protein controls ribulokinase synthesis, then a mutant C protein should have similar altered properties in in vivo and in vitro systems. A characteristic in vivo property of the l-arabinose system of E. coli is the interaction of D-fucose, an analogue of l-arabinose, with C protein that antagonizes induction of the arabinose operon. C mutations can be found which lie within the C gene and which allow cells to grow on l-arabinose in the presence of D-fucose. C mutants usually seem to be constitutive in vivo, but it was necessary to include l-arabinose in in vivo experiments with mutant C protein because, as we shall mention in more detail...
Fig. 3 Proof that it is a protein coded by ara C that is stimulating ribulokinase synthesis. 50 μl in vitro synthesis reactions contained 200 μg of C protein containing extract prepared from cells growing either the phage λdara C'B' (a) or the phage λdara C'B' (b). Synthesis and assay conditions were as described in the legend to Fig. 1, except that synthesis reactions contained the indicated amounts of α-D-fucose (Sigma). Assay backgrounds were defined by running synthesis reactions in the presence of chloramphenicol. The background is somewhat higher in (b) than in (a) because the phage λdara C'B' makes a small amount of ribulokinase during its growth. Ribulokinase assays were incubated for 6.5 h. Only C+ protein dependent ribulokinase synthesis is sensitive to D-fucose inhibition.

later, C protein does not induce either in vitro or in vivo in the absence of L-arabinose or some other inducer. An extract containing C protein was derived from cells producing C ribulokinase-negative λdara phage.

As expected from the in vivo behaviour of C+ cells, addition of D-fucose to a system containing wild type C protein and L-arabinose completely depresses ribulokinase synthesis (Fig. 3a). In the crucial experiment, however, ribulokinase synthesis stimulated by mutant C* protein in the presence of L-arabinose is resistant to D-fucose inhibition (Fig. 3b). The corresponding in vivo data on the behaviour of these C gene alleles are shown in Table 1. In agreement with the in vitro data, the C* allele is D-fucose resistant while the C* allele is normally inducible and D-fucose-sensitive. This correspondence proves that the λdara gene whose product is essential for ribulokinase synthesis in vitro is the E. coli gene ara C.

<table>
<thead>
<tr>
<th>ara C allele</th>
<th>L-arabinose</th>
<th>D-fucose</th>
<th>Monomers per cell x 10^1</th>
</tr>
</thead>
<tbody>
<tr>
<td>C+</td>
<td>+</td>
<td>+</td>
<td>0.09</td>
</tr>
<tr>
<td>C</td>
<td>+</td>
<td>+</td>
<td>4.4</td>
</tr>
<tr>
<td>C+</td>
<td>+</td>
<td>+</td>
<td>10.0</td>
</tr>
<tr>
<td>C</td>
<td>+</td>
<td>+</td>
<td>8.5</td>
</tr>
</tbody>
</table>

Arabinose isomerase levels in arabinose deletions containing λdara B' phage. These strains are spontaneous revertants to ara' of the strains containing the prophage λdara C'B' and λdara C'B' which were used in the in vitro experiments described in the text. The cells were grown in minimal medium containing 1% glycerol and 1% casamino-acids. Arabinose isomerase was assayed as described previously. The C* allele is L-arabinose inducible and D-fucose sensitive, while the C* allele is constitutive and D-fucose resistant. Note that the C* allele is induced to a higher enzyme level by L-arabinose.

**Table 1 Properties of C* and C Alleles used for in vitro Studies**

In vitro Regulation

C protein-dependent ribulokinase synthesis in vivo requires the presence of an inducing amount of L-arabinose. The corresponding in vitro induction experiment is shown in Fig. 4. Ribulokinase synthesis is completely L-arabinose dependent, requiring for half maximal induction an L-arabinose concentration of 10^-3 M. Thus the in vivo system has the characteristic in vivo property of being induced by L-arabinose. It has been shown that an intracellular L-arabinose concentration of 10^-4 M is required for full induction of the arabinose operon of E. coli K12 (ref. 12). The apparent Kms observed in vivo and in vitro could depend on the relative concentrations of C protein and C protein sites of action (DNA or, possibly, RNA polymerase). Also, the gene ara C of the phage λdara is derived from E. coli B/r,10, and the C protein of this strain of E. coli might have an even higher Kms for L-arabinose than E. coli K12. Therefore the apparent Kms measured in vivo and in vitro cannot yet be directly compared. The relatively high concentration of L-arabinose required for induction explains why it has not been possible to detect C protein by the radioactive sugar binding method used by Gilbert and Müller-Hill to detect lac repressor.13.
and in vitro of L-arabinose and D-fucose (Figs. 3, 4 and 6) confirm that in vitro induction and repression are a result of C protein action.

**New Property of C Protein**

The D-fucose resistant ara C mutants are usually constitutive with variable levels of arabinose operon enzymes in the absence of L-arabinose. It was a surprise, therefore, to discover that C protein prepared from the one high level C strain we have tested depends completely on the presence of L-arabinose for its in vitro induction of ribokinase synthesis. The concentration of L-arabinose giving half maximal induction with C protein is about $10^{-3}$ M, the same as with wild type C protein.

A second experiment prompted by the in vivo behaviour of the arabinose operon is to show that the C protein-dependent ribokinase synthesis requires the presence of 3',5'-cyclic AMP. The arabinose operon of λdara is catabolite repressible in vivo, implying that CAP factor and 3',5'-cyclic AMP are required for full expression of the operon. The in vitro protein synthesizing system already contains CAP factor, but the omission of 3',5'-cyclic AMP from the system should reduce or abolish ribokinase synthesis. Fig. 5 shows the effect of restoring 3',5'-cyclic AMP (chromatographically checked before use) to the system. Ribokinase synthesis is stimulated at least ten-fold at saturation. The concentration of 3',5'-cyclic AMP giving half maximal stimulation is approximately $5 \times 10^{-4}$ M, ten times higher than that reported most recently for the lac system, but about the same as the value reported in a preliminary communication. The inhibition of ribokinase synthesis by D-fucose (Fig. 3a) and the dependence of synthesis on 3',5'-cyclic AMP and L-arabinose provide strong evidence that the arabinose operon is being regulated correctly in vitro.

**In vitro Repression by C Protein**

Genetic data have suggested that, in addition to its inducing properties, C protein can also function as a repressor of arabinose enzyme synthesis. If a cell is diploid for its arabinose region with genotype C'/C', the phenotype is C' (inducible and D-fucose sensitive) rather than C (constitutive and D-fucose resistant). We have exploited this dominance of D-fucose sensitivity to test for C' repressor activity. C' protein was used to stimulate ribokinase synthesis in the presence of an inducing amount of L-arabinose. As was previously shown in Fig. 3, this synthesis is completely resistant to D-fucose. But as increasing quantities of C' protein were added and the in vitro system approached the equivalent of the C'/C' diploid state, the synthesis became increasingly sensitive to D-fucose inhibition (Fig. 6). Thus the C' phenotype is dominant to C, and the repressing properties possessed by C' protein in vivo are also elicited in vitro. The corresponding effects in vivo.

This result suggested that, at least for some D-fucose resistant ara C mutants, the changed C' protein can induce in vivo in the absence of L-arabinose, not because it is frozen in an inducing conformation, but because it is induced by some metabolite always present in growing cells. In an effort to find the unknown metabolite we tested various carbohydrates for ability to induce the C' protein in vivo. Among those tested were L-arabinose, D-fucose, D-galactose, D-glucose, D-ribose and D-xylose. The only in vivo inducer of C' protein we found, besides L-arabinose, was D-fucose, and D-fucose is probably not a naturally occurring compound in E. coli.

After this unexpected in vitro behaviour of C' protein was discovered, we examined the in vivo induction properties of several independently isolated D-fucose resistant ara C mutants (Table 2). The four low level C' mutants tested are induced by L-arabinose or D-fucose to higher arabinose enzyme levels. The two high level C' mutants tested, including the λdara C'B' used for in vitro studies, seem to be induced to a higher level only by L-arabinose and not by D-fucose. The data suggest that D-fucose does not induce as efficiently as either L-arabinose or the postulated natural inducing metabolite. Nevertheless, it is clear that in these mutants the structure of
C protein is modified so that a ligand, D-fucose, which inhibits the wild type C+ protein induces the mutant C- proteins. Although it is only useful for ara C- mutants, D-fucose is the only known gratuitous inducer of the arabinose operon.

We have used the cell-free DNA-dependent protein-synthesizing system developed by Zayub to synthesize ribulokinase, an enzyme of the L-arabinose operon of E. coli. Ribulokinase synthesis in vitro requires the presence of DNA containing the arabinose operon, 3',5'-cyclic AMP, L-arabinose, and the protein product of the gene ara C. The omission of any one of these results in little or no ribulokinase synthesis. We have therefore developed an assay for biologically active C protein with a system which can also be used to study the properties of C protein regulation of arabinose gene function.

Together with the dependence of ribulokinase synthesis on added L-arabinose DNA the experiments shown in Figs. 1-6 demonstrate that the observed ribulokinase synthesis is de novo synthesis and not an artefact caused, for example, by microbial growth during the ribulokinase assays. Furthermore, because two independently isolated ara B- mutants, one of them an amber mutant (Table 2), have been used in these experiments as a source of C protein (Fig. 3), the observed synthesis is unlikely to reflect an unusual property of the particular ara B- mutant used.

Table 2 D-Fucose and L-Arabinose Induction of D-Fucose Resistant Constituves

<table>
<thead>
<tr>
<th>Strain</th>
<th>Monomers isomerase per cell ( \times 10^4 )</th>
<th>Broth</th>
<th>Broth + 2 x 10^-3 M L-arabinose</th>
<th>Broth + 2 x 10^-3 M D-fucose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type K-12 Hfr H</td>
<td>&lt;0.1</td>
<td>2.3</td>
<td>&lt;0.1</td>
<td></td>
</tr>
<tr>
<td>High level C-</td>
<td>1.2</td>
<td>2.3</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Low level C-1</td>
<td>&lt;0.1</td>
<td>2.7</td>
<td>0.9</td>
<td></td>
</tr>
<tr>
<td>Low level C-2</td>
<td>0.15</td>
<td>2.7</td>
<td>1.2</td>
<td></td>
</tr>
<tr>
<td>Low level C-3</td>
<td>&lt;0.1</td>
<td>2.5</td>
<td>0.8</td>
<td></td>
</tr>
<tr>
<td>Low level C-4</td>
<td>&lt;0.1</td>
<td>2.5</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>Wild-type dara C* B-</td>
<td>0.01</td>
<td>2.1</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>High level dara C* B-</td>
<td>0.25*</td>
<td>0.49</td>
<td>0.10</td>
<td></td>
</tr>
</tbody>
</table>

Arabinose isomerase level (monomers per cell) in cells growing in broth plus the indicated sugars. All constitutives were derived from the wild type parent by nitroguanidine mutagenesis and grow at normal rates. The constitutives are fully D-fucose-resistant in arabinose minimal medium. D-Fucose does not induce as well as L-arabinose. The D-fucose was examined chromatographically and found to be free of L-arabinose (< 5%).

* The B- mutation in this phage is an amber mutation, suppressible by sucrose and 90% polar on the expression of the ara A (isomerase) gene.

The experiments depicted in Figs. 1-6 show that most of the known in vivo properties of C protein regulation also occur in the in vitro system. L-Arabinose is an inducer of the arabinose operon in vivo, and the same is found in vitro. D-Fucose, an antagonist of L-arabinose induction in vivo, also antagonizes induction in vitro. Furthermore, C- protein prepared from a strain resistant to D-fucose inhibition in vivo is insensitive to D-fucose inhibition in vitro. This last fact proves that the system is indeed responding to C protein, for the C- mutation is located within the gene ara C. Finally, in agreement with the fact that the arabinose operon is catabolite repressible in vivo, the in vitro synthesis of ribulokinase is dependent upon added 3',5'-cyclic AMP.

Genetic evidence has suggested that C protein can function in vivo as either a repressor or an inducer.1,14 This same functional duality has been demonstrated in vitro. In addition out in vitro experiments suggest that at least some D-fucose resistant ara C- mutants have constitutive synthesis of arabinose enzymes as a result of an increased sensitivity of mutant C- protein to some inducing metabolite always present in growing cells. In many of these D-fucose resistant mutants, C protein is modified so that D-fucose itself is an inducer, in sharp contrast to the properties of wild-type strains.

In recent years several negative regulators have been isolated, among them the lac repressor13 and the repressors of phages λ (ref. 17) and 434 (ref. 18). These act to prevent messenger transcription when bound to homologous operator DNA19,20. In the presence of galactoside inducers the lac repressor detaches from the lac operator and allows transcription to proceed into the lac operon.1,12 The mode of action of positive regulators, however, is much less clear. Several general positive-acting elements involved in the transcription process have been studied. Sigma factor is required for specific initiation of RNA synthesis by RNA polymerase23 and is therefore a positively acting control element. The same is true of CAP factor, a protein which binds 3',5'-cyclic AMP24 and is required for messenger synthesis from the catabolite repressible operons.25 The factor ψ is similar to CAP in that it also stimulates, at least in vitro, the synthesis of a specific class of RNA molecules, in this case ribosomal RNA24. The detection of C protein represents the first isolation of an operon specific positive regulator with an in vivo role which has been genetically defined. Although our initial attempts to purify C protein have not been successful, it should soon be possible to complete the purification and to study the biochemical basis of C protein control.

We thank Professor Walter Gilbert for critical advice. We also thank him and Professor James D. Watson for helpful criticism of the manuscript. This work was supported by grants from the US National Institutes of Health (to J. D. Watson) and from the L. Rosenstiel Fund (to R. F. S.). J. G. thanks the Medical Research Council (Canada) for a studentship.

Received May 6, 1971.