Molec. gen. Genet. 122, 131—136 (1973)

Synthesis of the Arabinose Operon Regulator Protein in a Cell-Free System

Huey-Lang Yang and Geoffrey Zubay

Department of Biological Sciences, Columbia University, New York City, New York

Received January 24, 1973

Summary. A DNA-directed cell-free system which synthesizes L-ribulokinase coded by the L-arabinose operon, *ara* OIBAD, has been developed. L-arabinose is required for the expression of this operon and, in addition, cyclic AMP and guanosine-tetraphosphate are needed for optimal synthesis. The *ara* C gene product is also required and can be supplied either by *de novo* synthesis in the cell-free system or added back from extracts of whole cells. Evidence is presented that the C gene product is highly susceptible to protease attack.

Introduction

The L-arabinose operon, OIBAD, in *Escherichia coli* consists of a cluster of three structural genes, *araA*, *araB* and *araD*; two controlling sites *araI* (initiator) and *araO* (operator); and one regulatory gene *araC* (Fig. 1). The structural genes code for three enzymes that convert L-arabinose into D-xylulose. Genetical and biochemical studies have indicated that the L-arabinose operon, OIBAD, is both negatively and positively controlled by the specific regulatory gene *araC* (Englesberg, 1971). A model has been proposed by Englesberg to explain the regulation of the *ara* operon. The product of the regulatory gene *araC* is hypothesized to exist in two functional states: P1, a repressor, and P2, an activator, which are presumed to be in equilibrium with each other; P1 and P2 attach to their respective controlling sites, *araO* and *araI*. The complex of P1 at the *araO* site prevents operon expression whereas the complex P2 at the *araI* site is essential to expression. The substrate L-arabinose functions as the specific effector of the operon presumably by binding to the regulator protein; this binding favours the conversion of P1 to P2.

Like many other catabolite sensitive genes, activation of the arabinose operon requires the catabolite gene activator protein (CAP) and 3′5′ cyclic adenosine monophosphate (cAMP) (Zubay, Gielow, and Englesberg, 1971). The CAP protein probably interacts at or near the *araI* locus.

Direct testing of this model of gene regulation would be facilitated by measurements of gene activity as a function of the concentration of each of the alleged regulating components. A cell-free system in which it is possible to mix the required components in any amounts and observe the effects on gene activity was developed for this purpose (Zubay, Gielow, and Englesberg, 1971). A similar system with somewhat different properties has also been developed by Greenblatt and Schlief (1971). The system comprises a preincubated and dialyzed cell-free extract of *E. coli*, which shall be referred to as an S-30 extract, DNA

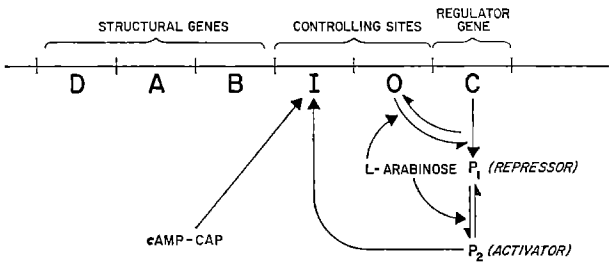

Fig. 1. Model for control of the arabinose operon (see Introduction for detailed explanation)

from the defective transducing virus *λdara* and the cofactors and substrates necessary for RNA and protein synthesis. The synthesis of L-ribulokinase (RK) coded by the *araB* gene is dependent upon the presence of *ara* operon-containing DNA and certain small molecule gene activators. The small molecule activators required include L-arabinose, the specific inducer of the *ara* operon, and cAMP, the small molecule inducer for many catabolite sensitive operons. These two requirements were predictable from whole cell studies and show that synthesis of RK in the cell-free system is subject to normal *in vivo* controls. In addition a 4 to 20 fold stimulation of RK synthesis is effected by guanosine 5′ pyrophosphate 3′ (or 2′) pyrophosphate (ppGpp). This could not have been predicted from whole cell studies. S-30 extract for these syntheses was prepared from strains which contain a deletion in the *B* and *A* genes so that any RK activity detected must result from *de novo* synthesis. Since according to the Englesberg model the C gene product is required for activating the operon, it was anticipated that effective S-30 extracts would have to be made from strains carrying the C gene. Therefore it came as a great surprise that equally effective S-30 extracts could be prepared from strains with or without the C gene. The main object of this paper is to explain this lack of requirement for preformed C gene product to activate the *ara* operon in the cell-free system.

Materials and Methods

For cell-free synthesis of the enzymes β-galactosidase (β-gal) and RK, the S-30 extracts used were prepared from strain SB 7219 containing C gene product or SB 7223 lacking C gene product (Englesberg, 1971). *λdlac* DNA was used for β-gal synthesis and *λdara* DNA was used for RK synthesis. Except for slight modifications, all procedures used for synthesis, enzyme assay, and preparation of bacterial S-30 extracts and DNA have been described in detail elsewhere (Zubay, Chambers, and Cheong, 1970). The incubation mixture contained per ml: 44 μmol Tris-acetate, pH 8.2; 1.37 μmol dithiothreitol; 55 μmol KAc; 27 μmol NH_4Ac; 14.7 μmol $Mg(Ac)_2$; 7.4 μmol $CaCl_2$; 0.22 μmole amino-acids; 2.2 μmol ATP; 0.55 μmol each GTP, CTP, UTP; 21 μmol phosphoenolpyruvic acid; 100 μg tRNA; 27 μg pyridoxine HCl; 27 μg triphosphopyridine nucleotide; 27 μg flavine adenine dinucleotide; 11 μg p-aminobenzoic acid; 0.5 μmol cyclic AMP. These were preincubated for 3 min at 37° C with 100 γ/ml of *λdlac* DNA with shaking before 6.5 mg S-30 protein extract was added. Incubation with shaking was continued for 60 min at 37° C. After synthesis, a 0.2 ml aliquot was removed and assayed for β-gal with O-nitrophenyl β-D-galactoside.

Conditions for the synthesis of RK were identical to those used for β-gal synthesis with the following exceptions: *λdara* DNA was used instead of *λdlac* DNA. 13 μmol of L-arabinose

and 0.05 μmol of ppGpp were added per ml of incubation mixture. The amount of *de novo* synthesized RK was quantified by a radiometric assay, in which the conversion of ^{14}C-L-ribulose to ^{14}C-L-ribulose-5-phosphate was measured. The specific activity of the ^{14}C-L-ribulose was 2.6×10^7 cpm/mol. 0.1 ml of reaction mix containing about 7×10^4 cpm of radioactive ribulose was incubated for 16 h at 28° C with 0.1 ml aliquot of the enzyme containing solution. The assay was linear for this period. After 16 h 1.6 ml of chilled ethanol and 0.2 ml 1 M barium acetate were added; the resulting precipitate was centrifuged and rewashed with 80% ethanol three times. The washed precipitate was dissolved in formic acid, plated, dried and counted in a gas-flow counter. Duplicate analyses usually agreed within 5%.

The partial purification of the *araC* protein was carried out as follows. Strain SB 7219 was grown as described previously, except for the addition of 2 g per liter of L-arabinose and 50 mg per liter of p-toluenesulfonylfluoride (pTSF). Cells were collected and frozen at −70° C in $^1/_4$″ thick sheets until ready for use. 50 g of thawed cells were allowed to soften at 5° C for 30 min. The cells were homogenized in 180 ml buffer I (0.01 M Tris-acetate, pH 8.2, 0.01 M $Mg(Ac)_2$, 0.06 M KCl, 0.01 M mercaptoethanol, 2 g/l L-arabinose, 50 mg/l pTSF and 5% glycerol). The resulting suspension was centrifuged at 10000 rpm for 30 min in a Serval GSA Rotor. The supernatant was discarded and the sediment was resuspended in buffer I and recentrifuged, again discarding the supernatant. The sediment was resuspended in 65 ml buffer I and lysed in a French pressure cell at pressures between 4000 and 8000 psi. The lysate was centrifuged at 15000 rpm for 30 min in a SS-34 Serval Rotor. The lysate was centrifuged for an additional 3.5 hours at 30000 rpm in a Spinco #30 Rotor. The resulting supernatant was dialyzed overnight against buffer II (0.01 M KH_2PO_4-KOH, pH 7.7, 2 g/l L-arabinose, 50 mg/l pTSF, 5% glycerol and 0.7 mM dithiothreitol). The solution was passed through a 140 ml DEAE-cellulose column (0.95 meq/g) previously equilibrated with buffer II. After the solution had passed into the column the column was rinsed with another 0.5 liters of buffer II. The cellulose-bound protein was eluted with buffer II + 0.25 M NaCl. Protein containing eluate was pooled and 2 volumes of saturated $(NH_4)_2SO_4$ were added to precipitate the bulk of the protein which was collected by a 5 min centrifugation at 10000 rpm in a Serval GSA Rotor. The sediment was resuspended in 25 ml of buffer III (0.01 M Tris-acetate, pH 8.2, 0.014 M $Mg(Ac)_2$, 0.06 M KAC, 50 mg/l pTSF, 2 g/l L-arabinose, 0.7 mM dithiothreitol and 5% glycerol) and dialyzed overnight against the same buffer. The resulting solution containing about 40 mg/ml of protein was used directly or stored indefinitely in nitrocellulose tubes in liquid N_2.

Results and Discussion

1. When S-30 Extract is Used as the Sole Source of Bacterial Protein the C Gene Must be Present on the DNA for RK Synthesis

Two *λdara* DNAs (designated 8830 and 8803 in Table 1) have been successfully employed for RK synthesis with S-30 preparations. Each of these contains the entire *ara* operon and the regulator gene *C*, COIBAD. These DNAs (see Table 1) are about equally effective in RK synthesis and require the small molecule factors cAMP, L-arabinose and ppGpp for optimal synthesis. These results are essentially the same as those reported previously (Zubay, Gielow, and Englesberg, 1971) and serve as a control on the further experiments reported here. When another *λdara* DNA containing the *ara* operon but not the C gene (designated C1020 in Table 1) is used, no RK synthesis is obtained. The requirement for C gene on the DNA is a strong indication that synthesis of the *ara* C protein is essential for RK synthesis. It seems likely that the sequence of events in the cell-free system involves synthesis of C gene product followed by activation of the operon and synthesis of RK. If this is the case there should be an unusual delay in the appearance of the first RK. It was found that more than 30 minutes is required before there is detectable RK (see Fig. 2). Synthesis of

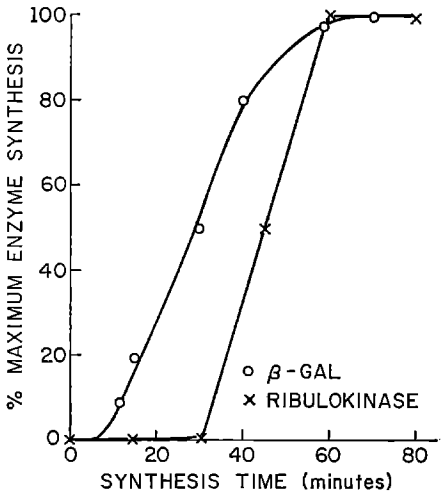

Fig. 2. Enzyme synthesized as a function of time. Conditions of synthesis described in Materials and Methods

Table 1. Synthesis of ribulokinase with different DNAs and different factor missing

DNA Source	L-Ribulokinase (CPM[a])
8830	
complete	690
—DNA	0[b]
— ppGpp	170
— ara	20
— cAMP, —ara, — ppGpp	20
8803	
complete	650
—cAMP, —ara, —ppGpp	60
C1020	
complete	0
— DNA	0

[a] 1 cpm equals 3.75×10^{-8} units of enzyme per ml of incubation mix.
[b] Background counts obtained without DNA added of about 15 cpm have been substrated from all data.

most proteins in the cell-free system is detectable after 8 to 10 minutes as indicated for the DNA-directed synthesis of β-galactosidase (β-gal). Thus the kinetics of appearance of RK is consistent with the notion that this synthesis is delayed awaiting production of C protein to activate the operon.

2. *Preparation of Active C Gene Product from Whole Cells Requires Special Precautions*

The S-30 for RK synthesis is equally effective whether it is prepared from C^+ or C^- strains. Since the C gene product is essential for RK synthesis this

indicates that something has happened so that the regulator protein is not active in C^+ derived extracts. Following this line of thought a variety of methods was used to make active C gene product that could be added back to an S-30, the ultimate criterion for active C protein being stimulation of RK synthesis using the *λdara* DNA lacking the C gene. A number of factors have been found to be important in obtaining optimum yields of active C gene product. Currently active C protein containing extracts are prepared from crude lysates by centrifugal removal of membranes and ribosomes. Other nucleic acid is removed from the extract by passage over a DEAE cellulose column. Active extracts require the addition during growth and isolation of two factors which are believed to protect the C protein from enzymatic degradation. These factors are L-arabinose and pTSF. When arabinose alone is present the extracts are about 20 per cent as active in stimulating RK synthesis; when pTSF alone is present the extracts are about 40 per cent as active. It seems likely that L-arabinose stabilizes the C protein by binding to it and changing its conformation so that it is less susceptible to enzyme attack. pTSF is a well-known serine protease inhibitor and probably helps by inhibiting the serine protease known to be present in *E. coli* (Prouty and Goldberg, 1972). In Table 2 results are presented in which parallel

Table 2. Effect of added ara C protein on the synthesis of ribulokinase

DNA Source	L-Ribulokinase (CPM)	
	without C protein added	with C protein added[a]
8830		
complete	650	1780
8803		
complete	650	2040
−ppGpp	20	1230
− cAMP	40	780
C1020		
complete	24	900

Conditions for cell-free synthesis are described in Materials and Methods.
[a] 2.5 mg C protein containing extract added per ml of incubation mixture.

syntheses were carried out with and without added C protein. The addition of C protein approximately triples the extent of RK synthesis when the DNAs containing the C gene are used (8830 or 8803). The kinetics of RK synthesis with added C protein is similar to that for β-gal synthesis (Fig. 2). This seems reasonable since the operon should be activated at once with preformed C protein present. The specificity of the C protein action was demonstrated in two ways: 1) Parallel extracts were made from C^+ and C^- strains; only the C^+ extract was active in stimulating RK synthesis. 2) The C^+ extracts had no effect on stimulating β-gal in a comparable DNA-directed system using *λdlac* DNA.

Synthesis of RK in the presence of added C protein is still completely dependent upon arabinose but not upon cAMP. Thus about one third of the maxi-

mum synthesis is obtained without added cAMP. The dependence of ppGpp also is less when C protein is added. When C protein is added ppGpp stimulates RK synthesis only 1.5 to 3 fold whereas when C protein is left out ppGpp stimulates RK synthesis 4 to 20 fold. The explanation for these differences is not clear at present. Our present efforts are directed towards using the cell-free system as an assay tool to obtain purified C protein which functions normally in the cell-free system.

Acknowledgements. We are grateful to J. Beckwith and E. Englesberg for strains, to A. L. Goldberg for suggesting the use of p-Toluenesulfonylfluoride and to Michael Cashel for his generous donation of ppGpp. This work was supported by a grant from the National Institutes of Health (2-R01-GM-16648-04).

References

Englesberg, E.: Regulation in the L-arabinose system, p. 257–296, in: Metabolic pathways (edit. by D. Greenberg and H. Vogel), 3rd ed., vol. V. New York: Academic Press 1971.

Greenblatt, J., Schlief, R.: Arabinose C protein: Regulation of the arabinose operon *in vitro*. Nature (Lond.) New Biol. **233**, 166–170 (1971).

Prouty, W. F., Goldberg, A. L.: Fate of abnormal proteins in *E. coli* accumulation in intracellular granules before catabolism. Nature (Lond.) New Biol. **240**, 147–150 (1972).

Zubay, G., Chambers, D. A., Cheong, L.: Cell-free studies on the regulation of the *lac* operon, p. 375–391, in: The *lac* operon (edit. by D. Zipser and J. Beckwith). Cold Spring Harbor Laboratory of Quantitative Biology, New York, 1970.

Zubay, G., Gielow, Englesberg, E.: Cell-free studies on the arabinose operon. Nature (Lond.) New Biol. **233**, 164–165 (1971).

Communicated by W. Maas

Dr. Huey-Lang Yang
Dr. Geoffrey Zubay
Department of Biological Sciences
Columbia University
New York, New York 10027
U.S.A.