

A EUKARYOTIC GENE IS FUNCTIONALLY EXPRESSED IN E. COLI

Kevin Struhl and Ronald W. Davis

Department of Biochemistry
Stanford University School of Medicine
Stanford, California 94305

ABSTRACT: The bacterium E. coli can transcribe and translate a cloned segment of eukaryotic DNA with sufficient fidelity to result in the synthesis of a protein which is enzymologically and physiologically similar to the predicted eukaryotic protein. The amount of enzyme activity produced as a result of functional genetic expression of this eukaryotic DNA in E. coli approaches the amount resulting from the analogous E. coli gene. This cloning of such a eukaryotic gene in E. coli represents a good model system for synthesis of polypeptides, such as insulin, which have clinical applications.

INTRODUCTION

Many medically important polypeptides such as insulin are currently isolated directly from animal sources. Some, such as somatostatin, are made in small amounts and consequently such agents are difficult to obtain. Others, such as insulin, though obtainable with reasonable ease at present, will become increasingly more difficult to obtain in the next ten years. Obviously, it would be of great practical benefit to produce large quantities of these rare clinically useful agents. For treatment of human conditions such as diabetes, it would clearly be optimal to obtain human insulin instead of related mammalian substitutes.

A promising approach for both these problems is to isolate the gene coding for a given polypeptide and to introduce it stably into E. coli, and to have the bacterium synthesize the desired polypeptide in large (essentially limitless) amounts. Such an approach would involve the recently developed recombinant DNA technology.

Recombinant DNA is the name given to molecules which result from the physical joining of two or more previously separate DNA molecules. If one of these molecules, when joined, is a bacterial replicon, be it plasmid or bacteriophage, such recombinant DNA can be propagated in E. coli; the replicon (or vector) and the foreign DNA. Therefore, this technology enables one to clone almost any segment of DNA from any organism and to preserve it, by replication in E. coli as recombinant DNA. Simply by virtue of the fact

that large quantities of a pure DNA segment are obtainable, structural studies of genes, chromosomes, and genomes are now under way.

The question to be addressed in this paper is whether cloned eukaryotic DNA can be functionally expressed in E. coli. In other words, can E. coli transcribe and translate a segment of foreign DNA with sufficient fidelity such that the protein coded for by this DNA function in an E. coli environment.

The basis of these experiments is that if DNA can be functionally expressed, then a genetic selection should allow for isolation of the hybrid of interest. Such a selection could be the complementation of a specific bacterial auxotroph. For example, a histidine auxotroph can not grow unless the media is supplemented with histidine, because an enzyme necessary for histidine biosynthesis is non-functional. If a cloned eukaryotic segment is expressed so that the analogous eukaryotic protein is synthesized in functional form, auxotrophic cells containing such a eukaryotic segment should be able to grow in the absence of added histidine. The approach therefore has been to make a collection of hybrids representing most of the genome of a particular species (in this case, yeast) with bacteriophage λ , and to genetically select for a specific hybrid from such a pool of hybrids. There are three basic reasons for these experiments. Firstly, they provide model systems for designing strains of E. coli which can overproduce proteins of practical benefit. Secondly, they provide for methods for selecting and cloning a wide variety of defined genes for which no other selection system is applicable. And thirdly, they allow for the synthesis of probes for studying specific and coordinate gene regulation in eukaryotes of typical structural genes.

In this paper, we address ourselves to three issues concerning the complementation of a histidine auxotroph by a segment of yeast DNA.

1. The question of functional genetic expression of eukaryotic DNA in E. coli.
2. Fidelity of expression as determined by the enzymological characterization of the activity in cells containing the expressed eukaryotic DNA segment.
3. Level of expression as determined by a physiological characterization of strains containing the expressed eukaryotic DNA fragment.

RESULTS

Cloning yeast DNA in bacteriophage λ ; construction of hybrid pools

The cloning vector for these experiments was developed by Thomas, Cameron, and Davis (1) and Murray and Murray (2). There are two sites in the vector (λ gt- λ B) which are recognized by EcoRI endonuclease. No genes essential for λ growth are located in the middle EcoRI fragment. To form hybrid molecules, one separately cleaves the vector and the yeast DNA with

EcoRI, mixes, and s
EcoRI endonuclease
E. coli strain which
Mandel and Higa (4).
cedure. A molecule
containing all the ess
and is not viable. A
an EcoRI fragment is
fore either hybrids c
using appropriate co
Following transfectio
then passaged once c
modified. Phage in
sensitive to λ infecti

This system f
between the size of l
large fragments will
be inserted at a low

GENETIC SELECTION

λ , being a ten
passively when inter
Selections can be p
both cases, a hybr
to isolate a speci
selections is that
We have developed
which when integra
cell to grow in the
are deleted for the
However, the hybr
infection with an i
pre-existing lysog
general recombina
technique requires
initiation site sinc
EXPRESSION OF

We employed
yeast DNA. The

EcoRI, mixes, and seals with DNA ligase via the cohesive ends generated by EcoRI endonuclease (3). These molecules are transfected for plaques into an E. coli strain which does not restrict foreign DNA (rK^-) by the method of Mandel and Higa (4). Various kinds of molecules are generated by this procedure. A molecule consisting solely of the left and right ends of λ gt, while containing all the essential genes for λ growth, is too short to be packaged and is not viable. A viable molecule (one that results in a plaque) must contain an EcoRI fragment inserted between the λ gt ends. The viable phage are therefore either hybrids containing yeast DNA or the original vector λ gt- λ B. By using appropriate conditions (1), greater than 90% of the plaques are hybrids. Following transfection, the plaques are scraped to form a hybrid pool which is then passaged once on a rK^-mK^+ strain to form a high titer stock which is K modified. Phage in such a pool can infect any strain of E. coli which is sensitive to λ infection.

This system for cloning DNA is not general since only EcoRI fragments between the size of 1 and 15 kilobase (kb) pairs can be cloned. In addition, large fragments will render the viable phage unstable, and small fragments will be inserted at a lower frequency.

GENETIC SELECTION TECHNIQUES

λ , being a temperate phage can replicate actively as a lytic virus or passively when integrated into the E. coli chromosome as a stable lysogen. Selections can be performed when λ is in either the lytic or lysogenic state. In both cases, a hybrid pool from yeast is put through a selective process in order to isolate a specific hybrid of interest. The essential feature of such genetic selections is that the inserted DNA be functionally expressed in the E. coli cell. We have developed a lysogenic selection technique to select for a hybrid phage which when integrated into the chromosome of an E. coli auxotroph allows the cell to grow in the absence of the auxotrophic requirement. All hybrid phage are deleted for the integration gene (int) and the phage attachment site (att). However, the hybrids can be integrated into the E. coli chromosome by co-infection with an integration helper phage which is int⁺att⁺ or by infection of a pre-existing lysogen. This double lysogen formation is mediated by bacterial general recombination and occurs at a frequency of 1% (5). This selection technique requires that the inserted piece of DNA contain a transcription initiation site since the λ promoters in λ gt are deleted or strongly repressed.

EXPRESSION OF YEAST DNA IN E COLI

We employed the lysogenic selection procedure on hybrid pools made with yeast DNA. The mutant used was a histidine auxotroph of E. coli (his B463).

It was chosen primarily because it reverts with a frequency of less than 10^{-11} . His B codes for a product which is a bifunctional enzyme. It catalyzes both imidazole glycerol phosphate (IGP) dehydratase and histidinol phosphate phosphatase. His B463 by enzymological and complementation criteria lacks only the dehydratase activity.

When either of two yeast hybrid pools were co-infected with an integration helper into his B463, his⁺ colonies were found at a frequency of 10^{-8} after 3 days. These are not revertants since no revertants of his B463 have ever been found even after treatment with strong mutagens.

One can imagine many possible artifacts besides the trivial case of bacterial revertants in a preliminary complementation result of this kind. The following experiments were done to rule out as many of these as possible (6). The his⁺ colonies contain λ prophages. If prophages are cured from the his⁺ colonies, the resulting strains are unable to grow without histidine and are indistinguishable from the original parent his B463. The double lysogens formed were normal by two criteria. First, induction of them resulted in equal quantities of two phage. Second, after passaging the his⁺ colonies once in non-selective medium, 5% of the colonies become his⁻. This segregation is expected since it occurs by the reverse of the rec mediated formation of double lysogens.

Following induction of a culture of a his⁺ colony, the presumptive complementing phage was purified away from the helper. This phage is called λ gt-Sc his. When λ gt-Sc his was integrated into the chromosome of his B463 with the integration helper, his⁺ colonies were obtained at a frequency of 10^{-4} . This represents an enrichment of 10,000 fold from the yeast hybrid pool. This is the approximate frequency one would expect to find λ gt-Sc his in the pool. Therefore, we have isolated a phage capable of complementing his B463. The complementation is not dependent on the double lysogen technique. It is possible to integrate λ gt-Sc his into his B463 without the integration helper and obtain his⁺ colonies. This occurs by abnormal integration (that is, not at the bacterial attachment site) and at very low frequency. Phage can be induced from these single lysogens. The excision is inefficient as expected but the viable phage induced are indistinguishable from λ gt-Sc his by functional and heteroduplex criteria.

Complementation by λ gt-Sc his is extremely specific. It complements the non-reverting his B463 and another his B auxotroph lacking only IGP dehydratase activity (his B2404) which is revertible. It does not complement any other histidine auxotroph including his B mutants which lack both the dehydratase and phosphatase activities. This is consistent with the fact that these two activities are separate enzymes in yeast, but are part of the same polypeptide in E. coli.

The specific complementation was made from the phage DNA fragment of 10.3 kb in size.

One important control experiment insert comes from its complementations with bacterial DNA. The phage pool has been purified and indeed contains yeast DNA as a contaminant. λ gt-Sc his DNA as a probe hybridized across a range of DNA in a technique developed in a different manner. This was used to minimize background in a radiograph from samples corresponding to a range of DNA, no hybridization was observed. This contains a fragment of DNA which is 10.3 kb in size. This fragment in total yeast DNA is rearranged with close proximity to the his gene.

Two conclusions can be drawn. First, the yeast DNA specific site effects on a class of histidine auxotrophs for λ prophages does not seem necessary for the complementation of DNA. Since the complementation is repressed. Strong complementation is observed with respect to λ gt-Sc his with DNA ligase. A phage with inverted orientation of λ gt-Sc his complementation of his B463. Therefore, transcription must be in the same direction as the his gene.

Expression of the complementation. The fragment has been transformed into his⁺ colonies. This complex

The specific complementing phage λ gt-Sc his is indeed a λ gt hybrid. DNA made from the phage when cleaved with EcoRI endonuclease reveals an EcoRI fragment of 10.3 kb inserted between the λ gt ends.

One important consideration in work of this kind is whether the DNA in the insert comes from its stated source. Contamination of eukaryotic DNA preparations with bacterial DNA is a particular worry. In addition, the hybrid phage pool has been passaged through essentially wild type E. coli. If the hybrid indeed contains yeast DNA, 32 P labelled complementary RNA made using λ gt-Sc his DNA as a template should hybridize to an EcoRI fragment of yeast DNA which is 10.3 kb in length. Accordingly, cRNA made to λ gt-Sc his was hybridized across a gel of total yeast DNA cut with EcoRI endonuclease by a technique developed by Southern (7). A new preparation of total yeast DNA made in a different manner, from a different strain, and in a different laboratory was used to minimize contamination. As published in reference 6, an autoradiograph from such an experiment reveals a single band of grains at a position corresponding to a length of 10-11 kb. When cRNA was made to the λ gt- λ B vector DNA, no hybridization was detected. Therefore, not only does λ gt-Sc his DNA contain a fragment of yeast DNA corresponding to a similar sized EcoRI fragment in total yeast DNA, but in addition, the fragment has probably not been rearranged with cloning and propagation in E. coli.

Two conclusions can be drawn from the nature of the selection system. First, the yeast DNA must code for a diffusible product. There can be no specific site effects on the E. coli histidine operon because only a very specific class of histidine auxotrophs is complemented and because the attachment site for λ prophages does not map near the histidine operon. Secondly, transcription necessary for the complementation is probably initiated in the segment of yeast DNA. Since the complementation occurs in a λ lysogen, λ promoters are strongly repressed. Strong support for these conclusions comes from the fact that the complementation is independent of the orientation of the fragment of yeast DNA with respect to λ . λ gt-Sc his was cleaved with EcoRI endonuclease and resealed with DNA ligase. A phage was isolated which contained the yeast fragment in an inverted orientation as determined by heteroduplex analysis. The inversion, λ gt-Sc his' complements his B463 with an equal efficiency as the original isolated λ gt-Sc his. Therefore, unless there are two as of yet unknown λ promoters, transcription must be initiated in the yeast fragment.

Expression of this fragment of yeast DNA is not dependent on a λ vector. The fragment has been cloned in pMB9 and pSC101. When these hybrid plasmids are transformed into his B463, all tetracycline resistant colonies tested were his⁺. This complementation was confirmed by Barry Ratxkin and John Carbon

at Santa Barbara. Using their pool of col-E1 yeast hybrids which were generated in a different manner, and a selection system depending on F' mediated transfer of the col E1 factor, they have isolated a hybrid which complements his B463.

/ Strains of E. coli able to grow in M9 minimal medium must be able to synthesize functional IGP dehydratase. Therefore, merely assaying derivatives of his B463 which contain the complementing yeast DNA in any particular cloning vehicle for IGP dehydratase is not relevant to the question of the nature of the complementation. Firstly, a level of expression sufficient for growth may not be detectable by a particular assay. And secondly, IGP activity in such cells does not distinguish between functional expression of a yeast structural gene and suppression of the original mutation of his B463.

To conclusively rule out suppression of the his B locus, the yeast DNA cloned in pMB9 was introduced into a strain of E. coli deleted for the entire histidine operon (his 461). Selection and maintenance of this event was via the plasmid coded trait for resistance to tetracycline since the resulting strain would not grow without added histidine. The strain, Δ his (pMB9- Sc his), had significant IGP activity in contrast to the parent his 461 (Δ his). Both his 461 and the derivative containing pMB9-Sc his had no histidinol phosphate phosphatase activity as expected from the genotype. The results with his B463 and the plasmid derivative were similar with the expected difference that both strains had wild type levels of histidinol phosphate phosphatase. This experiment provides direct evidence that the IGP activity dependent upon cloned yeast DNA is measurable by an enzyme assay of a crude extract and that the activity and the complementation is not the result of suppression of the E. coli his B protein.

LEVEL OF EXPRESSION

The level of functional genetic expression of eukaryotic DNA in E. coli was assayed in three ways. The first was by the growth rate in the absence of histidine (growth dependent upon the cloned yeast segment). The second was by the growth rate in the presence of a specific inhibitor of IGP dehydratase (aminotriazole). The third method was by measurement of enzyme activity levels in crude extracts. The basic result is that more enzyme activity is produced when the yeast segment is cloned in a plasmid vector than when it is cloned in a λ vector. This is not surprising since plasmids are present in many copies per cell, while λ in the lysogenic state is present in only one copy per cell. The second result is that the expression of the yeast segment when cloned in a plasmid vector approaches the level of the analogous E. coli gene. The growth rate of the plasmid derivatives in the absence of histidine is approximately equal to the wild type E. coli rate, while in the λ derivatives it is more than 50% slower. In the

presence of aminotriazole, the λ derivatives at concentrations of 10mM) of the wild type level. It should be pointed out that aminotriazole shows no inhibition of IGP dehydratase. When the strains were assayed for IGP activity they have 50-80% of the wild type level of IGP activity.

FIDELITY OF EX
The fidelity of E. coli was assessed by measuring histidine dehydratase activity found in E. coli histidine dehydratase. The "yeast in E. coli strains of yeast were very inactive and had a broader range of affinity for IGP. The concentration was about 20 times the activities were. The inhibition of the that for plasmid these determined is clear that the type yeast activity requirement of magnesium, calcium activity. From is quite good. H It is quite possible certain amino ac

SUMMARY

We have

presence of aminotriazole (a competitive and specific inhibitor of IGP dehydratase, the λ derivatives are strongly inhibited for growth even at relatively low concentrations of the inhibitor (1mM). The plasmid derivatives require higher dosages (10mM) of aminotriazole for inhibition to the same degree. It should be pointed out that selection of colonies able to grow in high concentrations of aminotriazole should result in strains which overproduce the eukaryotic IGP dehydratase. When extracts of various derivatives containing the yeast DNA are assayed for IGP dehydratase activity, it is found that the plasmid derivatives have 50-80% of the wild type E. coli levels. The λ derivatives have about 10% of the wild type level.

FIDELITY OF EXPRESSION

The fidelity of this functional genetic expression of eukaryotic DNA in E. coli was assessed by an enzymological characterization of the IGP dehydratase activity found in cells containing the Sc his fragment and lacking any of the E. coli histidine enzymes. The IGP dehydratase reaction is the Mn^{++} catalyzed dehydration of imidazole glycerol phosphate to imidazole acetol phosphate. The "yeast in coli" activity was compared to the analogous activity in wild type strains of yeast. Both activities have a pH optimum of around 7.5 and become very inactive as the pH becomes higher than 8. The E. coli activity has a broader range with an optimum of 8.1 and significant activity at pH 9. The affinity for IGP (as measured by the Km) for both activities was about 300 μ M. The concentration of Mn^{++} necessary for half maximal activity in both cases was about 20 micromolar. Both the "yeast in coli" and the wild type yeast activities were inhibited competitively by aminotriazole and phosphate ion. The inhibition constant (Ki) for aminotriazole was about 25 micromolar and the that for phosphate ion was about 2mM. Since the assay is not very accurate, these determinations could be as imprecise as factors of 3 to 5. However, it is clear that the similarities between the "yeast in coli" activity and the wild type yeast activity are quite striking. One final similarity is the specific requirement of Mn^{++} . Substitution of this requirement by the divalent cations of magnesium, calcium, or zinc are not effective in promoting IGP dehydratase activity. From these experiments, we conclude that the fidelity of expression is quite good. However, the exact nature of the product is still open to question. It is quite possible that the product may have additional amino acids or lack certain amino acids when compared to the bonafide yeast protein.

SUMMARY

We have isolated a segment of yeast DNA by molecular cloning which

complements a histidine auxotroph of E. coli. We conclude the following:

- 1) There is functional genetic expression of eukaryotic DNA in E. coli.
- 2) Transcription necessary for this expression is initiated in the segment of yeast DNA.
- 3) By virtue of this expression, a yeast coded IGP dehydratase is synthesized in E. coli.
- 4) The level of expression is quite respectable. The level of yeast enzyme activity approaches the level found in wild type strains of E. coli.
- 5) The fidelity of expression as determined by an enzymological characterization of the yeast coded activity is quite high.

REFERENCES

1. Thomas, M., Cameron, J. R., and Davis, R. W. (1974)
Proc. Nat. Acad. Sci. USA 71 4579-4583
2. Murray, N. E., and Murray, K. (1974)
Nature (London) 251 476-481
3. Mertz, J. E., and Davis, R. W. (1972)
Proc. Nat. Acad. Sci. USA 57 1514-1519
4. Mandel, M., and Higa, A. (1970)
J. Mol. Biol. 53 159-162
5. Gottesman, M. E., and Yarmolinsky, M. B. (1968)
J. Mol. Biol. 31 487-505
6. Struhl, K., Cameron, J. R., and Davis, R. W. (1976)
Proc. Nat. Acad. Sci. USA 73 1471-1475
7. Southern, E. (1975) J. Mol. Biol. 98 503-517

IN VITRO RECOMBINATION
AND SPECULATION

Lennart Philipson

Department of Microbiology
Uppsala, Sweden

Techniques have been developed for the synthesis of large molecules. The use of DNA ligase after gel electrophoresis and the addition of complementary DNA to almost entirely intact E. coli. The use of lambda (4). Evidence for the transfer of DNA to mammalian cells. Evidence for the development of the system. Cells as hosts for the development of systems containing the DNA of cells using S₁ nuclease.

The present system is useful for in vivo replication of vector-host systems. A virus-like system is useful for an efficient method of cloning.

Plasmid-like systems

The search for a host genome which can be covalently closed without affecting the structures in the machinery for replication. The system can be suitable as a virus genome and cccDNA in both instances. In the case of poliovirus under separate conditions formed by an intact host cell genome and cell DNA. In principle the course independent of the viral cccDNA ultimately determined for amplification achieved in the