Immunology, Epidemiology and Social Aspects of Leprosy
Pontifical Academy of Sciences, *Scripta Varia* 72, Vatican City 1988
www.pas.va/content/dam/accademia/pdf/sv72/sv72-young.pdf

RECOMBINANT DNA AND FUTURE VACCINES

RICHARD A. YOUNG
Department of Biochemistry
Stanford University School of Medicine
Stanford, California 94305, U.S.A.

INTRODUCTION

New advances in recombinant DNA technology have provided powerful tools for basic and applied research in human health problems. Two components of vaccine research can benefit significantly from current recombinant DNA technology: the identification and isolation of genes which specify antigens potentially relevant to immunity, and the large-scale production of the antigen in relatively pure form. Genes that encode antigens of interest can be isolated by using antibodies to identify the products of individual foreign genes in *E. coli* host cells. Thus, using this technology, genes which specify parasite antigens can be simultaneously identified and isolated by using antibodies from patients afflicted with the parasite. The isolated genes of interest can be expressed in procaryotic or eucaryotic hosts to produce (hopefully) large amounts of antigen, which can be used for research, diagnostic or vaccine purposes. This manuscript briefly reviews recent advances and current limitations in the use of recombinant DNA in vaccine research.

RECOMBINANT DNA STRATEGY FOR SURVEYING PARASITE ANTIGENS

A means to survey the protein components of a pathogen for potential immunogens would be a useful prerequisite to selecting a polypeptide vaccine candidate for further study. Recombinant DNA technology offers an effective strategy to thoroughly examine the

antigens encoded in a pathogen genome when antibodies are available for use as probes [1, 2]. Part of the power of this approach lies in the potential to express all possible coding sequences in the genome of interest, even those which may not always be expressed *in vivo*. Moreover, inherent in this recombinant DNA methodology is the ability to simultaneously identify immunogens — in this case, antigens against which humans have mounted a humoral antibody response — and clonally isolate the DNA which specifies the polypeptide.

The antigen coding capacity of a pathogen genome can be systematically examined by using a recombinant DNA expression library. In principle, fragments of the pathogen genome are expressed under the control of procaryotic gene signals located in the vector; sufficient numbers of these recombinant molecules should exist in a library to represent redundantly the pathogen genome. To construct such a library, DNA is mechanically sheared to produce random endpoints, and these DNA fragments are inserted into the expression vector λgtll. Sufficiently large numbers of recombinants are produced to obtain inserted DNA endpoints at each base pair within each gene in the pathogen genome. Antigens produced by the recombinant λ phage plaques are screened, using serum or monoclonal antibody probes, as shown in Figure 1. The complete library, and thus all antigen coding sequences, can be examined in a single experiment (10^6 recombinants can be probed in a single experiment). Recombinants identified by the antibody of interest can be clonally isolated.

Requirements for a recombinant DNA expression system

A systematic examination of foreign polypeptide antigens in *E. coli* is possible if the factors which influence production of detectable levels of each of many different kinds of proteins are adequately considered. There are three major problems associated with obtaining expression of foreign DNA as a stable antigen. The first problem is that most foreign DNA does not contain the transcription control signals required for expression in *E. coli*. Thus, the foreign gene must be placed under the control of an *E. coli* promoter that is efficiently recognized by *E. coli* RNA polymerase.

The second and probably most serious problem is that unusual polypeptides are efficiently recognized and degraded in *E. coli* [3-6]. The severity of the problem differs with each antigen; some foreign proteins are quite stable, some appear highly unstable. While the

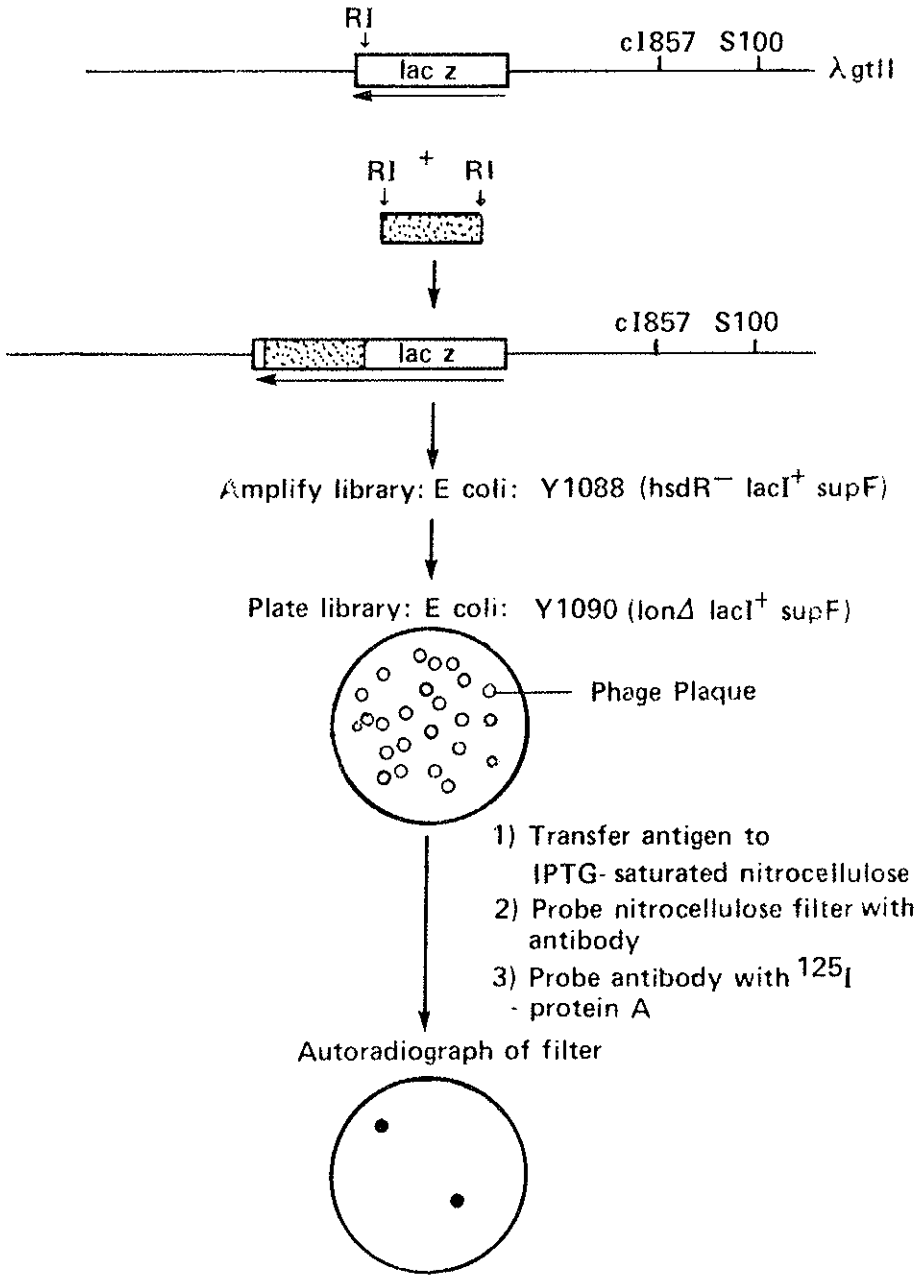
RI
lac z
cI857 S100
λgtII
RI + RI
cI857 S100
lac z
Amplify library: E coli: Y1088 (hsdR⁻ lacI⁺ supF)
Plate library: E coli: Y1090 (lonΔ lacI⁺ supF)
Phage Plaque
1) Transfer antigen to IPTG- saturated nitrocellulose
2) Probe nitrocellulose filter with antibody
3) Probe antibody with 125I - protein A
Autoradiograph of filter

FIG. 1

problem of antigen instability is rarely addressed directly in the literature, the difficulty in accumulating foreign proteins is revealed by the techniques used to detect the polypeptide. Thus, enzymatic activity or immunoprecipitation of pulse-labeled proteins are commonly used approaches [4-6].

The instability of foreign antigens can be reduced in most cases by fusing the antigen to a stable host protein and by using host mutants deficient in proteolysis [1]. Fusion of unstable foreign antigens to the carboxy-terminus of the stable *E. coli* protein β-galactosidase has been shown to enhance the stability of some foreign proteins [1, 6, 7]. More importantly, the stability of the fusion product of β-galactosidase and eucaryotic antigen can be markedly increased (>100-fold in some cases) in *lon* mutants of *E. coli*. *lon* mutant strains are deficient in one of the ATP-dependent proteases which are responsible for the destruction of abnormal proteins [3]. This particular protease deficiency is especially useful since the presence of the mutation does not appear to alter the normal growth properties of the cell and because the *lon* protease appears to have some specificity for the class of abnormal polypeptides of which β-galactosidase fusions are a member [1].

The third major problem with foreign antigen synthesis in *E. coli* is that the presence of these unusual proteins is often harmful or even lethal to the cell. Demanding high levels of gene expression can compound this problem, since constitutive high level expression of even normal components of the cell can often be lethal [8]. A suitable solution to this problem has been to ensure that expression of the foreign antigen is transient. Thus, the expression of the DNA encoding the foreign protein is repressed during early log-phase growth of the host cell culture. Near the end of this period, when the transcriptional and translational apparatus are still fully active, the expression of the foreign protein is induced, and satisfactory levels of the antigen are produced before cells become unviable.

These concepts, designed to improve the levels to which foreign antigens can accumulate in *E. coli*, have been incorporated into the λgtll expression vector-host system [1]. The λ phage expression vector was constructed to permit insertion of foreign DNA into the β-galactosidase structural gene *lac* Z under the control of the *lac* operator. The recombinant DNA can be propagated lytically or lysogenically (efficient lysogeny is obtained with *hfl*A [high frequency lysogeny] mutant hosts). In either case, expression of the foreign DNA is

repressed by the presence of the *lac*I gene product. Production of the foreign antigen fused to β-galactosidase can be rapidly induced by the addition of isopropyl-β-D-thiogalactoside (IPTG) to the culture medium. The presence of the *lon* mutation permits accumulation of otherwise unstable novel proteins to levels which facilitate detection by immunological or physical (i.e., polyacrylamide gel) analysis. Figure 1 outlines the experimental scheme which we have used to isolate eucaryotic DNA sequences which specify antigens of interest [1, 2]. Details of the procedure can be found in reference 2.

Surveying Mycobacterium tuberculosis *Antigens*

The λgtll expression system is being used to investigate specific antigens encoded by *M. tuberculosis*. *M. tuberculosis* DNA (isolated by Grosskinsky and Bloom from Erdman strain, TMC § 107) was sheared to generate random endpoints and was inserted into λgtll to produce an expression library of 10^7 unique recombinant phage. *M. tuberculosis* DNA insert lengths ranged from 2.5 to 8.5 kb with an average length of approximately 4 kb. The number of recombinant phage in this library exceeds the number of base pairs in the pathogen's genome by approximately 3-fold; it is, therefore, highly likely that the library has the capacity to express all *M. tuberculosis* DNA in both orientations and in all three translation frames.

Attention is now being directed to two questions: Can clones be isolated which express antigens that are bound by monoclonal antibodies with unique specificity to *M. tuberculosis* proteins? Against what protein antigens do humans produce serum antibodies when they successfully mount a protective immune response? Experiments which address the latter question may provide antigens which are useful vaccine candidates not only because they are immunogens, but also because the DNA which encodes them is identified and can be isolated directly.

Limitations to a Recombinant DNA Survey of Subunit Vaccine Candidates

The λgtll expression system provides a means to thoroughly and efficiently survey the potential immunogens of particular pathogens. The method, however, is designed to detect protein antigens which lie within a contiguous polypeptide chain. It obviously cannot be used to identify nonprotein components which might contribute to a

protective vaccine. Moreover, antigenic sites which are created through intermolecular interactions are excluded from such a survey using recombinant DNA technology. The importance of this latter consideration is unclear since the contribution of this type of antigen to an organism's immunogenicity is as yet unknown.

Large Scale Production of Foreign (Parasite) Gene Products

Many different gene products have been produced in foreign hosts through the use of recombinant DNA expression vectors. The successful production of foreign proteins is dependent upon overcoming the same problems encountered in the system discussed above for the expression and isolation of recombinant DNA clones. However, while these are general problems, their solutions are best recognized on a case by case basis; what works for the production of good yields of one foreign product may not work well for another.

Variations in the level of accumulated foreign protein can occur if (1) foreign DNA sequences interfere with efficient transcription or translation, (2) the mRNA is unstable (as is often the case if it is poorly translated), (3) the polypeptide product is rapidly proteolyzed and/or (4) the product is toxic to host cells. Some of these problems have been understood for some time, others have been recognized only recently. Solutions to all of them exist, even if they are primitive or are in the early stages of development. For example, transcriptive interference by foreign DNA may be minimized by changing vectors. If mRNA instability is caused by poor translatability, efficient translation initiation signals can be provided. If the protein product is toxic to the host cell, its production can be suppressed until a time propitious for obtaining satisfactory yields. Most importantly, it is now recognized that the major factor in obtaining poor yields from recombinant DNA in foreign hosts is the ability of most cells to efficiently recognize and degrade protein that is "unusual". Thus, an exciting new development is the isolation of *E. coli* strains which are defective in many different proteolytic functions (9). Preliminary evidence indicates that these strains are precisely what is needed to produce good yields of a wider variety of foreign proteins than previously possible.

REFERENCES

[1] YOUNG R. A. and DAVIS R. W., Proc. Natl. Acad. Sci. U.S.A. *80*, 1194-1198 (1983).

[2] YOUNG R. A. and DAVIS R. W., Science, *222*, 778-782 (1983).

[3] GOLDBERG A. L., and ST. JOHN A. C., Annu. Rev. Biochem. *45*, 747-803 (1976).

[4] CHARNAY P., GERVAIS M., LOUISE A., GALIBERT F. and TIOLLAIS P., Nature *286*, 893-895 (1980).

[5] EDMAN J. C., HALLEWELL R. A., VALENZUELA P., GOODMAN H. M. and RUTTER W. J., Nature *291*, 503-506 (1981).

[6] KUPPER H., KELLER W., KURTZ C., FORSS S., SCHALLER H., FRANZE R., STROHMAIER K., MARQUARDT O., ZASLAVSKY V. G. and HOFSCHNEIDER P. H., Nature *289*, 555-559 (1981).

[7] STANLEY K., Nucleic Acids Res. *11*, 4077-4092 (1983).

[8] BROSSIUS J., Personal communication.

[9] ROSENBERG M., Personal communication.