科技资料

Recombinant Systems in Protein Expression

RECOMBINANT SYSTEMS IN PROTEIN EXPRESSION

Proceedings of the Labsystems Research Symposium II on Recombinant Systems in Protein Expression, Imatra, Finland, 23-26 July 1989

Editors:
Kari K. Alitalo
Tumor Biology Unit,
University of Helsinki,
Helsinki, Finland

Marja-Liisa Huhtala
Labsystems Research Laboratories,
Helsinki, Finland

Jonathon Knowles
Biotechnical Laboratory,
Technical Research Center of Finland,
Helsinki, Finland

Antti Vaheri
Department of Virology,
University of Helsinki,
Helsinki, Finland

1990

ELSEVIER SCIENCE PUBLISHERS
AMSTERDAM • NEW YORK • OXFORD

ISBN 0 444 81129 x

This book is printed on acid-free paper.

Published by:
Elsevier Science Publishers B.V.
(Biomedical Division)
P.O. Box 211
1000 AE Amsterdam
The Netherlands

Sole distributors for the USA and Canada:
Elsevier Science Publishing Company Inc.
655 Avenue of the Americas
New York, NY 10010
USA

Printed in The Netherlands

PREFACE

Labsystems symposia were founded to meet the challenges of modern biotechnology and to advance knowledge on the rapid advances and molecular biology, which have occurred during the past few years. The first symposium, held in 1985, focused on the antigenicity and uses of synthetic peptides in biology and medicine. Subsequent developments have witnessed the importance of this topic as well as the novel progress within the biotechnology of protein, antigens, enzymes, growth factors and extracellular matrix, to mention some of the best studied systems, in which synthetic peptides have been instrumental.

The second symposium "Recombinant systems in protein expression", was held in summer 1989 in Eastern Finland, at the Imatra State Hotel, which provided a very pleasant setting for such a meeting. The topic was divided in six sessions dealing with the prokaryotic systems, lower eukaryotic systems, viral vectors and DNA viruses, retroviral vectors, gene expression in animal cells and, as the most complex topic, the expression in transgenic systems. Both basic research and practical aspects of the different expression vector systems were explored in the contributions. The symposium thus covered vector developments in a variety of biological systems, from micro-organisms to higher eukaryotic cells and organisms. Due to the collection of top biomedical scientists and protein biochemists contributing their work to the symposium, the decision was made to publish the proceedings volume in a format readily available for the scientific community.

This volume should be of interest and use for researchers in basic and applied science in the areas of modern medicine, biochemistry, genetics, microbiology, immunology, matrix biology, oncogene and growth factor research and endocrinology. Of special interest for the future are the challenges of recombinant DNA vectors in higher organisms, the production of proteins in transgenic animals, and the manipulation of genomes of animal cells by the use of homologous recombination. These novel developments should allow the reproduction of animal models for various monogenic human diseases, and for the study of the multistep nature of tumor development *in vivo*. The cloning and expression of polypeptide hormone and growth factor receptors and identification of their ligands has great potential in the treatment of various diseases, including the promotion of wound healing. Microbial proteins expressed by recombinant systems are increasingly used as diagnostic reagents and in vaccines. The editors hope that this volume can act as a further catalyst in these rapidly advancing research fields.

We wish to thank all authors of the present volume and various members of the Labsystems Research Department and especially Ms. Aira Halonen for the contributions to the sysmposium which made this volume possible.

Kari Alitalo, MD, PhD
Professor, Tumor Biology
University of Helsinki

Jonathan Knowles, PhD
Professor
Biotechnical Laboratory
Technical Research
Center of Finland

Marja-Liisa Huhtala, PhD
Director of Research
LABSYSTEMS Oy

Antti Vaheri, MD, PhD
Professor and Chairman
Department of Virology
University of Helsinki

CONTENTS

PROKARYOTIC SYSTEMS

LOWER EUKARYOTIC SYSTEMS

HIGHER EUKARYOTIC SYSTEMS
Viral vectors – DNA-viruses

Expression in animal cells

Retroviral vectors

TRANSGENIC SYSTEMS

PROKARYOTIC SYSTEMS

Recombinant Systems in Protein Expression
K.K. Alitalo, M. L. Huhtala, J. Knowles and A. Vaheri, editors

THE pEX VECTORS AND THEIR USES IN MOLECULAR AND CELL BIOLOGY

KEITH K. STANLEY

The Heart Research Institute, 145-147 Missenden Road, Camperdown, Sydney, New South Wales 2050 (Australia)

INTRODUCTION

The pEX family of bacterial expression vectors provides a flexible system for expressing open reading frame DNA. All the signals for high level expression are provided by the vectors, and variations are available which allow proteins to be expressed as fusions with E. coli β-galactosidase, fusions with a short segment of cro, or as native proteins using an *in vitro* transcription-translation system. Large libraries are easy to construct in the vectors using adaptors to insert DNA fragments and electroporation to transform *E. coli*. The vectors may be used for expression screening of cDNA libraries, synthesis of protein for functional studies or raising antibodies, and mapping of antibody and receptor binding sites.

VECTOR DESIGN

The pEX family of expression vectors[1] (Figure 1) was optimised for expression by fusing a DNA fragment containing the P_R promoter from bacteriophage λ and varying lengths of the downstream *cro* gene to a derivative of the lac Z gene encoding *E. coli* β-galactosidase. It was found that a minimum of 9 codons of the *cro* mRNA were necessary in order to obtain the highest levels of synthesis, presumably reflecting a requirement for appropriate folding at the 5' end of the mRNA[2]. This cro amino terminus is present in all the pEX family of vectors giving expression levels in the order of 10^5 molecules per cell, or 100 ng per colony. Since in most pEX vectors the DNA insert is located at a distant site it cannot affect the folding of the 5' end of the mRNA and expression should be independent of its sequence.

The inclusion of a gene encoding the temperature sensitive λ repressor (*cI857*) on the pUEX plasmids[3] ensures that the level of expression is low when the plasmids are grown at 30-34°C allowing DNA encoding potentially lethal fusion proteins to be cloned (pEX plasmids must be grown in strains expressing the *cI857* gene in order to obtain the same control of expression). When logarithmic phase cultures are then shifted to 42°C a 300 fold induction of expression occurs giving rise to fusion proteins which after 2 h can account for 30% of the cellular protein. This makes the pEX family of vectors very easy to use as the fusion protein bands are often the strongest bands to be seen on SDS gels of the total *E. coli* extract (Figure 2). In addition the pEX and pUEX plasmids are available with three different reading frames through the cloning linker, and all plasmids in the family contain stop codons for translation ensuring that almost any open reading frame DNA can be expressed in the vectors.

A variety of derivatives of the pUEX plasmids have been constructed (Figure 1). The pTEX plasmid is only available in one reading frame, but this plasmid supplies a longer cloning linker and a T7 promoter in between the lac Z gene fragment and the cloning site. Transcripts from the inserted DNA can therefore be synthesised *in vitro*, allowing non-fused proteins to be expressed in protein translation systems.

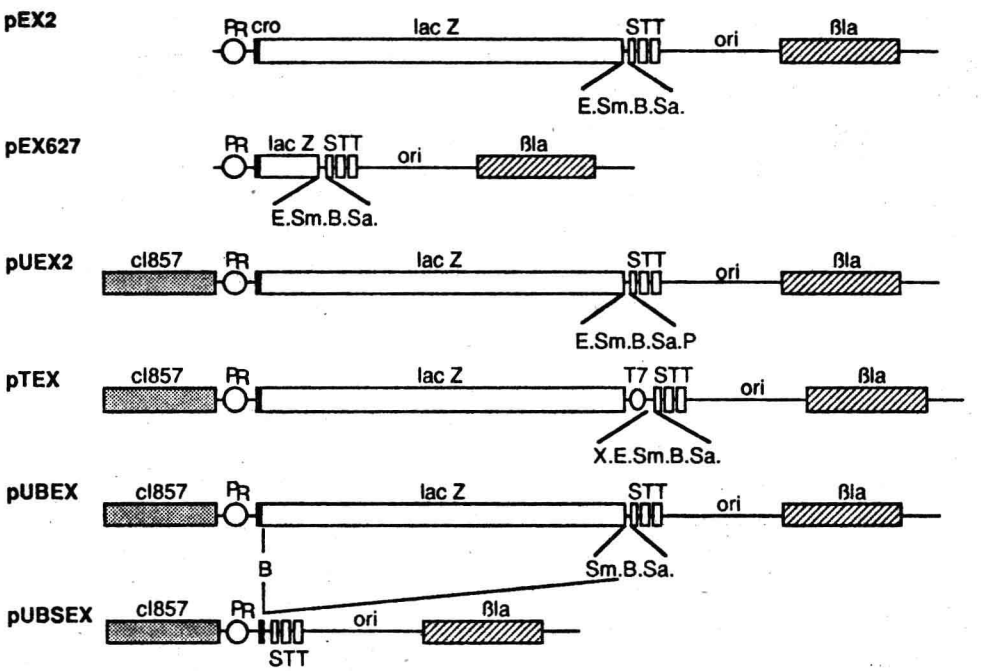

Fig. 1. The pEX vector family. For pEX and pUEX three different reading frames are available but only one version is shown. pTEX and pUBEX are derivatives of pUEX, pUBSEX is a derivative of pUBEX obtained by *Bam* HI digestion and intramolecular ligation. Restriction enzyme sites are abbreviated: E - *Eco* RI, Sm - *Sma* I, B - *Bam* HI, Sa - *Sal* I, P - *Pst* I, X - *Xho* I. S indicates the segment encoding stop codons in all three reading frames, T is a transcription termination segment from phage fd.

pUBEX[4] is a version of pUEX containing an additional Bam HI site in between the cro sequence and the lac Z gene fragment. This vector therefore has Bam HI sites flanking both sides of the lac Z gene enabling it to be deleted by a single restriction enzyme digestion and intramolecular ligation. This vector is particularly useful when antibodies against cDNA clones are required. A cDNA clone may be selected from a library in pUBEX as a stable fusion with *E. coli* β-galactosidase and then modified to produce a fusion protein containing only 9 amino acids of cro at the amino terminus.

In about 50% of cases this leads to a very high level of expression of the protein (Figure 3). In the remainder the cloned DNA sequences presumably effect the efficiency of translation initiation and reduce the level of expression. This can sometimes be rectified by small deletions at the fusion point using Bal 31. The shortened version of the vector (pUBSEX, Figure 1) can also be used for direct cloning and expression of DNA fragments, although only 2 reading frames for this vector are available at present.

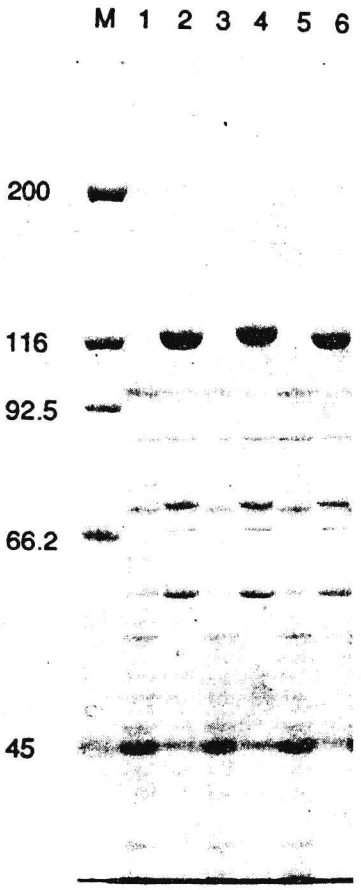

Fig. 2. Expression of β-galactosidase fusion protein by the pUEX vectors. Lanes 1, 2 pUEX1; lanes 3, 4 pUEX2; lanes 5, 6 pUEX3. Lanes 1, 3 and 5 show the SDS extractable proteins from cultures grown at 30°C, lanes 2, 4, and 6 show the same cultures after induction for 2 h at 42 °C. M indicates a lane of molecular weight markers with the sizes shown in kD.

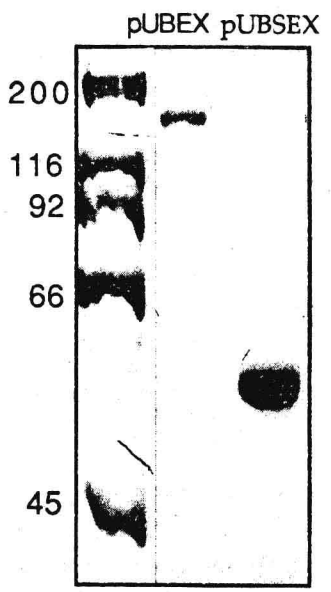

Fig. 3. Expression of a fusion protein in pUBEX and pUBSEX. Lane 1 - molecular weight markers; lane 2 - fusion protein in pUBEX; lane 3 - same construction after deletion of the lac Z by *Bam* HI cut and religation. Lanes 2 and 3 both show total SDS extractable proteins from cultures induced at 42°C for 2 h.

CONSTRUCTION OF cDNA LIBRARIES IN PLASMID VECTORS

A detailed protocol for making cDNA libraries in plasmid vectors has been described elsewhere[5,6]. In essence this involves the ligation of a 24 bp single stranded extension to each end of the vector and blunt-ended cDNA followed by annealing and ligation into the vector (Figure 4). The advantages of this method are its efficiency, allowing libraries of 10^6 to 10^7 clones per μg of cDNA to be achieved, the ability to follow each step biochemically giving quick trouble shooting, and the flexibility of being able to design adaptors for different purposes.

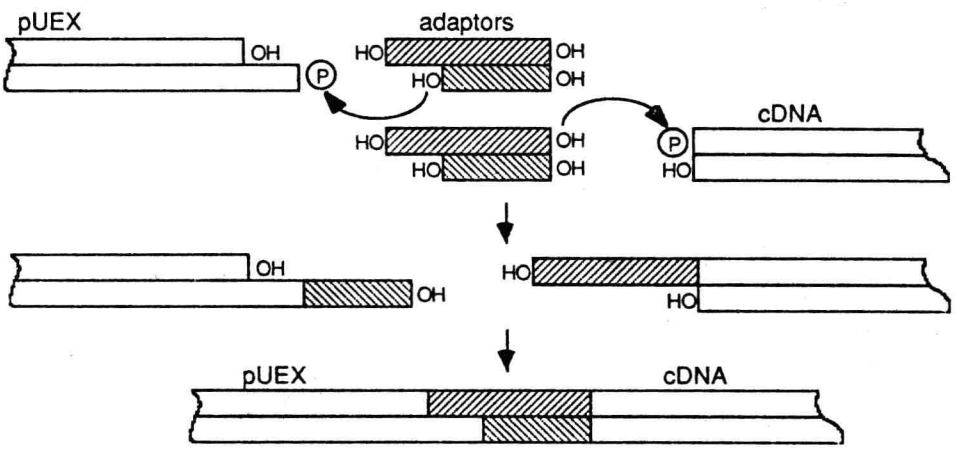

Fig. 4. Adaptor cloning strategy. 1) Unphosphorylated double-stranded adaptors are ligated to both vector and cDNA in a single tube reaction. 2) Non-covalently linked adaptor is removed by gel filtration at 70°C to reveal 24 bp sticky ends. 3) cDNA and vector are kinased, annealed and ligated in a single step to yield covalent circular DNA having a high transformation efficiency.

Almost any sequence of complementary oligonucleotides about 24 base pairs long with one sticky end and one blunt end should work. All our designs have incorporated peptide sequences which serve as flexible spacers between the β-galactosidase and foreign polypeptide giving better access to antibodies, and use codons found in abundant *E. coli* proteins so that there is no danger of premature termination between the β-galactosidase and expressed insert due to exhaustion of a rare aminoacyl-tRNA. Particular adaptor pairs encode special features, for example allowing cleavage of the fusion protein by factor Xa (Figure 5). Adaptors with *Sal* I sticky ends are required for cloning into pUBEX so that the lac Z can be subsequently deleted by *Bam* HI digestion without excision of the inserted DNA. We normally use random primers for first strand cDNA synthesis so that all parts of long mRNA molecules are represented in the

cDNA libraries. Since the double stranded cDNA is size selected[5] small mRNA molecules are represented by full length cDNA clones in the library while very large ones are present as overlapping sets of transcripts.

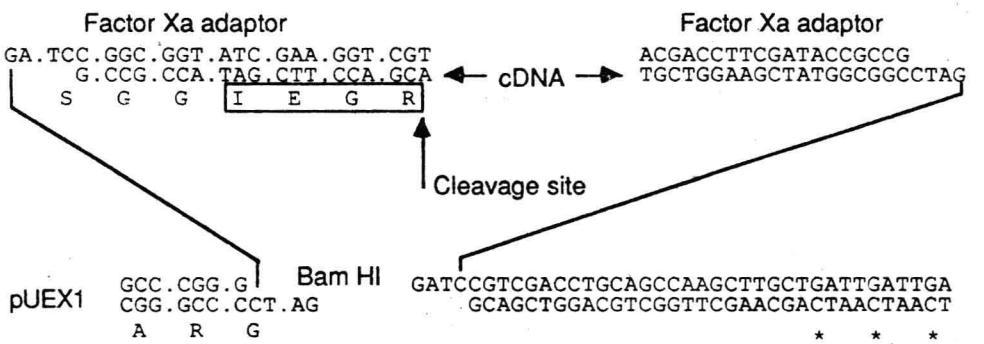

Fig. 5. Cloning of cDNA using factor Xa adaptors. Adaptors containing a factor Xa site flanking the cDNA fragment allow excision of unmodified polypeptide after expression by digestion with factor Xa. Asterisks indicates the stop codons in each reading frame.

EXPRESSION SCREENING OF cDNA LIBRARIES

A byproduct of the extremely high levels of expression achieved by the pEX vectors is that the fusion proteins generated are insoluble and precipitate as quasi-crystalline inclusion bodies in the E. coli cytoplasm[7]. Methods for detecting the expressed proteins must therefore include powerful solubilisation if the full potential of the system is to be realised. Both SDS and guanidinium HCl give complete solubilisation of the fusion proteins and may be used for expression screening. When screening with an antibody that has been tested by Western blotting it is usual to use SDS to solubilise the *E. coli* colonies on nitrocellulose filters since the antigen is then presented in a similar way to the antibody. This method has been described in detail[1,8] and has been used successfully for many cDNA clones.

When a library is being screened for a receptor by binding of its ligand to expressed proteins, then it is usually wise to avoid the use of SDS which often does not allow full renaturation of the fusion protein. In these cases solubilisation with 6M guanidinium hydrochloride followed by renaturation using capillary blotting has proven to be more effective.

EXPRESSION OF PROTEINS IN THE pEX PLASMIDS

Large amounts of protein may be generated from most pEX clones allowing simple isolation procedures. When affinity purifying a polyclonal antibody raised against a mammalian protein (which does not cross-react with any bacterial protein) it

is sufficient to lyse a lawn of whole E. coli colonies on a nitrocellulose filter and process the filter as for colony screening (Figure 6). From a lawn of 5000 colonies on an 8 cm filter several hundred μg of fusion protein is available. The antibody obtained is sufficiently pure for most analytical techniques like Western blotting and immune-fluorescence microscopy. When higher purity is required fusion proteins separated on SDS gels and transferred to nitrocellulose may be used as immune-absorbant.

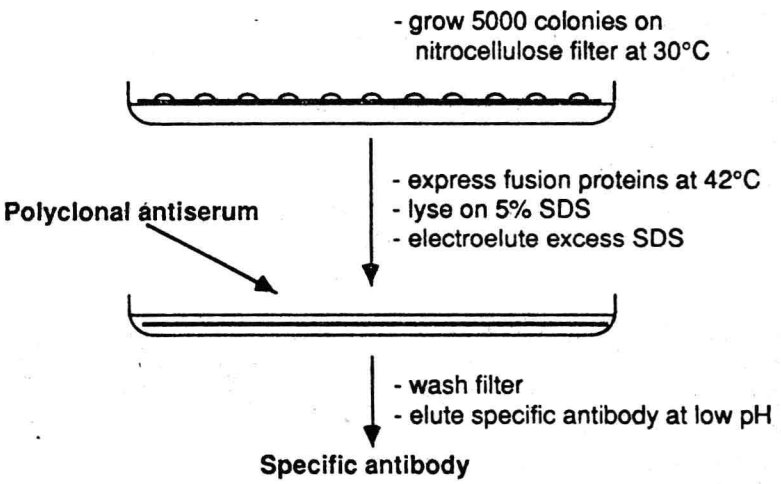

Fig 6. Affinity purification of polyclonal antisera on lawns of E. coli expressing fusion proteins in pEX.

Although affinity purified antibodies can be rapidly obtained it is often desirable to raise antibodies against fusion proteins. These antibodies are of predetermined specificity and, because of the clonal nature of the antigen, are without contamination by mammalian proteins giving rise to clean antisera from relatively crude preparations. For raising antibodies the fusion protein antigen can be quickly prepared from washed inclusion bodies[9] solubilised at high pH[10,11]. In our experience however washed inclusion bodies contain lipopolysaccharides which can cause polyclonal activation when injected into some animals. We therefore prefer to solubilise inclusion bodies in SDS and run them on SDS gels. The fusion protein band is excised and homogenised in PBS in order to isolate the purified protein for immunisation. This procedure gives antisera against the expressed insert in approximately 50% of cases, even when the insert is only a few amino acids in length. Unfortunately, however, the antibodies often only recognise denatured forms of the protein. Fusion proteins expressing short exposed segments of the protein should be used when antibodies recognising the native protein are required. Exposed segments may be determined by epitope mapping of a crude serum raised against the native protein (see below). Although the β-galactosidase

portion of the fusion protein often makes the expressed insert more stable, it also is very immunogenic limiting the amount of specific IgG obtained against the expressed DNA insert. We would therefore recommend using fusion proteins expressed in pUBSEX if these are stable.

In many circumstances it is desirable to express the native protein without any additional amino acid residues. The pTEX vectors yield high levels of transcript from the T7 promoter allowing the synthesis of polypeptide chains corresponding to the cloned cDNA in *in vitro* translation systems. An in-frame initiating methionine is supplied by the adaptor sequence allowing peptide fragments to be synthesised which do not contain the 5' end of the cDNA. Highly labelled proteins can be generated which are useful in establishing the validity of clones by comparison of proteolytic cleavage fragments. Both pTEX and pUBEX may be used like pUEX for expression screening of cDNA libraries allowing multiple functions within a single cDNA library.

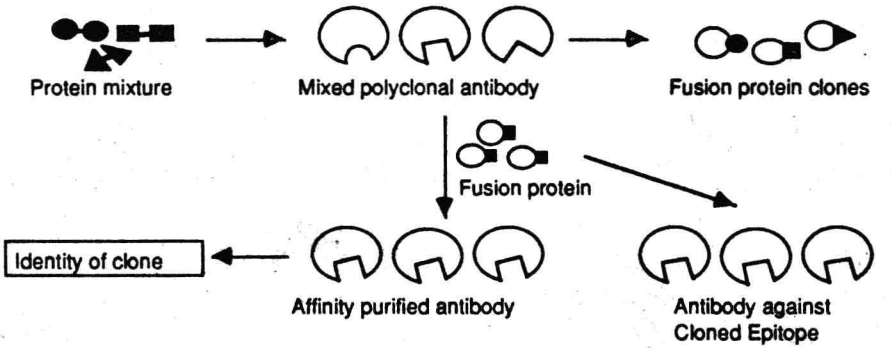

Fig. 7. Isolation of individual cDNA and antibody probes starting with a complex mixture of proteins by expression cloning and immune-selection.

IMMUNE-SELECTION OF cDNA CLONES

A consequence of the high levels of expression in the pEX plasmids is the ease of antibody affinity purification using fusion protein bound to nitrocellulose as a solid phase (Figure 6). We have used this as a step in the isolation of cDNA clones when using an antibody which recognises many proteins in a crude mixture. cDNA clones expressing proteins which bind the crude serum are first isolated from a cDNA library and colony purified. At this stage it is not known which clones correspond to which protein in the original antigen. The serum is then affinity purified on the first clone and used to screen the colony purified clones, thus identifying other clones with the same epitope. This step may then be repeated until all the isolated clones have been grouped into epitope classes (Figure 7). The affinity purified antibodies may then be used to characterise each cloned protein by Western blot and immune-fluorescence.