pALEX, a dual-tag prokaryotic expression vector for the purification of full-length proteins

(Recombinant DNA; fusion protein; protein purification)

Christos A. Panagiotidis and Saul J. Silverstein

Department of Microbiology, College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA. Tel. (1-212) 305-4137

Received by F. Barany, 29 March 1995; Accepted 8 May 1995; Received at publishers: 12 June 1995

SUMMARY

pALEX, a prokaryotic expression vector, was constructed in which the multiple cloning site (MCS, polylinker) is flanked by sequences encoding glutathione S-transferase (GST) at the 5’ end and a His$_6$ residue tag at the 3’ end. Open reading frames cloned into this vector can direct production of fusion proteins with GST at their N terminus and a His$_6$ tag at their C terminus. This allows for the purification of full-size fusion proteins by a sequential two-step procedure on glutathione-agarose and Ni$^{2+}$-agarose columns.

INTRODUCTION

The study of the biochemical properties of proteins has been greatly facilitated in recent years by the ability to clone and express protein-coding nt sequences in a variety of systems. A number of refined gene expression systems take advantage of the high affinities of certain proteins or protein sequences (tags) for their substrates, cofactors or other ligands to achieve rapid purification of re-proteins fused with such tags (Uhlen and Mocks, 1990). Fusion proteins can be purified relatively easily by affinity chromatography on supports containing immobilized ligands that specifically bind the tags.

Two of the most widely used protein fusion systems incorporate the enzyme GST (Smith and Johnson, 1988) or a small His$_6$ or His$_{16}$ tag (Smith et al., 1988; Hochuli, 1990). Fusion proteins are purified in one step by affinity chromatography on columns containing immobilized ligand (reduced glutathione for GST, or Ni$^{2+}$ for His$_6$ fusions).

In many instances production of a re-protein in bacteria can be a challenging task. Certain proteins fail to fold properly when overproduced in Escherichia coli forming insoluble aggregates, whereas others are degraded to variable extents by cell proteases or expressed at only very low levels or any combination of the above. While a denatured or partially degraded protein can still be quite useful, e.g., as antigen for the generation of antibodies, purification of full-length proteins is desirable and often required. Here we describe the construction of pALEX, an expression vector that places tags at both the N- and C-termini of the protein of interest and permits the dual selection and selective purification of full-length proteins.

EXPERIMENTAL AND DISCUSSION

(a) Vector construction

The GST-coding sequence was amplified from plasmid pGEX-3X (Smith and Johnson, 1988) by PCR using a forward primer (5’-TTACCATGGAAACAGTA-
TTCA1GTGTC) that creates a NcoI site overlapping the ATG start codon of the GST. The reverse primer (5'-CGAGGCGAGATCGTCACTGTC) was complementary to sequences located downstream from the polylinker of pGEX-3X. The PCR product was cloned in plasmid pCR II (Invitrogen, San Diego, CA, USA), using the TA cloning system. The insert, which includes the GST-coding sequence, was recovered from the resulting plasmid (pCPC-GST-Ni) after digestion with NcoI + EcoRI and it was cloned between the NcoI and EcoRI sites of plasmid pET-21d (Novagen, Madison, WI, USA). The resulting plasmid (pALEX, Fig. 1A) produces GST under the control of the bacteriophage T7 promoter in E. coli host strains that can provide T7 RNA polymerase upon induction. The system has all the advantages of the Studier et al. (1990) pET system with the added bonus of the double selection. The His6 tag in pALEX is in frame with the GST sequences (Fig. 1B) and the protein can be purified by chromatography on either glutathione- or Ni²⁺-agarose affinity supports (Fig. 1C). GST can be cleaved away from the rest of the protein by digestion with coagulation factor Xa, an endoproteinase whose cleavage sequence is located at the C terminus of GST (Fig. 1B).

(b) Expression and purification of a full-length GST-ICP-0-3H fusion protein

We tested the pALEX expression vector system for its utility for selectively purifying full-length proteins by cloning in the sequences encoding an 1-395 of ICP-0, a herpes simplex virus type-1 transcriptional regulator. This protein represented a good test for the system as it was our experience (unpublished observations) that it is normally severely degraded when produced in bacteria. The ICP-0 expressing plasmid (pCPC-X0-3H) contains the appropriate ICP-0 sequences placed in frame with both GST and the His6 tag. Purification of the fusion protein by affinity chromatography on glutathione-agarose yielded a preparation where partial degradation products constituted the majority of the purified protein and the full-size fusion was underrepresented (Fig. 1C). Further purification of the glutathione-agarose eluate on Ni²⁺-agarose resulted in the removal of most degradation products and the selective enrichment in the preparation of full-length fusion protein. The persistence of some partially degraded fusion proteins in the final preparation might be the result of GST dimerization (Mamervik et al., 1985). ICP-0 intermolecular interactions or the ability of ICP-0’s zinc-finger domain (Everett et al., 1993) to bind Ni²⁺ ions. In our experience less degradation products are present in the final preparation when binding and subsequent washing of the fusion proteins on Ni²⁺-agarose take place in the presence of 5-15 mM glutathione, presumably because of disruption of GST dimers. Another way to eliminate contamination problems resulting from GST dimerization is by cleavage of the fusion protein with factor Xa prior to Ni²⁺-agarose

![Fig. 1. Map of pALEX. its MCS and purification of fusion proteins under native conditions by double selection.](image-url)
purification. Thus, double selection of fusion proteins tagged at both ends provides rapid selection and enrichment for the full-length product.

(c) Potential applications for the purification of renatured insoluble proteins

As mentioned above many proteins do not fold properly when overproduced in bacteria and they form insoluble aggregates. Insoluble proteins produced using pALEX can be purified by Ni$^{2+}$-agarose affinity chromatography under denaturing conditions as they contain the His$_6$ tag. Because most functional assays require the use of soluble native proteins, proteins have to be renatured before use in these assays. Protein renaturation can be achieved, although to variable extents, by the gradual removal of the denaturing agent from dilute protein solutions. Renaturation of proteins synthesized using pALEX can be tested by their ability to bind to glutathione-agarose. Only properly folded GST domains will bind their immobilized substrate and be purified. Although renaturation of the GST domain of a fusion protein does not assure that the rest of the protein is renatured, it may be indicative of the folding state of the protein.

(d) Conclusions

We describe a new expression vector that combines the simplicity and specificity of the GST-fusion system with the ability of the His$_6$-tagged proteins to be purified under denaturing conditions. Furthermore, the system allows for the purification of full size re-proteins by flanking them with these two tags. Purification of full size fusion proteins by two sequential affinity chromatography steps is simpler and more rapid than conventional protein purification. Therefore it should provide higher protein yields than multistep protein purification because it minimizes the number of manipulations and the purification time.

ACKNOWLEDGEMENT

This study was supported by a United States Public Health Service grant AI 33952 to S.J.S.

REFERENCES


