THE FUNCTIONAL EXPRESSION OF ANTIBODY F_v FRAGMENTS IN *ESCHERICHIA COLI:* IMPROVED VECTORS AND A GENERALLY APPLICABLE PURIFICATION TECHNIQUE

Arne Skerra[1], Ilse Pfitzinger and Andreas Plückthun*

Genzentrum der Universität München, Max-Planck-Institut für Biochemie, D-8033 Martinsried, FRG. [1]Present address: Abt. Molekulare Membranbiologie, Max-Planck-Institut für Biophysik, D-6000 Frankfurt 70, FRG. *Corresponding author.

We have previously demonstrated that the expression of fully functional F_v and F_{ab} fragments in *E. coli* is possible by the simultaneous secretion of both chains to the periplasm. To increase production levels and facilitate engineering and random mutagenesis, we improved our previous vectors by introducing a resident repressor gene and a filamentous phage origin. We also developed a new purification strategy based on immobilized metal ion chromatography, with which a single-chain F_v fragment can be purified to homogeneity in a single step. We investigated the most efficient tail constructions and found that only a minimal structural change of three additional C-terminal amino acids is necessary. This modification has no deleterious effect on *in vivo* transport and folding or antigen affinity.

Besides facilitating antibody engineering, *E. coli* expression is particularly suited for the direct production of smaller fragments of the antibody (F_v or F_{ab})[1,2]. These fragments are interesting candidates for structural studies (such as NMR) requiring molecules of lower molecular weight, and for diagnostic and therapeutic applications requiring good tissue penetration and low antigenicity. The expression strategy we described[1,2] has proven to be useful in the construction of expression libraries of antibody fragments[3,4], and another extension of this methodology may be in the use of random mutagenesis. Further, antibody fragments expressed in *E. coli* can be an interesting alternative to monoclonal antibodies when they are used for the purification of or coupling to other proteins produced in *E. coli.* We report here improved vectors and purification procedures to facilitate mutagenesis, characterization and production of antibody fragments in *E. coli.*

The experiments for developing this methodology were carried out with the well characterized antibody McPC603, an IgA from mouse, which binds the hapten phosphorylcholine[5,6]. The availability of information about the binding properties of this antibody[7–9] and the three-dimensional structure of its F_{ab} fragment with the bound hapten[10,11] was particularly helpful. We demonstrated previously that the F_v fragment of McPC603 has the same intrinsic hapten affinity constant as the whole antibody from mouse or its F_{ab} fragment[1,12] and is therefore a good model for the study of antigen binding. Furthermore, we could show[12] that the two chains of the F_v fragment can be linked to form a secreted single-chain F_v fragment. Previously, such single-chain fragments had been obtained from inclusion bodies of *E. coli*[13,14], but such a fragment can also be expressed in functional form when it is obtained from a secretion system. It was found to possess unchanged hapten binding properties compared to the original F_v fragment[12].

In this paper, we describe improved vectors for the functional expression of these antigen binding immunoglobulin fragments in *E. coli.* These new plasmids carry a regulatable promoter and its cognate repressor gene, efficient ribosomal binding sites, structural genes encoding the corresponding antibody fragments fused to bacterial signal sequences, a transcription terminator, a filamentous phage replication origin and a high copy number plasmid replication origin. We also present a general purification scheme applicable to antigen binding frag-

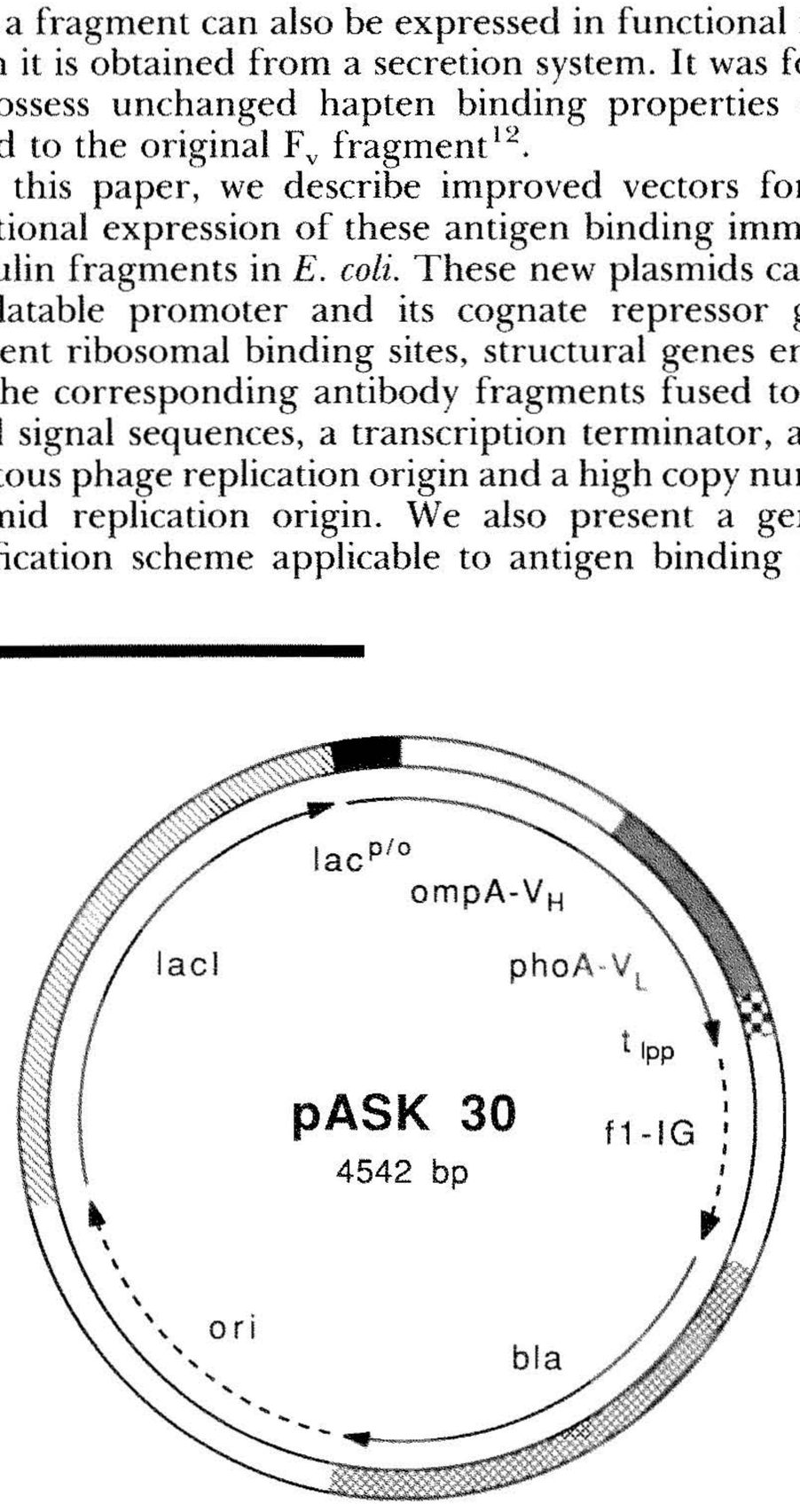
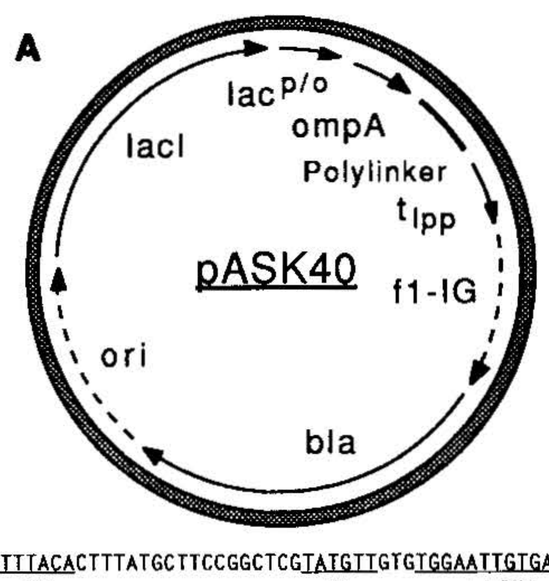


FIGURE 1 Schematic drawing of the expression vector pASK30. For an explanation of the symbols, see the text.

```
TTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCA  75
                -35                     -10        → mRNA
                                  XbaI
CACAGGAAACAGCTATGACCATGATTACGAATTTCTAGATAACGAGGGCAAAAAATGAAAAAGACAGCTATCGCG  150
     LacZ:  MetThrMetIleThrAsnPheEnd     OmpA-VH:  MetLysLysThrAlaIleAla
                                                   -21

ATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCCGAAGTTAAACTGGTAGAGTCTGGTGGTGGTCTG  225
IleAlaValAlaLeuAlaGlyPheAlaThrValAlaGlnAlaGluValLysLeuValGluSerGlyGlyGlyLeu
                                   +1

GTACAGCCGGGTGGATCCCTGCGTCTGTCTTGCGCTACCTCAGGTTTCACCTTCTCTGACTTCTACATGGAGTGG  300
ValGlnProGlyGlySerLeuArgLeuSerCysAlaThrSerGlyPheThrPheSerAspPheTyrMetGluTrp
                                                  ---- CDR I ----

GTACGTCAGCCCCCGGGTAAACGTCTCGAGTGGATCGCAGCTAGCCGTAACAAAGGTAACAAGTATACCACCGAA  375
ValArgGlnProProGlyLysArgLeuGluTrpIleAlaAlaSerArgAsnLysGlyAsnLysTyrThrThrGlu
                                  ---------- CDR II ----

TACAGCGCTTCTGTTAAAGGTCGTTTCATCGTTTCTCGTGACACTAGTCAATCGATCCTGTACCTGCAGATGAAT  450
TyrSerAlaSerValLysGlyArgPheIleValSerArgAspThrSerGlnSerIleLeuTyrLeuGlnMetAsn
----------

GCATTGCGTGCTGAAGACACCGCTATCTACTACTGCGCGCGTAACTACTATGGCAGCACTTGGTACTTCGACGTT  525
AlaLeuArgAlaGluAspThrAlaIleTyrTyrCysAlaArgAsnTyrTyrGlySerThrTrpTyrPheAspVal
                                  ---------- CDR III ----------

TGGGGTGCAGGTACCACCGTTACCGTTTCTTCTTGATAACATGGAGAAAATAAAGTGAAACAAAGCACTATTGCA  600
TrpGlyAlaGlyThrThrValThrValSerSerEnd     PhoA-VL:  MetLysGlnSerThrIleAla
                    +122                            -21

CTGGCACTCTTACCGTTACTGTTTACCCCTGTGACAAAAGCCGATATCGTTATGACCCAGTCTCCGAGCTCTCTG  675
LeuAlaLeuLeuProLeuLeuPheThrProValThrLysAlaAspIleValMetThrGlnSerProSerSerLeu
                                   +1

TCTGTATCTGCAGGTGAACGTGTTACCATGTCTTGCAAATCTTCTCAGTCTCTGCTGAACTCTGGTAACCAGAAA  750
SerValSerAlaGlyGluArgValThrMetSerCysLysSerSerGlnSerLeuLeuAsnSerGlyAsnGlnLys
                                  ---------- CDR I ----------

AACTTCCTGGCGTGGTATCAGCAAAAGCCTGGCCAACCGCCGAAACTGCTGATCTACGGTGCGTCGACCCGTGAA  825
AsnPheLeuAlaTrpTyrGlnGlnLysProGlyGlnProProLysLeuLeuIleTyrGlyAlaSerThrArgGlu
----------                                        ---------- CDR II ----

TCTGGTGTTCCGGACCGTTTTACCGGTAGCGGTAGCGGTACCGACTTCACTCTGACCATCTCTTCTGTACAGGCT  900
SerGlyValProAspArgPheThrGlySerGlySerGlyThrAspPheThrLeuThrIleSerSerValGlnAla
-

GAAGATCTGGCTGTTTACTACTGTCAAAACGACCACTCTTACCCGCTGACCTTTGGCGCCGGCACCAAACTGGAA  975
GluAspLeuAlaValTyrTyrCysGlnAsnAspHisSerTyrProLeuThrPheGlyAlaGlyThrLysLeuGlu
                                  ---------- CDR III ----------

            HindIII
CTGAAGCGCGCTTGATAAGCTTGACCTGTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTTTGTCTGCCG 1050
LeuLysArgAlaEnd                                  *        mRNA →
     +115
                  BamHI
TTTACCGCTACTGCGTCACGGATCC
```

FIGURE 2 DNA sequence of the F_v operon on pASK30. The horizontal arrows denote the beginning and the end of the mRNA (as found in the *lacZ* and *lpp* operons), respectively. Underlined are, from the beginning of the operon, the −35 and −10 region of the promoter, the *lac* operator, the two ribosome binding sites and the *lpp* transcriptional terminator. Underneath the protein sequences, +1 denotes the beginning of the mature protein. The complementarity determining regions (CDR) are also labeled.




```
B
TTAGGCACCCCAGGCTTTACACTTTATGCTTCCGGCTCGTATGTTGTGTGGAATTGTGAGCGGATAACAATTTCA
                -35                     -10        → mRNA
                                  XbaI
CACAGGAAACAGCTATGACCATGATTACGAATTTCTAGATAACGAGGGCAAAAAATGAAAAAGACAGCTATCGCG
     LacZ:  MetThrMetIleThrAsnPheEnd     OmpA:  MetLysLysThrAlaIleAla
                                                -21
                                        EcoRI     KpnI   BamHI XbaI
ATTGCAGTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCCGCGAATTCGAGCTCGGTACCCGGGGATCCTCTA
IleAlaValAlaLeuAlaGlyPheAlaThrValAlaGlnAla...
                                   -1
    PstI      HindIII
GAGTCGACCTGCAGGCATGCAAGCTTGACCTGTGAAGTGAAAAATGGCGCACATTGTGCGACATTTTTTTTGTCT
                                                *        mRNA →
                          BamHI
GCCGTTTACCGCTACTGCGTCACGGATCC
```

FIGURE 3 (A) Schematic drawing of pASK40. For an explanation of the symbols, see the text. (B) DNA sequence of polylinker region of pASK40. The labeling is consistent with Figure 2.

ments independent of their particular antigen binding properties. It is based on the principle of immobilized metal affinity chromatography (IMAC) and relies on a short stretch of histidine residues at the C-terminus of one antibody domain, enabling the chelation of metal ions bound to a solid carrier[15–18]. We show here that a single chain F_v fragment can be purified to homogeneity in a single step with this method yielding the fully functional recombinant protein.

RESULTS

Design of improved expression vectors for functional antibody fragments. The plasmid pASK30 (Fig. 1) was constructed both for the bacterial expression of the F_v fragment of McPC603 and the facile engineering of its genes. Our strategy for the functional expression of the F_v fragment in *E. coli* is based on the co-expression and co-secretion of both of its chains. Thus, as in the plasmid pASK22 described earlier[1] the structural genes coding for the V_L and the V_H domain are fused to DNA sequences encoding the *omp*A and *pho*A signal peptides, respectively, and are arranged in a dicistronic operon. The DNA sequence of this operon, which is under transcriptional control of the IPTG-inducible *lac* promoter, is shown in Figure 2. The promoter/operator is followed by the 5′-terminal coding region of the *lac*Z gene including its Shine-Dalgarno sequence. The *lac*Z coding sequence is terminated after the seventh codon. The corresponding stop codon is followed by the modified translation initiation site of the *omp*A gene[19] and the start codon of the *omp*A-V_H gene fusion, which forms the first cistron of the F_v operon (neglecting the *lac*Z minicistron mentioned above). In this way, a tandem ribosomal binding site[20] is created, which leads to the efficient expression of the downstream structural gene. The second cistron of the operon is formed by the *pho*A-V_L fusion gene. It is arranged such that the stop codon of the V_H gene is followed by the *pho*A ribosomal binding site and the initiation codon. Whereas the operon itself is identical to that on the plasmid pASK22[1], its transcription is terminated on pASK30 by the strong *rho*-independent *lpp*-terminator[21]. This terminator was introduced in order to enhance the stability of the recombinant mRNA. It should also prevent transcriptional read-through into the intergenic region of the single-stranded phage f1, which is located immediately downstream of the terminator on pASK30 (Fig. 2). The construction of this plasmid is described in more detail in the Experimental Protocol.

A particularly suitable host for the new series of plasmids (pASK30 and its derivatives) was found to be the *E. coli* K12 strain JM83[22], both for DNA manipulations and for the expression of the F_v and F_{ab} fragment of McPC603. In the absence of the inducer IPTG, we observed no deleterious growth effect of the plasmid on the *E. coli* cells. Single stranded DNA of pASK30 can be produced using the *dut*$^-$*ung*$^-$strain CJ236[23]. Thus, the site directed mutagenesis protocol of Kunkel et al.[24] can be applied directly to the expression plasmid. Typical yields after purification are about 0.5 $mg \cdot l^{-1} \cdot OD_{550}^{-1}$ for the F_v fragment and 0.2 to 0.3 $mg \cdot l^{-1} \cdot OD_{550}^{-1}$ for the single chain fragments. For more general purposes than the expression of F_v or F_{ab} fragments, or the facile cloning of other antibody domains, we also constructed the plasmid pASK40, a derivative of pASK30 in which the V_H gene and the *pho*A-V_L gene fusion are replaced by the polylinker from the plasmid pUC19 (Fig. 3).

Construction of the single-chain derivative and the affinity tail. The construction of the derivative pASK-lisc encoding the single chain version of the F_v fragment of

McPC603[12] was performed by site directed mutagenesis. The noncoding region downstream of the V_H gene of pASK30 together with the gene fragment encoding the *pho*A signal sequence—as part of the *pho*A-V_L gene fusion—was replaced[12] by a DNA sequence encoding the peptide linker $(Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser)_3$ (Fig. 4).

In order to introduce an oligo-histidine tail at the C-terminus of the light chain, additional site directed mutagenesis experiments with the plasmid encoding the single-chain F_v fragment were carried out. In the first approach, both ArgL114 and AlaL115 were changed to histidine residues and three additional histidine residues were added to the C-terminus of the light chain. Thus, this modified single-chain protein now carries 5 consecutive histidine residues at its carboxy-terminal end (Table 1). Our design was primarily guided by the concern to only minimally perturb the structure of the F_v fragment.

Expression of the single-chain $(his)_5$ F_v fragment. The single-chain F_v fragment carrying the penta-histidine tail was purified by phosphorylcholine affinity chromatography in yields indistinguishable from the original single-chain F_v fragment (see Experimental Protocol). No difference in cell growth, protein transport or protein folding caused by the $(his)_5$ tail was apparent. This is in sharp contrast to results obtained from the equivalent addition of five arginine residues. In this case, severe growth defects and a high proportion of revertants not expressing this tail were observed (data not shown). Arginine tails have also been reported as a tool to separate the proteins of interest[25] by ion-exchange chromatography, especially if the protein of interest already possesses a basic isoelectric point. From our results it has to be concluded, however, that they are not suitable for secretion[26,27], even when they are located at the C-terminus. Yet, the histidine affinity tail appears not to cause such deleterious effects.

Purification of the single chain $(his)_5$ F_v fragment by IMAC. The goal of the method described here was to establish a generally useful one-step purification of this recombinant protein directly from an *E. coli* lysate. Metal affinity chromatography depends on the complexation of the column-bound metal by suitably positioned histidine residues in the protein. The protein can be eluted by using salt or protons to compete with the metal or by using imidazole to compete with the protein ligands. We have chosen the latter strategy. To optimize the immobilized metal affinity chromatography (IMAC) procedure for the sc-Fv-$(his)_5$ fragment, we systematically varied several experimental parameters such as pH, salt concentration, the choice of the metal ion and the elution method. The resulting procedure is described in more detail in the Experimental Protocol. The commercial Chelating Sepharose resin loaded with Zn^{2+} at pH 7.0 with a high salt concentration in the running buffer gave the most satisfactory purification from *E. coli* proteins, and the specific elution was achieved with a gradient of imidazole. To demonstrate the efficiency of the separation, a chromatogram from a whole cell lysate is shown (Fig. 5), but the bulk of the protein can already be separated before the chromatography step by carrying out a periplasmic fractionation (see Experimental Protocol). Other metal ions caused significantly more binding of *E. coli* proteins to the column and gave less satisfactory purifications. The nitrilotriacetic acid based IMAC resin developed by Hochuli et al.[17] was also used with Ni^{2+} for the purification of the sc-$(his)_5$-F_v fragment (data not shown). It was found comparable in performance to Zn^{2+} in conjunction with the commercial nitrilodiacetic acid based Chelating Sepharose resin.

Structural requirements of the affinity tail. We also investigated several alternatives to the $(his)_5$ tail (Table 1) in order to examine whether any of these variants might show even tighter binding. It was found that none of the tails lead to any significant growth defects, and Western blots showed complete processing for all F_v fragments with affinity tails (data not shown). In the first variant, pro$(his)_5$pro, the histidine tail was framed by proline residues. This might alter the accessibility of the tail and protect it from proteolytic attack, as the efficiency of trypsin-like proteases of *E. coli* is probably low on lys-pro bonds. The chromatographic properties were, however, found indistinguishable from the $(his)_5$ variant. Smith et al.[15] had reported that the peptides his-gly-his and his-trp bound most tightly in IMAC with several metals, and these motifs were therefore also used in the affinity tail

FIGURE 4 Comparison of pASK-lisc (bottom) with pASK30 (top). The intergenic sequence with the second ribosome binding site and the signal sequence in the dicistronic operon (underlined) was replaced by the peptide linker (bottom, underlined)

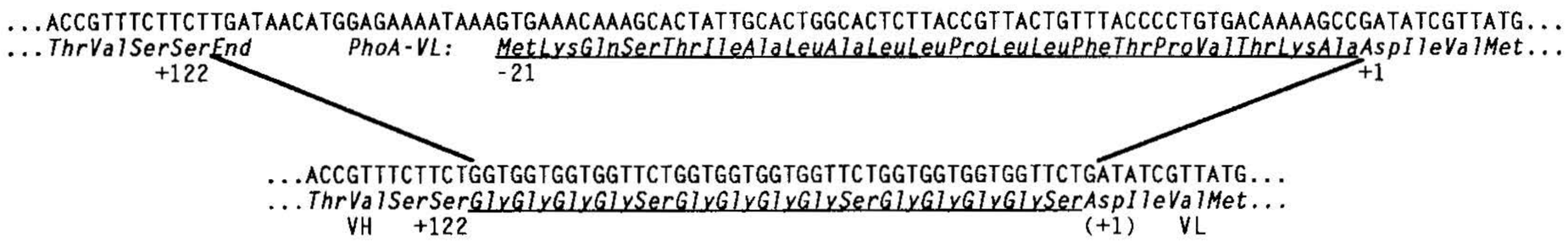


TABLE 1 Structural requirements of the affinity tails.

Variant	C-Terminal Sequence[a]	Separation from *E. coli* Proteins[b]	Elution of Pure Protein from IMAC[c]
W.T.	... ELKRA	−	n.d.[d]
$(His)_5$	... ELKHHHHH	+	(55), 105
Pro$(His)_5$Pro	... ELKPHHHHHP	+	(65), 115
$(GlyHis)_3$	... ELKLGHGHGH	−	0–20
His$(TrpHis)_3$	... ELKHWHWHWH	(+)	95
$(His)_9$	... ELKHHHHHHHHH	(+)	(0–20), (150), (190)[e]

[a] The C-terminal sequence of the single chain F_v fragment is shown, which corresponds to the end of the V_L domain. [b] A plus sign denotes a peak being eluted with high imidazole under standard IMAC conditions at the same position as in Figure 5 (see Experimental Protocol). [c] The protein was purified by hapten affinity chromatography and the elution profile of the purified protein was monitored under standard IMAC conditions. The imidazole concentration of the main peak is shown, minor peaks are shown in parentheses. A gradient from 0 to 300 mM imidazole (2 × 100 ml) was applied to a 5 ml Zn^{2+} column. [d] Not determined. [e] Trace.

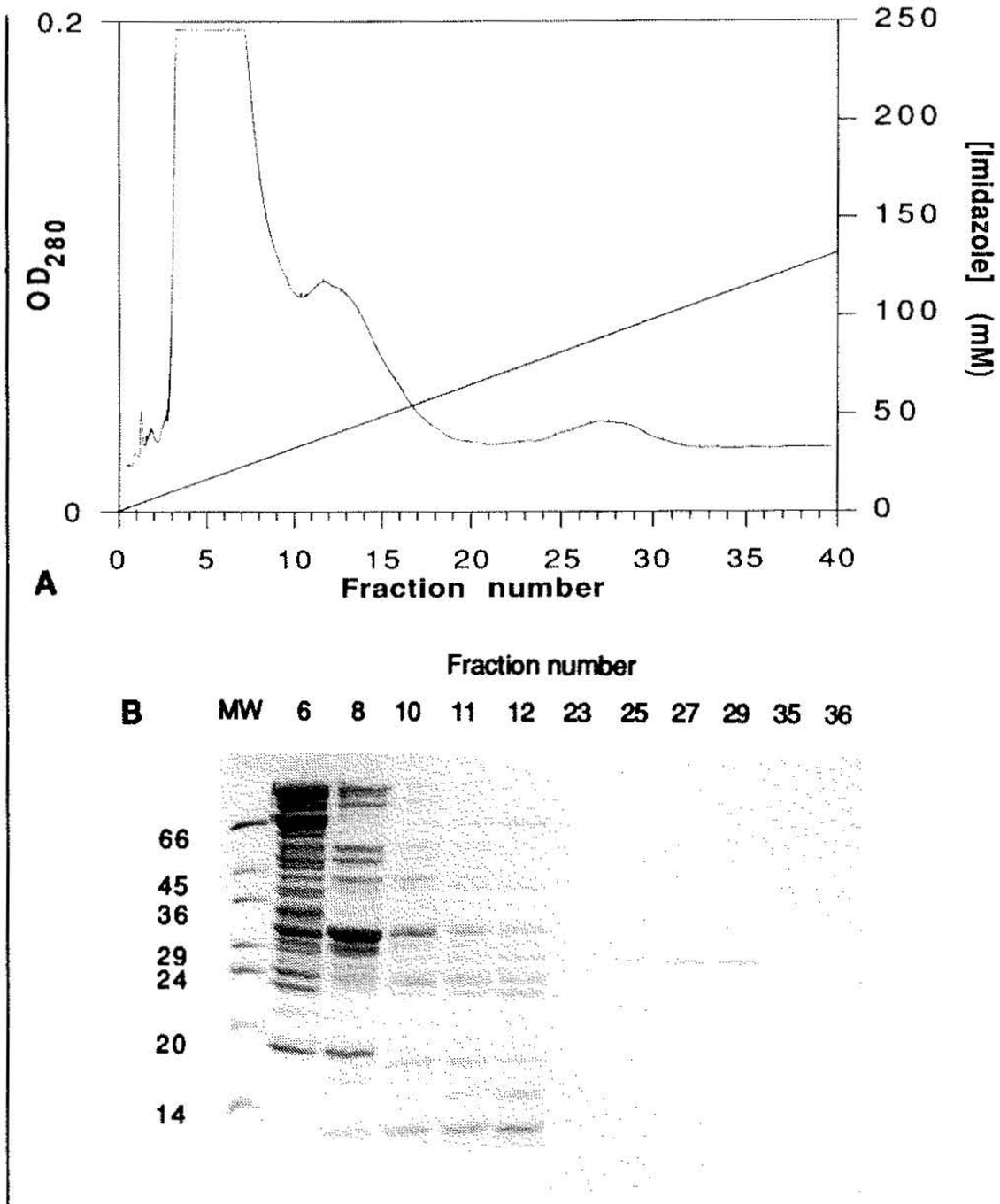


FIGURE 5 Purification of the $(his)_5$ single-chain antibody by IMAC. (A) Column elution profile as monitored by OD_{280}. (B) SDS-PAGE of relevant fractions from the chromatography shown in (A).

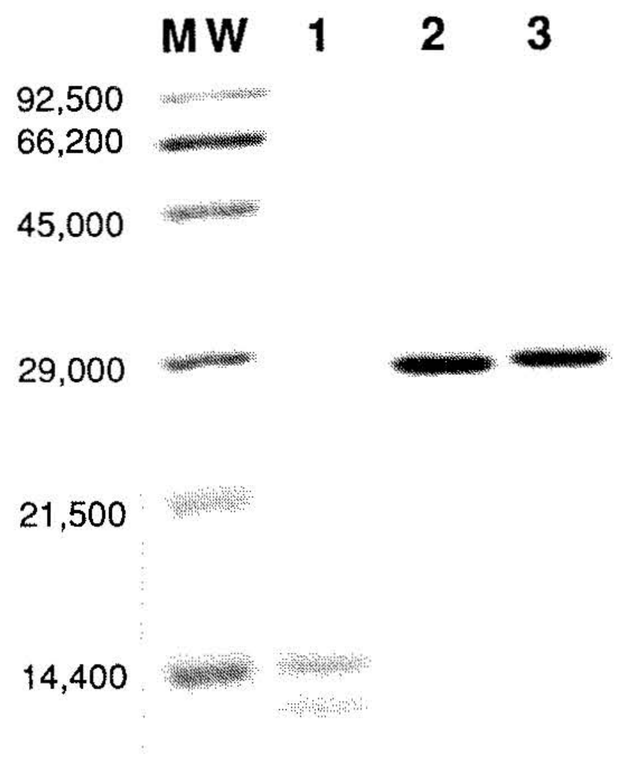


FIGURE 6 Purified antibody proteins from *E. coli:* F_v fragment (lane 1) purified by hapten affinity chromatography, single chain F_v fragment (lane 2) purified by hapten affinity chromatography, and $(his)_5$ single chain F_v fragment (lane 3) purified by IMAC.

arranged as shown in Table 1. Interestingly, both of the tails with his-X-his motifs bound more weakly to the column than the $(his)_5$ variant Zn^{2+} under these conditions. Finally, the length of the linker was extended, but the $(his)_9$ variant does not give rise to a satisfactory purification from *E. coli* proteins. The expression levels of all variants were similar to the single-chain protein carrying no affinity tail, with the exception of the $(his)_9$ variant and the his(trphis)$_3$ variant, which were obtained at about 50% and 30% of these levels, respectively.

We then investigated the binding of all variants more quantitatively by first purifying the proteins by hapten affinity chromatography and then monitoring the elution of pure protein under standard IMAC conditions (Table 1). The protein with the $(glyhis)_3$ tail and a portion of the protein with the $(his)_9$ tail elute too early to be useful. The other portion of the $(his)_9$ protein elutes over a wide range, but is produced in amounts too low to offer an advantage over the $(his)_5$ variant. (Table 1). The his(trp-his)$_3$ variant elutes slightly earlier than the $(his)_5$ variant and does not give as complete a separation from *E. coli* proteins (Table 1). Both the $(his)_5$ and the pro$(his)_5$pro variant give rise to an elution peak at around 90-120 mM imidazole, and it is this peak which is useful for a one step purification. In both cases, a fraction of the protein is already eluted at approximately 55-65 mM imidazole, but we have at present no explanation for this phenomenon. In conclusion, the five consecutive histidine residues give enough binding strength to allow a one step purification of the single-chain antibody protein. We therefore chose the $(his)_5$ variant, which has the best binding and expression properties and also causes the smallest structural perturbation, for further investigations.

Properties of the purified single-chain $(his)_5$ F_v fragment. An SDS analysis of the IMAC purified sc-$(his)_5$-F_v fragment is shown in Figure 6. When the IMAC-purified protein was subsequently passed over a hapten affinity column, all of the protein bound to the column and could be eluted with phosphorylcholine. This demonstrates its native structure. The hapten binding constant was 1.39×10^5 M^{-1} and is thus indistinguishable from the single chain F_v fragment[12] not carrying this tail and only very slightly lower than the whole antibody.

While many examples of affinity handles have appeared in the literature (summarized in ref. 28), the approach chosen by us is the most conservative in terms of adding amino acids, as it requires only three, and is sufficient for a one step purification to homogeneity. While C-terminal extensions with higher binding affinity have been reported[29], they would not be as suitable for many structural and mechanistic studies.

DISCUSSION

In the vectors described here, the immunoglobulin genes—either as a dicistronic transcription unit (the original F_v fragment) or as one structural gene (the single chain variants)—are under transcriptional control of the *lac* promoter/operator. The wild type *lac* promoter was found most useful for several reasons: it is repressible and has a relatively low leakage rate in the absence of induction, especially in the presence of glucose. This is of particular relevance in the case of the secretion of antibody fragments, in order to avoid toxic effects on the *E. coli* cells[2], possibly caused by incorrectly folded protein. We have shown in detail for the F_{ab} fragment that stronger promoters do not yield proportionally more correctly folded antibody fragment (A. Skerra and A. Plückthun, in preparation). The induction of the *lac* promoter is possible with IPTG in almost all growth media, and the presence of the gene encoding the *lac* repressor on the plasmid guarantees a sufficient overproduction of the repressor to make the expression plasmid essentially host-independent.

Further features of the improved plasmids described here include a strong transcriptional terminator of the synthetic operon. The plasmids all contain the high copy number origin of the pUC family and the ampicillin resistance gene[30]. In addition, they carry the intergenic region of the filamentous phage f1, which allows the production of single-stranded plasmid DNA upon superinfection of *E. coli* strains expressing F-pili with a helper

phage[31]. Thus, site-directed mutagenesis can be carried out directly with the plasmid. Finally, the plasmid size was kept minimal and multiple homologous sequences were avoided, in order to ensure a high genetic stability of the vectors.

During the development of the expression system, we found that induction at lower temperature for a limited period is beneficial to the yield of correctly folded protein as well as for the integrity of the host cells. Apparently, misfolded protein may permeabilize the outer membrane at higher temperatures and after long induction times[2,32]. Especially on larger scales, the work-up from periplasmic extracts or whole cell-lysates is more convenient than from cell supernatants. Thus, we favor conditions that maintain the integrity of the cell.

The antibody combining site of the F_v fragment is in close proximity to both N-termini. In contrast, the C-termini of the variable domains are located far away from the antigen binding site and normally connect to the constant domains of the light chain (C_L) or the heavy chain (C_H1). The presence of these domains has no effect on binding[1] and therefore, the C-termini rather than the N-termini appeared to be the optimal place for an "affinity tail".

We chose to use the technique of immobilized metal ion chromatography (IMAC)[15–18] as a general means for the purification of antibody fragments independent of their binding specificities. The results clearly show that the single chain F_v fragment of an antibody with an oligohistidine affinity tail can be purified to homogeneity from *E. coli*. This modified antibody fragment has a binding constant identical to the fragment lacking the affinity tail, and all of the recombinant protein purified by this method is able to bind to a hapten affinity column. In conjunction with the single chain strategy, the purification described here should be particularly useful in assuring the association of variable domains with weak affinity for each other[12], and provide a general methodology for the purification of antibody fragments from *E. coli*.

EXPERIMENTAL PROTOCOL

Strains and phages. *E. coli* K12 strain JM83[22] (*ara,Δ(lac-proAB), rps*L, Φ80, *lacZ*ΔM15), which was from our departmental collection, was used for plasmid propagation and expression of the antibody fragments. Strain CJ236[23] (*dut, ung, thi, relA1* [pCJ105 (*Cam*^{r})]) was purchased from Bio-Rad Laboratories and the helper phage M13KO7[31] was from Pharmacia.

Plasmids. pASK22 was from our own collection and has been described[1]. pINIII-*ompA*1[19] was obtained from M. Inouye and pTTQ181[33] was provided by M. Stark. The pUC-series plasmids[22] and pUC-f1 were purchased from Pharmacia.

Vector constructions. DNA manipulations were performed according to standard methodology[34]. Restriction enzymes were purchased from New England Biolabs or Boehringer Mannheim. Enzyme digests were performed as recommended by the supplier. The plasmid pASK30 (Fig. 1) was assembled as follows. The F_v operon, the sequence of which is shown in Figure 2, was obtained from plasmid pASK22[1]. The *lpp*-terminator was derived as a RsaI/BstUI fragment from the plasmid pIN-III-*ompA1*[19], whereas the f1-origin[35] was excised as a BamHI/DraI fragment from the plasmid pUC-f1. The f1 intergenic region is followed on pASK30 by the large AatII/AflIII fragment from the pUC series of plasmids[22], which carries the ampicillin resistance gene and the ColEI-origin of replication. The plasmid origin precedes the *lac*I gene coding for the *lac* repressor, which is placed on pASK30 in its authentic arrangement upstream of the *lac* promoter. Initially, the *lacI*^{q} allele of the repressor gene, which was excised as a RsaI/EaeI fragment (1.11 kb, partial digest) from the plasmid pTTQ181[33], was employed. However, since the *lac* promoter could no longer be fully induced in the resulting construct, the 5′-terminal part of this gene with the *lacI*^{q} promoter was then replaced by the corresponding PflMI/BstEII fragment of the wild type *lac*I gene from the plasmid pIN-III-*ompA1*. In the course of the construction, the 5′-protruding DNA ends of the AflIII site on the downstream end of the pUC-origin fragment and of the BamHI site upstream of the f1 intergenic region from pUC-f1 were filled in using Klenow-fragment (Boehringer). The 3′-protruding DNA-end of the AatII site upstream of the ampicillin resistance gene on the pUC-fragment was cleaved by mung bean nuclease (Promega Biotec).

***In vitro* mutagenesis.** Site directed mutagenesis was performed according to the method described by Kunkel et al.[24] using the Muta-Gene Kit from Bio-Rad Laboratories[23]. Oligonucleotides were synthesized by an Applied Biosystems 380A automated DNA synthesizer and purified by denaturing polyacrylamide gel electrophoresis. Single stranded uridine-doped template DNA was produced essentially as described by Vieira and Messing[31] from *E. coli* strain CJ236 transformed with the plasmid pASK30 or its derivatives upon superinfection with the helper phage M13KO7.

Cell growth, induction and harvest. Cells were grown at room temperature in LB medium, containing 100 μg/l ampicillin, to an OD_{550} of about 0.5 to 1.0, and IPTG was added to a final concentration of 1 mM. The induction was continued for 3 hours, the cells were collected by centrifugation, resuspended in BBS (0.2 M Na borate, pH 8.0, 0.16 M NaCl) for hapten affinity chromatography, or IMAC loading buffer (50 mM Na phosphate, pH 7.0, 1 M NaCl), and passed through a French pressure cell. The cell homogenate was centrifuged again and the soluble portion was passed through a sterile filter. As an alternative to the French press procedure and in order to achieve a prepurification of the antibody fragment, the periplasmic fraction can be prepared on a larger scale. In this case, the centrifuged cells (grown and induced as described above) were taken up in either BBS-EDTA (0.2 M Na borate, 0.1 M NaCl, 1 mM EDTA) for hapten affinity chromatography, or in 50 mM Na phosphate, 100 mM NaCl, 1 mM EDTA for IMAC. Per gram of cells, 2 ml buffer was used, and the cells were stirred for 30 min. to permeabilize the outer membrane. The cells were then centrifuged at 30000× g for 30 min. The supernatant in BBS-EDTA can be directly applied to hapten affinity chromatography. For IMAC, the EDTA must be complexed by adding $ZnCl_2$ to a final concentration of 1 mM, and NaCl is adjusted to a final concentration of 1 M.

Protein purification by hapten affinity chromatography. Hapten affinity chromatography was carried out essentially as described previously[34]. Typical yields are about 0.5 $mg \cdot l^{-1} \cdot OD_{550}{}^{-1}$ for the F_v fragment and 0.2 to 0.3 $mg \cdot l^{-1} \cdot OD_{550}{}^{-1}$ for the single chain fragments.

Immobilized metal ion chromatography (IMAC). Chelating Sepharose Fast Flow (Pharmacia; typically about 6 ml resin per l l of *E. coli* culture) was loaded with $ZnCl_2$ according to the manufacturer's instructions. The column was equilibrated with loading buffer (50 mM Na phosphate, pH 7.0, 1 M NaCl), the protein was loaded in this buffer and the column washed with additional 20 column volumes of the same buffer. A linear gradient of imidazole (0-250 mM in 2 × 80 ml or 0-300 mM in 2 × 100 ml of the same buffer per liter of *E. coli* culture) was applied to elute the protein. Typical yields were about 0.2 $mg \cdot l^{-1} \cdot OD_{550}{}^{-1}$.

Acknowledgments

We would like to thank H. Döbeli (Hoffman-La Roche, Basel) for kindly donating a sample of the nitrilo-triacetic acid column resin and J. Stadlmüller for carrying out the fluorescence measurements on the single chain antibody.

Received 10 September 1990; accepted 28 November 1990.

References

1. Skerra, A. and Plückthun, A. 1988. Assembly of a functional immunoglobulin F_v fragment in *Escherichia coli*. Science **240**:1038–1041.
2. Plückthun, A. and Skerra, A. 1989. Expression of functional antibody F_v and F_{ab} fragments in *E. coli*. Meth. Enzymol. **178**:497–515.
3. Ward, E. S., Güssow, D., Griffiths, A. D., Jones, P. T., and Winter, G. 1989. Binding activities of a repertoire of single immunoglobulin variable domains secreted from *Escherichia coli*. Nature **341**:544–546.
4. Huse, W. D., Sastry, L., Iverson, S. A., Kang, A. S., Alting-Mees, M., Burton, D. R., Benkovic, S. J. and Lerner R. A. 1989. Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science **246**:1275–1281.
5. Potter, M. 1977. Antigen-binding myeloma proteins of mice. Adv. Immunol. **25**:141–211.
6. Perlmutter, R. M., Crews, S. T., Douglas, R., Sorensen, G., Johnson, N., Nivera, N., Gearhart, P. J., and Hood, L. 1984. The generation of diversity in phosphorylcholine-binding antibodies. Adv. Immunol. **35**:1–37.
7. Leon, M. A. and Young, N. M. 1971. Specificity for phosphorylcholine of six murine myeloma proteins reactive with *Pneumococcus* C polysaccharide and β-lipoprotein. Biochemistry **10**:1424–1429.

8. Young, N. M. and Leon, M. A. 1977. The binding of analogs of phosphorylcholine by the murine myeloma proteins McPC603, MOPC167 and S107. Immunochemistry **14**:757–761.
9. Metzger, H., Chesebro, B., Hadler, N. M., Lee, J., and Otchin, N. 1971. Modification of immunoglobulin combining sites, p. 253–267. *In:* Progress in Immunology: Proceedings of the 1st Congress of Immunology. Amos, B. (Ed.). Academic Press, New York.
10. Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M., and Davies, D. R. 1974. The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin F_{ab} and the nature of the antigen binding site. Proc. Nat. Acad. Sci. USA **71**:4298–4302.
11. Satow, Y., Cohen, G. H., Padlan, E. A., and Davis, D. R. 1986. Phosphocholine binding immunoglobulin F_{ab} McPC603. J. Mol. Biol. **190**:593–604.
12. Glockshuber, R., Malia, M., Pfitzinger, I., and Plückthun, A. 1990. A comparison of strategies to stabilize immunoglobulin F_v fragments. Biochemistry **29**:1362–1367.
13. Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S. M., Lee, T., Pope, S. H., Riordan, G. S., and Whitlow, M. 1988. Single-chain antigen-binding proteins. Science **242**:423–426.
14. Huston, J. S., Levinson, D., Mudgett-Hunter, M., Tai, M. S., Novotny, J., Margolies, M. N., Ridge, R. J., Bruccoleri, R. E., Haber, E., Crea, R., and Oppermann, H. 1988. Protein engineering of antibody binding sites: Recovery of specific activity in an anti-digoxin single-chain F_v analogue produced in *Escherichia coli*. Proc. Nat. Acad. Sci. USA **85**:5879–5883.
15. Smith, M. C., Furman, T. C., and Pidgeon, C. 1987. Immobilized iminodiacetic acid metal peptide complexes. Identification of chelating peptide purification handles for recombinant proteins, Inorg. Chem. **26**:1965–1969.
16. Smith, M. C., Furman, T. C., Ingolia, T. D., and Pidgeon, C. 1988. Chelating peptide-immobilized metal ion affinity chromatography. J. Biol. Chem. **263**:7211–7215.
17. Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R., and Stüber, D. 1988. Genetic approach to facilitate purification of recombinant proteins with a novel metal chelate adsorbent. Bio/Technology **6**:1321–1325.
18. Hochuli, E., Döbeli, H., and Schacher, A. 1987. New metal chelate adsorbent selective for proteins and peptides containing neighbouring histidine residues. J. Chromatography **411**:177–184.
19. Ghrayeb, J., Kimura, H., Takahara, M., Hsiung, H., Masui, Y., and Inouye, M. 1984. Secretion cloning vectors in *Escherichia coli*. EMBO J. **3**:2437–2442.
20. Schoner, B. E., Belagaje, R. M., and Schoner, R. G. 1990. Enhanced translational efficiency with two-cistron expression system. Methods Enzymol. **185**:94–114.
21. Nakamura, K., Pirtle, R. M., Pirtle, I. L., Takeishi, K., and Inouye, M. 1980. Messenger ribonucleic acid of the lipoprotein of the *Escherichia coli* outer membrane. J. Biol. Chem. **255**:210–216.
22. Yanisch-Perron, C., Vieira, J., and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. Gene **33**:103–119.
23. Geisselsoder, J., Witney, F., and Yuckenberg, P. 1987. Efficient site-directed *in vitro* mutagenesis. Biotechniques **5**:786–791.
24. Kunkel, T. A., Roberts, J. D., and Zakour, R. A. 1987. Rapid and efficient site-specific mutagenesis without phenotypic selection. Methods Enzymol. **154**:367–382.
25. Sassenfeld, H. M. and Brewer, S. J. 1984. A polypeptide fusion designed for the purification of recombinant proteins. Bio/Technology **2**:76–81.
26. von Heijne, G. 1989. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature **341**:456–458.
27. Summers, R. G., Harris, C. R., and Knowles, J. R. 1989. A conservative amino acid substitution, arginine for lysine, abolishes export of a hybrid protein in *Escherichia coli*. J. Biol. Chem. **264**:20082–20088.
28. Sassenfeld, H. M. 1990. Engineering proteins for purification. Trends Biotechnol. **8**:88–93.
29. Ljungquist, C., Breitholtz, A., Brink-Nilsson, H., Moks, T., Uhlén, M., and Nilsson, B. 1989. Immobilization and affinity purification of recombinant proteins using histidine peptide fusions. Eur. J. Biochem. **186**:563–569.
30. Vieira, J. and Messing, J. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene **19**:259–268.
31. Vieira, J. and Messing, J. 1987. Production of single-stranded plasmid DNA. Methods Enzymol. **153**:3–11.
32. Takagi, H., Morinaga, Y., Tsuchiya, M., Ikemura, H., and Inouye, M. 1988. Control of folding of proteins secreted by a high expression secretion vector, pIN-III-*omp*A: a 16-fold increase in production of active subtilisin E in *Escherichia coli*. Bio/Technology **6**:948–950.
33. Stark, M. J. R. 1987. Multicopy expression vectors carrying the *lac* repressor gene for regulated high-level expression of genes in *Escherichia coli*. Gene **51**:255–267.
34. Maniatis, T., Fritsch, E. F., and Sambrook, J. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, New York.
35. Beck, E. and Zink, B. 1981. Nucleotide sequence and genome organisation of filamentous bacteriophages f1 and fd. Gene **16**:35–58.
36. Chesebro, B. and Metzger, H. 1972. Affinity labeling of a phosphorylcholine binding mouse myeloma protein. Biochemistry **11**:766–771.