Expression of tetanus toxin fragment C in yeast: gene synthesis is required to eliminate fortuitous polyadenylation sites in AT-rich DNA

Michael A.Romanos*, Andrew J.Makoff, Neil F.Fairweather, Katrina M.Beesley, Debbie E.Slater, Fred B.Rayment, Mike M.Payne[1] and Jeffrey J.Clare
Departments of Molecular Biology and [1]Protein Chemistry, Wellcome Biotech, Langley Court, Beckenham, Kent BR3 3BS, UK

Received January 11, 1991; Revised and Accepted March 1, 1991

ABSTRACT

Fragment C is a non-toxic 50kDa fragment of tetanus toxin which is a candidate subunit vaccine against tetanus. The AT-rich *Clostridium tetani* DNA encoding fragment C could not be expressed in *Saccharomyces cerevisiae* due to the presence of several fortuitous polyadenylation sites which gave rise to truncated mRNAs. The polyadenylation sites were eliminated by chemically synthesising the DNA with increased GC-content (from 29% to 47%). Synthesis of the entire gene (1400 base pairs) was necessary to generate full-length transcripts and for protein production in yeast. Using a *GAL1* promoter vector, fragment C was expressed to 2 – 3% of soluble cell protein. Fragment C could also be secreted using the α-factor leader peptide as a secretion signal. The protein was present at 5 – 10mg/l in the culture medium in two forms: a high molecular mass hyper-glycosylated protein (75 – 200kDa) and a core-glycosylated protein (65kDa). Intracellular fragment C was as effective in vaccinating mice against tetanus as authentic fragment C. The glycosylated material was inactive, though it was rendered fully active by de-glycosylation.

INTRODUCTION

Most current vaccines against tetanus are produced by formaldehyde-inactivation of tetanus toxin present in culture filtrates of *Clostridium tetani* (reviewed in 1). These toxoid preparations are only partially pure and, though extremely effective as vaccines, occasionally give rise to adverse reactions on hyper-immunisation. A subunit vaccine consisting of a non-toxic fragment of the toxin, expressed in a safe host organism, would eliminate the need to culture *C.tetani* or purify tetanus toxin. Fragment C, a 50kDa C-terminal papain-cleavage product of tetanus toxin (2), is both non-toxic and immunogenic and has therefore been suggested as a possible subunit vaccine (3).

We have previously reported the production in *Escherichia coli* of fragment C, its purification and immunogenicity in mice (4). Expression of fragment C in *E.coli* was shown to be limited by the unfavourable codon bias of the highly AT-rich *C.tetani* coding sequence (5). Using an almost completely synthetic gene containing codons favourable for high-level expression in *E.coli*, we were able to increase expression levels almost four-fold to 12 – 14% of cell protein (5). However, *E.coli* has a disadvantage as a host organism in that it contains toxic cell wall pyrogens and it would be preferable to use a non-toxic production organism such as the yeast *Saccharomyces cerevisiae*.

Here we describe the intracellular expression and secretion of fragment C in yeast, and the use of yeast-derived fragment C in immunising mice against tetanus. In the course of the work we identified a problem in obtaining full-length transcripts which was due to the presence of several fortuitous polyadenylation sites in the AT-rich *C.tetani* gene. These sites were eliminated in a synthetic gene with increased GC-content.

MATERIALS AND METHODS

Expression vectors

The construction of the *GAL7* expression vector pWYG7 (Fig.1) has been described previously (6); it contains a 260bp synthetic promoter with a *Bam* HI site at positions −15 to −20 relative to the RNA start site. Foreign genes are cloned between the *Bam* HI site and the *Bcl* I site upstream of the 2μ *FLP* transcriptional terminator using a synthetic linker that recreates the native *GAL7* RNA start site and untranslated leader. pWYG5 (Fig.1) is similar to pWYG7 but contains the *GAL1* promoter from pBM150 (7), and the *Bam* HI cloning site is *downstream* of the RNA start sites so that foreign genes are tailored to have *Bam* HI-compatible sites immediately upstream of their initiation codons. Four vectors were constructed, using each of the four versions of the fragment C gene (TET2, TET7, TET11, and TET15; Fig.2) which were described previously in plasmids pTETtac2, pTETtac7,

* To whom correspondence should be addressed

pTETtac11, and pTETtac115 (5), and based either on pWYG5 or pWYG7 (summarised in Table I). Note that the protein expressed from pWYG7-TET2 (Met-Val-fragment C) has a change in the second amino acid (Lys to Val), since an *Nco* I site (CCATGG) was placed at the initiating ATG to facilitate transfer to secretion vectors. pWYG5-TET7 and pWYG5-TET11 were constructed for rapid analysis and had upstream leaders derived from *E.coli* expression vectors, possibly giving rise to sub-optimally translated mRNAs. The α-factor prepro region containing a KEX2 cleavage site and two Glu-Ala spacers was used to direct secretion of fragment C. The DNA was synthesised with the natural sequence (8) except for a change of two nucleotides near the 3′ end to engineer a *Xho* I site, resulting in an amino acid change (Asp to Glu) preceeding the Lys-Arg. pWYG9-TET2 and pWYG59-TET15 (Table I) are similar to pWYG7-TET2 and pWYG5-TET15, respectively, but each has the α-factor prepro DNA sequence upstream of the fragment C gene.

Transformation, induction of yeast cells and protein analysis

Plasmids were introduced into a 2μ-free derivative of *S.cerevisiae* strain S150-2B (*a leu2 his3 ura3 trp1*; 9), cells were induced with galactose, and protein extracts made as described previously (6). Insoluble proteins (about 30% of total) could be removed by centrifugation at 10000g for 15min. Proteins present in culture supernatants were concentrated by ultrafiltration using Centricon 30 micro-concentrators (4000g, 45 min) or Amicon stirred cells with P30 membranes. For endoglycosidase H (endo H) digestions, 25μl of concentrate was mixed with 5μl of 0.2M NaH_2PO_4, 10mM 2-mercaptoethanol, 1% SDS, boiled for 5 min and cooled on ice, then incubated at 37°C for 18h with 9mU of endo H (Boehringer, Mannheim) in the presence of the same protease inhibitors used during breakage. Proteins were separated by electrophoresis and visualised by staining with Coomassie blue or Western blotting as described previously (4). Fragment C was quantified using a two-antibody sandwich ELISA (4). For N-terminal sequencing, Coomassie blue-stained polypeptides in SDS-polyacrylamide gels were electrophoretically transfered onto polyvinylidene difluoride Immobilon P membrane (Millipore), and analysed in an Applied Biosystems ABI 477A pulse liquid sequencer using the Normal-1 program.

RNA analysis

Total RNA was prepared from induced cells and analysed by Northern blotting as described previously (10). Probes were labelled using the random priming method (11). Fragment C-specific RNAs were detected by autoradiography and their size estimated by comparison to stained RNA size markers (Bethesda Research Labs, 0.16−1.77kb ladder, 5μg per track).

Purification of fragment C and immunisations

Fragment C present in cell extracts was purified by affinity chromatography using monoclonal antibody TTO8 (12) linked to cyanogen bromide-activated Sepharose 4B. Fragment C was eluted with 0.1M sodium citrate pH3.0 and neutralised by addition of one volume of 0.1M sodium phosphate pH7.0. Vaccines containing fragment C were prepared with 10% Alhydrogel (Trade Mark). Serial dilutions were made as in Table 2 and Balb/c mice were injected subcutaneously with 0.5ml. Mice were challenged four weeks later with approximately $100LD_{50}$ of tetanus toxin, and survivors after a further four days were counted.

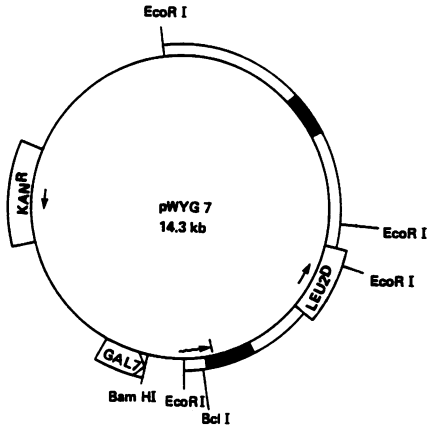

Fig. 1. Map of intracellular expression vector pWYG7. Bacterial plasmid (pMB9) DNA is in single lines and yeast 2μ plasmid DNA is boxed. Also indicated are the 2μ inverted repeats (shaded), kanamycin-resistance (*kan*^{r}) and *LEU2*^{d} markers, the *GAL7* promoter, and the FLP gene terminator (curly arrow). Foreign genes are inserted between the *Bam* HI and *Bcl* I sites. pWYG5 is the same but has the *GAL1* promoter in place of *GAL7*.

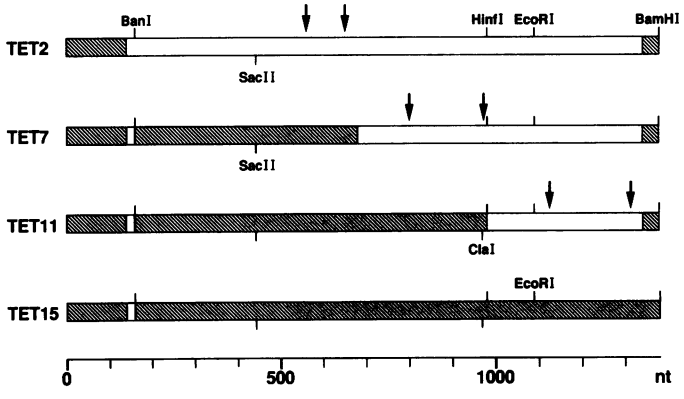

Fig. 2. Maps of native and synthetic fragment C genes. Coding regions for fragment C are boxed; regions that were chemically synthesised with codons optimal for translation in *E.coli* are hatched. The four versions of the gene, TET2, TET7, TET11 and TET15, had 12%, 50%, 73% and 99% synthetic DNA, respectively. Selected restriction sites are shown in TET2 and in the other genes where the DNA has been altered. The approximate positions of yeast polyadenylation sites found in the native sequence, estimated from the sizes of short transcripts in Northern blots (Fig.3), are indicated by arrows.

RESULTS

Expression vectors

Four versions of the fragment C gene had been constructed previously (5). The first, TET2, contained the largely unaltered *C.tetani* coding sequence for fragment C, adapted using synthetic oligonucleotides at its 5′ (161bp) and 3′ (42bp) ends. Three subsequent versions, TET7, TET11, and TET15, contained increasing amounts of synthetic DNA at the 5′ end (50%, 73%, and 99% synthetic, respectively; Fig.2). The synthetic DNA was altered to contain codons favoured in highly expressed *E.coli* genes and consequently had a much higher GC-content than the native *C.tetani* sequence. The codon usage in the synthetic DNA was also considerably more favourable than the original *C.tetani* DNA for expression in yeast (13). One indicator of this is the number of codons not found in highly-expressed yeast genes (13),

Table I. Summary of fragment C expression vectors

Vector[a]	Promoter	Sequence between *Bam* HI and initiating ATG	Encoded polypeptide
pWYG7-TET2	*GAL7*	*GAL7* leader[b]	Met-Val-fragment C[d]
pWYG5-TET7	*GAL1*	*E.coli* leader[c]	Met-fragment C
pWYG5-TET11	*GAL1*	*E.coli* leader[c]	Met-fragment C
pWYG5-TET15	*GAL1*	*P.pastoris AOX1*[e]	Met-fragment C
pWYG9-TET2	*GAL7*	*GAL7* leader[b]	prepro α-factor-Met-Val-fragment C[d]
pWYG59-TET15	*GAL1*	*AOX1*[e]	prepro α-factor-Met-Val-fragment C[d]

[a] Fragment C gene version (TET2, 7, 11 or 15) indicated.
[b] ACATGATAAAAAAAACAGTTGAATATTCCCTCAACC*ATG*G
[c] TAATCATCCACAGGAGACTTTCTG*ATG*
[d] Met-fragment C with second amino acid altered (Lys to Val) due to insertion of *Nco* I site (CCATGG) at initiating ATG.
[e] AAACG*ATG* derived from sequence upstream of *Pichia pastoris AOX1* gene (36).

which decreased progressively in the series TET2, TET7, TET11, and TET15 (290, 225, 193, and 143, respectively, out of 452). Perhaps more significant is the reduction in the number of very rare codons, for example in Ile ATA codons from 19 in TET2 to 0 in TET15. Expression of each of these genes was examined in the yeast intracellular expression vectors pWYG5 (*GAL1* promoter) or pWYG7 (*GAL7* promoter). We have found these two vectors to be very similar in their efficiency, as judged by their use in the expression of *E.coli* β-galactosidase (JJC, unpubl.). Secretion vectors based on pWYG5 (*GAL1* promoter) or pWYG7 (*GAL7* promoter) were constructed, using the α-factor leader peptide with Glu-Ala spacers, with either the TET2 or TET15 genes (Table I).

Transcription of fragment C genes in yeast

In initial experiments using the *S.cerevisiae* strain S150-2B, it was found that pWYG7-TET2, containing the predominantly native *C.tetani* gene, did not give rise to any fragment C detectable in Western blots. ELISA quantification gave an extremely low figure of 0.001% of cell protein. Similar results were obtained with different host strains including the protease-deficient strain 20B-12 (α *trp1 pep4-3*; 14). These results were surprising since high-level expression had been obtained in two other systems: *E.coli* and baculovirus (5; I.G.Charles *et al.*, in preparation). When RNA derived from pWYG7-TET2 was analysed by Northern blotting, two abundant fragment C-specific transcripts were detected, of 600 nucleotides (nt) and 700nt, instead of the expected full-length transcript of 1655nt (Fig.3).

Synthetic versions of the gene cloned in pWYG5 were next tested for transcription. pWYG5-TET7 was found to give rise to two transcripts of 900nt and 1100nt, while pWYG5-TET11 gave transcripts of 1200nt and 1400nt and also a small amount of the full-length transcript (Fig.3). Only the vector containing the almost fully synthetic gene, pWYG5-TET15, gave large amounts of the full-length transcript (1655nt) as the major species (Fig.3).

The probe used to detect the short transcripts of TET2, TET7, and TET11 was a 225bp 5′ terminal fragment of TET7 (Fig.3); a 3′ 280bp probe did not hybridise with the pWYG7-TET2 or pWYG5-TET7 transcripts (not shown). This is consistent with the transcripts all initiating in the *GAL7* or *GAL1* promoter with 3′ ends mapping at different positions in the unaltered *C.tetani* sequence. The approximate positions of the 3′ ends have been calculated from the sizes of the RNAs (Fig.2), taking into account the different lengths of untranslated leader in pWYG5 or pWYG7. We have found that all the short mRNAs described are found exclusively in the polyadenylated RNA fraction after oligo(dT)-cellulose chromatography (data not shown).

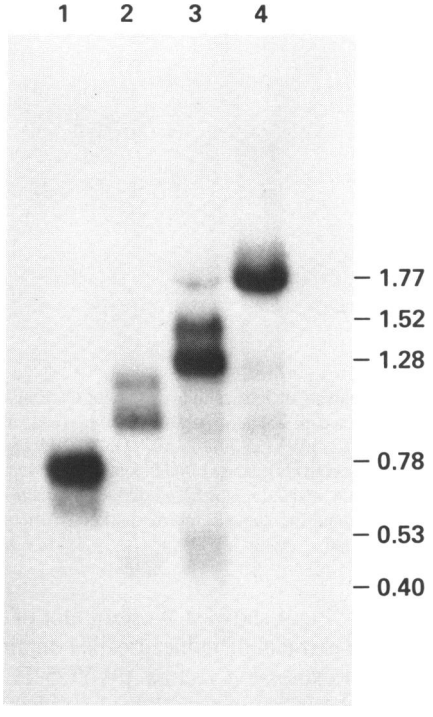

Fig. 3. Analysis of RNA derived from fragment C intracellular expression vectors. Total RNA (5μg) isolated from induced cultures containing each of the intracellular expression vectors was separated in a 1% agarose denaturing gel and analysed by Northern blotting. The probe used contained the 5′ 225bp of the TET7 gene. Lanes 1 to 4 contained RNA derived from pWYG7-TET2, pWYG5-TET7, pWYG5-TET11 and pWYG5-TET15, respectively. RNA size markers (positions indicated) were used to estimate the sizes of fragment C transcripts: lane 1 – 600 and 700nt; lane 2 – 900 and 1100nt; lane 3 – 1200, 1400nt and 1700nt (minor); lane 4 – 1700nt.

Intracellular expression of fragment C

As described above, fragment C could not be detected in induced cells containing pWYG7-TET2. Expression of fragment C from vectors containing the synthetic genes TET7, TET11, and TET15

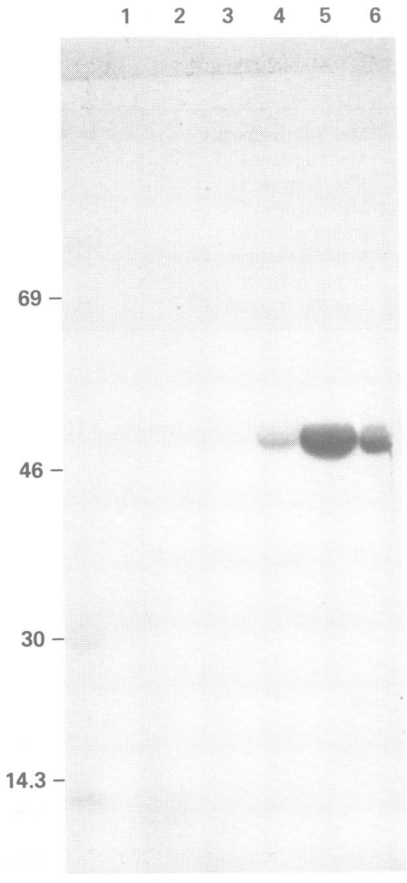

Fig. 4. Western blot analysis of intracellular fragment C. Fragment C-specific polypeptides were detected in 25μg of soluble proteins from induced cultures containing pWYG7-TET2, pWYG5-TET7, pWYG5-TET11 or pWYG-TET15 (lanes 2, 3, 4, and 5, respectively); lanes 1 and 6 contained proteins from induced cells containing no plasmid, and 0.25μg of *E.coli*-derived fragment C, respectively. The positions of pre-stained markers (Rainbow markers, Amersham) and their sizes in kDa are indicated.

was examined next. Fig.4 shows a Western blot of fragment C-specific proteins in extracts of induced cells containing each of the four intracellular expression vectors. pWYG5-TET7 gave rise to a very faint doublet migrating at 30kDa (not visible in the blot shown). The 30kDa polypeptides are presumably run-off translation products of the 900nt and 1100nt short transcripts. pWYG5-TET11 and pWYG5-TET15 gave full-length 50kDa fragment C though the levels were significantly higher with the latter, apparently paralleling the relative levels of full-length mRNA. The difference in untranslated leaders in the two vectors (Table I) could also have contributed to the difference in expression levels.

The levels of fragment C in 24h inductions of pWYG5-TET11 or pWYG5-TET15-transformed cells were quantified by ELISA and found to be 0.5% and 2.0–3.0% (60–90 mg/l) of soluble cell protein, respectively. In an induction time-course for pWYG5-TET15-transformed cells, levels reached a plateau after 8h (not shown). In all cases the product was fully soluble.

Secretion of fragment C

Secretion of fragment C polypeptides from cells harboring pWYG9-TET2 or pWYG59-TET15 was examined. When culture supernatants from induced pWYG9-TET2-transformed cells were

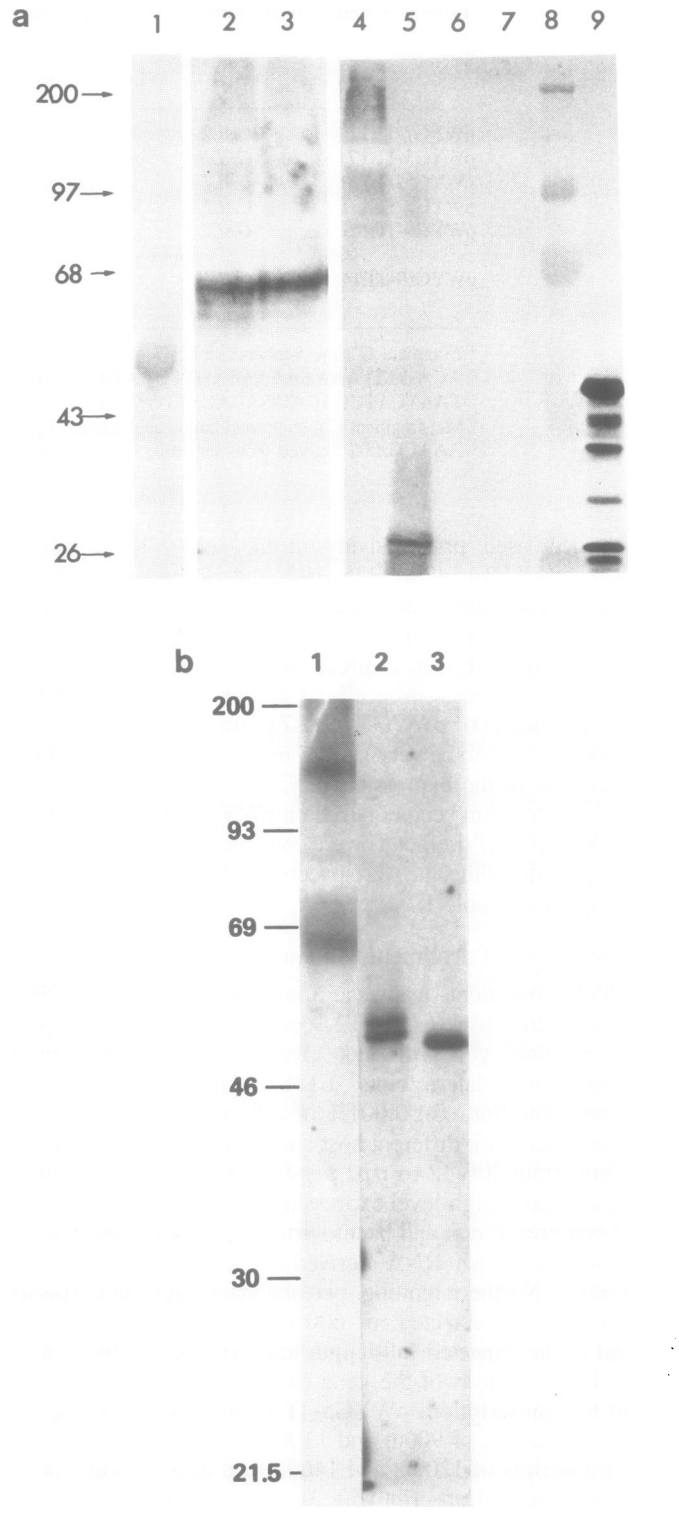

Fig. 5. SDS-polyacrylamide gel analysis of secreted fragment C polypeptides in concentrated culture supernatants. (a) Western blot analysis. Lane 1- endo H-treated culture supernatant from a pWYG59-TET15 transformant; lanes 2 and 3– untreated supernatants of two different pWYG59-TET15 transformants; lane 4–supernatant of pWYG9-TET2 transformant; lane 5–as lane 4 but endo H-treated; lanes 6 and 7–supernatants from untransformed cells; lane 8–size markers (Rainbow markers, Amersham; sizes indicated by arrows); lane 9–*E.coli*-derived fragment C. (b) Coomassie blue-stained gel. Lane 1–7μl of 26-fold concentrated culture supernatant of pWYG59-TET15 transformant; lane 2–as in lane 1 but treated with endo H; lane 3–1.25μg authentic fragment C. Size markers are indicated by arrows (Gibco BRL).

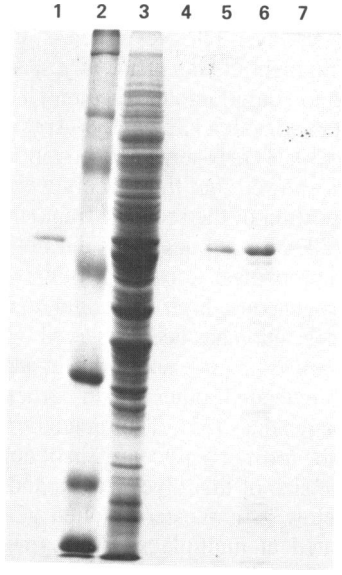

Fig. 6. Purification of intracellular fragment C by affinity chromatography. Fragment C present in soluble proteins from induced cells containing pWYG5-TET15 was bound to a TTO8 antibody column and eluted using 0.1M citrate buffer (see Materials and Methods). Fractions were analysed by SDS-gel electrophoresis and Coomassie blue staining. Lane 1 – 0.25μg *C.tetani* fragment C, lane 2 – pre-stained size markers (Rainbow markers, Amersham), lane 3 – 50μg soluble yeast proteins from a pWYG5-TET15 transformant containing fragment C, lanes 4 to 7-fragment C eluate fractions.

analysed in Western blots, a broad smear of reactive material was detected in the range 75 – 200kDa (Fig.5a). The molecular mass of this material was substantially reduced on treatment with endo H to a single major species of approximately 26kDa. The size of this polypeptide is consistent with it being a run-off translation product of short transcripts derived from the TET2 gene (Fig.2). Since the equivalent polypeptide was not detected with the intracellular expression vector pWYG7-TET2, the secreted form is presumably stabilised in the secretory pathway, perhaps by glycosylation.

Culture supernatants of induced cells containing pWYG59-TET15 contained two forms of fragment C detected by Western blotting and staining of SDS-polyacrylamide gels (Fig.5b): a diffuse band of very high molecular mass, and a major band migrating at 65kDa. A ladder of at least four fainter bands was also seen below the 65kDa band in Western blots, suggestive of glycosylation at multiple sites (Fig.5a). On treatment with endo H all the species reduced to two distinct bands of about 50kDa, near the size expected for correctly processed full-length fragment C (Fig.5b). N-terminal sequence analysis of these polypeptides indicated that the faster-migrating polypeptide was fully processed to Met-Val-fragment C (first 20 amino acids determined), whereas the slower-migrating polypeptide retained the N-terminal Glu-Ala-Glu-Ala spacer. No fragment C proteins were detected in cell extracts (Fig.5a), indicating that KEX2 processing and secretion were highly efficient.

Gel scanning of stained gels (Fig.5b) indicated that the fragment C products constituted the majority of proteins in the medium with about half hyper-glycosylated and half in the 65kDa form. The total amount was estimated to be 5 – 10mg/l fragment C, by gel scanning quantification of endo H-treated proteins compared to standards.

Table II. Protection of immunised mice against tetanus toxin challenge

	Survivors after challenge[a]				
Antigen (dose in μg)	2	0.5	0.125	0.03	0.008
E.coli fragment C	4[b]	2	0	0	0
Yeast intracellular	5	5	5	0	0
Yeast secreted	0	0	0	0	0
Yeast secreted, de-glycosylated	5	5	3	0	0
Saline	0				

[a] Groups of five mice were challenged with $100LD_{50}$ of tetanus toxin.
[b] Four mice in this group.

Immunisation of mice with yeast-derived fragment C

Yeast intracellular and secreted fragment C preparations were compared in their ability to immunise mice against tetanus. Four-fold serial dilutions of intracellular, secreted, and de-glycosylated secreted fragment C were tested. The intracellular fragment C was purified to a high level using immuno-affinity chromatography (Fig.6), while the secreted material was used without purification. The results (Table II) show that the intracellular fragment C was at least as effective as the *E.coli*-derived material, which has previously been shown to be equivalent to *C.tetani* fragment C (4). In contrast, the secreted product was totally inactive. However, de-glycosylation of secreted fragment C with endo H rendered it as protective as the intracellular product.

DISCUSSION

Truncated transcripts from AT-rich fragment C DNA

Production of fragment C in *S.cerevisiae* required the use of a synthetic gene, since the native *C.tetani* gene gave rise to short transcripts. We analysed transcripts derived from the native gene and a 5′ nested set of three versions containing progressively more synthetic DNA. The lengths of the transcripts produced increased with increasing amounts of synthetic DNA, so that the fully-synthetic gene gave full-length fragment C mRNA. Our data suggest the presence of at least six fortuitous polyadenylation sites in the *C.tetani* gene which were eliminated in the synthetic gene. The synthetic DNA, originally altered to have codons found in highly-expressed *E.coli* genes, had an increased GC-content: 47% for the fully-synthetic gene compared to 29% in the *C.tetani* sequence. While we have not precisely identified the sites in the native gene responsible for premature 3′ end formation, the increase in GC-content clearly correlates with the elimination of polyadenylation sites (which normally occur in AT-rich DNA; see below).

In higher eukaryotes mRNA 3′ end formation involves cleavage and polyadenylation of precursor mRNAs that extend at least several hundred nt beyond the coding region (reviewed in 15). In contrast, true termination, tightly coupled to polyadenylation, has been thought to occur near the 3′ ends of yeast genes (16 – 20). At least three different consensus sequences, found downstream of yeast genes, have been proposed as part of the mRNA terminator: the tripartite sequence TAG..(T-rich)..TA(T)GT..(AT-rich)..TTT (16); TTTTTATA (21); and TAAATAAA/T (22). Deletion analysis of the tripartite signal and TTTTTATA has confirmed their essential role in 3′ end formation (16, 21). The commonly found tripartite motif shows poor sequence conservation, and a study of the *CYC1* terminator

by linker mutagenesis suggests that large sequence alterations do not abolish 3′ end formation (19). Therefore a more general feature such as high AT-content may be critical. However this cannot be sufficient since not all AT-rich sequences contain terminators, and since terminators are frequently unidirectional (23) implying some kind of sequence specificity. Sequences downstream of the consensus sequence, near the mRNA 3′ end, are known to affect the efficiency and choice of site (17, 18).

More recent *in vitro* studies suggest that the mature 3′ ends of yeast mRNAs are formed by endonucleolytic processing of longer precursors followed by polyadenylation, and that sites previously designated as 'terminators' are involved in processing. However, such precursors have not been detected *in vivo* either because they are very short-lived or because termination and processing are tightly coupled (24, 25). The exact sequence of events in yeast mRNA 3′ end formation is still unclear and therefore yeast 'terminators' are perhaps more correctly denoted polyadenylation sites. Since the short fragment C transcripts described here are both polyadenylated and abundant, their 3′ ends are due to the presence of fortuitous polyadenylation sites and are probably formed by the normal mechanism for mRNA. Our finding of multiple short transcripts for each version of the fragment C gene would suggest that there are sites of varying efficiency. However, we cannot distinguish between the alternative possibilities of true premature termination or processing of a full-length precursor from our results.

Since sequence requirements for active 'terminators' are poorly defined (19), it is difficult to predict their presence or location in DNA of such high AT-content as the TET2 gene. At least a dozen sequences resembling a 'terminator' consensus, especially the tripartite motif, can be found in the TET2 sequence. However, several are present in the 5′ region known not to cause premature polyadenylation. A single TTTTTATA sequence occurs at nt506 to 513 and may account for the two short transcripts of TET2. However, given the number of 'terminator'-like sequences it would be difficult to assign particular 3′ ends to any of them, even if accurate 3′ mapping data were available, since the spacing between the consensus and the 3′ end is variable (16).

The problem encountered here with the TET2 gene may be a general one also occuring with other foreign AT-rich DNA or with DNA containing short stretches of AT-rich sequence. We are aware of failures to express several other AT-rich genes including various malarial cDNAs (JJC, unpublished; P.Ross-MacDonald and D.H.Williamson, pers. comm.) and genes of the *Kluyveromyces lactis* cytoplasmic plasmid k1 (26). If the problem is diagnosed, the simplest solution may be to re-synthesise the offending DNA with increased GC-content.

Intracellular expression and secretion of fragment C

We have used the *GAL1* promoter to express fragment C in yeast to levels of 2–3% of soluble protein (equivalent to 60–90 mg/l of culture in shake-flasks, or greater than 1g/l in high-density fermentations). Although the tight regulation of *GAL* promoters is desirable for proteins that show toxicity, the promoters are known to be limited by a shortage of transcriptional activators, including *GAL4* protein (27). However, use of the glucose-repressible *ADH2* promoter (28) did not give higher levels of fragment C (KMB and MAR, unpubl.). We have been able to express fragment C in another yeast, *Pichia pastoris*, to levels of over 25% of cell protein (JJC, FRB and MAR, in preparation), using the powerful, tightly-regulated *AOX1* promoter (29).

Tetanus toxin is not naturally secreted but is released from *C.tetani* following cell lysis (1). Nevertheless, using the α-factor leader peptide, fragment C was efficiently secreted from yeast cells to levels of 5 to 10mg/l into the medium. The leader peptide that we used contained both a KEX2 (Lys-Arg) cleavage site and Glu-Ala spacers (30, 31). N-terminal sequence analysis of the secreted products showed that the spacers were still present in a substantial proportion of the secreted fragment C, as found in other examples (32, 33), although this did not appear to affect its immunological properties. About 50% of the secreted fragment C was in a heterogeneous, high molecular mass, glycosylated form. Hyper-glycosylation has been observed with other proteins secreted from *S.cerevisiae* (34), and is due to addition to the core carbohydrate of extended outer chain structures containing multiple mannose residues (35). The remainder of the fragment C was mainly in the form of a glycoprotein of apparent molecular mass 65kDa. The size of this glycoprotein and the faint ladder of bands seen below it in Western blotted gels suggest that it is core-glycosylated at multiple sites; fragment C has seven potential sites for Asn-linked glycosylation.

Yeast-derived fragment C as a tetanus vaccine

Intracellular fragment C produced from *S.cerevisiae* has been shown to be as effective as *E.coli* or *C.tetani*-derived fragment C in immunising mice against tetanus toxin. In contrast the secreted, glycosylated fragment C was totally inactive as a vaccine. However the de-glycosylated protein appeared to be as active as intracellularly expressed fragment C, suggesting that the carbohydrate must sterically hinder the generation of neutralising antibodies to adjacent epitopes. The alternative possibility, that the glycosylation sites are within epitopes, is unlikely, since after de-glycosylation with endo H the remaining N-acetylglucosamine residues would probably still disrupt the epitopes. As it should be possible to secrete fragment C to levels in excess of 100mg/l in high-density fermentations, the de-glycosylated secreted product may provide a feasible production alternative to the intracellular protein.

ACKNOWLEDGEMENTS

We thank Stuart Ballantine for ELISA assays, John Bester for carrying out the mouse challenge experiments, and Carol Scorer for helpful discussions.

REFERENCES

1. Bizzini,B. (1984) In Germanier,R. (ed), Bacterial Vaccines, Academic Press, pp.37–67.
2. Helting,T.B. and Zwisler,O. (1977) *J. Biol. Chem.*, **252**, 187–193.
3. Helting,T.B. and Nau,H.H. (1984) *Acta Pathol. Microbiol. Scand. Sect. C* **92**, 59–63.
4. Makoff,A.J., Ballantine,S.P., Smallwood,A.E. and Fairweather,N.F. (1989) *Bio/Technology* **7**, 1043–1046.
5. Makoff,A.J., Oxer,M.D., Romanos,M.A., Fairweather,N.F. and Ballantine,S. (1989) *Nucleic Acids Res.* **17**, 10191–10202.
6. Beesley,K.M., Francis,M.J., Clarke,B.E., Beesley,J.E., Dopping-Hepenstal,P.J.C., Clare,J.J., Brown,F. and Romanos,M.A. (1990) *Bio/Technology* **8**, 644–649.
7. Johnston,M. and Davis,R.W. (1984) *Mol. Cell. Biol.* **4**, 1440–1448.
8. Kurjan,J. and Herskowitz,I. (1982) *Cell* **30**, 933–948.
9. McCleod,M., Volkert,F. and Broach,J. (1984) *Cold Spring Harbor Symp. Quant. Biol.* **49**, 779–787.
10. Romanos,M.A. and Boyd,A. (1988) *Nucleic Acids Res.* **16**, 7333–7350.
11. Feinberg,A. and Vogelstein,B. (1983) *Anal. Biochem.* **132**, 6–12.

12. Sheppard,A.J., Cussell,D. and Hughes,M. (1984) *Infect. Immun.* **43**, 710–714.
13. Bennetzen,J.L. and Hall,B.D. (1982) *J. Biol. Chem.* **257**, 3026–3031.
14. Hemmings,B.A., Zubenko,G.S., Hasilik,A. and Jones,E.W. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 435–439.
15. Platt,T. (1986) *Ann. Rev. Biochem.* **55**, 339–372.
16. Zaret,K.S. and Sherman,F. (1982) *Cell* **28**, 563–573.
17. Zaret,K.S. and Sherman,F. (1984) *J. Mol. Biol.* **176**, 107–135.
18. Henikoff,S. and Cohen,E.H. (1984) *Mol. Cell. Biol.* **4**, 1515–1520.
19. Osborne,B.I. and Guarente, L.(1989) *Proc. Natl. Acad. Sci. USA* **86**, 4097–4101.
20. Russo,P. and Sherman,F. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 8348–8352.
21. Henikoff,S., Kelley,J.D. and Cohen,E.H. (1983) *Cell* **33**, 607–614.
22. Bennetzen,J.L. and Hall,B.D. (1982) *J. Biol. Chem.* **257**, 3018–3025.
23. Murray,J.A.H. and Cesarini,G. (1986) *EMBO J.* **5**, 3391–3399.
24. Butler,J.S. and Platt,T. (1988) *Science* **242**, 1270–1274.
25. Butler,J.S., Sadhale,P.P. and Platt,T. (1990) *Mol. Cell. Biol.* **10**, 2599–2605.
26. Stark,M.J.R., Boyd,A., Mileham,A.J. and Romanos,M.A. (1990) *Yeast* **6**, 1–29.
27. Baker,S.M., Johnston,S.A., Hopper,J.E. and Jaehning,J.A. (1987) *Mol. Gen. Genet.* **208**, 127–134.
28. Russell,D.W., Smith,M., Williamson,V. and Young,E.T. (1983) *J. Biol. Chem.* **258**, 2674–2682.
29. Cregg,J.M., Tschopp,J.F., Stillman,C., Siegel,R., Akong,M., Craig,W.S., Buckholz,R.G., Madden,K.R., Kellaris,P.A., Davis,G.R., Smiley,B.L., Cruze,J., Torregrossa,R., Velicelebi,G. and Thill,G.P. (1987) *Bio/Technology* **5**,479–485.
30. Julius,D., Blair,L., Brake,A., Sprague,G. and Thorner,J. (1983) *Cell* **32**, 839–852.
31. Julius,D., Brake,A., Blair,L., Kunisawa,R. and Thorner,J. (1984) *Cell* **37**, 1075–1089
32. Brake,A., Merryweather,J.P., Coit,D.G., Heberlein,U.A., Masiarz,F.R., Mullerbach,G.T., Urdea,M.S., Valenzuela,P. and Barr,P.J. (1984) *Proc. Natl. Acad. Sci. USA* **81**, 4642–4646.
33. Zsebo,K.M., Lu,H.S., Fieschko,J.C., Goldstein,L., Davis,J., Duker,K., Suggs,S.V., Lai,P.-H. and Bitter,G.A. (1986) *J. Biol. Chem.* **261**, 5858–5865.
34. Innis,M.A.(1989) In Barr,P.J., Brake,A.J. and Valenzuela,P. (eds.), Yeast Genetic Engineering, Butterworths, Stoneham, MA.
35. Ballou,C.E. (1982) In Strathern,J.N., Jones,E.W. and Broach,J.R. (eds.) The Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression, Cold Spring Harbor Laboratory, New York.
36. Ellis,S.B., Brust,P.F., Koutz,P.J., Waters,A.F., Harpold,M.M. and Gingeras,T.R. (1985) *Mol. Cell. Biol.* **5**, 1111–1121.