Cloning and expression of a protective antigen from the cattle tick *Boophilus microplus*

(vaccine development/cDNA/epidermal growth factor-like repeats/membrane glycoprotein)

Keith N. Rand*, Terry Moore*, Alagacone Sriskantha*, Kevin Spring*, Ross Tellam[†], Peter Willadsen[†], and Gary S. Cobon*[‡]

*Biotechnology Australia Pty. Ltd., P.O. Box 20, Roseville, New South Wales, 2069, Australia; and [†]CSIRO Division of Tropical Animal Production, PMB3, Indooroopilly, Queensland, 4068, Australia

Communicated by Rutherford Robertson, August 18, 1989 (received for review June 26, 1989)

ABSTRACT Glycoproteins located on the luminal surface of the plasma membrane of tick gut epithelial cells, when used to vaccinate cattle, are capable of stimulating an immune response that protects cattle against subsequent tick infestation. One such tick gut glycoprotein, designated Bm86, has been purified to homogeneity and the amino acid sequences of peptide fragments generated by endoproteinase Lys-C digestion have been determined. We report here the isolation and characterization of a cDNA that encodes Bm86. The nucleotide sequence of the cDNA contains a 1982-base-pair open reading frame and predicts that Bm86 contains 650 amino acids including a 19-amino acid signal sequence and a 23-amino acid hydrophobic region adjacent to the carboxyl terminus. The main feature of the deduced protein sequence is the repeated pattern of 6 cysteine residues, suggesting the presence of several epidermal growth factor-like domains. A fusion protein consisting of 599 amino acids of Bm86 and 651 amino acids of β -galactosidase was expressed in *Escherichia coli* as inclusion bodies. Ticks engorging on cattle vaccinated with these inclusion bodies were significantly damaged as a result of the immune response against the cloned antigen.

The tick *Boophilus microplus* is a major ectoparasite of cattle in many parts of the world. A single female cattle tick takes up as much as 1.5 ml of bovine blood, increasing its body weight to ≈ 250 mg. It has been estimated that cattle in tropical areas of Australia may become infested with 1000 tick larvae per day, resulting in greatly reduced productivity. In addition, *B. microplus* is the vector of hematoprotozoal parasites such as *Babesia bovis*. Chemicals have been used extensively to control ticks and have been partially successful, but this approach suffers from certain drawbacks such as environmental and residue problems, the high incidence of acaricide resistance that has developed in tick populations in the field, the need for frequent administration, and high cost.

Recently it was shown that cattle immunized against a membrane-bound glycoprotein (Bm86) purified from cattle ticks are highly resistant to parasite challenge (1). A vaccine based on Bm86 would be a very attractive alternative to acaricide treatment and would overcome most of the difficulties associated with the use of chemicals. Available evidence suggests that immunized cattle produce antibodies that recognize Bm86 present on the surface of tick gut digest cells (1, 2). After the ticks ingest blood, these antibodies, together with other components of the immune system such as complement, cause lysis of the gut epithelial cells, leading either to tick death or to damaged ticks exhibiting reduced growth and egg-laying ability. The effects are very striking. As many as 90% of ticks failed to survive to adulthood on cattle

vaccinated with three doses of only 2 μ g of Bm86 (1). However, it was necessary to begin with 40,000–60,000 ticks in order to purify 20–100 μ g of Bm86. In order to produce the larger quantities of Bm86 that are needed to develop a commercial vaccine, we have cloned and expressed the Bm86 gene.[§] Recombinant Bm86, even in the form of a fusion protein produced in *Escherichia coli* as inclusion bodies, is capable of inducing a substantial degree of protection against tick infestation.[¶]

MATERIALS AND METHODS

Construction and Screening of cDNA Library. RNA was extracted (3) from adult *B. microplus* (picked from cattle 15 days after infestation), cDNA was synthesized from 4 μ g of poly(A)⁺ RNA (4), and cDNA fragments larger than 800 base pairs (bp) were ligated to λ gt11 (5) to generate a library of 8 × 10⁵ recombinant clones. Oligodeoxynucleotide probes (Table 1) were based on the sequences derived from peptides generated by endoproteinase Lys-C digestion of Bm86 (1).

Three nitrocellulose filters were prepared from five 150-mm plates each containing 10^5 plaques. After prehybridization in 0.6 M sodium pyrophosphate/0.005% heparin (Sigma) at 40°C for 4 hr, hybridization was carried out for 16 hr at 40°C in the same solution with the radioactive oligonucleotide probe. Two of the filters were hybridized with the 63-mer probe while the third was hybridized with a mixture of the 50-, 51-, and 72-mer probes. The filters were washed in 0.3 M NaCl/0.03 M sodium citrate, pH 7.5/0.1% SDS at 45°C and positive plaques were identified by autoradiography. Other molecular biology techniques were carried out basically as described (6).

DNA Sequencing. The 3.9-kilobase-pair (kb) EcoRI fragment in the λ gt11 clone was self-ligated to form circles and concatemers, then sonicated. Fragments were end-repaired and then size-fractionated by electrophoresis in a low-melting-temperature agarose gel. Fragments in the size range 200–500 bp were cloned into the phage M13mp18 and sequenced (7). The sequence was compiled from all of the subsequences and analyzed with the programs GEL and SEQ (IntelliGenetics). Some DNA sequence was also obtained by the exonuclease III/nuclease S1 method (8).

Construction of Expression Vector pBTA708. A Bm86 gene fragment was prepared by first inserting a 2058-bp Xmn I restriction fragment (positions 120–2178) into the Sma I site of vector M13um31 (International Biotechnologies). After digestion with Sac I and Pst I, a fragment encoding most of Bm86 was ligated into pBTA224 [derived from pUR292 (9) by

Abbreviations: EGF, epidermal growth factor; IPTG, isopropyl β -D-thiogalactopyranoside.

[‡]To whom reprint requests should be addressed.

[§]The sequence reported in this paper has been deposited in the GenBank data base (accession no. M29321).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" deposition accordance with 18 U.S.C. \$1734 solely to indicate this fact.

[¶]E. coli BTA1751, which contains the entire Bm86 gene, has been deposited in the American Type Culture Collection (ATCC designation 67548).

CTT

Table 1.	Oligodeoxynucleotide probes				
Probe	Sequence $(5' \rightarrow 3')$				
50-mer	TTACCAATGGATGTACAAATAGCTTCAAGGACACCATCTTCGTACCACTT				
51-mer	CTTCGACGGATTGGATTCGACGCATCTGCCATAGCTACATTCCCTCGTCTT				
72-mer	ΤΤΤΑGGTACAACCTCACATTCAGCATTCCTACAAAATTCATTACCGAAATCAAAACAAATACTACTCTC				
63-mer	CTTGCAATGGATTCCATCCTCGGCGACAGTGAAAGCTCTAGGGCAAGTGCACTCATAAGCCTT				

eliminating the *Eco*RI site that lies outside of the β -galactosidase-coding region] to give pBTA708, which was used to transform *E. coli* JM101 (10) to give BTA1752.

Expression Tests. BTA1752 was grown in tryptone soybroth (TSB) containing ampicillin (50 μ g/ml) overnight, diluted 1:10 in fresh TSB, and induced with 1 mM isopropyl β -D-thiogalactopyranoside (IPTG) for 30 hr. The cells were resuspended in 50 mM Tris, pH 7.5/1 mM EDTA/1 mM phenylmethylsulfonyl fluoride and disrupted by sonication. Crude inclusion bodies were solubilized in 2.7% SDS/0.8 M urea/1.6% 2mercaptoethanol), resolved in a SDS/8% polyacrylamide gel, and transferred electrophoretically to nitrocellulose (Schleicher & Schuell). The membrane was blocked overnight with 0.1% gelatin in 10 mM borate buffer (pH 8.0), washed three times in 10 mM Tris, pH 8.0/0.1 M NaCl/0.1% Tween 20 (TST) and incubated for 1 hr at room temperature in TST containing a 1:500 dilution of serum from cattle 32 and 34, which had been vaccinated with native Bm86 isolated from ticks (1). The membrane was washed three times with TST, incubated for 5 min with a 1:500 dilution in TST of a rabbit anti-bovine antibody conjugated to horseradish peroxidase (Dakopatts, Copenhagen), washed three times with TST, and incubated in a solution of chloronaphthol (0.5 mg/ml) and H_2O_2 (0.00025%) in 10 mM Tris (pH 7.5) until color developed.

 $[^{35}S]$ Cysteine Labeling. A culture of BTA1752 was grown overnight in minimal medium (11) plus ampicillin (50 µg/ml) and then diluted 1:20 in TSB and incubated until the optical density at 600 nm reached 0.5. Proteins were induced and labeled by addition of IPTG and $[^{35}S]$ cysteine to 1 mM and 2 µCi/ml, respectively (1 µCi = 37 kBq). Six hours after induction, labeled and unlabeled samples were removed and analyzed by SDS/PAGE. The gel was soaked in Amplify (Amersham), dried, and autoradiographed.

Cattle Vaccination and Challenge Trials. Vaccination and challenge trials were carried out essentially as described (1). In brief, groups of three cattle received three vaccinations at 4-week intervals, the first two intramuscularly in Freund's complete adjuvant, the third subcutaneously in 0.9% NaCl.

RESULTS AND DISCUSSION

Isolation of cDNA Clones. One λ gt11 plaque that hybridized with both the 63-mer and the mixture of 50-, 51-, and 72-mer probes was isolated from $\approx 5 \times 10^5$ recombinants. The clone was found to contain an insert of 5.6 kb, which could be digested with *Eco*RI to give 3.9-kb, 1.5-kb, and 0.2-kb fragments. All four probes hybridized to the 3.9-kb fragment and not to the others. The sequence (Fig. 1) of nucleotides 1-2235 of the 3.9-kb *Eco*RI fragment encodes Bm86. All of the peptide sequences used for designing probes (underlined) as well as the other sequences determined from the endoproteinase Lys-C peptide fragments (1) can be identified. The sequence Set-Gly-Ser at amino acid positions 235-237 is different from that determined by peptide sequencing (Arg-Ala-Phe). The sequence of a second cDNA clone suggests that this is due to polymorphism within Bm86.

It is likely that amino acid 20 is the amino-terminal residue of the mature Bm86 protein. (i) The amino-terminal 19-amino acid segment is hydrophobic, which is a characteristic of leader sequences (12). (ii) The amino acids at positions -1and -3 of the putative cleavage site are small and nonpolar, which is also characteristic of the amino acids in these positions of many signal peptides (12). (iii) Although peptides were generated from Bm86 by endoproteinase Lys-C digestion, only one of the 12 peptide sequences obtained, that starting at amino acid 20, is not preceded by a lysine.

A polyadenylylation sequence (AATAAA) is present at position 2203 and 17 nucleotides downstream is a stretch of 10 adenines. As the poly(A) stretch was short for a poly(A) tail and there was DNA downstream from this sequence, it was not clear whether this was the 3' end of the mature mRNA. Another Bm86 cDNA clone was isolated from a different cDNA library and was found to terminate in more than 50 adenines from position 2226. This suggests that an unrelated cDNA fragment was joined to the original Bm86 cDNA fragment at some stage in the construction of the library.

The carboxyl-terminal 23-amino acid region (amino acids 628–650) is very hydrophobic and resembles segments in other membrane proteins that are removed and then replaced by a glycosyl-phosphatidylinositol anchor (13). If this is the case, Bm86 has no cytoplasmic domain. The 27 amino acids that precede the hydrophobic sequence (amino acids 601–627) contain clusters of serine and threonine residues. This is a characteristic of several membrane-bound proteins, such as the low density lipoprotein receptor, which may have O-linked carbohydrate chains in such regions (14). There are also five sites of possible N-linked carbohydrate addition (Asn-Xaa-Ser/Thr; Fig. 1).

Comparison of the Bm86 amino acid sequence with sequences in the National Biomedical Research Foundation (January 1989) and Kyoto (July 1988) data banks revealed many similarities between regions of Bm86 and the epidermal growth factor (EGF) precursor and other proteins containing EGF-like repeats. The similarities are mainly due to the conservation of the 6 cysteine residues within the EGF-like region. Several regions of Bm86 (Fig. 2) fall into the pattern Cys-Xaa₄₋₈-Cys-Xaa₃₋₆-Cys-Xaa₈₋₁₁-Cys-Xaa₀₋₁-Cys-Xaa₅₋₁₅-Cys, where Xaa is any amino acid except cysteine. The pattern in EGF is Cys-Xaa₇-Cys-Xaa₅-Cys-Xaa₁₀-Cys-Xaa₁-Cys-Xaa₈-Cys. In EGF, 5 amino acids precede the first cysteine and there are 11 after the sixth, but we show only the sequence between the cysteine residues in Fig. 2 because it cannot be predicted where the regions begin and end in the Bm86 sequence. There are no obvious proteolytic cleavage sites that would release a single EGF repeat from Bm86 in a manner similar to the release of EGF from its precursor.

Several other extracellular proteins have been found to contain EGF-like regions (reviewed in ref. 15). These fall into two general categories: those involved in blood coagulation and complement cascades and those associated with the regulation of cell growth. Bm86 clearly resembles the latter group, which is characterized by multiple EGF repeats, transmembrane or carboxyl-terminal hydrophobic regions, and location on the extracellular surface. The function of Bm86 is not known but, by analogy with the other members of this group, one possibility is that Bm86 is a cell membranebound ligand transmitting positional or cell-type information to adjacent cells and perhaps even influencing the cell lineage of those adjacent cells in a manner similar to the function of the *Drosophila* Notch protein (16).

There are some striking internal sequence homologies at the carboxyl terminus of Bm86, which includes the EGF-like region H (Fig. 2). There is sufficient sequence similarity to

	1 CCGCGACAGCTGCGGTGGTTCGACGCAGTGAG	32				
Mer Ang Giy Ile Ala Leu Phe Val Ala Ala Val Ser Lea Ile Val Giu Giy The Ala Giu	Ser Ser Le Cys Ser Amp Pho Gly Ann Glu	30				
Ang Coti GGC Anc GCT TIG TIC GTC GCC GCC GTT TCA CTG ATT GTA GAGGGC ACA <u>GCA GAA</u>	TCA TCC ATT TGC TCT GAC TTC GGG AAC GAG	122				
Pho Cys Ang Ang Ang Ang Chu Cys Chu Vai Vai Pro Chy Ang Chu Ang Ang Pho Vai Cys Lys	Cys Pro Ang Asp Ann Mot Tyr Pho Ann Ala	60				
TTC TOT CGC AAC GCT GAA TGT GAA GTG GTG CCT GGT GCA GAG GAT GAT TTC GTG TGC AAA	TGT CCG CGA GAT AAT ATG TAC TTC AAT GCT	212				
Ale Giu Lys Gin Cys Giu Tyr Lys Ann Thr Cys Lys Thr Ang Giu Cys Sor Tyr Giy Ang GCT GAA AAG CAA TGC GAA TAT AAA GAC ACG TGC <u>AAG ACA AGG GAG TGC AGC TAT GGA CGT</u> .	Cys Val Glu Ser <u>Ann Pro Ser</u> Lys Ala Ser TGC GTT GAA AGT AAC CCG AGC AAG GCT AGC	90 302				
Cys Vai Cys Giu Ala Ser And And Lou Thr Lou Gin Cys Lys Be Lys Ann And Tar An	The Amp Cys Arg Ama Arg Giy Giy The Ala	120				
Toc gitc toc gaa gca tog gac gat cta agg ctar caa toc and att and aat gac tac gca	ACT GAC TOC CGA AAT CGA GGT GGC ACT GCT	392				
Lys Lea Arg Thr Ang Gly Pho Bo Gly Ala Thr Cys Ang Cys Gly Gla Thr Gly Ala Met AAG TTG CGC ACG GAT GGG TTT ATT GGC GCA ACG TGT GAC TGT GGT GAA TGG GGT GCG ATG	Ann. Met Thr Thr Arg Ann Cýs Val Pro Thr AAC ATG ACC ACC CGG AAC TGT GTC CCT ACC	150 482				
The Cys Lea Arg Pro Asp Lea The Cys Lys Asp Lea Cys Gia Lys Asn Lea Lea Gin Arg acg tot ctt cot cot cac and acc toc and gac ctc toc gag and aac ctg ctt can agg	Asp Ser Ang Cys Cys Gin Giy Trp Ann Thr GAT TCT COT TOT TOC CAG GOG TOG AAC ACA	180 572				
Ale And Cys See Ale Ale Pro Pro Ale And See Tyr Cys See Pro Gy See Pro Lys Gy GCAAAC TGT TCA GCC GCT CCT CCA GCT GAC TCC TAT TGC TCT GCGGAGC CCC AAA GGA	Pro Ano Giy Gin Cys lie Ann Ain Cys Lys CCG GAC GGA CAG TGT ATA AAT GCT TGC AAG	210 662				
The Lys Giu Ala Giy Pho Vai Cys Lys His Giy Cys Ang See The Giy Lys Ala Tye Gin	Cys The Cys Pro See Gly See The Val Ala	240				
ACG AAA GAA GCT GGG TTT GTC TOC AAG CAT GGA TOC AGG TOG ACC GGC <u>AAG GCG TAC GAG</u>	TGC ACG TGC CCG AGT GGC TCT ACC GTC GCC	752				
Giu And Giy Be Thr Cys Lys Ser Be Ser His Thr Val Ser Cys Thr Ala Giu Gha Lys	Gha Thr Cys Arg Pro Thr Glu Anp Cys Arg	270				
GAA GAT GGC ATT ACC TIC AAA AGT ATT TCG CAC ACA GTC AGC TIC ACT GCT GAG CAA AAA	CAG ACC TGC CGC CCA ACC GAA GAC TGT CGT	842				
Val His Lys Giy Thr Val Lou Cys Giu Cys Pro Trp Ann Gin His Lou Val Giy Anp Thr	Cys II: Ser Asp Cys Val Asp Lys Lys Cys	300				
GTG CAC AAA GGA ACT GTG TGT GTG TGT GAG TGC CCG TGG AAT CAA CAT CTA GTG GGGGAC ACG	TGC ATA AGT GAT TGC GTC GAC AAG AAA TGC	932				
His Giu Giu Pho Met Asp Cys Giy Vai Tyr Met Asn Ang Gin Ser Cys Tyr Cys Pro Typ	Lys Ser Ang Lys Pro Gly Pro Ann Val Ann	330				
CAC GAA GAA TTT ATG GAC TGT GGC GTA TAT ATG AAT OGA CAA AGC TGC TAT TGT CCA TGG	AAA TCA AGGAAG CCG GGC CCA AAT GTC AAC	1022				
$ \begin{tabular}{ll} Ibc Ann Giu Cys Lou Lou Ann Giu Tyr Tyr Tyr Thr Val Ser Pho Thr Pro Ann Be Ser ATC AAT GAA TGC CTA CTG AAT GAG TAT TAC TAC ACG GTG TCA TTC ACC CCA AAC ATA TCT ACC ACG GTG TCA TTC ACC CCA AAC ATA TCT ACC ACG GTG TCA TTC ACC CCA AAC ATA TCT ACC ACG GTG TCA TTC ACC CCA AAC ATA TCT ACC ACG GTG TCA TTC ACC ACG CTG TCA TTC ACC ACG ACG ATA TCT ACC ACG GTG TCA TTC ACC ACG ACG ATA TCT ACC ACG GTG TCA TTC ACC ACG ACG ATA TCT ACC ACG GTG TCA TTC ACC ACG ACG ATA TCT ACC ACG ACG ACG ACG ACG ACG ACG ACG ACG$	Phe Asp Ser Asp His Cys Los Trp Tyr Glu TIT GAT TCT GAC CAT TGC <u>AAA TGG TAT GAG</u>	360 1112				
ARP AND Val Lou Giu Ala Be Ang The See Be Giy Lys Giu Val Pho Lys Val Giu Be	Less <u>Ann Cys Thr</u> Gin Anp lie Lys Ala Ang	390				
GAT COT GTT TTG GAA GCG ATA CGG ACC AGT ATC GGA AAA GAA GTT TTT AAG GTT GAG ATA	CTT AAC TGC ACG CAG GAC ATT AAG GCA AGA	1202				
Lea Ile Ala Giu Lys Pro Lea: Ser Lys His Val Lea Ang Lys Lea Gel Ala Cys Giu His	Pro lie Gly Glu Trp Cys Met Met Tyr Pro	420				
CTC ATA GCA GAG AMA CCA CTG TCA AMA CAC GTG CTC AGG AMA CTA CAAGCA TGC GAG CAT	CCA ATC GOCGAA TGG TGC ATG ATG TAT CCG	1292				
Lys Lea Lea Lea Lys Ann Ser Ain ¹ Thr Gia Lie Gia Gia Gia Ann Lea Cys Ann Ser	Lea Lea Lys Asp Gin Gin Ala Ala Tyr Lys	450				
Ang thg chg anc ang ana anc tet goa aca a gaa ang ang ang ang cet tig gaa agt	CTG CTC AAG GAT CAG GAA GCT GCC TAC AAA	1382				
Gly Gin Are Lys Cys Val Lys Val Are Are Lon Pho Tro Pho Gin Cys Ala Are Gly Tyr	The The The Tyr Glu Met The Arg Gly Arg	480				
GGT CAA AAC AAA TOC GTC AAG GTC GAC AAC CTC TTC TGG TTC CAG TGC GCT GAT GGT TAC	ACA ACA ACT TAC GAG ATG ACA CGA GGT CGC	1472				
Leu Ang Ang Ser Vai Cys Lys Ala Giy Vai	Ala Ang Lys Gily Gin Les Phe Val Tyr Giu	510				
CTA COC CGC TCC GTG TGT AAA GCT GGA GTT TCT TGC AAC GAA AAC GAGCAG TCG GAG TGT	GCT GAC AAA GGG CAA ATA TTT GTT TAC GAA	1562				
Ann Giy Lys Ala Ann Cys Gin Cys Pro Pro Anp The Lys Pro Giy Gin Be Giy Cys Be	Gha Arg Thr Thr Cys Ann Pro Lys Gha Bo	540				
Anc GGC Ann GCG Ant TGC CAN TGC CCA CCA GAC ACT ANA CCT GGG GAG ATT GGC TGC ATT	GAG CGT ACC ACA TGC AAC CCT AAA GAA ATA	1652				
Gin Giu Cys Gin And Lys Lys Lou Giu Cys Vai Tyr Lys Ann His Lys Ain Giu Cys Giu Caa gaa tgc caa gaa tgc caa gaa tgc gag tgc ga	Cys Pro Asp Asp His Glu Cys Tyr Arg Glu TGT CCT GAT GAT CAC GAG TGT TAC AGG GAG	570 1742				
Pro Ala Lys Anp Ser Cys Ser Giu Giu Anp Ann Giy Lys Cys Gin Ser Ser Giy Gin Ang	Cys Val Be Glu Ann Gly Lys Ala Val Cys	600				
CCT GCC AAA GAC TCT TGC AGT GAA GAG GAT AAT GGT AAA TGT CAA AGC AGT GGG CAG CCT	TGT GTA ATA GAA AAC GGA AAG GCT GTT TGC	1832				
Lys Giu Lys See Giu Ala Thr Thr Ala Ala Thr Thr Thr Thr Lys Ala Lys Amp Lys Amp Ang Gaa ang Gca ang Gca aca aca aca gct gcg ang gca ang gcg an	Pro Anp Pro Gly Lys Ser Ser Ala Ala Ala CCA GAT CCT GGA AAG TCA AGT GCT GCA GCA	630 1922				
Val Ser Ala Thr Giy Lou Lou Lou Lou Ala Ala Thr Ser Val Thr Ala Ala Ser Lou GTA TCA GCT ACT GGG CTC TTG TTA CTG CTC GCA GCT ACT TCA GTC ACCGCA GCA TCG TTG	End TAA GGAAGATGTCCAACTTGAATACGGAACAG	2014				
CTIGAATATGTATATACATCACOCTTACATCGAACACCTAGCTTGGTTTTTGGAATTTCAATATTGCGCATTGGTACTCACOGCAACATGAATGTATTACTTTAAAATGA 212						

FIG. 1. The DNA sequence and the derived amino acid sequence of Bm86. Regions corresponding to the 72-, 51-, 63-, and 50-mer probes and the five sites for potential N-linked glycosylation are underlined.

suggest that the amino acid segment 492–529 is an EGF-like region, albeit with only 5 cysteines. Because the strongest feature of EGF-like regions is the conservation of the cysteine residues, we considered that a mutation could have occurred at some stage in the construction of the cDNA library. We therefore isolated a Bm86 clone from a different tick cDNA library and determined the DNA sequence in this region. Base 1552 in the new clone is a guanine instead of the thymine in the original clone, translating to cysteine instead of phenylalanine in the predicted amino acid sequence.

The significant sequence similarities shown in Fig. 2 suggest that amino acids 576–600 comprise a truncated EGF-like

region that contains only the first 4 cysteines. There are no cysteines after position 600. It has been determined that in EGF and in the EGF repeat of bovine factor X, disulfide bonds are formed between the cysteine pairs 1 and 3, 2 and 4, and 5 and 6 (17, 18). If the same occurs in Bm86, the 576–600 region could form the first two disulfide bonds of an EGF-like region. There are 4 cysteines between regions B and C (see Figs. 2 and 3). The pattern in this region, Cys-Xaa₁₈-Cys-Xaa₁-Cys-Xaa₁₂-Cys-Xaa₁₈, falls outside the pattern of Bm86 EGF-like regions. But the region fits the pattern if the cysteine residues are 2, 4, 5, and 6. The disulfide bond that would be missing in this case from this possible

	AMINO ACID POSITION	1		2	3	4 5	6
A	24 - 65 71 - 103	CSD	FGNEF-	CRNAE -	- C E V V P G A	E D D F V C K C P R D N M Y F I	NAAEKQC
2	113 - 147			CRNRGG	TAKLRTDG	F I G A T C D C G E W G A M N	MTTRNC
D	152 - 183 209 - 246	CKTI	KEAGEV	VCKHG	- C R S T G K A	YECTCPSGSTVA	E D G I T C
E	295 - 291 295 - 334	CVD	Е Q К Q Т - К К	- C R P T E D - C H E E F M	- C R V H K G T D C G V Y M N R	QSCECPWNQHLV QSCYCPWKSRKP	G D T C G P N V N I N E C
G	492 - 529	CNEI	NEQSE -	- C A D K G Q	I <u>F*VYENG</u> K	AN COCPPDTKPG	EIGC
H	535 - 567 576 - 600	CNPI	KEIQE- EDNGK-	- CQDKKL - CQSSGQ	ECVYKNHK RCVIENGK	A E C E C P D D H E A V C	[C]

FIG. 2. Comparison of the amino acid sequences of the regions of Bm86 whose sequences fit (or partially fit) the pattern Cys-Xaa₄₋₈-Cys-Xaa₃₋₆-Cys-Xaa₈₋₁₁-Cys-Xaa₅₋₁₅-Cys (where Xaa is any amino acid except for cysteine). Dashes were inserted to align the cysteine residues (shaded). The regions containing 5 or 6 cysteine residues are labeled A-H for reference (see also Fig. 3). In the carboxyl-terminal region, residues identical in two or three of the aligned sequences are enclosed in boxed outlines. At amino acid position 507, the asterisk highlights the phenylalanine identified from the sequence of the first Bm86 cDNA clone that is predicted as a cysteine from the sequence of a second cDNA isolated from a separate library.

"partial EGF repeat" is that which forms in EGF and bovine factor X between cysteines 1 and 3. A diagram of Bm86 that illustrates the above features (Fig. 3) shows that most of the protein consists of EGF-like regions or regions with only 4 cysteines that may be "partial EGFs."

Expression of Recombinant Bm86 in E. coli. pBTA708 encodes a fusion protein of 143 kDa consisting of the first 651 amino acids of E. coli β -galactosidase, 599 amino acids of Bm86 (amino acids 31-629), and 19 amino acids that are encoded by other parts of the vector, such as the cloning-site regions. After induction of cultures of BTA1752 with IPTG, a band of \approx 143 kDa could be seen on polyacrylamide gels (Fig. 4A). Because about 10% of Bm86 consists of cysteine residues it was expected that Bm86 would be highly labeled by adding $[^{35}S]$ cysteine to the growth medium after induction. A strong band was seen on the autoradiograph in the region of 143 kDa after induction of cultures with IPTG (Fig. 4B). Further evidence that the 143-kDa band was related to Bm86 was obtained by demonstrating that the 143-kDa protein is recognized by antiserum from a cow vaccinated with native Bm86 (Fig. 4C) but not by prevaccination serum (data not shown).

Partial Protection of Cattle by Vaccination with β -Galactosidase-Bm86 Inclusion Bodies. Most, if not all, of the β galactosidase-Bm86 fusion protein was present as inclusion bodies (19) in the insoluble fraction of the E. coli extract. A crude preparation of inclusion bodies was used to vaccinate three cattle which were subsequently challenged with ticks. The most dramatic aspect of the results (Table 2) was the high proportion of damaged ticks that were recovered from the vaccinated cattle. Damage was recognized by the striking change in coloration of ticks from the normal gray to red, due to the leakage of bovine erythrocytes through the damaged tick gut wall into the tick hemolymph. In this experiment, there was only a modest decrease in the number of ticks that survived on the vaccinated animals (24%). However, the majority of the surviving ticks were damaged and the average weight of ticks dropping from vaccinated animals after engorgement was significantly reduced. These damaged ticks





FIG. 3. Structure of Bm86. Proposed EGF-like regions are designated A-H. The numbers in the ovals and rectangles refer to the number of cysteines in each region; region G is shown as having 5 cysteines, the number deduced from the original clone. As described in the text, a second cDNA clone was found to encode 6 cysteines in this region. The sites for potential N-linked glycosylation as defined by Asn-Xaa-Ser/Thr are indicated (-<). The membrane is shaded and possible sites for O-linked carbohydrate addition are shown as short horizontal lines.

were greatly impaired in their egg-laying ability. In this preliminary experiment, the overall effects of vaccination were to reduce the reproductive ability of the ticks on vac-

> FIG. 4. Characterization of recombinant Bm86. (A and B) BTA1752 and JM101 containing pBTA224 were grown in tryptone soy broth (A) or in tryptone soy broth containing [³⁵S]cysteine (B). Crude inclusion body preparations were resolved by SDS/8% PAGE and stained with Coomassie brilliant blue (A) or autoradiographed (B). Samples were isolated from JM101/pBTA224 not induced with IPTG (lane i), JM101/pBTA224 induced with IPTG (lane ii), BTA1752 induced with IPTG (lane iii), BTA1752 not induced with IPTG (lane iv). (C) Samples were also transferred to nitrocellulose and incubated with bovine serum from cattle vaccinated with native Bm86 isolated from ticks. Lane i, JM101/pBTA224; lane ii, BTA1752. Bio-Rad high molecular weight markers were used (lane M, A and C): myosin heavy chain, M_r 200,000; β -galactosidase, M_r 116,000; phosphorylase $b, M_r 92,500$; bovine serum albumin, M_r 66,200; ovalbumin, M_r 45,000.

Table 2.	Results of	cattle	vaccination	and	challenge	trial
----------	------------	--------	-------------	-----	-----------	-------

Animal	Cumulative tick number	Cumulative tick weight, g	Mean tick weight, mg	% ticks damaged	Egg/tick weight ratio	Cumulative egg weight, g
Controls						
181	979	272	278	1	0.57	155
183	1269	319	251	4	0.57	182
185	1365	381	279	1	0.57	217
Mean ± SD	1204 ± 201	324 ± 55	269 ± 16	2 ± 2	0.57	185 ± 31
Total	3613	972				553
Vaccinates						
180	1191	247	207	81	0.21	52
182	662	111	168	80	0.22	25
188	880	161	183	74	0.32	52
Mean ± SD	911 ± 266	182 ± 84	186 ± 20	78 ± 4	0.25 ± 0.06	43 ± 16
Total	2733	519				129
Reduction	24%	47%	31%		56%	77%

Three cattle were vaccinated with $\approx 400 \ \mu g$ of the β -galactosidase–Bm86 fusion polypeptide isolated as inclusion bodies from BTA1752. Due to the high degree of purity of the antigen preparation, control cattle were vaccinated with adjuvant alone. The cattle were subsequently challenged with ≈ 1000 tick larvae on each of three successive days. The total number and total weight of ticks >4 mm (all females) collected from each animal over an 8-day period beginning 18 days after the challenge commenced was used to calculate the mean weight per tick. The number of ticks that were visibly damaged was determined on each day and the sum of these is expressed as a percentage of the total number of ticks collected from each animal. On days 18, 19, 20, and 22 following the commencement of the challenge, 20–60 representative ticks from each animal were separated, weighed, and incubated until egg production had ceased. The weight of the eggs produced was determined and used to calculate the proportion of the tick weight that was converted into eggs. This figure was then used to estimate the total weight of eggs that all of the surviving ticks would have produced under ideal conditions. The percent reduction in the measurements for the experimental groups is indicated relative to the control animals. Differences in total weight, mean weight, percent damage, and total egg production between the two groups are significantly different (>95% confidence level).

cinated animals by 77% compared with controls (as measured by the weight of eggs laid by survivors). These results demonstrate that protective epitopes are present in the protein portion of Bm86, and we are confident that it will be possible to produce an effective recombinant anti-tick vaccine. Mathematical models (20) predict that if the effects obtained in this preliminary trial could be maintained in the field for a complete tick season, pasture contamination would be reduced to such an extent that productivity losses due to ticks would be minimal.

We express our appreciation to Drs. Neil Willetts, David Irving, and Vince Atrache for their helpful contributions to the manuscript. We thank Mr. George Riding, Mr. Robby McKenna, and Miss Janine Nielsen for assistance during the cattle vaccination and challenge. Mr. Bernie McInernie and Miss Megan Ash are also gratefully acknowledged. This research was supported in part by a grant from the Industrial Research and Development Board under the National Biotechnology Grant Scheme.

- Willadsen, P., Riding, G. A., McKenna, R. V., Kemp, D. H., Tellam, R. L., Nielsen, J. N., Lahnstein, J., Cobon, G. S. & Gough, J. M. (1989) J. Immunol. 143, 1346–1351.
- Kemp, D. H., Pearson, R. D., Gough, J. M. & Willadsen, P. (1989) Exp. Appl. Acarol. 7, 43-58.
- 3. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter,

W. J. (1979) Biochemistry 18, 5294-5299.

- 4. Gubler, U. & Hoffman, B. J. (1983) Gene 25, 263-269.
- 5. Young, R. A. & Davis, R. W. (1983) Science 222, 778-782.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- 7. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) Proc. Natl. - Acad. Sci. USA 74, 5463-5467.
- 8. Henikoff, S. (1984) Gene 28, 351-359.
- 9. Ruther, U. & Muller-Hill, B. (1983) EMBO J. 2, 1791-1794.
- 10. Messing, J. (1983) Methods Enzymol. 101, 20-78.
- 11. Millar, J. H., ed. (1972) Experiments in Molecular Genetics (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- 12. Briggs, M. S. & Gierasch, L. M. (1986) Adv. Protein Chem. 38, 109-180.
- 13. Low, M. G. & Saltiel, A. R. (1988) Science 239, 268-275.
- Russell, D. W., Schnieder, W. J., Yamamoto, T., Luskey, K. L., Brown, M. S. & Goldstein, J. L. (1984) Cell 37, 577-585.
- Rees, D. J. G., Jones, I. M., Handford, P. A., Walter, S. J., Esnouf, M. P., Smith, K. J. & Brownlee, G. G. (1988) *EMBO J.* 7, 2053–2061.
- 16. Bender, W. (1985) Cell 43, 559-560.
- 17. Savage, C. R., Jr., Hash, J. H. & Cohen, S. (1973) J. Biol. Chem. 248, 7669-7672.
- 18. Hojrup, P. & Magnusson, S. (1987) Biochem. J. 245, 887-892.
- 19. Kane, J. F. & Hartley, D. L. (1988) Tibtech. 6, 95-101.
- 20. Sutherst, R. W. & Maywald, G. F. (1985) Agric. Ecosyst. Environ. 13, 281-299.