Volume 10 Number 7 1982 Nucleic Acids Research

Molecular cloning and nucleotide sequence of cDNA coding for calf preprochymosin

T.J.R.Harris*, P.A.Lowe*, A.Lyons*, P.G.Thomas*, M.A.W.Eaton+, T.A.Millican+, T.P.Patel+, C.C.Bose+, N.H.Carey** and M.T.Doel*

*Department of Molecular Biology, and +Department of Chemistry, Celltech Ltd., 250 Bath Road, Slough SL1 4DY, Berkshire, UK

Received 8 March 1982; Accepted 25 March 1982

ABSTRACT

DNA complementary to calf stomach mRNA has been synthesised and inserted into the Pst1 site of pAT153 by G-C tailing. Clones containing sequences coding for prochymosin were recognised by colony hybridisation with cDNA extended from a chemically synthesised oligodeoxynucleotide primer, the sequence of which was predicted from the published amino acid sequence of calf prochymosin[1]. Two clones were identified which together contained a complete copy of prochymosin mRNA. The nucleotide sequence is in substantial agreement with the reported amino acid sequence of prochymosin[2] and shows that this protein has a mol.wt. of 40431 and chymosin a mol.wt. of 35612. The sequence also indicates that prochymosin is synthesised as a precursor molecule, preprochymosin, having a 16 amino acid hydrophobic leader sequence analogous to that reported for other secreted proteins[3].

INTRODUCTION

Chymosin (Rennin E.C.3.4.23.4) is the major milk clotting enzyme in the fourth stomach of the unweaned calf. The enzyme, an aspartate proteinase[2], is secreted as a zymogen, prochymosin (365 amino acids, mol.wt. *c*40,000) and is irreversibly converted under acid conditions into active enzyme (mol.wt. *c*35,600) by cleavage of 42 amino acids at the NH_2-terminus[4]. Chymosin is also the active constituent of cheese rennet and so is an enzyme of some commercial importance. A reduction in the veal calf market has led to a shortage of bovine chymosin and to the introduction of several substitutes, notably of fungal origin, the so called mucor rennins[5]. The scarcity and potential value of bovine chymosin in the dairy industry suggests the attractive possibility of producing the enzyme by recombinant DNA techniques. In this paper we report the cloning and the nucleotide sequence of calf preprochymosin cDNA.

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

0305-1048/82/1007-2177$2.00/0

MATERIALS AND METHODS

Purification of mRNA

Total nucleic acids were extracted from the mucosal layer of the fourth stomach of a freshly slaughtered suckling calf by phenol-m-cresol extraction[6] and mRNA purified by 3M sodium acetate precipitation and oligo(dT)-cellulose chromatography. About 400μg of mRNA was obtained from 13mg of total RNA.

In vitro translation and immunoprecipitation

One μg of mRNA was translated *in vitro* in 10μl of a commercial reticulocyte lysate (Amersham International) supplemented with 12μCi of [^{35}S]-methionine (Amersham) at 30°C for 2h. Polypeptides were analysed by SDS polyacrylamide gel electrophoresis[7] and fluorography[8]. Immunoprecipitation of *in vitro* translation products was done in 0.15M NaCl, 0.005M EDTA, 0.05M Tris-HCl pH 7.5 0.05%(v/v) NP-40 and bovine serum albumin (1mg/ml) as described previously[9] using an appropriate dilution of sheep antichymosin serum (Dept. of Immunodiagnostic Research, University of Birmingham) and *S.aureus* ghosts (Immunoprecipitin BRL) to precipitate immune complexes.

Isolation of cDNA clones

Double stranded cDNA was synthesised from mRNA by oligo(dT)-primed reverse transcription[10], S_1 treated and tailed with dC residues[11]. The tailed DNA was annealed to dG tailed Pst1 digested pAT153[12] and used to transform *E.coli* HB101 to tetracycline resistance[13]. Ampicillin sensitive transformants were picked into Microtiter trays[14] and colony hybridisation done on nitrocellulose filters[15] (Schleicher & Schuell) in plastic bags, overnight at 42°C, in 40%(v/v) formamide, 0.5M NaCl, 10mM Pipes-NaOH pH 6.4, 0.5% SDS, 100μg/ml salmon sperm DNA (Sigma), 2 x Denhardt's solution and 50μg/ml *E.coli* tRNA. The probe used was either alkali fragmented [^{32}P] end labelled mRNA from calf stomach (or calf liver), or [^{32}P] cDNA. After hybridisation the filters were washed 3 times in 2 x SSC, 0.1% SDS for 30 min. at 37°C and once with 1 x SSC, 0.1% SDS, before drying and autoradiography. The recombinant designated G1 was isolated from a set of cDNA clones prepared without the hairpin priming of second strand synthesis[17]. For this batch of cDNA, first strands were fractionated on alkaline agarose gels and DNA of 800-1200 nucleotides electroeluted before tailing with dC residues. Second strand synthesis was primed by oligo(dG)$_{12-18}$ (Collaborative Research) and double stranded DNA of 700-1200 nucleotides isolated from a neutral agarose gel, before dC-tailing and annealing to dG tailed, Pst1 cut pAT153. G1 was

recognised by hybridisation to the central BamH1 fragment of clone A36 DNA (Fig. 2), [^{32}P] end labelled using polynucleotide kinase.

Primer Extension

The primers ATTCATCATATT, ATTCATCATGTT, GTTCATCATATT and the primer GTTCATCATGTT used to prepare [^{32}P] cDNA, for use in colony hybridisation, were synthesised manually by the solid phase phosphotriester method[16]. The primer GGTGATCTCAGCGCC, used to obtain the sequence at the 5' end of prochymosin mRNA, was synthesised by an improved phosphotriester approach on a Celltech automated DNA synthesiser. For hybridisation, primer and mRNA were coprecipitated from ethanol, resuspended in 10mM Tris, 1mM EDTA pH 7.4 and heated at 100°C for 2 min. before quick cooling adjusting to 20mM Pipes-NaOH (pH 6.4) 0.6M NaCl and leaving at 20°C for 16h. Reverse transcription was done as described[10].

DNA sequence determination

Plasmid DNA was prepared from E.coli by the alkaline-SDS procedure[18] followed by equilibrium density gradient centrifugation in a Beckman VTi 60 rotor. Restriction enzyme digestions were done under the conditions recommended by the suppliers (New England Biolabs or Boehringer Mannheim) and nucleotide sequences determined by the chemical procedure of Maxam and Gilbert[19]. At least 50% of the sequence was obtained from both DNA strands and all restriction sites used for 5' end labelling (except the right hand BamH1 site, Fig. 2) were read through from another site.

RESULTS

In vitro translation of calf stomach mRNA

Nucleic acids were extracted from the mucosa of the abomasum of a freshly slaughtered suckling calf by the phenol-m-cresol procedure[6] and mRNA purified by oligo(dT)-cellulose chromatography. Fig. 1 shows the polypeptides translated from this mRNA when added to a rabbit reticulocyte lysate. Several low molecular weight species were made but a polypeptide of mol.wt. about 40,000, migrating slightly more slowly than the chymosin marker (mol.wt. *c*35,000), was the major product. Sucrose gradient centrifugation and in vitro translation showed that the mRNA coding for this polypeptide sedimented at about 16S, indicating that the mRNA was largely undegraded. Immunoprecipitation, with sheep antichymosin antibody, of the polypeptides synthesised in vitro showed that only the protein of mol.wt. 40,000 was precipitated (Fig. 1 lane b). Moreover, the addition of increasing amounts of chymosin to the reticulocyte lysate inhibited

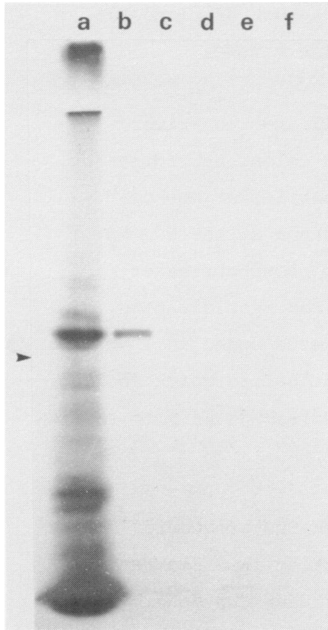

Figure 1
Polyacrylamide gel of an in vitro translation of mucosal mRNA. Lane (a) polypeptides translated from unfractionated mucosal mRNA. Lane (b) polypeptides precipitated from the in vitro translation by antichymosin serum. Lanes (c-f) polypeptides precipitated from the in vitro translation in the presence of (c) 2.5µg; (d) 5µg; (e) 12.5µg; (f) 25µq of chymosin (Sigma) further purified by DEAE-cellulose chromatography[20]. The arrow alongside lane (a) marks the position of migration of purified chymosin.

the precipitation of this polypeptide (Fig. 1 lane c-f). As chymosin and prochymosin share antigenic determinants[21] it is reasonable to assume that the 40,000 mol.wt. polypeptide is prochymosin. These results indicate that prochymosin mRNA makes up a considerable proportion of the mRNA species present in the mucosa and confirm the data of Beppu et al[22,23].

Isolation of clones containing chymosin specific cDNA

Molecular cloning of calf stomach mRNA sequences was achieved by standard techniques; DNA was prepared by reverse transcriptase, tailed with dC residues, annealed to dG tailed Pst1 digested pAT153 DNA, and used to transform E.coli HB101 to tetracycline resistance. Chymosin specific clones were identified by colony hybridisation. An initial selection was obtained by comparing the hybridisation of the clones to [^{32}P] labelled alkali fragmented mRNA prepared from either calf stomach or calf liver mRNA. From about 200 recombinants, 35 were found to hybridise specifically to mRNA from calf stomach but not to the mRNA from calf liver. These recombinants were analysed further by a second hybridisation screen with [^{32}P] labelled cDNA prepared from mucosal mRNA by reverse transcriptase, using a chemically synthesised oligodeoxynucleotide primer. The nucleotide sequence of the primer was predicted from the known amino acid sequence

of chymosin at positions 183-186[1,2] as follows:

	183 184 185 186		
NH_2 ---	Asn Met Met Asn	--- COOH	Protein
5' ---	AAU/C AUG AUG AAU/C	--- 3'	mRNA
3' ---	TTA/G TAC TAC TTA/G	--- 5'	Primer

Preliminary experiments established that only one of the four possible dodecanucleotides (GTTCATCATGTT) primed specifically the synthesis of cDNA from calf stomach mRNA. Nucleotide sequence analysis of unfractionated cDNA extended from this primer confirmed that priming was specific, since the sequence obtained was consistent with the amino acids preceeding amino acid 183[1,2]. In addition the sequence showed the presence of three infrequent restriction enzyme sites within 40 nucleotides of each other (BamH1, $EcoR_1$, SmaI; see Fig. 2). A computer prediction of the possible nucleotide sequence and restriction enzyme sites, made from a knowledge of the amino acid sequence of prochymosin, was compatible with this finding. Furthermore the predicted sequence showed that only one SmaI site was possible in the coding sequence for prochymosin (CCCGGG, pro-gly, aa 152-153). This suggested that contiguous SmaI, $EcoR_1$ and BamH1 restriction sites would be diagnostic of chymosin related sequences. One of the 35 calf stomach specific clones showed unequivocal hybridisation with the extended primer. Single colony lysate agarose gels[24] showed that the plasmid in this clone contained an insert of about 1000 base pairs and restriction enzyme digestion of plasmid DNA prepared from 1ml of culture[18] confirmed the presence of closely linked BamH1, $EcoR_1$ and SmaI sites. A restriction map of the cDNA insert in this clone (A36) is shown in Fig. 2. From the alignment of the restriction map with the unique SmaI site it was evident that the A36 insert contained the sequence coding for the -COOH terminus of the enzyme but did not extend to the region coding for the junction between prochymosin and chymosin.

Cloning of 5' sequences

To obtain clones containing the 5' end of the mRNA a second batch of cDNA was prepared using the procedure of Land *et al*[17]. After first strand synthesis single stranded DNA was tailed with dC residues and oligo(dG) used to prime second strand synthesis. This avoids hairpin priming of second strand synthesis and the subsequent loss of hairpin sequences during S1 digestion. Ampicillin sensitive recombinants obtained from this cDNA, annealed to G tailed Pst1 cut pAT153, were screened by colony

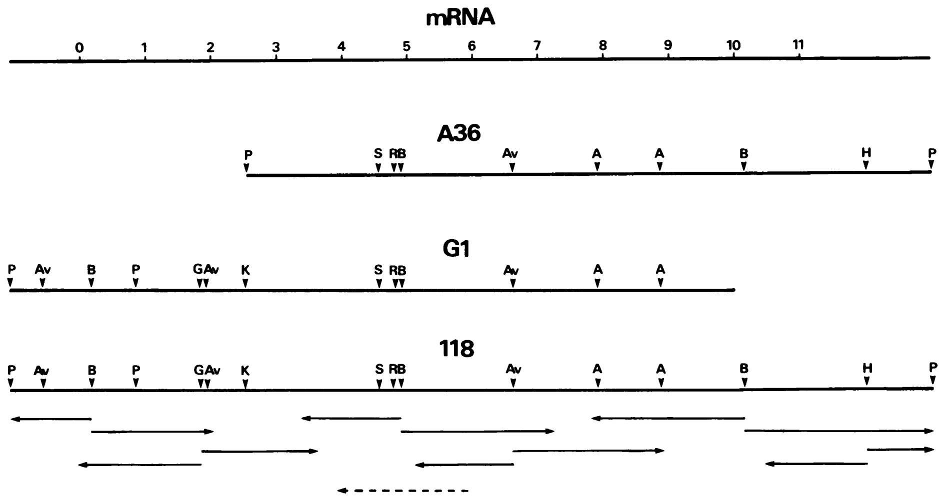

Figure 2
Alignment of partial restriction maps of the cDNA inserts of prochymosin recombinants with prochymosin mRNA. Capital letters refer to enzymes as follows: A = AluI; B = BamH1; H = Hinf1; G = BglII; K = KpnI; P = Pst1; R = $EcoR_1$; S = SmaI; Av = AvaII. The PstI site at the 3' end of G1 was not regenerated so the precise 3' limit of the insert is not known. Numbers above the mRNA designate the nucleotide number ($x10^{-2}$) starting at the first amino acid of prochymosin, 14 nucleotides upstream from the left hand BamH1 site. The sites used for labelling and the direction of sequence determination are denoted by arrows beneath clone 118. The broken arrow shows the sequence determined from cDNA made from mRNA using the primer GTTCATCATGTT (see text).

hybridisation using the central BamH1 fragment of A36 DNA (Fig. 2) as a probe. Recombinant G1 was obtained from this experiment and the restriction map of the insert is shown in Fig. 2.

Nucleotide sequence determination

As it was clear that clones A36 and G1 covered all the coding sequence of prochymosin, the nucleotide sequence of the complete cDNA insert was determined. Sequencing was facilitated by the construction of a recombinant containing a complete cDNA insert (designated 118). This was made by reconstruction through the unique $EcoR_1$ site in the two inserts. The restriction map of this insert and the sequencing strategy is shown in Fig. 2. The nucleotide sequence is presented in Fig. 3.

DISCUSSION

The nucleotide sequence presented here indicates that we have cloned the

mRNA coding for the more predominant species of chymosin, chymosin B since glycine rather than aspartate is the amino acid at position 286 (Ref. 1). Apart from the assignment of asparagine for aspartate at amino acids 202 and 214 the nucleotide sequence is in agreement with the amino acid sequence of prochymosin obtained by Foltmann et al[1].
However the nucleotide sequence upstream from that coding for the NH_2 terminus of prochymosin (alanine) shows that there are 45 nucleotides before an ATG triplet. The possibility that this ATG is the initiation codon for a 16 amino acid leader peptide sequence preceeding prochymosin, is strengthened by the observation that the predicted amino acid residues are largely hydrophobic with a -COOH terminal glycine[3,25]. Moreover, the nucleotide sequence surrounding the ATG is like that surrounding other initiation codons in having a purine (A) three bases before and a purine one base after it[26]. These date indicate that prochymosin, like other secreted proteins, is synthesised in a precursor form, preprochymosin, and suggests that the product of translation of calf stomach mRNA in the reticulocyte lysate is preprochymosin, since the lysate would not be expected to process the precursor in the absence of microsomal enzymes. Codon usage (Table 1) shows no unexpected preferences, there being the bias towards C and G in the third position typical of eukaryotic mRNAs[27]. There seems to be little suggestion of any conservation in the nucleotide sequence coding for the similar amino acid sequences near the aspartate residues at the active site of the enzyme (amino acids 76-85, and 258-267)[28]. The overall length of the mRNA coding for preprochymosin, a protein of 381 amino acids (mol.wt. = 42131) is 1300 nucleotides.
The length of the 5' untranslated region was ascertained by using a synthetic deoxyoligonucleotide complementary to nucleotides 46-60 (GTTGATCTCAGCGCC) (see Fig. 3) to prime cDNA synthesis from mRNA. A product of about 90 nucleotides was obtained and its nucleotide sequence was consistent with that found at the 5' end of recombinant G1. Furthermore it was apparent that there were only 2 or 3 nucleotides at the 5' end of the mRNA not contained in recombinant G1, indicating that there is a 5' untranslated region of about 25 nucleotides. A termination codon after the -COOH terminal amino acid marks the beginning of the 3' untranslated region of 136 nucleotides, which contains one further in phase termination codon and several repeated sequences based on the doublet CA. It also includes the AATAAA sequence thought to be the signal for polyadenylation, 14 nucleotides from the poly A tail[29].

```
                             1                                                                                   50
AGCAGCGGCTGGACCCAGATCCAAG ATG AGG TGT CTC GTG GTG CTA CTT GCT GTC TTC GCT CTC TCC CAA GGC / GCT GAG ATC ACC AGG
                          MET ARG CYS LEU VAL VAL LEU LEU ALA VAL PHE ALA LEU SER GLN GLY / ALA GLU ILE THR ARG
                                                  -10                                        1

                                            100
ATC CCT CTG TAC AAA GGC AAG TCT CTG AGG AAG GCG CTG AAG GAG CAT GGG CTT CTG GAG GAC TTC CTG CAG AAA CAG CAG TAT
ILE PRO LEU TYR LYS GLY LYS SER LEU ARG LYS ALA LEU LYS GLU HIS GLY LEU LEU GLU ASP PHE LEU GLN LYS GLN GLN TYR
                 10                                  20                                      30

150                                                          200
GGC ATC AGC AGC AAG TAC TCC GGC TTC // GGG GAG GTG GCC AGC GTG CCC CTG ACC AAC TAC CTG GAT AGT CAG TAC TTT GGG
GLY ILE SER SER LYS TYR SER GLY PHE // GLY GLU VAL ALA SER VAL PRO LEU THR ASN TYR LEU ASP SER GLN TYR PHE GLY
                         40      42 // 43                         50                                       60

                        250                                                          300
AAG ATC TAC CTC GGG ACC CCG CCC CAG GAG TTC ACC GTG CTG TTT GAC ACT GGC TCC TCT GAC TTC TGG GTA CCC TCT ATC TAC
LYS ILE TYR LEU GLY THR PRO PRO GLN GLU PHE THR VAL LEU PHE ASP THR GLY SER SER ASP PHE TRP VAL PRO SER ILE TYR
                                     70                                  80

                                            350
TGC AAG AGC AAT GCC TGC AAA AAC CAC CAG CGC TTC GAC CCG AGA AAG TCG TCC ACC TTC CAG AAC CTG GGC AAG CCC CTG TCT
CYS LYS SER ASN ALA CYS LYS ASN HIS GLN ARG PHE ASP PRO ARG LYS SER SER THR PHE GLN ASN LEU GLY LYS PRO LEU SER
     90                                     100                                 110

    400                                                     450
ATC CAC TAC GGG ACA GGC AGC ATG CAG GGC ATC CTA GGC TAT GAC ACC GTC ACT GTC TCC AAC ATT GTG GAC ATC CAG CAG ACA
ILE HIS TYR GLY THR GLY SER MET GLN GLY ILE LEU GLY TYR ASP THR VAL THR VAL SER ASN ILE VAL ASP ILE GLN GLN THR
            120                                 130                                     140

                        500                                                                  550
GTA GGC CTG AGC ACC CAG GAG CCC GGG GAC GTC TTC ACC TAT GCC GAA TTC GAC GGG ATC CTG GGG ATG GCC TAC CCC TCG CTC
VAL GLY LEU SER THR GLN GLU PRO GLY ASP VAL PHE THR TYR ALA GLU PHE ASP GLY ILE LEU GLY MET ALA TYR PRO SER LEU
                    150                                 160                                     170

                                        600
GCC TCA GAG TAC TCG ATA CCC GTG TTT GAC AAC ATG ATG AAC AGG CAC CTG GTG GCC CAA GAC CTG TTC TCG GTT TAC ATG GAC
ALA SER GLU TYR SER ILE PRO VAL PHE ASP ASN MET MET ASN ARG HIS LEU VAL ALA GLN ASP LEU PHE SER VAL TYR MET ASP
                            180                                 190                                     200

650                                                             700
AGG AAT GGC CAG GAG AGC ATG CTC ACG CTG GGG GCC ATC AAC CCG TCC TAC TAC ACA GGG TCC CTG CAC TGG GTG CCC GTG ACA
ARG ASN GLY GLN GLU SER MET LEU THR LEU GLY ALN ILE ASN PRO SER TYR TYR THR GLY SER LEU HIS TRP VAL PRO VAL THR
    ***                             210             ***                     220
```

```
                    750                                                              800
GTG CAG CAG TAC TGG CAG TTC ACT GTG GAC AGT GTC ACC ATC AGC GGT GTG GTT GTG GCC TGT GAG GGT GGC TGT CAG GCC ATC
VAL GLN GLN TYR TRP GLN PHE THR VAL ASP SER VAL THR ILE SER GLY VAL VAL VAL ALA CYS GLU GLY GLY CYS GLN ALA ILE
    230                                     240                                     250

                                            850                                                             900
TTG GAC ACG GGC ACC TCC AAG CTG GTC GGG CCC AGC AGC GAC ATC CTC AAC ATC CAG CAG GCC ATT GGA GCC ACA CAG AAC CAG
LEU ASP THR GLY THR SER LYS LEU VAL GLY PRO SER SER ASP ILE LEU ASN ILE GLN GLN ALA ILE GLY ALA THR GLN ASN GLN
            260                                     270                                 280

                                                                950
TAC GGT GAG TTT GAC ATC GAC TGC GAC AAC CTG AGC TAC ATG CCC ACT GTG GTC TTT GAG ATC AAT GGC AAA ATG TAC CCA CTG
TYR GLY GLU PHE ASP ILE ASP CYS ASP ASN LEU SER TYR MET PRO THR VAL VAL PHE GLU ILE ASN GLY LYS MET TYR PRO LEU
                    290                                     300                                     310

                    1000                                                             1050
ACC CCC TCC GCC TAT ACC AGC CAA GAC CAG GGC TTC TGT ACC AGT GGC TTC CAG AGT GAA AAT CAT TCC CAG AAA TGG ATC CTG
THR PRO SER ALA TYR THR SER GLN ASP GLN GLY PHE CYS THR SER GLY PHE GLN SER GLU ASN HIS SER GLN LYS TRP ILE LEU
                            320                                 330                                         340

                                        1100                                                             1150
GGG GAT GTT TTC ATC CGA GAG TAT TAC AGC GTC TTT GAC AGG GCC AAC AAC CTC GTG GGG CTG GCC AAA GCC ATC TGATCACATCG
GLY ASP VAL PHE ILE ARG GLU TYR TYR SER VAL PHE ASP ARG ALA ASN ASN LEU VAL GLY LEU ALA LYS ALA ILE
                                    350                                     360                 365

                                            1200                                                   1250
CTGACCAAGAACCTCACTGTCCCCACACACCTGCACACACACATGCACACATGTACATGAGCACATGTGCACACACACAGATGAGGTTTCCAGACAGATGATTCTCAATAA

ACGTTGTCTTTCGTCAAAAAAAA .............
```

Figure 3
Nucleotide sequence coding for calf preprochymosin. The nucleotide sequence and amino acid sequence is numbered from the beginning of prochymosin. Asterisks mark amino acids which are different to those reported by Foltmann et al[1].
The single bar indicates the junction between preprochymosin and prochymosin and the double bar the junction between prochymosin and chymosin.

Table 1. Codon usage of calf preprochymosin mRNA

Amino acid	Codon	Number
Arg	CGA	1
	CGC	1
	CGG	0
	CGT	0
	AGA	1
	AGG	6
Leu	CTA	2
	CTC	7
	CTG	21
	CTT	2
	TTA	0
	TTG	1
Ser	TCA	1
	TCC	10
	TCG	4
	TCT	4
	AGC	13
	AGT	4
Ile	ATA	1
	ATC	19
	ATT	2
Met	ATG	9
Thr	ACA	5
	ACC	13
	ACG	2
	ACT	4
Ala	GCA	0
	GCC	15
	GCG	1
	GCT	3
Gly	GGA	1
	GGC	16
	GGG	13
	GGT	3
Val	GTA	2
	GTC	8
	GTG	16
	GTT	3
Pro	CCA	1
	CCC	3
	CCG	11
	CCT	1
Lys	AAA	6
	AAG	9
Asn	AAC	12
	AAT	4
Glu	GAA	2
	GAG	12
Asp	GAC	19
	GAT	2
Gln	CAA	3
	CAG	23
His	CAC	4
	CAT	2
Tyr	TAC	17
	TAT	5
Cys	TGC	3
	TGT	4
Phe	TTC	14
	TTT	6
Trp	TGG	4

The isolation and analysis of recombinant clones containing calf prochymosin sequences, allowing the construction of a full length preprochymosin gene, extends the preliminary data of Nishimori et al (1981)[23]. It also opens the way to a characterisation of the structure of the preprochymosin gene. The construction of plasmid vectors for the expression of this gene will be reported elsewhere.

ACKNOWLEDGEMENTS
We would like to thank David Luther at the Institute for Research on Animal Diseases, Compton for the provision of calf stomachs and Peter Jones (University of Reading) for the introduction to him. We would also like to thank Roger Staden at the Laboratory of Molecular Biology, Cambridge for access to computer facilities while ours were being

organised. We are also grateful to Spencer Emtage and Brian Jenkins for much advice and encouragement and to Neil Perrett for technical assistance.

**Research Director.

REFERENCES

1. Foltmann, B., Pedersen, V.B., Kauffman, D. and Wybrandt, G. (1979) J. Biol. Chem. 254 8447-8456.
2. Foltmann, B., Pedersen, V.B., Jacobsen, H., Kauffman, D. and Wybrandt, G. (1977) Proc. Natl. Acad. Sci. U.S.A. 74 2321-2324.
3. Davis, B.D. and Tai, P.C. (1980) Nature 283 433-438.
4. Pedersen, V.B., Christensen, K.A. and Foltmann, B. (1979) Europ. J. Biochem. 94 573-580.
5. Scott, R. (1979) Topics in Enzyme & Fermentation Biotechnology 3 109-169.
6. Kirby, K.S. (1968) Methods in Enzymology Vol. XII 87-99.
7. Laemmli, U.K. (1970) Nature 227 680-685.
8. Bonner, W.M. and Laskey, R.A. (1974) Europ. J. Biochem. 46 83-88.
9. Harris, T.J.R., Brown, F. and Sangar, D.V. (1981) Virology 112 91-98.
10. Retzel, E.F., Collett, M.S. and Faras, A.J. (1980) Biochemistry 19 513-518.
11. Hoeijimakers, J.H.J., Borst, P., Van den Burg, J., Weissman, C. and Cross, G.A.M. (1980) Gene 8 391-417.
12. Otsuka, A. (1981) Gene 13 339-346.
13. Morrison, D.A. (1979) Methods in Enzymology 68 326-331.
14. Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) Nucleic Acids Res. 7 2115-2136.
15. Hanahan, D. and Meselson, M. (1980) Gene 10 63-67.
16. Edge, M.D., Green, A.R., Heathcliffe, G.R., Meacock, P.A., Schuch, W., Scanlon, D.B., Atkinson, T.C., Newton, C.R. and Markham, A.F. (1981) Nature 292 756-762.
17. Land, H., Grez, M., Hauser, H., Lindenmaier, W. and Schutz, G. (1981) Nucleic Acids Res. 9 2251-2266.
18. Ish-Horowicz, D. and Burke, J.F. (1981) Nucleic Acids Res. 9 2989-2998.
19. Maxam, A. and Gilbert, W. (1980) Methods in Enzymology 65 499-560.
20. Foltmann, B. (1970) Methods in Enzymology Vol XIX 421-436.
21. Uchiyama, H., Uozumi, T., Beppu, T. and Arima, K. (1981) J. Biochem. 90 483-487.
22. Uchiyama, H., Uozumi, T., Beppu, T. and Arima, K. (1980) Agric. Biol. Chem. 44 1373-1381.
23. Nishimori, K., Kawaguchi, Y., Hideka, M., Uozumi, T. and Beppu, T. (1981) J. Biochem. 90 901-904.
24. Boothroyd, J.C., Highfield, P.E., Cross, G.A.M., Rowlands, D.J., Lowe, P.A., Brown, F. and Harris, T.J.R. (1981) Nature 290 800-802.
25. Quinto, C., Quiroga, M., Swain, W.F., Nikovits, W.C., Standring, D.N., Pictet, R.L., Valenzuela, P. and Rutter, W.J. (1982) Proc. Natl. Acad. Sci. U.S.A. 79 31-35.
26. Kozak, M. (1981) Nucleic Acids Res. 9 5233-5252.
27. Wain-Hobson, S., Nussinov, R., Brown, R.J. and Sussman, J.L. (1981) Gene 13 355-364.
28. Foltmann, B. and Pedersen, V.B. (1977) Adv. Exp. Biol. Med. 95 3-22.
29. Proudfoot, N.J. and Brownlee, G.G. (1976) Nature 263 211-214.