

Nucleotide sequence of the coding region of two actin genes in *Bombyx mori*

Nicole Mounier, Janine Gaillard and Jean-Claude Prudhomme

Laboratoire de Biologie Cellulaire, CNRS UA-92, Université Lyon-1, 43 boulevard du 11 Novembre 1918, 69622 Villeurbanne Cedex, France

Submitted December 30, 1986

The 2 *Bombyx* actin genes A1 and A2 have been isolated from a genomic library and very probably code for muscle actins (1). We present here the complete nucleotide sequence of their coding region determined according to the method of Sanger (2). They do not contain any intron, contrary to another *Bombyx* actin gene (1). The coding region of these 2 genes presents a 89% homology at the nucleotide sequence level: there are 109 silent changes among the 124 nucleotides differing between the 2 genes. The deduced amino acid sequences are very similar (98% homology): the 8 varying amino acid residues are located in position 5 (Ala), 6 (Gly), 43 (Val), 103 (Ile), 129 (Cys), 228 (Gln), 298 (Leu) and 359 (Gln).

```

1      10      20
Met Cys Asp Asp Asp Val Arg Ala Leu Val Val Asp Asn Gly Ser Gly Met Oys Lys Ala Gly Phe Ala Gly Asp Asp Ala Pro Arg Ala
1  ATG TCC GAC GAC GAT CTT GGT GGG CTT GTA GTC GAC AAT GGC TCC GGC ATG TCC AGC GGC GGT TTC GGC GGG GAC GAC GGC GGC GGG GGC
   T   CQ G C   C   C   A   T   A   T   A   C   T   A
30      40      50
Val Phe Pro Ser Ile Val Gly Arg Pro Arg His Gln Gly Leu Met Val Gly Met Gly Gln Lys Asp Ser Tyr Val Gly Asp Glu Ala Gln
91  GTC TTC CCG TCC ATC CTG GGT GGC CCG GGC GAG GAG GGG CTG ATG GTC GGT ATG GGC GAG AAA GAC TCC TAC CTG GGT GAG GAG GGC GAG
   T   A   C   C   G   G   C   C   C   C   C   A   T   G   T   A   A
60      70      80
Ser Lys Arg Gly Ile Leu Thr Met Lys Tyr Pro Ile Glu His Gly Ile Ile Thr Asn Trp Asp Asp Met Glu Lys Ile Trp His His Thr
181 ACC AAG ACG GGT ATC CTC ACT CTG AAG TAC CCG ATC GAG CAC GGT ATC ATC ACC AAG TCC GAT GAC ATC GAG AAG ATC TCG CAC ACC
   A   A   G   A   G   A   G   A
90      100     110
Phe Tyr Asn Glu Leu Arg Val Ala Pro Glu His Pro Val Leu Leu Thr Glu Ala Pro Leu Asn Pro Lys Ala Asn Arg Glu Lys Met
271 TTC TAC AAT GAG CTC GGT CTC CCG CCG GAG CAG CAC CCG CTC CTC ACC GAA GGC CCG CTC ACC CCG AAG CCG AAC ACG GAG AAG ATG
   C   T   A   A   C   G   G   A   A   C
120     130     140
Thr Gln Ile Met Phe Glu Thr Phe Asn Ser Pro Ala Met Tyr Val Ala Ile Gln Ala Val Leu Ser Leu Tyr Ala Ser Gly Arg Thr
361 ACC CAG ATC ATG TTT GAG ACC TTC AAC TCC CCG CCG ATG TAC CTC GCG ATC CAG GCG CTC TCG CTG TAC CCG TCC GGT ACC ACC
   C   G   T   G
150     160     170
Gly Ile Val Leu Asp Ser Gly Asp Gly Val Ser His Thr Val Pro Ile Tyr Glu Gly Tyr Ala Leu Pro His Ala Ile Leu Arg Leu Asp
451 GGT ATC CTG CTC GAC TCC GGA GAT GGT TTC TCC CAC ACC GTA CCG ATC TAC GAA GGT TAC CCG CTC CCG CAC GGC ACC ATC CTC CCG
   C   T   G
180     190     200
Leu Ala Gly Asp Asp Leu Thr Asp Tyr Leu Met Lys Ile Leu Thr Glu Arg Gly Tyr Ser Phe Thr Thr Ala Glu Arg Glu Ile Val
541 TTG GCT GGT CCG GAC TTC ACC GAC TAC CTC ATG AAG ATC CTC ACC GAG ACG GGT TAC TCG TTC ACC ACC CCG CAG ACG AAG ATA CTT
   C   A   A   T   C   C   T   C
210     220     230     234a
Arg Asp Ile Lys Glu Lys Lys Oys Tyr Val Ala Leu Asp Phe Glu Gln Glu Met Ala Thr Ala Ala Ser Thr Ser Leu Glu Lys Ser
631 GGT GAC ATC AAG GAG AAG CTC TCC TAT CTC CCG CTC TTC GAG CAG GAG ATG GGC ACC GGT GGC CCG TCC ACC TCC CTC GAG AAG TCC
   T   G   T   T   C   C   T   A   C   A   C   A   T   G   A
240     250     260
Tyr Glu Leu Pro Asp Gly Gln Val Ile Thr Ile Gln Asn Glu Arg Phe Arg Oys Pro Glu Ala Leu Phe Gln Pro Ser Phe Leu Gly Met
721 TAC GAA CTT CCG GAC CCG CAG GTC ATC ACC ATC ACC GAG ACC TTC CCG TCC GGT CCG CAG GCT TCC GAC GCT TCC CTC GGT ATG
   G   A   A   A   C   C   C   C   G
270     280     290
Glu Ser Oys Gly Ile His Glu Thr Val Tyr Asn Ser Ile Met Lys Oys Asp Val Asp Ile Arg Lys Asp Leu Tyr Ala Asn Thr Val Met
811 GAA TCG TCC GGC ATC CAC GAG ACC CTG TAC AAC TCC ATC ATG AAG TCC GAC CTC GAC ATC CCG ATG CAG CTG TAC CCG AAC ACC CTC ATG
   G   T   G   T   C
300     310     320
Ser Gly Gly Thr Thr Met Tyr Pro Gly Ile Ala Asp Arg Met Gln Lys Glu Ile Thr Ala Leu Ala Pro Ser Thr Ile Lys Ile Lys Ile
901 TCC GGT GGT ACC ACC ATG TAC CCG GGT ATC CCG GAC ACG ATG CAG AAG GAA ATC ACC CCG CTC GCG CCG TCC ACC ATT AAG AAC ATC
   C   C   G
330     340     350
Ile Ala Pro Pro Glu Arg Lys Tyr Ser Val Trp Ile Gly Gly Ser Ile Leu Ala Ser Leu Ser Thr Phe Gln Gln Met Trp Ile Ser Lys
991 ATC GCT CCG CCG GAA ACG AAG TAC TCC GTA TGG ATC GGT GGA TCC ATC CTG GCT TCC CTG TCC ACC TTC CAG CAG ATG TCC ATC TCC AAG
   G   G   T   C   C   G   G   G
360     370     374
Glu Glu Tyr Asp Glu Ser Gly Pro Gly Ile Val His Arg Lys Oys Phe Ser
1081 CAG GAA TAC CAC GAC TCC GGC CCG GGC ATC GTC CAC CCG AAG TCC TTC TAA
   C   G   G   C   G   C   C   T

```

Figure legend: The first line corresponds to the codon number, the 2nd line to the deduced amino acid sequence of the gene A1, the 3rd line to the nucleotide sequence of the gene A1, and the 4th line to the nucleotide sequence of the gene A2 but only the changes are indicated.

1. Mounier N. and Prudhomme J.C., 1986, *Biochimie*, **68**, 1053-1061.
 2. Sanger F. et al, 1977, *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5447.