Genomic sequence of the Chinese hamster MT I gene

Deborah L. Grady*, Donna L. Robinson and C. Edgar Hildebrand
Los Alamos National Laboratory, Genetics Group, Los Alamos, NM 87545, USA

Submitted October 29, 1990

The metallothionein (MT) system has been used to analyze the structural features associated with the control of gene expression (1). A pre-requisite for this type of analysis is knowledge of the primary DNA sequence. We have isolated a clone containing the Chinese hamster MT I gene from a Charon 4A EcoRI total digest prepared from a Chinese hamster cell line amplified for the MT genes (XCH 4A-HA-20). The clone consists of 13 kb of DNA and contains the entire MT I gene and a portion of the MT II gene (Figure A). We have previously cloned the complete MT II gene (Figure A, see X590-HA-MT52). Combined, these two clones constitute the entire hamster MT locus. Southern blot analysis has indicated that these genes are separated by approximately 6 kb of DNA. Sequence analysis of the 1558 nucleotide Bgin/BgUI (Figure A) fragment containing the MT I gene was performed. The coding region of exon 1 runs from nucleotides 11 to 38. Exon II consists of nucleotides 507 to 571. Exon III spans nucleotides 786 to 986, with the coding region ending at nucleotide 875 (2). The size of introns I and II are 467 nucleotides and 214 nucleotides, respectively. The deduced amino acid sequence is shown below the nucleotide sequence. Three sets of simple repetitive sequences are indicated by underlining (nucleotides 1192 to 1297) 3' to the polyadenylation site (3).

REFERENCES

1 AGATCGGAAA ATGGACCCCA ACTGCTCTTG CTCACCAGCT AAGAGATAAC CAGATTAGGA CCGCAGAGAC CTCCCGCAGT TGTAGAGGAT
2 M D P N C S C S T
3 GTAGTGTAAG ATCTGCGGGG GAATAATGCT TCTGTGGGGG ATTCATTCTA GGGTTTTGTTA GCGTCCCCCT TTGCAAGCAG CTCCATCTAT
4 TCTTGGACTA TTAATATGTT CGGAAGAGTC CTCTGCGGGG GACATTAGGA ATCTGATGAT ATCTGATGAT GCAGGCGAG TCGCGCGGGC
5 61 GACCGACACG CAGGCGACCT CCGAGGTTAT TTGTGTTTCT CGACGACCTG ATCTGCTCTG ATCTGCTCTG CTCCGGCCAG TCGCGCGGGC
6 TCTTGCAGTA TTAATATTGT CCGAACGATC CTTTGCGGGG GCTGTGGGGG ATTCATTCTA GTTCTTTTTA GCGTCCCCCT TTGCAAGCAG CTCCATCTAT
7 TAGTGAGGAC ATCTGCGGGG GAATAATGCT TCTGTGGGGG ATTCATTCTA GGGTTTTGTTA GCGTCCCCCT TTGCAAGCAG CTCCATCTAT
8 GMDC KCT SCKK
9 CCGACAGATC CAGGGCGCCT CCGCGCTCTG CGCGCGCGCG GCAGCGCGAG TCTGCGCGGT CTGCGCGCGG CTGCGCGCGG CTGCGCGCGG
10 TCTTGGACTA TTAATATGGT CCGAACGATC CTTTGCGGGG GCTGTGGGGG ATTCATTCTA GTTCTTTTTA GCGTCCCCCT TTGCAAGCAG CTCCATCTAT
11 541 CAAGAGACGC AAGTGCACCT CCTGCAAGAA GAGTGAGGTG GGAGGGATAC CTGGGGTGTT GCGTAGAGT TGGCGGGAAC ACCCACAGGC
12 631 CCGACAGATC CAGGGCGCCT CCGCGCTCTG CGCGCGCGCG GCAGCGCGAG TCTGCGCGGT CTGCGCGCGG CTGCGCGCGG CTGCGCGCGG
13 TCTTGGACTA TTAATATGGT CCGAACGATC CTTTGCGGGG GCTGTGGGGG ATTCATTCTA GTTCTTTTTA GCGTCCCCCT TTGCAAGCAG CTCCATCTAT
14 721 CCGACAGATC CAGGGCGCCT CCGCGCTCTG CGCGCGCGCG GCAGCGCGAG TCTGCGCGGT CTGCGCGCGG CTGCGCGCGG CTGCGCGCGG
15 K D C K C T T S C K K
16 TCTTGGACTA TTAATATGGT CCGAACGATC CTTTGCGGGG GCTGTGGGGG ATTCATTCTA GTTCTTTTTA GCGTCCCCCT TTGCAAGCAG CTCCATCTAT
17 811 GCTGCTCCAA GTGGCGCCAG GTGCCTCTCT CGAAAGGGCG ATCGGAAAGT ATCGCGCGGT CTGACGGTCT GTGACGGGCT CTGACGGGCT
18 901 AGAGTGCCTA AGCTGGGTAG TGGAGGGGTG AGAGGAGGAT GCTGTGCCTA CTGACGGTCT GTGACGGGCT CTGACGGGCT CTGACGGGCT
19 991 GCTGCTCCAA GTGGCGCCAG GTGCCTCTCT CGAAAGGGCG ATCGGAAAGT ATCGCGCGGT CTGACGGTCT GTGACGGGCT CTGACGGGCT
20 1081 AGAGTGCCTA AGCTGGGTAG TGGAGGGGTG AGAGGAGGAT GCTGTGCCTA CTGACGGTCT GTGACGGGCT CTGACGGGCT CTGACGGGCT
21 1171 AGAGTGCCTA AGCTGGGTAG TGGAGGGGTG AGAGGAGGAT GCTGTGCCTA CTGACGGTCT GTGACGGGCT CTGACGGGCT CTGACGGGCT
22 1261 AGAGTGCCTA AGCTGGGTAG TGGAGGGGTG AGAGGAGGAT GCTGTGCCTA CTGACGGTCT GTGACGGGCT CTGACGGGCT CTGACGGGCT
23 1351 GCTGCTCCAA GTGGCGCCAG GTGCCTCTCT CGAAAGGGCG ATCGGAAAGT ATCGCGCGGT CTGACGGTCT GTGACGGGCT CTGACGGGCT
24 1441 AGAGTGCCTA AGCTGGGTAG TGGAGGGGTG AGAGGAGGAT GCTGTGCCTA CTGACGGTCT GTGACGGGCT CTGACGGGCT CTGACGGGCT
25 1531 TCTAATCTTA TTTTGGAAAA TCTGAGCT

* To whom correspondence should be addressed