Volume 222, number 1, 139–143 FEB 05142 September 1987

Sequence conservation in the 3′-untranslated regions of neurone-specific enolase, lymphokine and protooncogene mRNAs

I.N.M. Day, M.T.E.P. Allsopp, D.C.McN. Moore and R.J. Thompson

Department of Clinical Biochemistry, School of Clinical Medicine, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QR, England

Received 22 June 1987; revised version received 27 July 1987

The C-terminal protein-coding and the entire 3′-untranslated regions of a cDNA corresponding to human neurone-specific enolase mRNA have been sequenced. The 3′-untranslated region is 892 bases long and shows a high degree of homology with the 3′-untranslated region of rat neurone-specific enolase mRNA. This sequence conservation is not seen in non-neuronal enolase mRNAs. Features of the conserved sequence include an A-rich region approx. 250 bases from the stop codon at a point corresponding to the polyadenylation signal site in non-neuronal enolase mRNA, and a repeating ATTT sequence. This unusual motif in eukaryotic mRNAs has previously been reported in the 3′-untranslated regions of lymphokine and protooncogene mRNAs.

Neurone-specific enolase; cDNA sequence; Sequence conservation

1. INTRODUCTION

Enolase (EC 4.2.1.11) occurs in the cytoplasm of mammalian cells in three dimeric isoenzyme forms [1]. These are $\alpha\alpha$ (found in many tissues including liver), $\beta\beta$ (typical of skeletal muscle) and $\gamma\gamma$ (found in brain). Within the central nervous system $\gamma\gamma$ enolase is localized in neurones and $\alpha\alpha$ enolase in astrocytes, hence these two isoenzymes are usually referred to as neurone-specific enolase (NSE) and non-neuronal enolase (NNE), respectively [2]. NSE is also found in cells of the diffuse neuroendocrine system which include pancreatic islet cells, gut endocrine cells, pituitary endocrine cells, etc. [1]. The cDNAs corresponding to the mRNAs coding for both NSE and NNE in the rat and for NNE in the human have recently been sequenced [3–6]. The amino acid sequences of rat NSE and rat NNE show 82% identity [5], human NNE and rat NNE are 94% identical in amino acid sequence [6] and both NSE and NNE show sequence similarity (approx. 56%) with yeast enolase isoenzymes [4]. Despite the striking sequence conservation of the protein-coding regions of the mRNAs for rat NSE and rat NNE, the 5′- and 3′-untranslated regions show no significant homology [3]. The 3′-untranslated region of the rat NSE mRNA, however, is considerably longer (848 bases) than the equivalent region for rat NNE (312 bases) or human NNE (359 bases). Such long 3′-untranslated regions have been claimed to be a feature of brain-specific mRNAs [7]. The role of 3′-untranslated regions in eukaryotic mRNAs is

Correspondence address: I.N.M. Day, Department of Clinical Biochemistry, School of Clinical Medicine, Addenbrooke's Hospital, Hills Road, Cambridge CB2 2QR, England

The nucleotide sequence presented here has been submitted to the EMBL/GenBank database under the accession number Y00691

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/87/$3.50

currently unknown; however, sequence homologies have been found in these regions in mRNAs coding for conserved proteins expressed in the same tissue in different species but not in mRNAs coding for related proteins expressed in different tissues in the same species [8]. This, together with the demonstration that the presence of specific sequences in the 3′-untranslated region may determine mRNA stability [9], has led to the suggestion that 3′-untranslated regions may play a role in the post-transcriptional control of tissue-specific gene expression [8]. We have sequenced the 3′-untranslated region of the mRNA for human NSE and found (in contrast to the non-neuronal forms of the enzyme) a high degree of conservation with rat NSE mRNA. Additionally, a repeated sequence motif is present which has only previously been reported in lymphokine and protooncogene mRNAs and which has been implicated in selective mRNA degradation [9].

2. MATERIALS AND METHODS

A human retinal cDNA library of 10^6 independent recombinants constructed in λgt10 [10] was screened [11] with a synthetic oligonucleotide probe corresponding to positions 1403–1437 of rat NSE mRNA [3]. Initial screening was at a stringen-

A

```
putative human NSE   G H N F R N P S V L stop
rat NSE              G H N F R N P S V L stop
rat NNE              G R N F R N P L A K stop
human NNE            G R N F R S P L A K stop
```

B

```
                                                    TGA TTCCTCTG CTTGCCTGGA
19 GACGTGGAAC CTCTGTCTCA TCCTCCTGGA ACCTTGCTGT CCTGATCTGT GATAGTTCAC
79 CCCTGAGATC CCCTGAGCCC CAGGGTGCCC AGAACTTCCC TGATTGACCT GCTCCGCTGC
139 TCCTTGGCTT ACCTGACCTC TTGCTGTCTC TGCTCGCCCT CCTTTCTGTG CCCTACTCAT
199 TGGGGTTCCG CACTTTCCAC TTCTTCCTTT CTCTCTCTCT CTTCCCTCAG AAACTAGAAA
259 TGTGAATGAG GATTATTATA AAAGGGGGTC TGTGGAAGAA TGATCAGCAT CTGTGATGGG
319 AGCGTCAGGG TTGGTGTGCT GAGGTGTTAG AGAGGGACCA TGTGTCACTT GTGCTTTGCT
379 CTTGTCCCAC GTGTCTTCCA CTTTGCATAT GAGCCGTGAA CTGTGCATAG TGCTGGGATG
439 GAGGGGAGTG TTGGGCATGT GATCACGCCT GGCTAATAAG GCTTTAGTGT ATTTATTTAT
499 TTATTTATTT TATTTGTTTT TCATTCATCC CATTAATCAT TTCCCCATAA CTCAATGGCC
559 TAAAACTGGC CTGACTTGGG GGAACGATGT GTCTGTATTT CATGTGGCTG TAGATCCCAA
619 GATGACTGGG GTGGGAGGTC TTGCTAGAAT GGGAAGGGTC ATAGAAAGGG CCTTGACATC
679 AGTTCCTTTG TGTGTACTCA CTGAAGCCTG CGTTGGTCCA GAGCGGAGGC TGTGTGCCTG
739 GGGAGTTTCC TCTATACATC TCTCCCCAAC CCTAGGTTCC TGTTCTTCCT CCTGCTGCAC
799 CAGAGCAACC TCTCAATCCC CATGCCACGT TCCACAGTTG CCACCACCTC TGTGGCATTG
859 AAATGAGCAC CTCCATTAAA GTCTGAATCA GTGAAA
```

Fig.1. Sequence of a human NSE cDNA clone. (A) Comparison of the C-terminus sequence of human NSE deduced from this putative human NSE cDNA clone, with other previously published enolase sequences (see text). (B) Complete sequence of human NSE cDNA 3′-noncoding region, from the stop codon to the poly(A) tail. An internal A-rich stretch and an unusual repeating ATTT motif, discussed in the text, are underlined. The stop and polyadenylation signals are also underlined.

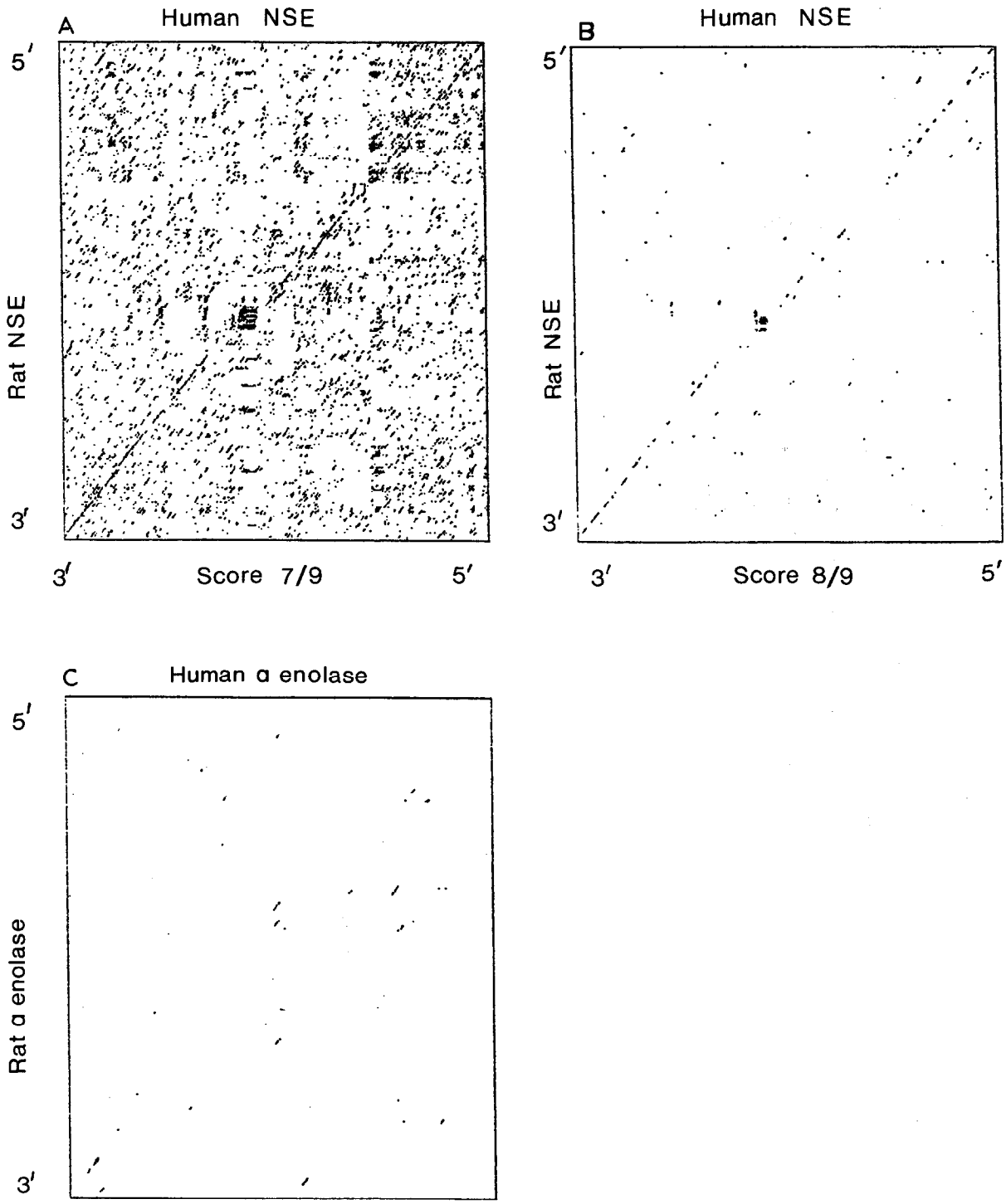

Fig.2. Diagon comparison [14] of the 3′-untranslated regions of rat NSE and human NSE (A,B) and of rat and human NNE (C). Plots A and C represent comparisons of stretches of 9 nucleotides scoring a dot for 7 or more matches out of 9; B is at the higher stringency of 8 or 9 matches out of 9.

cy corresponding to 70% probe/target match [12], rescreening being at a stringency of 85% probe/target match. A clone with a 1.2 kb insert was selected and the insert was subcloned into the *Eco*RI and *Bam*HI sites of mp8 after *Eco*RI and *Sau*3A digestion [11]. Sequencing was by the dideoxy nucleotide chain termination method [13].

Computer comparisons for sequence homology were by the Diagon method [14].

3. RESULTS

Screening of the human retinal cDNA library yielded approx. 1/1000 positives, ten of which were plaque purified and the insert sizes determined [11]. These ranged from 1 to 2 kb, and a 1.2 kb insert, which by end sequencing was shown to cover the 3′-noncoding region, was completely sequenced on both strands. The sequence of the insert is shown in fig.1. The ten amino acids preceding the stop codon (fig.1A) were identical to those in rat NSE [3] and different from those in rat and human NNE [5,6]. The preceding 60 amino acids (not shown) also corresponded to the sequence of rat NSE [3] and not to human NNE [6], showing that the cDNA selected was indeed for the neuronal form of the enzyme. The sequence of the entire 3′-untranslated region is shown in fig.1B. This spans 831 bases from the stop codon to the start of the poly(A) tail, compared with 848 bases for rat NSE and contrasting with the 359 bases seen in human NNE [6]. The mRNA for mouse NSE has also been shown to be longer than that for mouse NNE but no sequence data are available [15]. Fig.2 shows the Diagon comparisons of the 3′-untranslated regions of the mRNAs to human NSE (this paper) vs rat NSE [3] and of human NNE [6] vs rat NNE [5]. While there is no significant homology (fig.2C) between the 3′-untranslated regions of rat and human NNE mRNAs (although the lengths of these regions are similar) there is clear homology between the sequences of the 3′-untranslated regions of the mRNAs coding for the neuronal forms of the enzyme (fig.2A). More stringent Diagon comparison (fig.2B) indicates that the highest degree of conservation between the rat and human sequences is in the 200 nucleotides following the stop codon and the 200 nucleotides preceding the polyadenylation signal, while the intervening nucleotides show less sequence conservation but maintain overall length.

4. DISCUSSION

The functional reason for the markedly increased length of the 3′-untranslated regions in the mRNAs coding for the neuronal form of the enzyme is obscure but is probably not solely related to sequence conservation, since examples are known of homologous genes which preserve the length of the 3′-untranslated region without conserving its sequence [8]. The conservation of the sequence of the 3′-untranslated regions in the mRNAs for the neuronal form of the enzyme but not in the mRNAs for the non-neuronal form is reminiscent of the mRNAs coding for creatine kinase isoenzymes [16]. Here the brain form of the enzyme (creatine kinase BB) shows a high degree of conservation of the sequence of the 3′-untranslated region of the mRNA for the B subunit between species, but the mRNAs for the muscle-type enzyme (creatine kinase MM) show no such sequence conservation. The mRNA for brain-type creatine kinase enzyme, however, does not

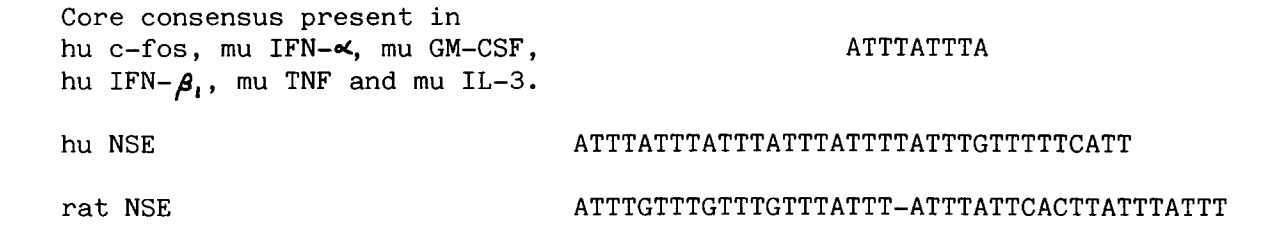

Fig.3. Comparison of a structural motif present in the 3′-untranslated region and possibly conferring instability on lymphokine and protooncogene mRNAs, with an identical motif present in human and rat neurone-specific enolase mRNAs. hu, human; mu, murine; IL, interleukin; IFN, interferon; GM-CSF, granulocyte-monocyte colony-stimulating factor; TNF, tumour necrosis factor; NSE, neurone-specific enolase.

show the markedly increased length of the 3'-untranslated region seen with neurone-specific enolase.

Apart from general sequence homology the 3'-untranslated regions of the mRNAs for rat and human NSE show two short conserved sequences worthy of comment. The first of these is a poly(A) stretch of 15 residues in the rat mRNA which appears as a 12-nucleotide sequence with eight A residues in the same position in the human mRNA. This position corresponds to the poly(A) addition site in the mRNAs for the yeast enolase isoenzymes and for non-neuronal enolase [5,6] and suggests that during evolution the lengthening of the neurone-specific transcript may have involved an insertional event. The second feature is an ATTT sequence which is repeated 9 times in 40 residues in the rat sequence (with G replacing A in three repeats) and is seen in a similar position in the human sequence where it is repeated 7 times in 28 residues with G replacing A once (fig.3). A similar sequence (fig.3), which is unusual among eukaryotic mRNAs, has been reported in the 3'-untranslated regions of the mRNAs for numerous lymphokine, cytokine, and protooncogene mRNAs where it usually spans 50 or so nucleotides [17,18]. It has been suggested that these gene products (which are not structurally related in protein-coding sequences) are transiently expressed and that this repeated sequence may be part of a mechanism for selective degradation of mRNAs containing it [18]. Insertion of this repeated sequence into the 3'-untranslated region of the stable β-globin message resulted in a dramatic increase in mRNA instability [18]. Neurone-specific enolase, however, is a major (1–2%) protein in adult neuronal cytoplasm and its mRNA appears to be abundant [4]. Detailed study of the rates of NSE mRNA turnover may throw light on the function of this unusual repeated sequence in the 3'-untranslated region.

ACKNOWLEDGEMENTS

We are grateful to Dr J.O. Thomas for providing the NSE oligonucleotide probe. This work was supported by a grant from the Wellcome Trust.

REFERENCES

[1] Marangos, P.J. and Schmechel, D.E. (1987) Annu. Rev. Neurosci. 10, 269–295.
[2] Schmechel, D., Marangos, P.J. and Brightman, M. (1978) Nature 276, 834–836.
[3] Sakimura, K., Kushiya, E., Obinata, M., Odami, S. and Takahashi, Y. (1985) Proc. Natl. Acad. Sci. USA 82, 7453–7457.
[4] Forss-Petter, S., Danielson, P. and Sutcliffe, J.G. (1986) J. Neurosci. Res. 16, 141–156.
[5] Sakimura, K., Kushiya, E., Obinata, M. and Takahashi, Y. (1985) Nucleic Acids Res. 13, 4365–4378.
[6] Giallongo, A., Feo, S., Moore, R., Croce, C.M. and Showe, L.C. (1986) Proc. Natl. Acad. Sci. USA 83, 6741–6745.
[7] Milner, R.J. and Sutcliffe, J.G. (1983) Nucleic Acids Res. 5, 1729–1739.
[8] Yaffe, D., Nudel, U., Mayer, Y. and Neuman, S. (1985) Nucleic Acids Res. 13, 3723–3737.
[9] Shaw, G. and Kamen, R. (1986) Cell 46, 659–667.
[10] Nathans, J., Thomas, D. and Hogness, D.S. (1986) Science 232, 193–202.
[11] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[12] Lathe, R. (1985) J. Mol. Biol. 183, 1–12.
[13] M13 Cloning and Sequencing Handbook (1984) Amersham International, England.
[14] Staden, R. (1982) Nucleic Acids Res. 10, 2951–2961.
[15] Lazar, M., Lucas, M., Lamande, N., Bishop, J.G., Gros, F. and Legault-Demare, L. (1986) Biochem. Biophys. Res. Commun. 141, 271–277.
[16] Billadello, J.J., Kelly, D.P., Raman, D.C. and Strauss, A.W. (1986) Biochem. Biophys. Res. Commun. 138, 392–398.
[17] Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Shimer, S. and Cerami, A. (1986) Proc. Natl. Acad. Sci. USA 83, 1670–1674.
[18] Shaw, G. and Kamen, R. (1986) Cell 46, 659–667.