FEBS Letters 512 (2002) 350–352 FEBS 25737

# Correspondence

## A second melanotransferrin gene (MTf2) and a novel protein isoform: explanation for the membrane-bound and soluble forms of melanotransferrin?

Eric Sekyere, Michael R. Food, Des R. Richardson*

First published online 18 January 2002

Melanotransferrin (MTf) has high sequence homology to serum transferrin (Tf) (for review see [1]). However, unlike other Tf family members, MTf is bound to the cell membrane by a glycosyl-phosphatidylinositol (GPI) anchor [1]. For the purpose of the present article, this well-described molecule will be called MTf1. Tfs contain two high affinity iron (Fe)-binding sites, and play a vital role in Fe transport [1]. However, human MTf1 contains a single N-terminal Fe-binding site and the membrane-bound molecule plays little role in Fe uptake by melanoma cells [1].

Recent studies have suggested that MTf1 can be found free in the serum of patients with various disease states, including Alzheimer's disease [2]. This important finding has prompted the question of whether there is a relationship between the soluble form and the well-characterised membrane-bound protein. It could be suggested that an alternative transcript of the MTf1 gene could encode a soluble molecule, or the membrane-bound form may partition from the membrane into the serum. In addition to these possibilities, a second MTf-like gene could exist that encodes a related protein that is soluble. Considering this, we examined the GenBank sequence database for evidence of human MTf sequences that could be indicative of a second gene or transcript.

Examination of this database led to the identification of several nucleotide depositories of human MTf (see Table 1). Review of both the nucleotide and protein sequence alignments revealed the presence of three different MTf cDNAs that encode two protein products. In the present report, these cDNAs have been denoted as the 'long MTf cDNA' (2364–2607 bp; also called variant 1 by the National Centre for Biotechnology Information (NCBI)) and the 'short MTf cDNA' (1629–1664 bp; also called variant 2 by the NCBI) (Table 1).

The long cDNA encodes the well-characterised membrane-bound MTf1 molecule that is composed of 16 exons spanning 26 kb of genomic sequence (Fig. 1). Within the GenBank database there are four entries for the long cDNA (Table 1). Three of these are 2.36 kb in length and are composed of 16 exons, while another cDNA (GenBank XM_028281) contains an additional 243 bp of 3′-untranslated region (UTR) (Table 1). Review of the MTf1 genomic sequence showed that this 243 bp region corresponds to an unreported exon that we have called '16b' (Fig. 1). Hence, the two long transcripts differ from one another by the presence of an additional 3′-UTR in exon 16b, but both encode membrane-bound MTf1.

In addition to the long transcripts, there are five entries in GenBank that are recently discovered short MTf cDNAs (Table 1). These cDNAs have nucleotide sequence lengths of 1629–1664 bp, the small differences between these being due to multiple transcription start sites and variation in length of the poly(A) tail. All five cDNAs contain the first six exons of the long MTf cDNA (Fig. 1) which encodes a large proportion of the N-terminal lobe of MTf1. Further, in all of the short MTf cDNAs, there is an additional 3′-sequence of 876 bp which is absent in the long cDNAs. Of this latter sequence, only 195 bp are encoding amino acids. This additional 876 bp of sequence was found to be a region within intron 6 of the MTf1 gene between exons 6 and 7. We have denoted this new exon as '6b' (Fig. 1). This shows that both the long and short MTf transcripts are the products of the well-known MTf1 gene, and could be due to alternative splicing. The protein sequence derived from exon 6b results in a C-terminal that shows limited conservation to part of the C-terminal of the long MTf1 transcript (Fig. 2).

Examination of the protein product of the short MTf1 cDNA (Fig. 2) showed it to consist of 302 amino acids, while the long MTf1 protein isoform was composed of 738 amino acids. The short MTf1 isoform contains the first four of five Fe-binding residues (DYRY at positions 78, 107, 136, and 210) found in the N-terminal lobe [1] (Fig. 2). The fifth coordinating residue H (279) resides in exon 7, that does not form part of the short MTf isoform (Fig. 2). This raises the question of whether this latter molecule can bind Fe. In addition, the consensus thermolysin metalloprotease sequence present within the long MTf1 protein isoform [1] is absent in the short isoform (Fig. 2).

Some molecules involved in Fe uptake (e.g. the Tf receptor) are regulated by the binding of the iron regulatory proteins (IRPs) to the iron-responsive elements (IREs) within the UTRs [3]. Considering this, we examined the MTf1 transcripts for the presence of the consensus IRE (5′-CAGUGN-3′) that forms the loop of this motif [3]. In the long 2.6 kb MTf1 transcript and the short MTf1 transcript, the consensus IRE

Table 1
MTf nucleotide and protein sequences in the GenBank database

| GenBank accession no. | | bp | aa | Tissue type |
|---|---|---|---|---|
| Nucleotide ID | Protein ID | | | |
| Long MTf transcripts | | | | |
| M12154 | AAA59992 | 2368 | 738 | SK-MEL 28 melanoma cell |
| NM_005929 | NP_005920 | 2368 | 738 | melanoma cells |
| XM_028282 | XP_028282 | 2364 | 738 | human |
| XM_028281 | XP_028281 | 2607 | 738 | human |
| Short MTf transcripts | | | | |
| NM_033316 | NP_201573 | 1664 | 302 | skin, melanotic melanoma |
| BC 022623 | AAH02623 | 1659 | 302 | uterus, endometrium-adenocarcinoma |
| BC 001875 | AAH01875 | 1664 | 302 | skin, melanotic melanoma |
| BC 007550 | AAH07550 | 1664 | 302 | skin, melanotic melanoma |
| XM_003133 | XP_003133 | 1629 | 302 | human |


PII: S0014-5793(02)02248-2

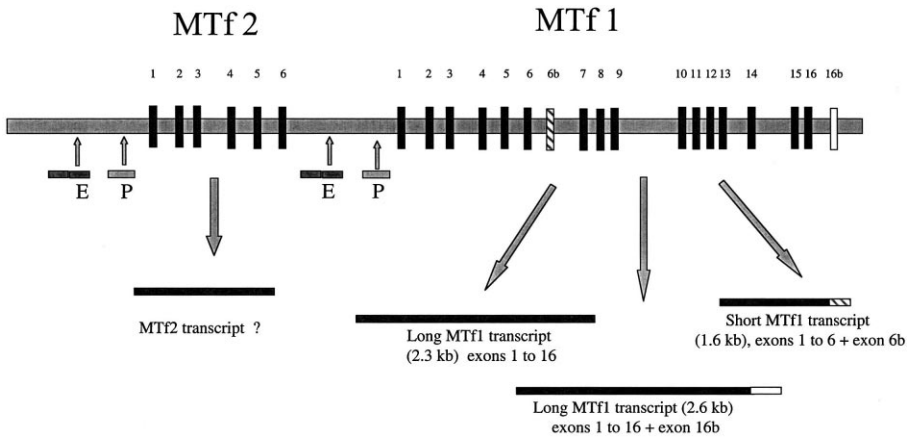


Fig. 1. Schematic illustration showing the relative positions of the originally identified MTf1 gene compared to the newly identified MTf2 gene. Three transcripts have been identified that originate from the MTf1 gene, that is, two 'long' transcripts and one 'short' transcript. Both long transcripts (16 exons) encode the membrane-bound MTf1 molecule and differ from one another by an additional exon (16b). The short MTf1 transcript (six exons and exon 6b) may encode the soluble form of the protein that is found in Alzheimer's disease [2]. The MTf2 gene consists of six putative exons that are identical to those found in the MTf1 gene. Furthermore, the MTf2 gene has identical minimal enhancer (E) and promoter (P) elements to those identified in MTf1, although a MTf2 transcript has not been reported.

```
AAA59992 Long  1 MRGPSGALWLLLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSA
NP201573 Short   MRGPSGALWLLLALRTVLGGMEVRWCATSDPEQHKCGNMSEAFREAGIQPSLLCVRGTSA
                 ************************************************************

AAA59992      61 DHCVQLIAAQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVT
NP201573         DHCVQLIAAQEADAITLDGGAIYEAGKEHGLKPVVGEVYDQEVGTSYYAVAVVRRSSHVT
                 ****************************************************************

AAA59992     121 IDTLKGVKSCHTGINRTVGWNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETS
NP201573         IDTLKGVKSCHTGINRTVGWNVPVGYLVESGRLSVMGCDVLKAVSDYFGGSCVPGAGETS
                 ************************************************************

                                                                          #
AAA59992     181 YSESLCRLCRGDSSGEGVCDKSPLERYYDYSGAFRCLAEGAGDVAFVKHSTVLENTDGKT
NP201573         YSESLCRLCRGDSSGEGVCDKSPLERYYDYSGAFRCLAEGAGDVAFVKHSTVLENTD---
                 ****************************************************************

AAA59992     241 LPSWGQALLSQDFELLCRDGSRADVTEWRQCHLARVPAHAVVVRADTDGGLIFRLLNEGQ
NP201573         ------------------------------------------------------------

AAA59992     301 RLFSHEGSSFQMFSSEAYGQKDLLFKDSTSELVPIATQTYEAWLGHEYLHAMKGLLCDPN
NP201573         ------------------------------------------------------------

AAA59992     361 RLPPYLRWCVLSTPEIQKCGDMAVAFRRQRLKPEIQCVSAKSPQHCMERIQAEQVDAVTL
NP201573         ------------------------------------------------------------

AAA59992     421 SGEDIYTAGKKYGLVPAAGEHYAPEDSSNSYYVVAVVRRDSSHAFTLDELRGKRSCHAGF
NP201573         ------------------------------------------------------------

AAA59992     481 GSPAGWDVPVGALIQRGFIRPKDCDVLTAVSEFFNASCVPVNNPKNYPSSLCALCVGDEQ
NP201573         ------------------------------------------------------------

AAA59992     541 GRNKCVGNSQERYYGYRGAFRCLVENAGDVAFVRHTTVFDNTNGHNSEPWAAELRSEDYE
NP201573         -----------ESPSRRQTWTRSEEEEGECPAHEEARRTMRSSAGQAWKWAPVHRPQDES
                           . . * ::    *: *: .  ..:    .:.. ::  **.  *.:* .

AAA59992     601 LLCPNGARAEVSQFAACNLAQIPPHAVMVRPDTNIFTVYGLLDKAQDLFGDDHNKNGFKM
NP201573         DKGEFGKRAKSRDMLG-----------------------------------------------
                      * **:  :: .

AAA59992     661 FDSSNYHGQDLLFKDATVRAVPVGEKTTYRGWLGLDYVAALEGMSSQQCSGAAAPAPGAP
NP201573         ------------------------------------------------------------

AAA59992     721 LLPLLLPALAARLLPPAL
NP201573         ------------------
```

Fig. 2. Sequence alignment of the short and long MTf1 protein isoforms from GenBank. The identical regions correspond to the first six exons of the MTf1 gene, the end of which is indicated by #. The boxes indicate conserved Fe-binding residues and the consensus thermolysin metalloprotease site (LG**HE**YL**H**AMK) [1]. Notably, the fifth Fe-binding residue H (279) and the thermolysin metalloprotease site are absent in the short isoform. It is also of interest that the C-terminals of the short and long isoforms share some conservation. (*) identical, (:) conservation of residue type, (.) partial conservation of residue type. Classification of amino acids is the same as described by the CLUSTAL W Multiple Sequence Alignment Program v. 1.8 (www.clustalw.genome.ad.jp/SIT/clustalw.html).

motif was present within the 3′-UTR. However, there was no evidence that a stem could form to produce the typical stem-loop structure of the IRE that is essential to bind IRPs [3].

Interestingly, using the computer program Sosui from the Expert Protein Analysis System (ExPASy) [4], the short MTf1 transcript was predicted to produce a soluble protein. In addition, two programs from ExPASy (the big-PI Predictor and DGPI) indicated there was no evidence that this short isoform would be GPI-anchored. Hence, it can be suggested that the short isoform may correspond to soluble MTf found in the serum of patients with Alzheimer's disease [2]. As a relevant control, the long MTf1 transcripts were also assessed and found to correctly predict a membrane-bound molecule attached via a GPI anchor.

A 'blast' search of the human genome using the MTf1 transcripts described above revealed a surprise, namely a second MTf gene (tentatively called MTf2), directly adjacent to MTf1 on chromosome 3. Diagrammatic representation of the two genes is shown in Fig. 1. MTf2 is remarkably similar to MTf1 in many respects, and may be the result of partial gene duplication. This tandem duplication was not observed for serum Tf or lactoferrin. The MTf1 and MTf2 genes share the same 56 and 110 bp minimal enhancers and minimal promoter from positions −204 to −1 [5]. In addition, the MTf2 gene contains exons 1–6, which are identical to the first six exons of MTf1. Furthermore, the intervening intronic regions are exactly the same. The region following exon 6 in MTf2 is composed of 702 bp of intronic sequence which is the same as that in MTf1. After this latter intron, the sequence then strongly deviates from that found in MTf1.

Similar to the short MTf1 transcript, only four of the five high affinity Fe-binding residues are present within the MTf2 gene, while the thermolysin metalloprotease consensus sequence found in the long MTf1 transcript [1] is absent from MTf2. The lack of conservation of the Fe-binding site and metalloprotease site suggests their redundancy.

The functions of the MTf1 protein isoforms and the putative MTf2 protein remain a mystery. While these molecules show high homology to Tf, evidence to date has not implicated a role for MTf in efficient cellular Fe transport [1]. In-

deed, the fact that there is no conservation of the MTf Fe-binding sites in the short MTf1 transcript, in MTf2, or in MTf identified from other organisms, could suggest that it has other functions [1]. However, the strong conservation of the sequence in the first six exons suggests that this may be vital for the biological function of MTf.

## References

[1] Sekyere, E. and Richardson, D.R. (2000) FEBS Lett. 483, 11–16.
[2] Kennard, M.L., Feldman, H., Yamada, T. and Jefferies, W.A. (1996) Nat. Med. 2, 1230–1235.
[3] Hentze, M.W. and Kühn, L.C. (1996) Proc. Natl. Acad. Sci. USA 93, 8175–8182.
[4] The Expert Protein Analysis System (ExPASy)-Proteomics Tools, www.expasy.org.
[5] Duchange, N., Ochoa, A., Plowman, G.D., Rozé, A., Amdjadi, M. and Zakin, M.M. (1992) Nucleic Acids Res. 20, 2853–2859.

*Corresponding author. Fax: (61)-2-9550 3302.
*E-mail address:* d.richardson@hri.org.au (D.R. Richardson).

*Heart Research Institute, The Iron Metabolism and Chelation Group, 145 Missenden Road, Sydney, NSW 2050, Australia*

PII: S0014-5793(02)02248-2