Nucleic Acids Research, 1997, Vol. 25, No. 3 455–457

Conceptual translation of *timeless* reveals alternative initiating methionines in *Drosophila*

Ezio Rosato, Annalisa Trevisan[1], Federica Sandrelli[1], Mauro Zordan[1], Charalambos P. Kyriacou* and Rodolfo Costa[1,2]

Department of Genetics, University of Leicester, Leicester LE1 7RH, UK, [1]Dipartimento di Biologia, Università di Padova, 35121 Padova, Italy and [2]Dipartimento di Biologia, Università di Lecce, 73100 Lecce, Italy

Received December 3, 1996; Accepted December 10, 1996

DDBJ/EMBL/GenBank accession nos Y09930–Y09937

ABSTRACT

We have sequenced genomic fragments which encode the N-terminus of the TIMELESS (TIM) clock protein in *Drosophila simulans* and *D.yakuba*. We observe that in these two species, the initiating methionine appears to lie downstream of the one proposed to encode the translational start in *D.melanogaster*, thereby truncating the N-terminus by 23 amino acids. We then sequenced the corresponding 5′ fragment in a number of *D.melanogaster* individuals from different strains. We observed a polymorphism which strongly suggests that the originally proposed start site cannot be utilised in some individuals, and that these flies will initiate translation of TIM at the downstream ATG. Given the current interest in TIM regulation in *D.melanogaster*, it is important to correctly define the N-terminus in this species.

INTRODUCTION

Circadian rhythmicity is a widespread and ancient adaptation of living organisms to the rotation of the earth. Mutations affecting circadian rhythms have been identified in several organisms ranging from prokaryotes (1) to plants (2) and mammals (3). In *Drosophila melanogaster* two clock genes which encode an integral component of the pacemaker have been cloned and characterized, *period* (*per*) (4–7) and *timeless* (*tim*) (8,9). It has been demonstrated that PER and TIM physically interact (9–11) and both proteins are involved in a mutual negative feedback regulation of their own expression (12). Moreover, *tim*$^+$ activity is necessary for the nuclear localization of a PER reporter (13).

The *D.melanogaster tim* cDNA sequence was obtained by sequencing several clones from independently derived head specific cDNA libraries as well as genomic DNA (8). The data reveal that the sequence contains two open reading frames (ORFs) coding for two conceptual proteins of 1122 or 1389 amino acids (8). Consequently, there may be two forms of the TIM protein. The two ORFs share the residues 1–1104 but have different C-termini depending on the removal or not of a particular intron.

As part of our continuing comparative analyses of clock genes in diptera (14–16), we designed a series of oligoprimers in order to amplify fragments of the *tim* gene from other species. Primers which focused on the region encoding the N-terminus, amplified a fragment in the *D.melanogaster* sibling species, *D.simulans*, and a more distantly related member of the *melanogaster* subgroup of species, *D.yakuba*. We observe in both species a different translational start from the one identified by Myers *et al*. (8) in *D.melanogaster*. Furthermore, the same alternative translational start is also observed in some *D.melanogaster* individuals. Given the current interest in PER and TIM (17,18) and the ongoing biochemical analysis of these two proteins in *D.melanogaster* (10,11), particularly with regard to TIM degradation, the N-terminus of TIM may play an important role in the regulation of the two proteins.

MATERIALS AND METHODS

Drosophila strains

Single *D.melanogaster* individuals were obtained from the following laboratory strains: Canton-S, *per*$^+$, *per*^{S} and *per*^{01} mutant lines. In addition, isofemale lines were collected from a natural population from Cognac in France (see ref. 19 for details). Individual male descendants were crossed to attached-*X* females to generate self perpetuating iso-*X* chromosome lines. Two of these lines, *Co-TG14* and *Co-TG20*, were used. Consequently the second chromosomes in these lines will carry either the natural Cognac-derived or the attached-*X* strain-derived *tim* allele.

A *D.yakuba* strain was obtained from J. David (University of Paris) and the *D.simulans* strain was obtained from the Bowling Green *Drosophila* Species Stock Center, USA.

PCR amplification and DNA sequencing

For the amplification of the *tim* fragment from *D.melanogaster* the 5′ primer used was 5′- CACAATCACATCTGGAATAA-3′ corresponding to nucleotide positions 264–283 in the sequence from Myers *et al*. (8). The 3′ primer corresponded to positions 384–403 (5′-GCATTCGGGTTGACCACATA-3′). For the amplification from *D.simulans* and *D.yakuba* the following degenerate primers were used: 5′ primer was 5′- CAYAAYCAY-ATHTGGAAYAA-3′ (264–283); 3′ primer was 5′-GCRTTSGG-RTTNACNACR-3′ (384–403).

Single fly DNA was prepared using the method of Gloor and Engels (20). DNA amplification by PCR was carried out according to Jeffreys *et al*. (21) in a MJ Research thermocycler

* To whom correspondence should be addressed. Tel: +44 116 2523430; Fax: +44 116 2523378; Email: cpk@leicester.ac.uk

```
D.mel.
(cDNA)      AATTTCATCA GTGCATATAA CAGCACTGAA ACTATAACAC GATCTATTCT GCAAAGAAAC  60

(cDNA)      CCAAAAAGTG CTCAGAAAAG CTCAATTGCT TAGAAACATA AACAATCAGC TTTAATTGTT 120

(cDNA)      GATTGCAATT CGGCTAAAAC TAAAACTAAA ACAGTAAAAT TGTCTGCGAT AGAAAAAATT 180

(cDNA)      TAAATAATTG TTACAGATAC CGCGCAAATG GCTAAGAAGT ACCTCAATGT TCGCAGTCGA 240

(cDNA)      CAATGAGCAG AGTTAGGCAG CTCCACAATC ACATCTGGAA TAATCAGAAC TTTGATAAAG 300
Canton-S                          ....... .......... .......... ..........
Co-TG14                           ....... .......... .......... ..........
Co-TG20                           ....... .......... .......... ..........
per+                              ....... .......... .......... ...-..C...
perS                              ....... .......... .......... ..........
per01                             ....... .......... .......... ...-..C...
D.sim                             ....... .......... .......... ...-..C...
D.yak                             ....... .......... .......G.. ...-..C...

(ref)       tgaAATCGGT TATGGACTGG TTACTAGCAA CTCCGCAGTT GTACAGCGCG TTCTCCTCCT 360
Canton-S    .......... .......... .......... .......... .......... ..........
Co-TG14     .......... .......... .......... .......... .......... ..........
Co-TG20     .......... .......... .......... .......... .......... ..........
per+        .......... .......... .......... .......... .......... ..........
perS        .......... .......... .......... .......... .......... ..........
per01       .......... .......... .......... .......... .......... ..........
D.sim.      .......... .......... .......... .......... .......... ........A.
D.yak.      .......... .......... .......... .C........ .C........ ........T.

(ref)       TGGGTTGCTT GGAGGGCGAT ACCTATGTGG TCAACCCGAA TGC  403
Canton-S    .......... .......... .......... .......... ...
Co-TG14     .......... .......... .......... .......... ...
Co-TG20     .......... .......... .......... .......... ...
per+        .......... .......... .......... .......... ...
perS        .......... .......... .......... .......... ...
per01       .......... .......... .......... .......... ...
D.sim       .......... .......... .......... .......... ...
D.yak       ....C..... ...A..G... .......... .......... ...
```

Figure 1. Nucleotide sequences of a *tim* fragment (139–140 bp) from six different laboratory lines of *D.melanogaster* and from one strain of *D.simulans* and *D.yakuba* (see Material and Methods). The fragments were obtained by PCR with oligos amplifying from positions 264 to 403 in the cDNA sequence of Myers*et al.* (8) which is shown above for comparison from nucleotide 1 to 403. The ATG motif in bold is the downstream translation initiation codon for*D.simulans*, *D.yakuba* and for the *D.melanogaster per*^{01} and *per*$^{+}$ sequences. The ATG underlined represents the upstream initiating motif previously proposed by Myers*et al.* (8). Two further upstream ATG sequences which are represented in bold italics generate premature translational stops. The tga motif in lower case represents the stop codon which is generated by the G deletion in position 294. The boxed area corresponds to the position of the Kozak sequence (see text).

for 30 cycles: 95°C for 1 min, 65°C for 1 min, 72°C for 1 min for *D.melanogaster* and 95°C for 1 min, 48°C for 1 min 10 s, 72°C for 1 min 20 s, for *D.simulans* and *D.yakuba*. BioTaq polymerase from Bioline was used.

The PCR amplified fragments to be sequenced were cloned into the pMOSBlueT-vector (Amersham). The recombinant DNAs were used to transform MOSBlue competent cells (Amersham). A minimum of three clones was sequenced from two independent PCR amplifications for each fly's DNA. The DNA sequencing was carried out either using the ThermoSequenase cycle sequencing kit US78500 (Amersham) or by an automatic system using the Perkin-Elmer Turbo-Catalyst robotic station and ABI 377 sequencing apparatus. The M13R and –40 oligos were used to sequence the cloned fragment.

RESULTS

Figure 1 illustrates the genomic sequences obtained from the 139 or 140 bp *tim* fragment. The 5′ stretch (from nucleotide 1 to 403) of the *tim* cDNA sequence of *D.melanogaster* obtained by Myers *et al.* (8) is also reported for comparison (GenBank accession number U37018). Note that there are a number of ATG motifs. The first two (represented in bold italics in Fig. 1, positions 208–210 and 227–229) give rise to premature translational stop codons, and the third (positions 243–245, underlined in Fig. 1), has been suggested by Myers *et al.* (8) to represent the initiating motif.

The *D.melanogaster* genomic *tim* sequences show both nucleotide and length polymorphisms. In particular, the *tim* sequences in *D.melanogaster* lines *per*$^{+}$ and *per*^{01} show a substitution (A→C) in position 297 and a single base deletion (G) in position 294, when compared with the other *D.melanogaster* genomic and cDNA sequences. The *D.simulans* sequence differs from the *D.melanogaster* cDNA adopted as a reference, with nucleotide substitutions in position 297 (A→C) and in position 359 (C→A). The same single base deletion (G) in position 294 observed in the *D.melanogaster per*$^{+}$ and *per*^{01} lines, is also found in *D.simulans*. Finally, the *D.yakuba* sequence shows substitutions in positions 288 (A→G), 297 (A→C), 332 (T→C), 342 (T→C), 359 (C→T), 365 (T→C), 374 (G→A), 377 (C→G) and, again, the single base deletion (G) in position 294. Among the intra- and interspecific nucleotide polymorphisms observed in the sequences reported in Figure 1, only one produces an amino acid substitution (position 342), giving the Y→H substitution present in the putative N-terminus region of the *yakuba* TIM protein (Fig. 2).

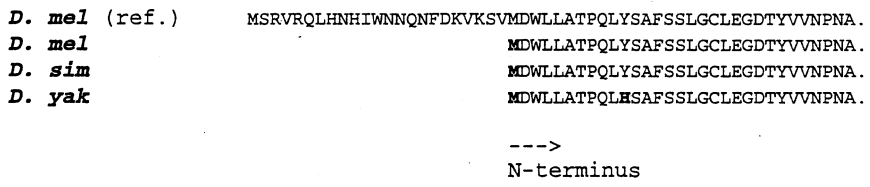

Figure 2. Putative N-terminus of the TIM protein for three Drosophila species (*D.melanogaster*, *D.simulans* and *D.yakuba*). The *D.melanogaster* TIM sequence indicated as a reference (ref.) is the one proposed by Myers et al. (8) from cDNA analysis. The other protein sequences are derived from the data in Figure 1 which suggest a different initiating methionine (in bold) for the TIM protein in the *D.melanogaster* strains per^{+} and per^{01}, and *D.simulans* and *D.yakuba*. The amino acid substitution observed in *D.yakuba* (Y to H) is also highlighted in bold.

Consequently, the G deletion in position 294 and the A→C substitution in position 297 are shared by *D.simulans* and *D.yakuba*, as well as the *tim* sequences in the *D.melanogaster* per^{+} and per^{01} strains. Moreover, the first and second ATG motifs in the cDNA (in bold italics in Fig. 1) generate precocious stop codons. The sequence of the *D.melanogaster* TIM putative protein published by Myers *et al.* (8) represents the translation of an ORF which starts from the third ATG motif (nucleotide positions 243–245, underlined in Fig. 1) present in the cDNA sequence. If we consider the reading frame generated by this ATG, it is dramatically affected by the G deletion in position 294, causing a frame-shift and a consequent premature translational stop codon (TGA, nucleotides 301–303, lower case in Fig. 1). Consequently, the next downstream in-frame ATG is found in positions 312–314 (bold in Fig. 1).

DISCUSSION

The intraspecific and interspecific pattern of sequences described above, strongly suggest that the first methionine of the TIM protein is coded by the ATG motif (positions 312–314, represented in bold in Fig. 1). This is located downstream from position 294 where the polymorphism for the deletion of the corresponding base (G) has been observed. This scenario would maintain the same reading frame initiated by the upstream ATG motif (position 243–245) suggested by Myers *et al.* (8), but would truncate the N-terminus by 23 amino acids (Fig. 2).

The sequences surrounding the downstream ATG (positions 306–315, boxed in Fig. 1) are in agreement with the most stringent requirements dictated by Kozak's consensus translation initiation sequence (GCC**R**CCATG**G**) (22,23). In fact, in this case, a purine (G) is present three bases before the ATG motif, and a G immediately following it. These two bases are the most important in the consensus sequence as they can influence the efficiency of translation 10-fold, while the other positions have much smaller effects (22). The upstream ATG motif favoured by Myers *et al.* (8) does not have the same characteristics, since a purine (A) is present three bases before it, but there is an A instead of a G immediately downstream. We cannot exclude that TIM protein synthesis occurs with more than one N-terminus, through a context-dependent leaky scanning mechanism (23) in those *D.melanogaster* individuals which do not show the G deletion at position 294. However, for those individuals of *D.melanogaster* and *D.simulans* and *D.yakuba* carrying the deletion, the putative TIM protein(s) should be 23 amino acids 'shorter' than previously predicted, as shown in Figure 2.

The regulation of PER and TIM is currently of great interest (10,11,17,18) and the stability and degradation of these two proteins has major implications for the negative feedback model of circadian timing (24). A precise definition of the N-terminus of TIM may therefore be of considerable importance.

ACKNOWLEDGEMENTS

We acknowledge a NERC grant to CPK, a CEC grant from the Biotechnology programme (ERBBIO4CT960096), and a MURST-British Council grant to CPK and RC for Anglo-Italian cooperation, and a MURST studentship to FS.

REFERENCES

1 Kondo,T., Strayer,C.A., Kulkarni,R.D., Taylor,W., Ishiura,M., Golden,S.S., Johnson,C.H. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 5672–5676.
2 Hicks,K.A., Millar,A.J., Carre,I.A., Somers,D.E., Straume,M., Meeks-Wagner,D. R and Kay,S.A. (1996) *Science*, **274**, 790–792.
3 Takahashi,J.S., Pinto,L.H. and Vitaterna,M.H. (1994), *Science*, **264**, 1724–1733
4 Bargiello,T.A. and Young,M.W. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 2142–2146.
5 Reddy,P., Zehring,D.A., Wheeler,D.A., Pirrotta,V., Hadfield,C., Hall,J.C., Rosbash,M. (1984) *Cell*, **38**, 701–710.
6 Jackson,F.R., Bargiello,T.A., Yun,S.H. and Young,M.W. (1986) *Nature*, **320**, 185–188.
7 Citri,Y., Colot,A.C., Jacquier,A.C., Yu,Q., Hall,J.C., Baltimore,D., Rosbash,M. (1987) *Nature*, **326**, 42–47.
8 Myers,M.P., Wager-Smith,K., Wesley,C.S., Young,M.W., Sehgal,A. (1995) *Science*, **270**, 805–808.
9 Gekakis,N., Saez,L., Delahaye-Brown,A.M., Myers,M.P., Sehgal,A., Young,M.W., Weitz,C.J. (1995). *Science*, **270**, 811–815.
10 Zeng,H., Qian,Z., Myers,M.P. and Rosbash,M. (1996) *Nature*, **380**, 129–135.
11 Lee,C., Parikh,V., Itsukaichi,T., Bae,K. and Edery,I. (1996). *Science*, **271**, 1740–1744.
12 Seghal,A., Rothenfluh-Hilfiker,A., Hunter-Ensor,M., Chen,Y., Myers,M. and Young,M.W. (1995). *Science*, **270**, 808–810
13 Vosshall,L., Price,J., Sehgal,A., Saez,L., Young,M. (1994) *Science*, **263**, 1603–1606.
14 Peixoto,A.A., Campesan,S., Costa,R. and Kyriacou,C.P. (1993). *Mol. Biol. Evol.*, **10**, 127–139.
15 Nielsen,J., Peixoto,A.A., Piccin,A., Costa,R., Kyriacou,C.P. and Chalmers,D. (1994). *Mol. Biol. Evol.*, **11**, 839–853.
16 Rosato,E., Peixoto,A.A., Gallippi,A., Kyriacou,C.P., Costa,R. (1996) *J. Mol. Evol.*, **42**, 392–408.
17 Myers,M.P., Wager-Smith,K., Rothenfluh-Hilfiker,A. and Young,M.W. (1996). *Science*, **271**, 1736–1740.
18 Hunter-Ensor,M., Ousley,A. and Sehgal,A. (1996) *Cell*, **84**, 677–685.
19 Costa,R., Peixoto,A.A., Barbujani,G., Kyriacou,C.P. (1992) *Proc.R. Soc. Lond. Biol.*, **250**, 43–49.
20 Gloor,G. and Engels,W. (1990) *Dros. Inf. Newsletter*, Vol. 1.
21 Jeffreys,A.J., Wilson,V., Neumann,R. and Kyte,J. (1988) *Nucleic Acids Res.*, **16**, 10953–10971.
22 Kozak,M. (1987) *Nucleic Acids Res.*, **15**, 8125–8132.
23 Kozak,M. (1991) *J. Cell Biol.*, **115**, 887–903.
24 Goldbeter,A. (1995) *Proc. Roy. Soc. Lond.B.*, **261**, 319–324.