Nucleic Acids Research, 2002, Vol. 30, No. 1 335–340

UTRdb and UTRsite: specialized databases of sequences and functional elements of 5′ and 3′ untranslated regions of eukaryotic mRNAs. Update 2002

Graziano Pesole*, Sabino Liuni[1], Giorgio Grillo[1], Flavio Licciulli[2], Flavio Mignone, Carmela Gissi and Cecilia Saccone[2]

Dipartimento di Fisiologia e Biochimica Generali, Università di Milano, via Celoria 26, 20133 Milano, Italy, [1]Centro di Studio sui Mitocondri e Metabolismo Energetico del Consiglio Nazionale delle Ricerche (CNR), via Amendola 165/A, 70126 Bari, Italy and [2]Dipartimento di Biochimica e Biologia Molecolare, Università di Bari, via Orabona 4, 70126 Bari, Italy

Received September 21, 2001; Accepted October 2, 2001

ABSTRACT

The 5′- and 3′-untranslated regions (5′- and 3′-UTRs) of eukaryotic mRNAs are known to play a crucial role in post-transcriptional regulation of gene expression modulating nucleo-cytoplasmic mRNA transport, translation efficiency, subcellular localization and stability. UTRdb is a specialized database of 5′ and 3′ untranslated sequences of eukaryotic mRNAs cleaned from redundancy. UTRdb entries are enriched with specialized information not present in the primary databases including the presence of nucleotide sequence patterns already demonstrated by experimental analysis to have some functional role. All these patterns have been collected in the UTRsite database so that it is possible to search any input sequence for the presence of annotated functional motifs. Furthermore, UTRdb entries have been annotated for the presence of repetitive elements. All Internet resources we implemented for retrieval and functional analysis of 5′- and 3′-UTRs of eukaryotic mRNAs are accessible at http://bighost.area.ba.cnr.it/BIG/UTRHome/.

INTRODUCTION

The completion of the sequencing of human and of other organism genomes has opened new avenues for understanding the basic mechanisms of cell function. These processes mostly rely on a spatial–temporal coordinated expression of genes mediated by regulatory elements embedded in the non-coding part of the genomes. Among non-coding regions, the 5′- and 3′-untranslated regions (5′- and 3′-UTRs) of eukaryotic mRNAs have often been experimentally demonstrated to contain sequence elements crucial for many aspects of gene regulation and expression (1–7).

Table 1. Number of entries (N) and nucleotide length (L) of UTRdb collections (release 15.0) after redundancy cleaning

Collection			Redundancy	
	N	L	%N	%L
5′-UTR				
Fungi	2223	275 886	27.76	16.74
Human	30 922	4 515 966	40.06	22.33
Invertebrate	19 947	2 987 661	28.06	18.60
Other_mammal	5751	852 910	35.51	14.36
Other_vertebrate	7327	792 573	8.61	15.00
Plant	17 819	1 490 067	25.64	13.51
Rodent	19 759	2 518 594	36.76	20.22
Virus	14 663	3 402 809	81.71	73.82
Total	118 411	16 836 466	–	–
3′-UTR				
Fungi	2304	465 396	14.84	11.30
Human	36 015	18 906 357	41.72	29.83
Invertebrate	18 230	5 151 363	36.74	20.92
Other_mammal	6927	2 548 434	29.78	17.55
Other_vertebrate	8528	3 351 751	21.22	13.37
Plant	21 526	4 328 226	16.93	13.28
Rodent	21 489	9 113 464	37.32	23.30
Patent	14 118	3 359 534	77.40	69.42
Total	129 137	47 224 525	–	–

UTRdb 15.0 was generated from EMBL release 67. Relevant redundancy percentages calculated with respect to the number of entries (%N) and to the nucleotide length (%L) are also indicated.

*To whom correspondence should be addressed. Tel: +39 02 58354915; Fax +39 02 58354912; Email: graziano.pesole@unimi.it

```
ID   5HSA012029 standard; DNA; HUM; 253 BP.
XX
AC   BB046362;
XX
DT   14-OCT-1998 (Rel. 9, Created)
DT   14-OCT-1998 (Rel. 9, Last updated, Version 1)
XX
DE   5'UTR in Homo sapiens chromosome 7q22 sequence, complete sequence.
XX
DR   EMBL; AF053356;
DR   UTR; CC052750;
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria;
OC   Primates; Catarrhini; Hominidae; Homo.
XX
UT   5'UTR; Complete; 2 exon(s)
XX
FH   Key             Location/Qualifiers
FH
FT   5'UTR           join(complement(EMBL::AF053356:101757..101920),
FT                   complement(EMBL::AF053356:99415..99503))
FT                   /product="GNB2"
FT   5'TOP           complement(EMBL::AF053356:101916..101920)
FT                   /evidence="Pattern Similarity"
FT                   /standard_name="Ribosomal mRNA 5'terminal oligopyrimidine Element"
FT                   /db_xref="UTRsite:U0010"
FT   repeat_region   complement(EMBL::AF053356:101820..101861)
FT                   /evidence="Pattern Similarity"
FT                   /repeat_type="(CCG)n"
FT                   /repeat_family="Simple_repeat"
XX
SQ   Sequence 253 BP; 29 A; 127 C; 82 G; 15 T; 0 other;
     ctccgctctg gggaggcagc gctggcggcg gggctggggc cactgaggaa atccatccgc        6
     gccgccgccg ccgccgccgc cgccgccgcc gcctccgccg cggaggaaga cagcgccgcc       12
     cgcgcaccgc cagcgacccc cgccgcagag tcccaccgcc acaggcctcg ggccagcggc       18
     caggagctgc ctcccccagc ccccgtcccg cggcccccag ccgcccccaa ccctgcccca       24
     cgggcccggc gcc                                                         25
//
```

Figure 1. Sample entry of UTRdb. Specialized information not present in the primary EMBL/GenBank database is shown in bold with active crosslinks with other databases underlined. The 'UT' line reports information regarding the relevant UTR entry (e.g. complete or partial) as well as the number of spanned exons in the case of genomic DNA sequences. The presence in this sequence entry of a '5′ ribosomal mRNA TOP' (33–35) (UTRsite entry: U0010) and of a microsatellite element has been also annotated.

The main functional roles so far demonstrated for 5′- and 3′-UTR sequences are: (i) control of mRNA cellular and subcellular localization (4,7–9); (ii) control of mRNA stability (1,10,11); (iii) control of mRNA translation efficiency (12–14).

Several regulatory signals have already been identified in 5′- or 3′-UTR sequences, usually corresponding to short oligonucleotide tracts, also able to fold in specific secondary structures, which are protein binding sites for various regulatory proteins.

The analysis of large collections of functionally equivalent sequences (15,16), such as 5′- and 3′-UTR sequences, could indeed be very useful for defining their structural and compositional features as well as for searching the alleged function-associated sequence patterns (17–19). For this reason we constructed UTRdb, a specialized sequence collection, deprived from redundancy, of 5′- and 3′-UTR sequences from eukaryotic mRNAs.

UTRdb entries have been enriched with specialized information, not present in the primary databases, including the presence of sequence patterns demonstrated by experimental evidence to play some functional role. Additionally, because ~10% of mammalian mRNAs contain repetitive elements in their UTRs (20) but they are usually not annotated in the original records, we decided to add this information into our database as well.

We also created UTRsite, a collection of functional sequence patterns located in the 5′- or 3′-UTR sequences which could prove very useful for automatic annotation of anonymous sequences generated by sequencing projects as well as for finding previously undetected signals in known gene sequences.

UTRdb GENERATION

The specialized database of UTR sequences was generated by UTRdb_gen, a computer program we devised for this task. Eight sequence collections were generated for both 5′- and 3′-UTR sequences, one for each of the eukaryotic division of the EMBL/GenBank nucleotide database, namely: (i) human, (ii) rodent, (iii) other mammal, (iv) other vertebrate, (v) invertebrate, (vi) plant, (vii) fungi and (viii) virus.

UTRdb_gen, performing an accurate parsing of the Feature Table of the relevant EMBL entries is able to automatically generate the various UTRdb collections. Although the feature keys '5′UTR' and '3′UTR' is a valid feature for the EMBL/GenBank entries,

Table 2. Functional patterns included so far in UTRsite. For each pattern the number of hits with non-redundant UTRdb entries is also reported

Functional patterns	Reference	Hits found in UTRdb 15.0
Iron-responsive element (IRE)	(24)	110
Histone 3′-UTR stem–loop structure	(25)	38
AU-rich class II destabilising element	(26)	66
Tra-2 and GLI translational regulation element (TGE)	(27)	81
Selenocysteine insertion sequence (SECIS)	(28–30)	2002
Amyloid precursor protein 3′-UTR stability control element	(31)	15
Cytoplasmatic polyadenylation element (CPE)	(32)	5184
Nanos translation control element		1
Ribosomal protein mRNA 5′ TOP	(33–35)	269
TNF mRNA translation repression element	(36)	8
Vimentin 3′-UTR mRNA element	(37)	6
GLUT1 mRNA stabilising element	(38)	66
Internal ribosome entry site (IRES)	(39)	7353
5′-UTR Msl-2	(40)	5
3′-UTR Msl-2	(40)	18
RpmS12 translational control element	(41)	2
Bruno responsive element (BRE)	(42,43)	196
Barley yellow dwarf virus (BYDV) element	(44)	6
ADH 3′-UTR down-regulation control element	(45)	61
15-LOX-DICE	(46)	90
Upstream ORFs	(47)	27 897
Repetitive elements		38 823

only a small percentage of the entries are adequately annotated. Indeed, of the about 250 000 primary entries where UTRdb_gen was able to extract 5′- or 3′-UTR sequences, only 12% contained the 5′UTR or 3′UTR feature key in the corresponding EMBL entry. UTRdb_gen is able to define UTRs, even when these keys are not reported in the primary entry by using a predefinite syntactic parsing of other relevant feature keys, such as mRNA, CDS, exon, intron, etc.

UTRdb_gen automatically annotates generated UTR entries by adding some specialized information such as completeness or incompleteness of the UTR, number of spanned exons and cross-referencing to the primary database entry. A cross reference between 5′- and 3′-UTR sequences from the same mRNA has also been established.

A further interface between the UTRdb_gen and the BLAST engine (parameters: expect < 10^{-5}, minimum length = 50 nt, percentage identity > 95%) adds information about the position and the identity of any vector that may contaminate UTR entries.

The generation of UTR entries cleaned from redundancy has been obtained by using CLEANUP program (21) which is able to generate automatically, and very quickly, cleaned collections by removing entries that have a similarity and overlapping degree with longer entries present in the database above a user-fixed threshold. In this case, the cut-off parameters we used for the CLEANUP application were 95% for similarity and 90% for overlapping.

The specialized information included in UTR entries is generated by using two programs: (i) UTRnote including information about the location of experimentally defined patterns collected in UTRsite and (ii) UTRrepeat (which uses RepeatMasker) including repetitive elements present in the Repbase database (19). The UTRsite entries describe the various regulatory elements present in UTRs whose functional role has been established on an experimental basis. Each UTRsite entry is constructed on the basis of information reported in the literature and revised by distinguished scientists experimentally working on the functional characterization of the relevant UTR regulatory element.

CONTENT OF UTRdb

Table 1 reports a summary description of UTRdb (release 15.0) which in total contains 247 548 entries and 64 060 991 nt. On average >35% of entries resulted to be redundant and were then removed from the database. Vector contamination was found in 188 and 196 entries of 5′- and 3′-UTRs, respectively.

5′-UTR sequences were defined as the mRNA region spanning from the cap site to the starting codon (excluded), whereas 3′-UTR sequences were defined as the mRNA region spanning from the stop codon (excluded) to the poly(A) starting site.

A sample entry of UTRdb is shown in Figure 1. The UTRdb entries have been formatted according to the EMBL database format.

Table 2 reports functional patterns and repetitive elements included in UTRsite. More entries will be included in further releases. A sample UTRsite entry is reported in Figure 2. Functional patterns, defined on the basis of the information reported in the

<Entry>
IRON-RESPONSIVE ELEMENT; U0002
<Description>
The "iron-responsive element" (IRE) is a particular hairpin structure located in the 5′-untranslated region (5′-UTR) or in the 3′-untranslated region (3′-UTR) of various mRNAs coding for proteins involved in cellular iron metabolism. The IREs are recognized by trans-acting proteins known as Iron Regulatory Proteins (IRPs) that control mRNA translation rate and stability. Two closely related IRPs, denoted as IRP-1 and IRP-2, have been identified so far which bind IREs and become inactivated (IRP-1) or degradated (IRP-2) when the iron level in the cell increases. IRPs show a significant degree of similarity to mitochondrial aconitase (EC 4.2.1.3). It has been shown that under high iron conditions IRP-1, which contains a 4Fe-4S cluster that possibly acts as a cellular iron biosensor, has enzymatic activity and may act as a cytosolic aconitase.
Cellular iron homeostasis in mammalian cells is maintained by the coordinate regulation of the expression of "Transferrin receptor", which determines the amount of iron acquired by the cell, and of "Ferritin", an iron storage protein, which determines the degree of intracellular iron sequestration. Thus if the cell requires more iron, the level of transferrin receptor has to increase and conversely the level of ferritin has to decrease.
Ferritin, in vertebrates, consists of 24 protein subunits of two types, type H with Mr of 21 kDa and type L with Mr of 19-20 kDa. The apoprotein (Mr 450 kDa) is able to store up to 4500 Fe (III) atoms.
The 5′-UTR of H- and L ferritin mRNAs contain one IRE whereas multiple IREs are located in the 3′-UTR of transferrin receptor mRNA.
In the case of low iron concentration, IRPs are able to bind the IREs in the 5′-UTR of H- and L-Ferritin mRNAs repressing their translation and the IREs in the 3′-UTR of transferrin mRNA increasing its stability. Conversely, if iron concentration is high, IRP binding is diminished, which increases translation of ferritins and downregulate expression of the transferrin receptor.
IREs have also been found in the mRNAs of other proteins involved in iron metabolism like "erythroid 5-aminolevulinic-acid synthase (eALAS) " involved in heme biosynthesis, the mRNA encoding the mitochondrial aconitase (a citric acid cycle enzyme) and the mRNA encoding the iron-sulfur subunit of succinate dehydrogenase (another citric acid cycle enzyme) in *Drosophila melanogaster*.
Two alternative IRE consensus (type A or type B) have been found. In certain IREs the bulge is best drawn with a single unpaired cytosine, whereas in others the cytosine nucleotide and two additional bases seem to oppose one free 3′ nucleotide. Some evidences also suggest a structured loop with an interaction between nucleotide one and nucleotide five (in boldcase).

```
 G  W        G  W
A    G      A    G
 C  H        C  H
  NN          NN
  NN          NN
  NN          NN
  NN          NN
  NN          NN
 C           C
  NN         N   N
  NN          N
  NN          NN
  NN          NN
  NN          NN
```

The lower stem can be of variable length and is AU-rich in transferrin mRNA. W=A,U and D=not G.
<Pattern>
r1={au,ua,gc,cg,gu,ug} ! r1 represents pairing rules
(p1=2...8 c p2=5...5 CAGWGH r1~p2 r1~p1 | p1=2...8 nnc p2=5...5 CAGWGH r1~p2 n r1~p1)
!(type A|type B)
<Bibliography>
Hentze MW and Kuhn LC (1996) Molecular control of vertebrate iron metabolism: mRNA based regulatory circuits operated by iron, nitric oxide, and oxidative stress. Proc. Natl. Acad. Sci. USA 93: 8175-8182.

Figure 2. Sample entry of UTRsite describing the 'Iron responsive element (IRE)' (24). The IRE functional pattern, which consists of both primary and secondary structure information, is described in the 'Pattern' section according to the format adopted by PatSearch program (22).

literature and/or advice by the scientists expert in the field, were described by using the pattern description syntax used in PatSearch program (22).

AVAILABILITY OF UTRdb

UTRdb and UTRsite are publicly available by anonymous FTP (ftp://area.ba.cnr.it/pub/embnet/database/utr/). All internet resources we implemented for retrieval and functional analysis of 5′- and 3′-UTR sequences are accessible at http://bighost.area.ba.cnr.it/BIG/UTRHome/. These include SRS retrieval (23) of UTRdb and UTRsite, also available at the EBI World Wide Web server (http://srs.ebi.ac.uk:80/), UTRscan and UTRblast. The UTRscan utility allows the enquirer to search user submitted sequences for any of the patterns collected in UTRsite. The UTRblast utility allows database searches against fully annotated UTRdb entries.

CONCLUSIONS AND PERSPECTIVES

The important role that UTRs of eukaryotic mRNAs may play in gene regulation and expression is now widely recognized. Indeed, experimental studies have demonstrated that sequence motifs located in the UTRs are involved in crucial biological functions.

The huge amount of functionally equivalent sequences stored in UTRdb now makes possible the study of their structural and compositional features and the application of

statistical methods for the identification of significant signals. Previous cleaning-up of databases is however necessary to avoid artefacts caused by redundant sequences. Even if statistical significance does not necessarily mean biological significance, it may provide useful indication for further experimental work, such as site-directed mutagenesis.

UTRdb will be updated with the new EMBL database releases and UTRsite will be continuously updated by adding new entries describing functional patterns whose biological role has been experimentally demonstrated.

ACKNOWLEDGEMENTS

For revision of UTRsite entries we would like to thank Jim Malter (APP 3′-UTR stability control element), Alain Krol (SECIS), Matthias Hentze (IRE, 15-LOX DICE and msl-2), Bill Marzluff (histone stem–loop structure), Ann-Bin Shyu (ARE), Arturo Verrotti (CPE), Elizabeth Goodwin (TGE), Roger Kaspar (ribosomal protein mRNA TOP), Danuta Radzioch (TNF mRNA translation repression element), Ruben Boado (GLUT1 mRNA stabilizing element), Zendra E. Zehner (Vimentin 3′-UTR mRNA element), Shu-Yun Le (IRES), Anne Ephrussi (BRE), Howy Jacobs (rpmS12), Allen Miller (BYDV), John Parsch (adh DRE). This work was supported by Ministero dell'Istruzione e Ricerca, Italy [projects: Bioinformatics and Genomic Research (COFIN99), Programma 'Biotecnologie' (legge 95/95 – 5%), Programma 'Studio di geni di interesse biomedico e agroalimentare' (CEGBA)].

REFERENCES

1. Decker,C.J. and Parker,R. (1994) Mechanism of mRNA degradation in eukaryotes *Trends Biochem. Sci.*, **19**, 336–340.
2. Kaufman,R.J. (1994) Control of gene expression at the level of translation initiation. *Curr. Opin. Biotechnol.*, **5**, 550–557.
3. Klausner,R.D., Rouault,T.A. and Harford,J.B. (1993) Regulating the fate of mRNA: the control of cellular iron metabolism. *Cell*, **72**, 19–28.
4. Singer,R.H. (1992) The cytoskeleton and mRNA localization. *Curr. Opin. Cell Biol.*, **4**, 15–19.
5. Wilhelm,J.E. and Vale,R.D. (1993) RNA on the move: the mRNA localization pathway. *J. Cell Biol.*, **123**, 269–274.
6. McCarthy,J.E.G. and Kollmus,H. (1995) Cytoplasmic mRNA–protein interactions in eukaryotic gene expression. *Trends Biochem. Sci.*, **20**, 191–197.
7. Bashirullah,A., Cooperstock,R.L. and Lipshitz,H.D. (1998) RNA localization in development. *Annu. Rev. Biochem.*, **67**, 335–394.
8. Johnston,D. (1995) The intracellular localization of messenger RNAs. *Cell*, **81**, 161–170.
9. Jansen,R.P. (2001) mRNA localization: message on the move. *Nat. Rev. Mol. Cell. Biol.*, **2**, 247–256.
10. Beelman,C.A. and Parker,R. (1995) Degradation of mRNA in eukaryotes. *Cell*, **81**, 179–183.
11. Mitchell,P. and Tollervey,D. (2001) mRNA turnover. *Curr. Opin. Cell Biol.*, **13**, 320–325.
12. Curtis,D., Lehman,R. and Zamore,P.D. (1995) Translational regulation in development. *Cell*, **81**, 171–178.
13. Sonenberg,N. (1994) mRNA translation: influence of the 5′ and 3′ untranslated regions. *Curr. Opin. Genet. Dev.*, **4**, 310–315.
14. Macdonald,P. (2001) Diversity in translational regulation. *Curr. Opin. Cell Biol.*, **13**, 326–331.
15. Mengeritsky,G. and Smith,T.F. (1987) Recognition of characteristic patterns in sets of functionally equivalent DNA sequences. *Comput. Appl. Biosci.*, **3**, 223–227.
16. Konopka,A.K. (1994) In Smith,D.W. (ed.), *Informatics and Genome Projects*. Academic Press, San Diego, CA.
17. Pesole,G., Liuni,S., Grillo,G. and Saccone,C. (1997) Structural and compositional features of untranslated regions of eukaryotic mRNAs. *Gene*, **205**, 95–102.
18. Pesole,G., Grillo,G. and Liuni,S. (1996) Databases of mRNA untranslated regions for Metazoa. *Comput. Chem.*, **20**, 141–144.
19. Pesole,G., Fiormarino,G. and Saccone,C. (1994) Sequence analysis and compositional properties of untranslated regions of human mRNAs. *Gene*, **140**, 219–225.
20. Makalowski,W., Zhang,J. and Boguski,M. (1996) Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences. *Genome Res.*, **6**, 846–857.
21. Grillo,G., Attimonelli,M., Liuni,S. and Pesole,G. (1996) CLEANUP: a fast computer program for removing redundancies from nucleotide sequence databases. *Comput. Appl. Biosci.*, **12**, 1–8.
22. Pesole,G., Liuni,S. and D'Souza,M. (2000) PatSearch: a pattern matcher software that finds functional elements in nucleotide and protein sequences and assesses their statistical significance. *Bioinformatics*, **16**, 439–450.
23. Etzold,T., Ulyanov,A. and Argos,P. (1996) SRS: information retrieval system for molecular biology data banks. *Methods Enzymol.*, **266**, 114–128.
24. Hentze,M.W. and Kuhn,L.C. (1996) Molecular control of vertebrate iron metabolism: mRNA-based regulatory circuits operated by iron, nitric oxide, and oxidative stress. *Proc. Natl Acad. Sci. USA*, **93**, 8175–8182.
25. Williams,A.S. and Marzluff,W.F. (1995) The sequence of the stem and flanking sequences at the 3′ end of histone mRNA are critical determinants for the binding of the stem–loop binding protein. *Nucleic Acids Res.*, **23**, 654–662.
26. Chen,C. and Shyu,A. (1995) AU-rich elements: characterization and importance in mRNA degradation. *Trends Biochem. Sci.*, **20**, 465–470.
27. Goodwin,E.B., Okkema,P.G., Evans,T.C. and Kimble,J. (1993) Translational regulation of tra-2 by its 3′-untranslated region controls sexual identity in *C. elegans*. *Cell*, **75**, 329–339.
28. Hubert,N., Walczak,R., Sturchler,C., Schuster,C., Westhof,E., Carbon,P. and Krol,A. (1996) RNAs mediating cotranslational insertion of selenocysteine in eukaryotic selenoproteins. *Biochimie*, **78**, 590–596.
29. Walczak,R., Westhof,E., Carbon,P. and Krol,A. (1996) A novel RNA structural motif in the selenocysteine insertion element of eukaryotic selenoprotein mRNAs. *RNA*, **2**, 367–379.
30. Fagegaltier,D., Lescure,A., Walczak,R., Carbon,P. and Krol,A. (2000) Structural analysis of new local features in SECIS RNA hairpins. *Nucleic Acids Res.*, **28**, 2679–2689.
31. Zaidi,S.H.E. and Malter,J.S. (1994) Amyloid precursor protein mRNA stability is controlled by a 29-base element in the 3′-untranslated region. *J. Biol. Chem.*, **269**, 24007–24013.
32. Verrotti,A., Thompson,S., Wreden,C., Strickland,S. and Wickens,M. (1996) Evolutionary conservation of sequence elements controlling cytoplasmic polyadenylylation. *Proc. Natl Acad. Sci. USA*, **93**, 9027–9032.
33. Amaldi,F. and Pierandrei-Amaldi,P. (1997) TOP genes: a translationally controlled class of genes including those coding for ribosomal proteins. *Prog. Mol. Subcell. Biol.*, **18**, 1–17.
34. Kaspar,R.L., Kakegawa,T., Cranston,H., Morris,D.R. and White,M.W. (1992) A regulatory *cis* element and a specific binding factor involved in the mitogenic control of murine ribosomal protein L32 translation. *J. Biol. Chem.*, **267**, 508–514.
35. Morris,D.R., Kakegawa,T., Kaspar,R.L. and White,M.W. (1993) Polypyrimidine tracts and their binding proteins: regulatory sites for posttranscriptional modulation of gene expression. *Biochemistry*, **32**, 2931–2937.
36. Hel,Z., Di Marco,S. and Radzioch,D. (1998) Characterization of the RNA binding proteins forming complexes with a novel putative regulatory region in the 3′-UTR of TNF-α mRNA. *Nucleic Acids Res.*, **26**, 2803–2812.
37. Zehner,Z.E., Shepherd,R.K., Gabryszuk,J., Fu,T.F., Al-Ali,M. and Holmes,W.M. (1997) RNA–protein interactions within the 3′ untranslated region of vimentin mRNA. *Nucleic Acids Res.*, **25**, 3362–3370.
38. Boado,R.J. and Pardridge,W.M. (1998) Ten nucleotide *cis* element in the 3′-untranslated region of the GLUT1 glucose transporter mRNA increases gene expression via mRNA stabilization. *Brain Res. Mol. Brain Res.*, **59**, 109–113.
39. Le,S.Y. and Maizel,J.V.,Jr (1997) A common RNA structural motif involved in the internal initiation of translation of cellular mRNAs. *Nucleic Acids Res.*, **25**, 362–369.
40. Gebauer,F., Corona,D.F., Preiss,T., Becker,P.B. and Hentze,M.W. (1999) Translational control of dosage compensation in *Drosophila* by sex-lethal: cooperative silencing via the 5′ and 3′ UTRs of msl-2 mRNA is independent of the poly(A) tail. *EMBO J.*, **18**, 6146–6154.
41. Mariottini,P., Shah,Z.H., Toivonen,J.M., Bagni,C., Spelbrink,J.N., Amaldi,F. and Jacobs,H.T. (1999) Expression of the gene for mitoribosomal

protein S12 is controlled in human cells at the levels of transcription, RNA splicing, and translation. *J. Biol. Chem.*, **274**, 31853–31862.
42. Castagnetti,S., Hentze,M.W., Ephrussi,A. and Gebauer,F. (2000) Control of oskar mRNA translation by Bruno in a novel cell-free system from *Drosophila* ovaries. *Development*, **127**, 1063–1068.
43. Kim-Ha,J., Kerr,K. and Macdonald,P.M. (1995) Translational regulation of oskar mRNA by bruno, an ovarian RNA-binding protein, is essential. *Cell*, **81**, 403–412.
44. Guo,L., Allen,E. and Miller,W.A. (2000) Structure and function of a cap-independent translation element that functions in either the 3′ or the 5′ untranslated region. *RNA*, **6**, 1808–1820.
45. Parsch,J., Stephan,W. and Tanda,S. (1999) A highly conserved sequence in the 3′-untranslated region of the drosophila Adh gene plays a functional role in Adh expression. *Genetics*, **151**, 667–674.
46. Ostareck-Lederer,A., Ostareck,D., Standart,N. and Thiele,B. (1994) Translation of 15-lipoxygenase mRNA is inhibited by a protein that binds to a repeated sequence in the 3′ untranslated region. *EMBO J.*, **13**, 1476–1481.
47. Kozak,M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene*, **234**, 187–208.