Reinitiation of protein synthesis in *Escherichia coli* can be induced by mRNA *cis*-elements unrelated to canonical translation initiation signals

Alessandra Andrè[a], Antimina Puca[a], Federica Sansone[a], Anna Brandi[b], Giovanni Antico[a], Raffaele A. Calogero[a,b,*]

[a] *Dipartimento di Genetica, Biologia Generale e Molecolare, Università Federico II di Napoli, via Mezzocannone 8, 80134 Naples, Italy*
[b] *Dipartimento di Scienze Cliniche e Biologiche, Università di Torino, Ospedale S. Luigi, Regione Gonzole 10, 10043 Orbassano, Italy*

Received 12 January 2000

Edited by G. Cesareni

Abstract In Eubacteria, de novo translation of some internal cistrons may be inefficient or impossible unless the 5′ neighboring cistron is also translated (translational coupling). Translation reinitiation is an extreme case of translational coupling in which translation of a message depends entirely on the presence of a nearby terminating ribosome. In this work, the characteristics of mRNA *cis*-elements inducing the reinitiation process in *Escherichia coli* have been investigated using a combinatorial approach. A number of novel translational reinitiation sequences (TRSs) were thus identified, which show a wide range of reinitiation activities fully dependent on a translational coupling event and unrelated to the presence/absence of secondary structure or mRNA stability. Moreover, some of the isolated TRSs are similar to intercistronic sequences present in the *E. coli* genome.

Key words: Combinatorial library; Reinitiation; Protein synthesis; Eubacterium

1. Introduction

In all organisms, the initiation step represents the main control point in the translational regulation of gene expression [1–5]. Unlike initiation, termination of protein synthesis is not generally considered of particular importance in the control of gene expression; in Eubacteria, however, initiation of protein synthesis can be affected by termination events occurring near the initiation codon or, more generally, within the translation initiation region (TIR) [6]. 'Translational reinitiation' can be regarded as a special case of this type of phenomena commonly designated as 'translational coupling' [6–9]. More precisely, translation reinitiation is the event in which translational coupling between two cistrons is mediated by the same ribosome which terminates translation at a stop codon and restarts at an initiation codon which is very close (in space if not in the primary sequence) or even partially overlaps (e.g. **UAAUG**, **UGAUG**, **UAGUG**, **AUGA**) the termination codon of the 5′ neighboring cistron [10–12]. During translational reinitiation, the terminating ribosome is believed to move bidirectionally on mRNA [11] and the restart efficiency is supposedly influenced by the sequences present in the neighborhood of the reinitiation triplet [6]. Most of the mechanistic aspects of translational reinitiation remain ill-defined, however, and the effects of mRNA *cis*-elements as well as the influence of translational initiation factors on this process are not clearly understood. The presence of a Shine–Dalgarno (SD)-like structure in front of the downstream cistron is believed to favor the translational reinitiation efficiency, although this process can occur, with lower efficiency, without SD [10].

To investigate the characteristics of mRNA *cis*-elements which in the absence of an SD sequence may induce/favor the reinitiation process in *Escherichia coli*, we have constructed a molecular repertoire by randomization of the region between two translationally coupled genes just upstream of the stop codon (*malE*-(T,A,C)$_{12}$**T^{A}/$_{G}$**ATG-*lacZ′*). This approach allowed the isolation of a number of novel translational reinitiation regions (TRSs), which show a wide range of reinitiation activities fully dependent on a translational coupling event and unrelated to the presence/absence of secondary structure or mRNA stability. Moreover, some of the isolated TRSs are similar to intercistronic sequences present in the *E. coli* genome.

2. Materials and methods

2.1. Construction of the reinitiation molecular repertoire and isolation of TRSs

The molecular repertoire was built by inverse PCR [13] using plasmid p*malE*-c2 (Biolabs, USA) cut by *Bam*HI and the following oligodeoxynucleotides (20 pmol each):

Forward primer: 5′-TTTT**GAATTC**(A,T,C)$_{12}$T^{G}/$_{A}$ATGTCTAG-AGTCGACCTGCAG-3′
Reverse primer: 5′-TTTT**GAATTC**TGAAATCCTTCCCTCG-AT-3′

PCR was performed in a Perkin Elmer 2400 thermal cycler (Perkin Elmer, USA), using the Expand long template PCR thermostable system (Boehringer, Germany) and the conditions suggested by the manufacturer. The PCR was performed using the following protocol: one cycle: 120 s 94°C; 10 cycles: 10 s 94°C, 30 s 60°C, 30 s 68°C; 20 cycles: 10 s 94°C, 30 s 60°C, 30 s 68°C (increasing 20 s at each elongation cycle). PCR products were cut by *Eco*RI (Biolabs, USA) and the ~6.5 kbp PCR product was purified as described [14]. The purified PCR product (200 ng) was ligated as described by Lund and coworkers [15], the ligation mix was desalted as described by Atrazhev and Elliot [16] and electroporated in electrocompetent TG1 *E. coli* cells using the Bio-Rad (USA) electroporation apparatus and conditions suggested by the manufacturer. Transformants were grown in liquid medium (9 ml of LB/ampicillin 100 μg/ml) at 37°C for 18 h, plasmid DNA was prepared by standard techniques [14] and cut by *Eco*RI. The ~6.5 kbp fragment was purified [14], self-ligated and electroporated in TG1 *E. coli* cells. The electroporated cells were then plated on IPTG/X-Gal LB plates containing ampicillin (100

*Corresponding author. Dipartimento di Scienze Cliniche e Biologiche, Università di Torino, Ospedale S. Luigi, Regione Gonzole 10, 10043 Orbassano, Italy. Fax: (39)-011-9038639.
E-mail: raffaele.calogero@unito.it

0014-5793/00/$20.00
PII: S0014-5793(00)01198-4

μg/ml). The A, U and C contents of the randomized sequences were respectively 38%, 37% and 35% and they were defined randomly selecting 200 colonies from the $\sim 10^5$ colonies counted on plates. Among the $\sim 10^5$ plated colonies, 3% showed a lacZ$^+$ phenotype (blue color). 200 lacZ$^+$ colonies were randomly selected. 75% of the 200 lacZ$^+$ clones showed by Western blotting the presence of a 42.7 kDa wild-type Mbp protein (data not shown), which indicates a reinitiation event in the translationally coupled *malE-lacZ′* system. Reinitiation positive clones were sequenced allowing the identification of 17 signals inducing the reinitiation phenomenon (Table 1).

2.2. Computer analysis of reinitiation sequences isolated from the molecular repertoire and potential E. coli reinitiation sequences

Potential reinitiation sequences based on the **U^{A}/$_{G}$AUG** termination/initiation signal were extracted from the TransTerm database [17] using the X-tract program (Iazzetti and Calogero, unpublished program); the program identifies in the *E. coli* TransTerm database all the potentially translationally coupled genes where the termination codon (UAA or UGA) of a cistron partially overlaps the initiation codon (AUG) of an other cistron. 117 Potential reinitiation sequences based on **UAA**AUG signal and 60 based on **UGA**UG signal were isolated from the 5406 gene sequences present in the TransTerm database. Both the potential natural reinitiation sequences (Fig. 2, right top and bottom panels) and de novo reinitiation signals isolated from the molecular repertoire (Fig. 2, left top and bottom panels) were aligned using the Slogos program [18]. The *E. coli* genome in the GenBank database was also searched for the TRSs described in Table 1 using BLASTN program [19]. This search yielded:

1. a TRS 21-like sequence (AUUUC**UAA**AUG) located in the gene cluster ECAE000288 between AAC75037.1 (putative reductase) and AAC75038.1 (orf, unknown function),
2. a TRS 34-like (ACAUCUC**AUGA**) sequence located in the gene cluster ECAE000229 between AAC74391.1 (putative polysaccharide hydrolase) and AAC74392.1 (putative transport periplasmic protein),
3. a TRS 41-like (A- - - - - -A- - -ACAC**UAA**UG) in the nik operon [20] between genes nikC and nikD.

Potential secondary structures at 37°C were computed for the *gst*-TRS-*lacZ′* mRNAs using the RNA mfold server [21].

Comparison between the TRSs and the 16S RNA was made using the Blast2 server [22].

2.3. Construction of the translational gst-tet coupled system

The *Tet* gene was amplified from pBR322 plasmid (Biolabs, USA) using the following oligodeoxynucleotides:

Forward primer: 5′-TTTAT**AGATCT**CAGGCACCGTGTATG-AAA-3′
Reverse primer: 5′-TATAT**GAATTC**TCAGGTCGAGGTGGC-CCG-3′

PCR was performed in a thermal cycler 2400 (Perkin Elmer) using Taq polymerase (Promega, USA) and conditions suggested by the manufacturer. PCR was performed using the following protocol: 40 cycles: 40 s 94°C, 40 s 59°C, 40 s 72°C. The ~1.2 kbp PCR product was cut by *Bgl*II and *Eco*RI and cloned in pGEX-3X (Pharmacia) cut by *Bam*HI and *Eco*RI (pGEX-TET, Fig. 3A). pGEX-41-TET (Fig. 3B), pGEX-neg-TET (Fig. 3C) were constructed starting from pGEX-TET via inverse PCR using the following oligodeoxynucleotides (20 pmol):

Forward primer pGEX-41-TET: 5′-TTCCC**AAGCTT**AAAATA-AATATACACTAATGAAATCTAACAATGCGCTC-3′
Forward primer pGEX-neg-TET: 5′-TTCCC**AAGCTT**AAAAT-AAATATACACTAAAGAAATCTAACAATGCGCTC-3′
Reverse primer: 5′-TATA**AAGCTT**CCCACGACCTTCGAT-3′

PCR was performed as previously described. The amplified plasmids were self-ligated. pGEX-noBstBI-41-TET (Fig. 3D) was obtained by filling-in [14] the 5′ protruding ends generated by cutting pGEX-41-TET with BstBI (Biolabs, USA) and self-ligating the resulting plasmid. pGEX-no*Hin*dIII-41-TET (Fig. 3E), was obtained by filling-in the 5′ protruding ends generated by cutting pGEX-41-TET with *Hin*dIII (Biolabs, USA) and self-ligating the resulting plasmid.

2.4. Construction of the translational coupled system gst-lacZ′

pGEX-3X (Pharmacia) cut by *Bam*HI was filled-in and plasmid was self-ligated (pGEX-3X-noBamHI). pGEX-3X-noBamHI was cut by *Eco*RI/*Mlu*I (Biolabs, USA); the 2.2 kbp fragment was ligated with the 4.3 kbp fragment obtained by cutting p*malE*-41-*lacZ′* (Table 1) with *Eco*RI/*Mlu*I (Biolabs, USA), to produce the pGEX-41-*lacZ′*. All the other plasmids used to evaluate the functionality (Fig. 4) of reinitiation sequences were produced as described for clone 41.

2.5. Reinitiation activity in the gst-lacZ′ coupled system

Reinitiation levels in bacteria harboring *gst*-TRS-*lacZ′* constructs were determined measuring the Gst and β-galactosidase activity in the crude lysate of each clone. Cells harboring the pGEX-TRS-*lacZ′* coupled system were grown (37°C) to 0.5 $OD_{600\,nm}$, 1 mM IPTG was added and incubation continued for 3 h. Eight $OD_{600\,nm}$ cell culture pellets were lysed by sonication in PBS (1 ml) and the lysate was cleared by centrifugation. The GST activity was measured with the following protocol: 10 μl of lysate was added to 990 μl Z buffer (60 mM Na_2HPO_4, 10 mM KCl, 1 mM $MgSO_4$, 40 mM NaH_2PO_4) containing reduced glutathione (1 mM) and 1-chloro-2,4-dinitrobenzene (1 mM). The absorbance at 340 nm was measured every minute for 10 min and $\Delta OD_{340\,nm} \times 1$ $OD_{600\,nm}$ cell culture was determined with the following equation: $\Delta OD_{340\,nm}$ min^{-1} ml^{-1} = $[OD_{340\,nm}(t2) - OD_{340\,nm}(t1)/(t2-t1)] \times 100/8$. β-Galactosidase activity was measured as described by Sambrook [14]: 150 μl crude lysate was added to 850 μl Z buffer containing 1.2 mg *o*-nitrophenyl-β-galactopyranoside. The absorbance at 420 nm was measured every 3 min for 30 min and $\Delta OD_{420\,nm} \times 1$ $OD_{600\,nm}$ cell culture was determined with the following equation: $\Delta OD_{420\,nm}$ min^{-1} ml^{-1} = $[OD_{420\,nm}(t2) - OD_{420\,nm}(t1)/(t2-t1)] \times 6.6/8$. The ratio between the *lacZ′* expression (reinitiation reporter gene) and the Gst expression (reinitiation driving force), expressed as their enzymatic activities ($\Delta OD_{420\,nm}/\Delta OD_{340\,nm}$) in 1 $OD_{600\,nm}$ of cell culture lysate, defines the arbitrary reinitiation units (rUs) used to compare the reinitiation level of the reinitiation *cis*-acting elements. The β-galactosidase and glutathione-*S*-transferase activity ratio was also calculated for a *gst*::*lacZ′* fusion product (~18 rUs) as a reference for the condition when each ribosome finishing the Gst polypeptide synthesis will reinitiate the synthesis of *LacZ′* peptide. Data shown in Fig. 4 were obtained by induction experiments repeated at least three times for each clone.

2.6. Determination of chemical stability of gst-lacZ′ mRNAs

The mRNA stability was measured for mRNAs encompassing the reinitiation signals inserted in the *gst-lacZ′* translationally coupled system as previously described by Brandi [23].

3. Results

3.1. Identification of TRSs

The identification of mRNA *cis*-elements favoring the reinitiation of protein synthesis in *E. coli* was approached designing a translationally coupled system (Fig. 1) consisting of

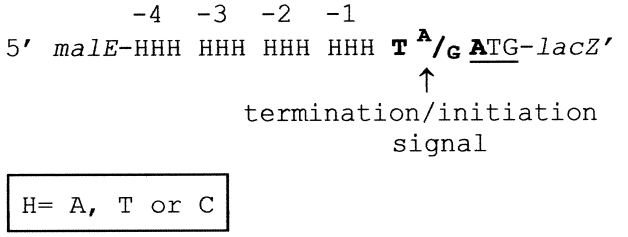

Fig. 1. Molecular repertoire used for the isolation of mRNA *cis*-elements inducing the process of reinitiation of protein synthesis in *E. coli*.

a 5′ cistron (*malE*) ending with a stop codon (UAA or UGA) partially overlapping the AUG start codon of the downstream cistron (*lacZ′*). Furthermore, the 12 nucleotides upstream the stop codon of *malE* were randomized avoiding the use of guanine (G) in order to minimize the formation of a SD sequence which could lead to de novo initiation of *lacZ′* translation. When cells harboring this reinitiation library (10^5 clones) were plated on medium containing IPTG/X-Gal, synthesis of LacZ′ was observed in about 3000 clones corresponding to approximately 3% of the colonies.

To determine if the LacZ′ synthesis was due to a real reinitiation event or to a genetic rearrangement leading to a *malE::lacZ′* fusion, the presence of either the *malE* polypeptide Mbp (42.7 kDa) or an Mbp-LacZ′ fusion (48.4 kDa) polypeptide was evaluated in 200 potential reinitiation clones by sodium dodecyl sulfate–polyacrylamide gel electrophoresis analysis followed by Western blotting and immunodetection. At least 150 clones were found to produce the wild-type 42.7 kDa Mbp (not shown), indicating that approximately 75% of the clones were able to induce a translational reinitiation event.

One hundred randomly chosen clones showing reinitiation activity were then selected for sequencing. A total of 17 TRSs were identified, seven of them contained the **UAAUG**- and 10 contained the **UGAUG**-based termination/initiation motifs. As seen in Table 1, these TRSs were not equally represented within the reinitiation repertoire as shown by their isolation frequencies; although sequences of clones randomly selected from the full library did not show any sequence bias (not shown).

3.2. Sequence requirements for reinitiation

Overall, the alignments, generated using the Slogos program [18] for both the **UAAUG**- and **UGAUG**-based reinitiation sequences identified using the combinatorial approach, suggest that reinitiation does not require specific sequence elements other than a U or an A at position −8 (Fig. 2, left panels, black arrow). Moreover, the **UAAUG**-based reinitiation signals show a strong bias in favor of U in position −12 (Fig. 2, left top panel) while those based on **UGAUG** show a preference for a C or a U in position −5 (Fig. 2, left bottom panel). In addition, the **UAAUG**-based reinitiation sequences contain only 15 of the 27 triplets available in the repertoire.

Table 1
TRSs using $\mathbf{U^A/_G}$**AUG** termination/initiation signals

TRSs	Sequences	Frequency
6	CUU CCA AAA AAA **UAA**UG	8
19	ACU CUU AAC UUC **UAA**UG	8
21	UAU UUA UCU UUU **UAA**UG	4
23	UAU UUA UCU UUU **UAA**UG	8
29	UUC UUU UCU CCA **UAA**UG	4
37	UUA AAU UUC CAU **UAA**UG	4
41	AAA AUA AAU AUA CAC **UAA**UG	4
1	AAU CUU AUA CCC **UGA**UG	8
2	AAU AUC AAA CCC **UGA**UG	4
3	CCU UCA UUA UUA **UGA**UG	4
4	UCA AUA UCC CUA **UGA**UG	12
5	AAU CAC UAA UCA **UGA**UG	4
8	CUC CAA ACU UAU **UGA**UG	4
12	AAC UUU UUC CUU **UGA**UG	8
26	UAU CAA AAC UCA **UGA**UG	4
34	AUU UAC AUC UCA **UGA**UG	8
36	ACA UAA UUC CUU **UGA**UG	4

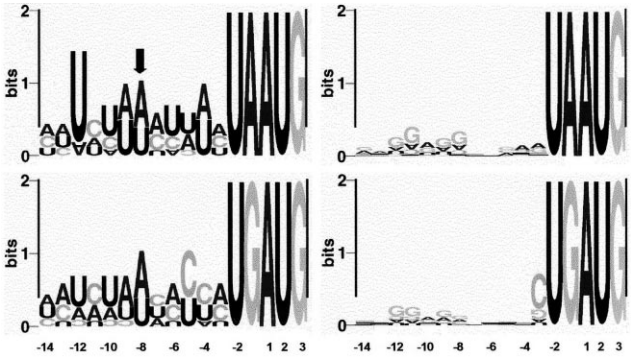

Fig. 2. Top and bottom left panels show the Slogos alignment of the reinitiation sequences isolated with the combinatorial approach. Top and bottom right panels show the Slogos alignment for *E. coli* translationally coupled genes based on the $\mathbf{U^G/_A}$**AUG** termination/initiation signal.

The **UGAUG**-based reinitiation signals, on the other hand, contain almost all (23/27) the available triplets.

To extract from the *E. coli* TransTerm database [17] all translationally coupled genes based on the **UAAUG** and **UGAUG** termination/initiation signals, a computer program (X-tract) was written by Iazzetti and Calogero (unpublished program); a total of 117 translationally coupled genes with the termination/initiation **UAAUG** signal and 60 with the **UGAUG** signal were extracted by this program from 5406 gene sequences.

Considering that the TransTerm database is largely based on canonical initiation regions which were deliberately avoided in our repertoire, it is not surprising that the Slogos alignment of the translationally coupled genes extracted from the *E. coli* TransTerm database revealed a consensus SD sequence upstream the $\mathbf{U^A/_G}$**AUG** termination/initiation signal (Fig. 2, top and bottom right panels) and that the natural sequences did not show extensive similarity with the TRSs identified by the combinatorial approach. Nevertheless, it is remarkable that we identified three TRS-like sequences located between translationally coupled genes in the *E. coli* genome; as seen in Table 2, these TRS-like sequences resemble the TRS 21, 34 and 41 identified by the combinatorial approach.

3.3. The TRSs can promote translational coupling by reinitiation events also in a different gene context

To evaluate, at least from a qualitative point of view, the effect of the sequence context on the reinitiation process, some of the reinitiation signals (TRSs) identified were placed in a different translationally coupled system (*gst-tet*). For this purpose, the *tet* gene was excised from pBR322 and inserted with

Table 2
TRS-like sequences in the *E. coli* genome

TRSs	E.coli gene clusters	Sequences
21	-	CAU AUU UCU AAA **UAA**UG
-	ECAE000288	--- AUU UCU --- **UAA**UG
34	-	AUU UAC AUC UCA **UGA**UG
-	ECAE000229	--- -AC AUC UCA **UGA**
41	-	AAA AUA AAU AUA CAC **UAA**UG
-	nik operon	--A --- -A- --A CAC **UAA**UG

its own TIR downstream from the *gst* gene in pGEX-3X (Promega) (Fig. 3A). When the *tet* TIR was exchanged with reinitiation signal 41 (Fig. 3B), or with reinitiation signals 3, 21, 23, 34 (not shown), a *tet*$^+$ phenotype, monitored as increase in cell density ($OD_{600\,nm}$) after 5 h growth in liquid medium supplemented with tetracycline, was observed. No differences, at least from the qualitative point of view, could be appreciated comparing the cells harboring the *tet* gene containing its natural TIR and those in which *tet* devoid of

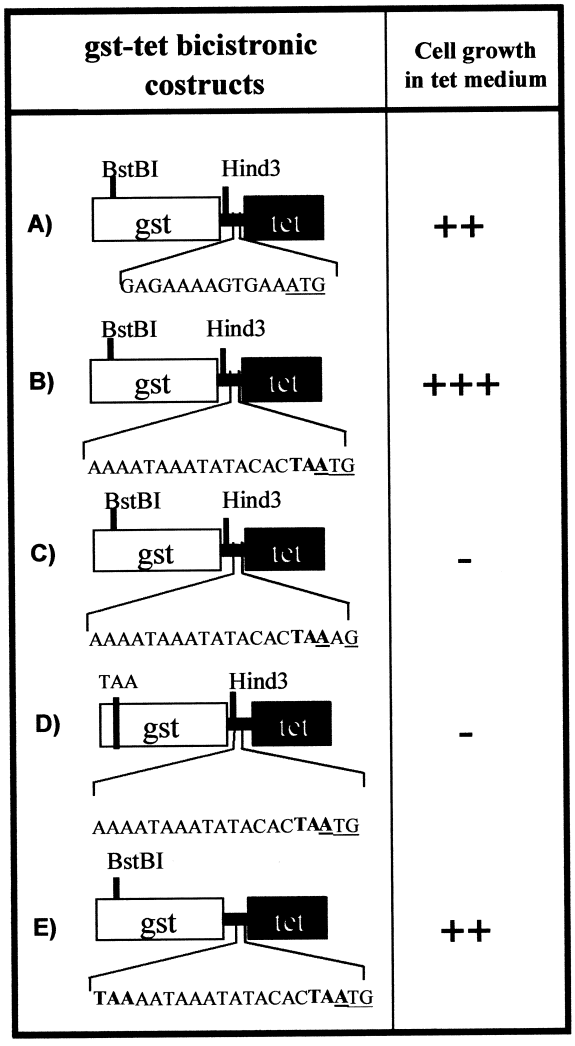

Fig. 3. *gst-tet* bicistronic vector was used to evaluate the mRNA *cis*-acting elements ability to induce reinitiation in a genetic environment different from the *malE-lacZ′* used for their isolation. A: *gst-tet* bicistronic construct (pGEX-TET), the *tet* expression is controlled by its natural TIR. B: *gst-tet* translationally coupled by reinitiation signal 41 (pGEX-41-TET). C: *gst-tet* translationally coupled by reinitiation signal 41, the AUG at the reinitiation site has been mutated to AAG (pGEX-neg-TET), a codon permitting levels of translation initiation too low to be quantified [24]. D: *gst-tet* translationally coupled by reinitiation signal 41, a stop codon has been inserted in frame with the Gst coding sequence 305 nucleotides upstream of the reinitiation sequence (pGEX-noBstBI-41-TET). E: *gst-tet* translationally coupled by reinitiation signal 41, a stop codon has been inserted in frame with the Gst coding sequence 18 nucleotides upstream of the reinitiation sequence (pGEX-noHindIII-41-TET). Cell growth was measured as adsorbance at 600 nm after 5 h in tet$^+$ medium (12 μg/ml). Growth levels are shown as: − = 0.0–0.1 $OD_{600\,nm}$, + = 0.2–0.4 $OD_{600\,nm}$, ++ = 0.4–0.6 $OD_{600\,nm}$, +++ = 0.6–0.8 $OD_{600\,nm}$.

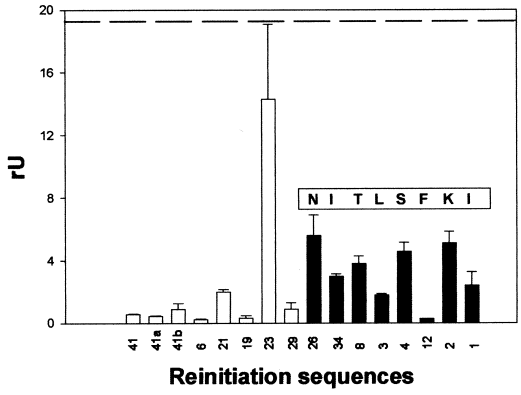

Fig. 4. *gst-lacZ′* translationally coupled system was used to compare the level of reinitiation for mRNA *cis*-acting element isolated with the combinatorial approach. Reinitiation activity is expressed as rUs, see Section 2. Open bars indicate the reinitiation activity of *cis*-acting elements based on the **UAAUG** termination/initiation signal and closed bars indicate the reinitiation activity of *cis*-acting elements based on the **UGAUG** signal. The inset box on **UGAUG**-based reinitiation signals shows the amino acids encoded by codon −2 upstream the termination/initiation signal. Dashed line indicates a 1:1 *LacZ′*/Gst expression ratio (∼18 rUs), the condition when each ribosome finishing the Gst polypeptide synthesis will reinitiate the synthesis of *LacZ′* peptide.

its TIR was preceded by the reinitiation signals. On the other hand, substitution of the AUG triplet of the reinitiation signal with a class IIB initiation codon (AAG), which does not permit a detectable level of translation [24], did not allow cell growth in tetracycline-supplemented medium (Fig. 3C). This result indicates that the AUG at the reinitiation site is necessary for Tet expression. A *tet*$^-$ phenotype was also observed when a UAA termination codon was inserted, in frame with the Gst coding sequence, 305 nucleotides upstream of the reinitiation signal 41 (Fig. 3D), or 3, 21, 23, 34 (not shown), indicating that Tet synthesis strictly depends on a translational termination event occurring in the proximity of the *tet* AUG. On the other hand, in agreement with the reinitiation scanning model proposed by Adhin [11], a termination event occurring in closer proximity to the Tet AUG start codon can promote reinitiation as indicated by the presence of the *tet*$^+$ phenotype observed when a UAA termination codon was inserted 18 nucleotides upstream the reinitiation AUG (Fig. 3E). This finding indicates that overlapping of the termination and initiation triplets of the upstream and downstream cistrons is not a prerequisite for reinitiation. Similar results were also obtained using the reinitiation signals 41, 3, 21, 23, 34 in a *gst-malE* translationally coupled system (not shown).

3.4. Reinitiation efficiency of the cis-acting elements isolated with the combinatorial approach

To evaluate quantitatively the activity of the various TRSs identified, the individual *cis*-acting reinitiation elements were inserted in a *gst-lacZ′* translationally coupled system to obtain a series of *gst*-TRS-*lacZ′* plasmids. Both glutathione-*S*-transferase and the β-galactosidase activities were then measured in crude extracts of cells harboring these constructs. To evaluate the functional efficiency of the various reinitiation sequences, we compared the rUs expressed by each cell extract arbitrarily defined as the ratio between β-galactosidase and glutathione-

Table 3
Chemical stability of *gst*-TRS-*lacZ′* mRNAs

TRSs	Half life	Reinitiation activity	Termination/ initiation signal
2	+	(M)	UGAUG
26	+	(M)	UGAUG
4	+	(M)	UGAUG
1	+	(L)	UGAUG
41	++	(VL)	UAAUG
6	++	(VL)	UAAUG
3	++	(L)	UGAUG
8	++	(M)	UGAUG
23	++	(H)	UAAUG
19	++	(VL)	UAAUG
21	++	(VL)	UAAUG

Reinitiation activities are expressed as VL = ≤1 rUs, L = ≤3, >1 rUs, M = ≤8, >3 rUs and H = >8 rUs. mRNA half lives are shown as + = 1–2 min, ++ = 2–4 min, +++ = 4–6 min.

S-transferase activities. Similar results were obtained when the enzymatic activities were determined spectrophotometrically as in Fig. 4 or when, in a *gst-malE* translationally coupled system, the expression levels of Gst or Mbp were evaluated by enzyme immuno assay (not shown). Although the reinitiation efficiency of the various TRSs was found to vary considerably from case to case, all TRS sequences tested were found to promote reinitiation (Fig. 4). This finding reinforces the previous conclusion that the reinitiation activity promoted by the TRSs identified by our approach does not depend on a specific genetic context. For the sake of simplicity, we classified the reinitiation efficiency as: high (H) (>8 rUs); medium (M) (≤8, >3 rUs); low (L) (≤3, >1 rUs) and very low (VL) (≤1 rU).

We then decided to evaluate whether the different levels of reinitiation efficiency expressed by the individual TRSs could stem from differences in either secondary structure or stability of the mRNAs. When the Zucker algorithm [21] was applied to the *gst*-TRS-*lacZ′* mRNAs, the TRSs were predicted to assume stems or looped conformations without any relation to their activity (not shown), thus leading us to conclude that no correlation exists between the efficiency of reinitiation and the presence (or absence) of any particular secondary structure of the mRNA.

Finally, when the half lives of the different *gst*-TRS-*lacZ′* transcripts were experimentally determined, no correlation could be found between the chemical stability of the various transcripts and their reinitiation efficiency (Table 3).

4. Discussion

Translational reinitiation can be regarded as a specialized type of translational coupling in which translation of a downstream cistron is ensured by the ribosome which is terminating on an adjacent termination codon belonging to the upstream cistron. Although the number of documented reinitiation events is relatively limited, the importance of these phenomena stems from their regulatory functions and/or from the possibility that the cell may use them to escape the lethal effects potentially elicited by nonsense mutation within essential genes.

The mechanistic aspects of reinitiation are so far poorly understood but it is generally assumed that translation termination in Eubacteria may be followed by bidirectional movement of the ribosome [11] and that the presence of an SD-like sequence upstream of a potential start codon may favor a reinitiation event. Since the presence of an SD element is not strictly necessary for a de novo initiation process [25], there is no a priori reason to assume that a SD element might be necessary to ensure reinitiation. Thus, in this work, we sought to identify the characteristics of mRNA *cis*-elements which might induce, in the absence of a canonical ribosomal binding site, an in vivo translational reinitiation process in *E. coli*. For this purpose, we have used a combinatorial approach which allowed us to screen for in vivo reinitiation 10^5 clones representative of a library of 10^6 sequence elements. Upon isolation of nearly 2300 clones in which a reinitiation event had taken place, a pool of 17 different reinitiation signals was identified (Table 1). An important characteristic of these sequences is that their functioning does not depend on a specific genetic system since they were found to be active with three types of translational coupling. The isolated TRSs are also able to induce reinitiation of protein synthesis although translational initiation signals like an SD region and any other sequence element displaying potential complementarity with 16S RNA (e.g. the downstream box; [26]) are not present. Furthermore, the reinitiation event was found to depend on the presence of an AUG located in the vicinity of the translational termination signal of the upstream cistron (Fig. 3B) and to be independent from the mRNA sequences surrounding the reinitiation signal (as mentioned above, all the TRSs isolated with the *malE-lacZ′* coupled system were able to induce a reinitiation event when inserted in the *gst-tet*, *gst-mbp* and *gst-lacZ′* translationally coupled systems).

Unlike the majority of the sequences involved in the translational coupling between genes in *E. coli*, which are often endowed with secondary structures masking the TIR of the downstream cistron [2,27], the TRSs identified here do not seem to have any common secondary structure or sequence. However, we observed that in the TRSs based on the UGA termination signal, the increasing hydrophobicity of the amino acid in position −2 (F > L = I > K,S,T,N) seems to correlate with a reduced reinitiation activity (Fig. 4). We also observed that the location of the AUG upstream or downstream the termination codon as well as the termination codon itself (UAA or UGA) seems not to be directly involved in TRS functionality, at least for the reinitiation signal 41 where the **UAAUG** can be modified to **UGAUG** (clone 41a: AAA AUA AAU AUA CAC **UGAUG**) or **AUGA** (clone 41b: AAA AUA AAU AUA CA**A UGA**) without any effect on the TRS activity (Fig. 4, reinitiation signals 41a, 41b). Furthermore, the various degrees of reinitiation efficiencies observed with different TRSs did not seem to correlate with the stability of the corresponding mRNAs (Table 3).

The identification of sequences inducing reinitiation independently from the presence of an SD signal (this work) as well published data on sequences able to induce de novo initiation without SD (e.g. pyrimidine rich sequences [28]; non-SD sequences [29]) suggest that TIRs can be characterized by signals other than the SD sequence. Although it is unlikely that TRSs can play an important role in vivo during normal conditions of growth, it is possible that these reinitiation events can be used by *E. coli* as an extreme escape from mutations inactivating conventional TIRs or they might be involved in control of gene expression under particular environmental and/or stress conditions (e.g. experiments not shown demonstrated that reinitiation caused by TRS 19 in-

creases 4-fold in heat-shocked cells). Finally, from the practical point of view, the TRSs identified in the present work could be used for the construction of an expression vector in which an efficient reinitiation signal (e.g. TRS 23) could ensure the expression of a protein of interest by coupling its translation to that of the upstream cistron, thus bypassing the use of gene fusion or the need for TIR optimization for the expression of heterologous proteins [29].

Acknowledgements: We thank Prof. M. De Marchi for critical reading of this manuscript. This project was supported by CNR Progetto Finalizzato Biotecnologie and in part by Ministero per l'Università e la Ricerca Scientifica, project 'Protein-nucleic acid interactions'.

References

[1] Gualerzi, C. and Pon, C.L. (1990) Biochemistry 29, 5881–5889.
[2] McCarthy, J.E.G. and Gualerzi, C. (1990) Trends Genet. 6, 78–85.
[3] McCarthy, J.E.G. and Brimacombe, R. (1994) Trends Genet. 10, 402–407.
[4] Jackson, R.J. (1996) in: Translational Control (Hershey, J.W.B., Mathews, M.B. and Sonenberg, N., Eds.), pp. 71–112, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[5] Sprengart, M.L. and Porter, A.G. (1997) Mol. Microbiol. 24, 19–28.
[6] Matteson, K.J., Bisswas, S.J. and Steege, D.A. (1991) Nucleic Acids Res. 13, 69955–69967.
[7] Ganem, D., Miller, J.H., Files, J.G., Platt, T. and Weber, K. (1973) Proc. Natl. Acad. Sci. USA 70, 3165–3169.
[8] Files, J.G., Weber, K. and Miller, J.H. (1974) Proc. Natl. Acad. Sci. USA 71, 667–670.
[9] Steege, D.A. (1977) Proc. Natl. Acad. Sci. USA 74, 4163–4167.
[10] Spanjaard, R.A. and van Duin, J. (1989) Nucleic Acids Res. 17, 5501–5507.
[11] Adhin, M.R. and van Duin, J. (1990) J. Mol. Biol. 213, 811–818.
[12] Peijnenburg, M., Venema, G. and Bron, S. (1990) Mol. Gen. Genet. 221, 267–272.
[13] Bianchi, E., Folgori, A., Wallace, A., Nicotra, M., Acali, S., Phalipon, A., Barbato, G., Bazzo, R., Cortese, R., Felici, F. and Pessi, A. (1995) J. Mol. Biol. 247, 154–160.
[14] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
[15] Lund, A.H., Duck, M. and Pedersen, S. (1996) Nucleic Acids Res. 24, 800–801.
[16] Atrarhev, A.M. and Elliot, J.E. (1996) BioTechniques 21, 8–9.
[17] Dalphin, M.E., Brown, C.M., Stockwell, P.A. and Tate, W.P. (1998) Nucleic Acids Res. 12, 335–337.
[18] Gorodkin, J., Heyer, L.J., Brunak, S. and Stormo, G.D. (1997) CABIOS 13, 583–586.
[19] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 17, 3389–3402.
[20] Navarro, C., Wu, L.F. and Mandrand-Berthelot, M.A. (1993) Mol. Microbiol. 9, 1181–1191.
[21] Jaeger, J.A., Turner, D.H. and Zucker, M. (1990) Methods Enzymol. 183, 281–306.
[22] Tatusova, T.A. (1999) FEMS Microbiol. Lett. 174, 247–250.
[23] Brandi, A., Pietroni, P., Gualerzi, C.O. and Pon, C.L. (1996) Mol. Microbiol. 19, 231–240.
[24] Sussman, J.K., Simons, L.E. and Simons, R.W. (1996) Mol. Microbiol. 21, 347–360.
[25] Calogero, R.A., Pon, C.L., Canonaco, M.A. and Gualerzi, C.O. (1988) Proc. Natl. Acad. Sci. USA 85, 6427–6431.
[26] O'Connor, M., Asai, T., Squires, C.L. and Dahlberg, A.E. (1999) Proc. Natl. Acad. Sci. USA 96, 8973–8978.
[27] Cone, K.C. and Steege, D.A. (1985) J. Mol. Biol. 186, 725–732.
[28] Tzareva, N.V., Makhno, V.I. and Boni, I.V. (1994) FEBS Lett. 337, 189–194.
[29] Mironova, R.S., Xu, J., AbouHaidar, M.G. and Ivanov, I.G. (1999) Microbiol. Res. 154, 35–41.