The gene of an archaeal α-L-fucosidase is expressed by translational frameshifting

Beatrice Cobucci-Ponzano¹, Fiorella Conte¹, Dario Benelli², Paola Londei², Angela Flagiello³, Maria Monti³, Piero Pucci³, Mosè Rossi¹,⁴ and Marco Moracci¹,∗

¹Institute of Protein Biochemistry—Consiglio Nazionale delle Ricerche, Via P. Castellino 111, 80131 Naples, Italy, ²Dipartimento di Biochimica Medica e Biologia Medica (DIBIM), Università di Bari—Piazzale Giulio Cesare, 70124 Bari, Italy, ³CEINGE Biotecnologie Avanzate s.c.a r.l., Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II, Complesso Universitario di Monte S. Angelo, via Cinthia 4, 80126 Napoli, Italy and ⁴Dipartimento di Biologia Strutturale e Funzionale, Università di Napoli 'Federico II', Complesso Universitario di Monte S. Angelo, Via Cinthia 4, 80126 Naples, Italy

Received April 19, 2006; Revised July 21, 2006; Accepted July 22, 2006

ABSTRACT

The standard rules of genetic translational decoding are altered in specific genes by different events that are globally termed recoding. In Archaea recoding has been unequivocally determined so far only for termination codon readthrough events. We study here the mechanism of expression of a gene encoding for a α-L-fucosidase from the archaeon Sulfolobus solfataricus (fucA1), which is split in two open reading frames separated by a −1 frameshifting. The expression in Escherichia coli of the wild-type split gene led to the production by frameshifting of full-length polypeptides with an efficiency of 5%. Mutations in the regulatory site where the shift takes place demonstrate that the expression in vivo occurs in a programmed way. Further, we identify a full-length product of fucA1 in S. solfataricus extracts, which translate this gene in vitro by following programmed −1 frameshifting. This is the first experimental demonstration that this kind of recoding is present in Archaea.

INTRODUCTION

Translation is optimally accurate and the correspondence between the nucleotide and the protein sequences are often considered as an immutable dogma. However, the genetic code is not quite universal: in certain organelles and in a small number of organisms the meaning of different codons has been reassigned and all the mRNAs are decoded accordingly. More surprisingly, the standard rules of genetic decoding are altered in specific genes by different events that are globally termed recoding (1). In all cases, translational recoding occurs in competition with normal decoding, with a proportion of the ribosomes not obeying to the ‘universal’ rules. Translational recoding has been identified in both prokaryotes and eukaryotes. It has crucial roles in the regulation of gene expression and includes stop codon readthrough, ribosome hopping and ±1 programmed frameshifting [for reviews see (2–4)].

In stop codon readthrough a stop codon is decoded by a tRNA carrying an unusual amino acid rather than a translational release factor. Specific stimulatory elements downstream to the stop codon regulate this process (5). Hopping, in which the ribosome stops translation in a particular site of the mRNA and re-start few nucleotides downstream, is a rare event and it has been studied in detail only in the bacteriophage T4 (6). In programmed frameshifting, ribosomes are induced to shift to an alternative, overlapping reading frame 1 nt 3′-wards (+1 frameshifting) or 5′-wards (−1 frameshifting) of the mRNA. This process is regulated and its frequency varies in different genes. The ±1 programmed frameshifting has been studied extensively in viruses, retrotransposons and insertion elements for which many cases are documented (7–9). Instead, this phenomenon is by far less common in cellular genes. A single case of programmed +1 frameshifting is known in prokaryotes (10,11) while in eukaryotes, including humans, several genes regulated by this recoding event have been described previously [(4) and references therein]. Compared to +1 frameshifting, −1 frameshifting is less widespread with only two examples in prokaryotes (12–14) and few others in eukaryotes (15–17).

The programmed −1 frameshifting is triggered by several elements in the mRNA. The slippery sequence, showing the X-XY-YYZ motif, in which X can be any base, Y is usually A or U, and Z is any base but G, has the function of favouring the tRNA misalignment and it is the site where the shift takes place (3,18). Frameshifting could be further stimulated by other elements flanking the slippery sequence: a codon for a low-abundance tRNA, a stop codon, a Shine–Dalgarno sequence and an mRNA secondary structure. It has been
reported that these elements, alone or in combination, enhance frameshifting by pausing the translating ribosome on the slippery sequence (4,18).

Noticeably, known cases of recoding in Archaea [recently reviewed in (19)] are limited to termination codon readthrough events that regulate the incorporation of the 21st and 22nd amino acids selenocysteine and pyrrolysine, respectively (20–23).

No archaeal genes regulated by translational programmed frameshifting and ribosome hopping have been identified experimentally so far; therefore, if compared with the others domains of life, the study of translational recoding in Archaea is still at its dawn.

We showed that the α-L-fucosidase gene from the crenarchaeon Sulfolobus solfataricus is putatively expressed by programmed −1 frameshifting (24). This gene, named fucAI, is organized in the open reading frames (ORFs) SSO11867 and SSO3060 of 81 and 426 amino acids, respectively, which are separated by a −1 frameshifting in a 40 base overlap (Figure 1A). We have reported previously that the region of overlap between the two ORFs had the characteristic features of the genes expressed by programmed −1 frameshifting including a slippery heptanucleotide A-AAA-AAT (codons are shown in the zero frame) flanked by a putative stem–loop and the rare codons CAC (Figure 1A) resembling the prokaryotic stem–loops/hairpins and the Shine–Dalgarno-like sites (24). We showed that the frameshifting, obtained by mutating by site-directed mutagenesis the fucAI gene exactly in the position predicted from the slippery site, produced a full-length gene, named fucAI\(^A\), encoding for a polypeptide of 495 amino acids (Figure 1B). This mutant gene expressed in Escherichia coli a fully functional α-L-fucosidase, named Ss\(\alpha\)-fuc, which was thermophilic, thermostable and had an unusual nonanumeric structure (24,25). More recently, we determined the reaction mechanism and the function of the residues of the active site of the mutant enzyme (26,27).

The functionality of the product of the mutant gene fucAI\(^A\) does not provide direct experimental evidence that programmed −1 frameshifting occurs in vivo and in S.solfataricus. To address these issues, we report here the study of the expression of the wild-type split gene fucAI and of its mutants in the slippery sequence. We demonstrate here that fucAI is expressed by programmed −1 frameshifting in both E.coli and S.solfataricus. This is the first experimental demonstration that this kind of recoding is present in the Archaea domain of life. The relevance of programmed −1 frameshifting in Archaea is also discussed.

Figure 1. The α-fucosidase gene. (A) Region of overlap in the wild-type split fucAI gene. The N-terminal SSO11867 ORF is in the zero frame, the C-terminal SSO3060 ORF, for which only a fragment is shown, is in the −1 frame. The slippery heptanucleotide is underlined; the rare codons are boxed and the arrows indicate the stems of the putative mRNA secondary structure. The amino acids involved in the programmed −1 frameshifting and the first codon translated after this event in the −1 frame are shadowed. (B) Fragment of the full-length mutant fucAI\(^A\) gene. The small arrows indicate the mutated nucleotides.
MATERIALS AND METHODS

Analysis of the α-fucosidase expression

*S. solfataricus* cells were grown, and cell extracts obtained, as described previously (24,28).

The expression in the *E. coli* strain BL21(RB791) of the wild-type gene *fucA1* and of the mutant genes *fucA1*\textsuperscript{A} [previously named FrameFuc in (24), *fucA1*\textsuperscript{B}, *fucA1*\textsuperscript{sm} and *fucA1*\textsuperscript{m} as fusions of glutathione S-transferase (GST) and the purification of the recombinant proteins were performed as reported previously (23). The nomenclature used in this paper for the different α-fucosidase genes is listed in Table 1.

For the western blot studies, equal amounts of *E. coli* cultures expressing the wild-type and mutant *fucA1* genes, normalized for the OD\textsubscript{600}, were resuspended in SDS–PAGE loading buffer containing 0.03 M Tris–HCl buffer, pH 6.8, 3% SDS (w/v), 6.7% glycerol (w/v), 6.7% 2-mercaptoethanol (w/v) and 0.002% blue bromophenol (w/v). The samples were incubated at 100°C for 5 min (unless otherwise indicated) and were directly loaded on to the gel. Western blot analyses were performed by blotting SDS–PAGEs of the concentrations indicated on Hybond-P polyvinylidenfluorid filters (Amersham Biosciences, Upsalla, Sweden); polyclonal anti-\(α\)-fuc antibodies from rabbit (PRIMM, Milan, Italy) and anti-GST antibodies (Amersham Biosciences) were diluted 1:5000 and 1:40 000, respectively. Thewestern blots were washed and incubated with the ImmunoPure anti-rabbit IgG antibody conjugated with the horseradish peroxidase (HRP) from Pierce Biotechnology (Rockford, IL, USA). Filters were washed and incubated with the ECL streptavidin–HRP conjugate (Amersham Biosciences).

The protein concentration of the samples was measured with the method of Bradford (29) and the amounts of sample loaded on to the SDS–PAGEs are those indicated. The quantification of the bands identified by western blot was performed by using the program Quantity One 4.4.0 in a ChemiDoc EQ System (Bio-Rad, Hercules, CA, USA) with the volume analysis tool. The frameshifting efficiency was calculated as the ratio of the intensity of the bands of the frameshifted product/frameshifted product + termination product.

The mutants in the slippery sequence of the wild-type gene *fucA1* were prepared by site-directed mutagenesis from the vector pGEX-11867/3060, described previously (24,27).

The synthetic oligonucleotides used (PRIMM) were the following: *FucA1sm*-rev, 5'-TTTAGGtightATGGTTTGCTCTGTTCTATCTCT-3'; *FucA1sm*-fwd, 5'-GAACACCAATATCACCCTAAAAGAACAAATTCGGGCCCA-3'; *FucA1tm*-rev, 5'-AGGTGGAATGGTCTGTGCTATCGGTG-3'; *FucA1tm*-fwd, 5'-CCAGAACCACCAAATACCACTCTCAAGAAGATCTGGCCCA-3'.

The expressed genes were underlined. Direct sequencing identified the plasmids containing the desired mutations and the mutant genes, named *fucA1*\textsuperscript{m} and *fucA1*\textsuperscript{m}, were completely re-sequenced.

Expression and characterization of Sso-fuc\textsuperscript{B}

The mutant Sso-fuc\textsuperscript{B} was prepared by site-directed mutagenesis from the vector pGEX-11867/3060, by using the same site-directed mutagenesis kit described above. The synthetic oligonucleotides used were *FucA1sm*-rev (described above) and the following mutagenic oligonucleotide: *Fuc-B*, 5'-GAACACCAATATCACCCTAAGAAGATCTGGCCCA-3', where the mismatched nucleotides are underlined. Direct sequencing identified the plasmid containing the desired mutations and the mutant gene, named *fucA1*\textsuperscript{B}, was completely re-sequenced. The enzymatic characterization of Sso-\(α\)-fuc\textsuperscript{B} was performed as described previously (24,27).

Mass spectrometry experiments

Samples of the proteins expressed in *E. coli* from the wild-type gene *fucA1* and the mutants *fucA1*\textsuperscript{A} and *fucA1*\textsuperscript{m}, purified as described, were fractionated on an SDS–PAGE. Protein bands were excised from the gel, washed in 50 mM ammonium bicarbonate, pH 8.0, in 50% acetonitrile, reduced with 10 mM DTT at 56°C for 45 min and alkylated with 55 mM iodoacetamide for 30 min at room temperature in the dark. The gel pieces were washed several times with the buffer, resuspended in 50 mM ammonium bicarbonate and incubated with 100 ng of trypsin for 2 h at 4°C and overnight at 37°C. The supernatant containing peptides was analysed by MALDIMS on an Applied Biosystem Voyager DE-PRO mass spectrometer using \(α\)-cyano-4-hydroxycynamic acid as matrix. Mass calibration was performed by using the standard mixture provided by manufacturer.

Liquid chromatography online tandem mass spectrometry (LCMSMS) analyses were performed on a Q-TOF hybrid mass spectrometer (Micromass, Waters, Milford, MA, USA) coupled with a CapLC capillary chromatographic system (Waters). Peptide ions were selected in the collision cell and fragmented. Analysis of the daughter ion spectra led to the reconstruction of peptide sequences.

Experiments of translation in vitro

Genomic DNA from *S. solfataricus* P2 strain was prepared as described previously (24). A DNA fragment of 1538 nt containing the complete *fucA1* gene, was prepared by PCR, by using the following synthetic oligonucleotides (Genenco, Florence, Italy): *FucA1*-fwd, 5'-CTGGAGGCGCGCTTAA-TAGCAGCACTATAGGTCACTAAATGTCAACAAAATCTCT-3'; *FucA1*-rev, 5'-GACTTTGGCCCGCTATCTATACATTAGGATAACCTTATT-3', in which the sequence corresponding to the genome of *S. solfataricus* is underlined. In the *FucA1*-fwd primer, the sequence of the promoter of the *T7* RNA polymerase is in boldface and the sequence of the BssHII site is shown in italics. The PCR amplification was performed as described previously (24) and the amplification products were cloned in the BssHII site of the plasmid pBluescript II KS+. The *fucA1* gene was completely re-sequenced to check if undesired mutations were introduced by PCR and the recombinant vector obtained, named pBlu-*FucA1*, was used for translation in vitro experiments.

The plasmids expressing the mutant genes *fucA1*\textsuperscript{A}, *fucA1*\textsuperscript{sm} and *fucA1*\textsuperscript{m} for experiments of translation in vitro were prepared by substituting the KpnI–NcoI wild-type fragment, containing the slippery site, with those isolated from the mutants. To check that the resulting plasmids had the correct sequence, the mutant genes were completely re-sequenced.
The mRNAs encoding wild-type fucA1 and its various mutants were obtained by in vitro run-off transcription. About 2 µg of each plasmid was linearized with BssHIII and incubated with 50 U of T7 RNA polymerase for 1 h at 37°C. The transcription mixtures were then treated with 10 U of RNaseA (RNase free) for 30 min. The transcribed RNAs were recovered by extracting the samples twice with phenol (pH 4.7) and once with phenol/chloroform 1:1 followed by precipitation with ethanol. The mRNAs were resuspended in DEPC-treated H2O at the approximate concentration of 0.6 pmol/µl.

In vitro translation assays were performed essentially as described by Condò et al. (28). The samples (25 µl final volume) contained 5 µl of S.olfataricus cell extract, 10 mM KCl, 20 mM Tris–HCl, pH 7.0, 20 mM Mg acetate, 3 mM ATP, 1 mM GTP, 5 µg of bulk S.olfataricus tRNA, 2 µl of [35S]methionine (1200 Ci/mmol at 10 mCi/ml) and ~10 pmol of each mRNA. The mixtures were incubated at 70°C for 45 min. After this time, the synthesized proteins were resolved by electrophoresis 12.5% acrylamide–SDS gels and revealed by autoradiography of the dried gels on an Instant Imager apparatus.

Transcriptional analysis of fucA1

Cells of S.olfataricus, strain P2, were grown in minimal salts culture media supplemented with yeast extract (0.1%), casamino acids (0.1%), plus glucose (0.1%) (YGM) or sucrose (0.1%) (YSM). The extraction of total RNA was performed as reported previously (24). Total RNA was extensively digested with DNase (Ambion, Austin, TX, USA) and the absence of DNA was assessed by the lack of PCR amplification with each sets of primers described below. The RT–PCR experiments were performed as reported previously (24) by using the primers described previously that allowed the amplification of a region of 833 nt (positions 1–833, in which the A of the first ATG codon is numbered as one) overlapping the ORFs SSO11867 and SSO3060 (24).

For real-time PCR experiments total cDNA was obtained using the kit Quantitect RT (Qiagen GmbH, Hilden, Germany) from 500 ng of the same preparation of RNA described above. cDNA was then amplified in a Bio-Rad LightCycler using the DyNAmo HS Syber Green qPCR Kit (Finzymes Oy, Espoo, Finland). Synthetic oligonucleotides (PRIMM) were used for the amplification of a region at the 3' of the ORF SSO3060 were as follows: 5'-Real: 5'T-TAAATGGC-GAACGAGTTTTC-3'; 3'-Real: 5'ATATGCTTTTGTGCG-GGATA-3' for the gene fucA1. 5'GAATGGGTTGATA-CCTGTCG-3' and 5'TTACACCAGCGGCATACAGG-3' for the 16S rRNA gene.

For each amplification of the fucA1 gene was used ~2500-fold more cDNA than that used for the amplification of the 16S rRNA. Controls with no template cDNA were always included. PCR conditions were 15 min at 95°C for initial denaturation, followed by 40 cycles of 10 s at 95°C, 25 s at 56°C and 35 s at 72°C, and a final step of 10 min at 72°C. Product purity was controlled by melting point analysis of setpoints with 0.5°C temperature increase from 72 to 95°C. PCR products were analysed on 2% agarose gels and visualized by ethidium bromide staining.

The expression values of fucA1 gene were normalized to the values determined for the 16S RNA gene. Absolute expression levels were calculated as fucA1/16S ratio in YSM and YSM cells, respectively. Relative mRNA expression levels (YSM/16S ratio) were calculated as (fucA1/16S ratio in YSM cells)/(fucA1/16S ratio in YSM). Each cDNA was used in triplicate for each amplification.

RESULTS

Expression of fucA1 in E.coli

The wild-type fucA1 gene, expressed in E.coli as a GST-fused protein, produced trace amounts of α-fucosidase activity (2.3 × 10−3 units mg−1 after removal of GST), suggesting that a programmed −1 frameshifting may occur in E.coli (24). The enzyme was then purified by using the GST purification system and analysed by SDS–PAGE revealing a major protein band (Figure 2A). The sample and control bands were excised from the gel, digested in situ with trypsin and directly analysed by matrix-assisted laser desorption/ionization mass spectrometry (MALDIMS). As shown in Figure 2B and C, both spectra revealed the occurrence of an identical mass signal at m/z 1244.6 corresponding to a peptide (Peptide A) encompassing the overlapping region of the two ORFs. This result was confirmed by liquid chromatography online tandem mass spectrometry (LCMSMS) analysis of the peptide mixtures. The fragmentation spectra of the two signals showed the common sequence Asn-Phe-Gly-Pro-Val- Thr-Asp-Phe-Gly-Tyr-Lys in which the amino acid from the ORF SSO11867 is underlined. These results unequivocally demonstrate that the protein containing the Peptide A is produced in E.coli by a frameshifting event that occurred exactly within the slippery heptamer predicted from the analysis of the DNA sequence in the region of overlap between the ORFs SSO11867 and SSO3060 (Figure 1A).

Remarkably, the MALDIMS analysis of the products of the wild-type fucA1 gene revealed the presence of a second Peptide B at m/z 1258.6 that is absent in the spectra of the Ssα-fuc control protein (Figure 2B and C). The sequence of Peptide B obtained by LCMSMS (Figure 2D) was Lys-Phe-Gly-Pro-Val- Thr-Asp-Phe-Gly-Tyr-Lys. This sequence differs only by one amino acid from Peptide A demonstrating that the interrupted gene fucA1 expresses in E.coli two full-length polypeptides originated by different −1 frameshifting events. Polypeptide A results from a shift in a site A and it is identical to Ssα-fuc prepared by site-directed mutagenesis (24), suggesting that the expression occurred with the simultaneous P- and A-site slippage. Instead, polypeptide B, named Ssα-fucB, is generated by frameshifting in a second site B as the result of a single P-site slippage (Figure 2E).

To measure the global efficiency of frameshifting in the two sites of the wild-type gene fucA1 we analysed the total extracts of E.coli by western blot using anti-GST antibodies (Figure 2F). Two bands with marked different electrophoretic mobility were observed: the polypeptide of 78.7 ± 1.1 kDa migrated like GST-Ssα-fuc fusion and was identified as originating from frameshifting in either site A or B of fucA1. The protein of 38.1 ± 1.2 kDa, which is not expressed by the mutant gene fucA1A (not shown), had an electrophoretic mobility compatible with GST fused to the polypeptide encoded by the
ORF SSO11867 solely (27 and 9.6 kDa, respectively). This polypeptide originated from the translational termination of the ribosome at the OCH codon of the fucA1 N-terminal ORF (Figure 1A). The calculated ratio of frameshifting to the termination products was 5%.

Preparation and characterization of Ssα-fucB

To test if the full-length α-fucosidase produced by the −1 frameshifting event in site B (Ssα-fucB), resulting from the single P-site slippage has different properties from Ssα-fuc, whose sequence arises from the simultaneous P- and A-site slippage, we prepared the enzyme by site-directed mutagenesis. The slippery sequence in fucA1 A-AAA-AAT was mutated in A-AA T where mutations are underlined. The new mutant gene was named fucA1B. The first G, producing the conservative mutation AAA → AAG, was made to disrupt the slippery sequence and hence reducing the shifting efficiency. The second G was inserted to produce the frameshifting that results in the amino acid sequence of Peptide B. Therefore, the sequence of the two full-length mutant genes fucA1A and fucA1B differs only in the region of the slippery sequence: A-AAG-AAT-TTC-GGC and A-AAG-AAG-TTC-GGC, respectively (the mutations are underlined, the nucleotides in boldface were originally in the −1 frame) (Table 1).

Table 1. Nomenclature and characteristics of the α-fucosidase genes

| Gene name | Status | Name of the recombinant protein | Slippery heptamer* |
|-----------|--------|---------------------------------|--------------------|
| fucA1 wild type | −1 frameshifted | — | A-AAA-AAT |
| fucA1A mutant | Full-length | Ssα-fuc | A-AAG-AAT |
| fucA1B mutant | Full-length | Ssα-fucB | A-AAG-AAG |
| fucA1mut mutant | −1 frameshifted | — | A-AAG-AAT |
| fucA1mut mutant | −1 frameshifted | — | C-AAG-T-AAC |

*Nucleotides modified by substitution and insertion mutations are underlined and in boldface, respectively.

The recombinant Ssα-fucB was purified up to ~95% (Materials and Methods). Gel filtration chromatography demonstrated that in native conditions Ssα-fucB had the same non-amer structure of Ssα-fuc with an identical molecular weight of 508 kDa (data not shown). In addition, Ssα-fucB had the same high substrate selectivity of Ssα-fuc. The two enzymes have high affinity for 4-nitrophenyl-α-L-fucoside (4NP-Fuc) substrate at 65°C; the Km is identical within the experimental error (0.0287 ± 0.005 mM) while the kcat of Ssα-fucB (137 ± 5.7 s⁻¹) is ~48% of that of Ssα-fuc (287 ± 11 s⁻¹). In addition, 4-nitrophenyl-α-L-arabinoside, -rhamnose, 4-nitrophenyl-α-D-glucoside, -xyloside, -galactoside and -mannoside were not substrates of Ssα-fucB as shown previously for Ssα-fuc (24). This suggests that the different amino acid sequence did not...
significantly affect the active site. Both enzymes showed an identical profile of specific activity versus temperature with an optimal temperature higher than 95°C (data not shown). The heat stability and the pH dependence of Ssα-fuc and Ssα-fucB are reported in Figure 3. At 80°C, the optimal growth temperature of S.solfataricus, the half-life of Ssα-fucB is 45 min, almost 4-fold lower than that of Ssα-fuc (Figure 3A). The two enzymes showed different behaviour at pH < 6.0 at which Ssα-fucB is only barely active and stable (Figure 3B); however, the two enzymes showed similar values of specific activity at pHs above 6.0, which is close to the intracellular pH of S.solfataricus (30).

Characterization of the slippery sequence of fucA1 in E.coli

The experimental data reported above indicate that the predicted slippery heptanucleotide in the region of overlap between the ORFs SSO11867 and SSO3060 of the wild-type gene fucA1 could regulate in cis the frameshifting events observed in E.coli. To test this hypothesis, we mutated the sequence A-AAA-AAT into A-AA-G-AAT (mutations are underlined) obtaining the fucA1 single mutant (fucA1sm) and triple mutant (fucA1tm) genes, respectively. It is worth noting that the mutations disrupt the slippery sequence, but they maintain the −1 frameshift between the two ORFs (Table 1).

Surprisingly, the expression of fucA1sm in E.coli produced a full-length polypeptide that, after purification by affinity chromatography and removal of the GST protein, showed the same electrophoretic migration of Ssα-fuc and Ssα-fucB (Figure 4A). This protein was then characterized by mass spectrometry analyses following in situ tryptic digestion. Interestingly, the MALDI spectra revealed the presence of a single peptide encompassing the overlapping region between the two ORFs with a mass value of 1259.7 Da (peptide C; Figure 4B). The sequence of peptide C, determined from the fragmentation spectra obtained by LCMSMS analysis, was Glu-Phe-Gly-Pro-Val-Thr-Asp-Phe-Gly-Tyr-Lys (Figure 4C). Remarkably, apart from the Glu residue, this sequence is identical to that of peptide B produced from fucA1, indicating that in the mutant gene fucA1sm only one of the two frameshifting events observed in the wild-type fucA1 gene had occurred. The presence of a Glu instead of Lys was not unexpected. The mutation A-AAA-AAT→A-AAG-AAT in fucA1sm was conservative in the zero frame of the ORF SSO11867 (AAA→AAG, both encoding Lys), but it produced the mutation AAA→GAA (Lys→Glu) in the −1 frame of the ORF SSO3060.

It is worth noting that the frameshifting efficiency of the gene fucA1sm, calculated by western blot as described above, was 2-folds higher (10%) if compared to fucA1 (5%) (Figure 4D). This indicates that the mutation cancelled the frameshifting site A and, in the same time, enhanced the frameshifting efficiency of site B.

In contrast, the triple mutant fucA1tm produced in E.coli only the low molecular weight band resulting from translational termination (Figure 4D). No full-length protein could be detected in western blots probed with either anti-GST (Figure 4D) or anti-Ssα-fuc antibodies (Figure 4E). These data show that the disruption of the heptameric slippery sequence completely abolished the frameshifting in E.coli confirming that this sequence has a direct role in controlling the frameshifting in vivo.

Expression of fucA1 in S.solfataricus

To test whether fucA1 is expressed in S.solfataricus we analysed the extracts of cells grown on yeast extract, sucrose and casaminoacids medium (YSM). Accurate assays showed that S.solfataricus extracts contained 3.4 × 10⁻⁴ units mg⁻¹ of α-fucosidase activity. These very low amounts hampered the purification of the enzyme. The extracts of S.solfataricus cells grown on YSM revealed by western blot a band of a molecular mass >97 kDa and no signals were detected with the pre-immune serum confirming the specificity of the anti-Ssα-fuc antibodies (Figure 5A). The different molecular mass may result from post-translational modifications occurred in the archaeon or from the incomplete denaturation of a protein complex. In particular, the latter event is not unusual among enzymes from hyperthermophilic archaea (31,32). To test which hypotheses were appropriate, cellular extracts of S.solfataricus were analysed by western blot extending the incubation at 100°C to 2 h. Interestingly, this

Figure 3. Comparison of the stability and pH dependence of Ssα-fuc and Ssα-fucB. (A) Thermal stability of Ssα-fuc (open circles) and Ssα-fucB (closed circles) at 80°C. (B) pH dependence of Ssα-fuc (open circles) and Ssα-fucB (closed circles) at 65°C.
treatment shifted the high-molecular mass band to 67.6 ± 1.2 kDa (Figure 5B and C), which still differs from that of the recombinant Ssα-fuc, 58.9 ± 1.2 kDa, leaving the question on the origin of this difference unsolved. To try to shed some light we immunoprecipitated extracts of S.solfataricus with anti-Ssα-fuc antibodies and we analysed the major protein band by MALDIMS. Unfortunately, we could not observe any peptide compatible with the fucosidase because the heavy IgG chain co-migrated with the band of the expected molecular weight (data not shown).

To test if the scarce amounts of the α-fucosidase in S.solfataricus extracts was the result of reduced expression at transcriptional level, we performed a northern blot analysis of total RNA extracted from cells grown either on YSM or YGM media. We could not observe any signal by using probes matching the 3' of the ORF SSO3060 (data not shown). These results suggest that fucA1 produced a rare transcript; therefore, we analysed the level of mRNA by RT–PCR and by real-time PCR. A band corresponding to the region of overlap between the ORFs SSO11867 and SSO3060 was observed in the RNA extracted from cells grown on YSM

**Figure 5.** Analysis of the expression of the α-fucosidase in S.solfataricus. (A) Western blot analysis of recombinant Ssα-fuc (lanes 1, 2, 5 and 6, 0.14 µg) and of extracts of S.solfataricus cells grown on YSM (lanes 3, 4, 7 and 8, 153 µg). Samples in lanes 1, 3, 5 and 7 were not denatured before loading. The left panel shows the blot probed with anti-Ssα-fuc antibodies; the right panel was probed with the pre-immune serum diluted 1:5000. (B) Western blot analysis: recombinant Ssα-fuc (lanes 1, 2, 3, 5 and 7) incubated at 100°C for 5 min, 1 h and 2 h, respectively; extracts of S.solfataricus cells (lanes 4, 5 and 6, 1 mg) incubated at 100°C for 5 min, 1 h and 2 h, respectively. (C) Western blot analysis of recombinant Ssα-fuc (lane 1, 0.1 µg) incubated at 100°C for 5 min and of extracts of S.solfataricus cells (lane 2, 1 mg) incubated at 100°C for 2 h, respectively. The molecular weight markers were: phosphorylase b (97 000), albumin (66 000), ovalbumin (45 000), carbonic anhydrase (31 000) and trypsin inhibitor (20 1000).
and YGM media, demonstrating that under these conditions the two ORFs were co-transcribed (Figure 6A).

The experiments of real-time PCR shown in Figure 6B demonstrated that rRNA16S was amplified after ~17 cycles while the amplification of fucA1 mRNA was observed after 38 cycles, despite the fact that we used ~2500-fold more cDNA for the amplification of fucA1. This indicates that the gene fucA1 is transcribed at very low level. No significant differences in the fucA1 mRNA level were observed in cells grown in YSM or YGM media. This is further confirmed by the analysis by western blot of the extracts of the same cells of *S.solfataricus* used to prepare the total RNAs, which revealed equal amounts of α-fucosidase in the two extracts (Figure 6C). Therefore, the low α-fucosidase activity observed under the conditions tested is the result of the poor transcription of the fucA1 gene.

**Analysis of the expression of fucA1 in S.solfataricus by in vitro translation**

To determine whether, and with what efficiency, the −1 frameshifting could be performed by *S.solfataricus* ribosomes, mRNAs obtained by in vitro transcription of the cloned wild-type fucA1 gene and the mutants thereof were used to program an in vitro translation system prepared as described by Condó *et al.* (28). To this aim, a promoter of T7 polymerase was inserted ahead of the gene of interest to obtain RNA transcripts endowed with the short 5′-untranslated region of 9 nt observed for the natural fucA1 mRNA (24). Autoradiography of an SDS–PAGE of the translation products (Figure 7) revealed that the wild-type fucA1 transcript produced a tiny but clear band whose molecular weight corresponded to that of the full-length Ssa-fuc obtained by site-directed mutagenesis (24); the latter was translated quite efficiently in the cell-free system in spite of being encoded by a quasi-leaderless mRNA. Judging from the relative intensity of the signals given by the translation products of the wild-type fucA1 and the full-length mutant fucA1A, the efficiency of the −1 frameshifting in the homologous system was ~10%. No signals corresponding to the polypeptides expected from the separated ORFs SSO11867 and SSO3060 (9.6 and 46.5 kDa, respectively) were observed. However, it should be noted that the product of SSO11867, even if synthesized, is too small to be detected in the gel system employed for this experiment. The larger product of ORF SSO3060, on the other hand, is certainly absent. These data unequivocally demonstrate that the ribosomes of *S.solfataricus* can decode the split fucA1 gene by programmed −1 frameshifting with considerable efficiency producing a full-length polypeptide from the two ORFs SSO11867 and SSO3060.

Remarkably, under the same conditions at which fucA1 drives the expression of the full-length protein, we could not observe any product from the fucA1tm and fucA1sm constructs. These data demonstrate that the integrity of the heptanucleotide is essential for the expression of the fucA1 gene in *S.solfataricus*, thus further confirming that the gene is decoded by programmed −1 frameshifting in this organism. In addition, the lack of expression of fucA1tm by translation in vitro in *S.solfataricus* contrasts with the efficient expression of this mutant in *E.coli*, indicating that the two

![Figure 6. Analysis of the expression of fucA1 in different media.](https://academic.oup.com/nar/article-abstract/34/15/4258/3111895)
organsms recognize different sequences regulating the translational frameshifting.

DISCUSSION

The identification of genes whose expression is regulated by recoding events is often serendipitous. In the framework of our studies on glycosidases from hyperthermophiles, we identified in the genome of the archaeon *S. solfataricus* a split gene encoding a putative α-fucosidase, which could be expressed through programmed −1 frameshifting (24).

We tackled this issue by studying the expression of *fucAl* in *S. solfataricus* and in *E. coli* to overcome the problems connected to the scarcity of expression of the α-fucosidase gene and to the manipulation of hyperthermophiles. As already reported by others, in fact, it is a common strategy to study recoding events from different organisms in *E. coli* (23, 33).

The expression in *E. coli* of the wild-type split gene *fucAl* led to the production by frameshifting of two full-length polypeptides with an efficiency of 5%. This is a value higher than that observed in other genes expressed by translational frameshifting in a heterologous system such as the proteins gpG and gpGT (0.3–3.5%) (33).

The gene *fucAl* is expressed in *S. solfataricus* at very low level under the conditions tested. In particular, the transcriptional analysis of the gene revealed that it is expressed at very low level in both YSM and YGM media. Similarly, no differences in the two media could be found by western blot probed with anti-SSa-fuc antibodies, indicating that the low expression of the enzyme in *S. solfataricus* is the result of scarce transcription rather than suppressed translation.

Western blots allowed us to identify a specific band ∼8.7 kDa heavier than that of the recombinant Ssα-fuc and experiments of translation in *vitro* showed that the wild-type gene expresses a full-length polypeptide exhibiting the same molecular mass of the recombinant protein. This demonstrates that the translational machinery of *S. solfataricus* is fully competent to perform programmed frameshifting. It seems likely that the observed discrepancy in molecular mass might arise from post-translational modifications that cannot be produced by the translation in *vitro*. Further experiments are required to characterize the α-1-fucosidase identified in *S. solfataricus*.

MALDIMS and LCMSMS analyses of the products in *E. coli* of the wild-type split gene *fucAl* demonstrated that two independent frameshifting events occurred in *vitro* in the proposed slippery site. In particular, the sequences obtained by LCMSMS demonstrate that peptide A results from a simultaneous backward slippage of both the P- and the A-site tRNAs (Figure 8A). Instead, the sequence of peptide B is the result of the re-positioning on the −1 frame of only the P-site tRNA; in fact, the next incorporated amino acid is specified by the codon in the new frame (Figure 8B). Therefore, the expression by −1 frameshifting of the wild-type gene *fucAl* in *E. coli* follows the models proposed for ribosomal frameshifting (34). We confirmed the significance of the slippery heptanucleotide in promoting the programmed frameshifting in *vitro* by mutating the putative regulatory sequence. The triple mutant *fucAl*tm gave no full-length products; presumably, the mutations in both the P- and in the A-site of the slippery sequence dramatically reduced the efficiency of the −1 frameshifting as observed previously in metazoans (35). This result confirms that the intact slippery sequence in the wild-type gene *fucAl* is absolutely necessary for its expression in *E. coli*. In contrast, surprisingly, the single mutant *fucAl*sm showed an even increased frequency of frameshifting (10%) if compared to the wild-type and produced only one polypeptide by shifting specifically in site B. We explained this result observing that the mutation in the P-site of the slippery sequence A-AAA-AAT→A-AAG-AAT created a novel slippery sequence A-AAG identical to that controlling the expression by programmed −1 frameshifting of a transposase gene in *E. coli* (36). Therefore, apparently, the single mutation inactivated the simultaneous P- and A-site tRNA re-positioning and, in the same time, fostered the shifting efficiency of the tRNA in the P-site. It is worth noting that, instead, in *S. solfataricus*, only the simultaneous slippage is effective (Figure 8B) and even the single mutation in the slippery sequence of *fucAl*tm completely annullated the expression of the gene. This indicates that this sequence is essential in the archaeon and that programmed frameshifting in *S. solfataricus* and *E. coli* exploits different mechanisms. Furthermore, since the only difference between the enzymes produced by the frameshifting sites A and B, Ssα-fuc and Ssα-fucB, respectively, is the stability at 80°C, which is the *S. solfataricus* physiological temperature, the functionality of Ssα-fucB in the archaeon appears questionable.

The reason why *fucAl* is regulated by programmed −1 frameshifting is not known. However, the physiological significance of programmed frameshifting has been assigned to a minority of the cellular genes while for most of them it is still uncertain [see (4) and reference therein; (16)]. This mechanism of recoding is exploited to set the ratio of two polypeptides such as the τ and γ subunits of the DNA polymerase III helozyme in *E. coli* (12). Alternatively, programmed frameshifting balances the expression of a protein, as the bacterial translational release factor 2 and the eukaryotic ornithine decarboxylase antizyme [see (4) and (18) and references therein]. In the case of *fucAl*, the polypeptide encoded by the smaller ORF SSO11867 could never be detected by western blots analyses. In addition, the modelling of Ssα-fuc on the high-resolution crystal structure of the α-1-fucosidase from *Thermotoga maritima* (25, 37) showed that the *fucAl* N-terminal polypeptide is not an independent domain. Moreover, we have shown recently that SSO11867 includes essential catalytic residues (27), excluding the possibility that a functional α-fucosidase can be obtained from the ORF SSO3060 alone. Therefore, several lines of evidence allow us to exclude that
programmed −1 frameshifting is used to set the ratio of two polypeptides of the α-fucosidase from *S. solfataricus*. More probably, this translational mechanism might be required to control the expression level of fucA1.

Noticeably, this is the only fucosidase gene expressed by programmed −1 frameshifting. Among carbohydrate active enzymes, the only example of expression through this recoding mechanism is that reported for a gene encoding for a α(1,2)-fucosyltransferase from *Helicobacter pylori* that is interrupted by a −1 frameshifting (38). In this case, the expression by programmed frameshifting would lead to a functional enzyme synthesizing components of the surface lipopolysaccharides to evade the human immune defensive system. It is hard to parallel this model to fucA1. Nevertheless, the monosaccharide fucose is involved in a variety of biological functions (39). Therefore, the α-L-fucosidase might play a role in the metabolism of fucosylated oligosaccharides; experiments are currently in progress to knockout the wild-type *fucA1* gene and to insert constitutive functional mutants of this gene in *S. solfataricus*.

*fucA1* is the only archaeal α-L-fucosidase gene identified so far; hence, it is probably the result of a horizontal gene transfer event in *S. solfataricus*. However, since there are no α-fucosidases genes regulated by programmed frameshifting in Bacteria and Eukarya, it is tempting to speculate that this sophisticated mechanism of translational regulation pre-existed in *S. solfataricus* and it was applied to the fucosidase gene for physiological reasons. The identification of other genes interrupted by −1 frameshifting in *S. solfataricus* would open the possibility that they are regulated by programmed −1 frameshifting. Recently, the computational analysis of prokaryotic genomes revealed that seven Archaea harbour interrupted coding sequences, but *S. solfataricus* is not included in this study (40). A computational analysis on several archaeal genomes revealed that 34 interrupted genes are present in the genome of *S. solfataricus*, 11 of these genes are composed by two ORFs separated by −1 frameshifting and could be expressed by recoding (B. Cobucci-Ponzano, M. Rossi and M. Moracci, manuscript in preparation).

We have experimentally shown here, for the first time, that programmed −1 frameshifting is present in the Archaea domain. This finding is the missing piece in the puzzle of the phylogenetic distribution of programmed frameshifting demonstrating that this mechanism is universally conserved.

ACKNOWLEDGEMENTS

We are grateful to Maria Carmina Ferrara for technical assistance in the real-time PCR experiments and to Maria Ciaramella and Massimo Di Giulio for useful discussion. We thank the TIGEM-IGB DNA sequencing core for the sequencing of the clones. The IBP-CNR belongs to the Centro Regionale di Competenza in Applicazioni Tecnologico-Industriali di Biomolocole e Biosistemi. P.P., M. R. and M. M. were supported by MIUR project ‘Folding di proteine: l’altra meta del codice genetico’ RBAU015B47. This work was partially supported by the ASI project MoMa n. 1/014/06/0. Funding to pay the Open Access publication charges for this article was provided by MIUR project RBAU015B47.

Conflict of interest statement. None declared.

REFERENCES

1. Gesteland, R.F., Weiss, R.B. and Atkins, J.F. (1992) Recoding: reprogrammed genetic decoding. *Science*, 257, 1640–1641.
2. Gesteland, R.F. and Atkins, J.F. (1996) Recoding: dynamic reprogramming of translation. *Annu. Rev. Biochem.*, 65, 741–768.
3. Farabaugh, P.J. (1996) Programmed translational frameshifting. *Annu. Rev. Genet.*, 30, 507–528.
10. Craigen, W.J. and Caskey, C.T. (1986) Expression of peptide chain
4268 Nucleic Acids Research, 2006, Vol. 34, No. 15
13. Tsuchihashi, Z. and Brown, P.O. (1992) Sequence requirements for
11. Baranov, P.V., Gesteland, R.F. and Atkins, J.F. (2002) Release factor
20. Wilting, R., Schorling, S., Persson, B.C. and Bock, A. (1997)
23. Blight, S.K., Larue, R.C., Mahapatra, A., Longstaff, D.G., Chang, E.,
Zhao, G., Kang, P.T., Green-Church, K.B., Chan, M.K. and Kryzky, J.A.
(2004) Direct charging of tRNA(CUA) with pyrroline in vitro and
in vivo. Nature, 431, 333–335.
24. Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M. and
Moracci, M. (2003) Identification of an archaeal alpha-L-fucosidase
encoded by an interrupted gene. Production of a functional enzyme by
mutations mimicking programmed –1 frameshifting. J. Biol. Chem.,
278, 14622–14631.
25. Rosano, C., Zaccotti, S., Cobucci-Ponzano, B., Mazzone, M., Rossi, M.,
Moracci, M., Petoukhov, M.V., Svergun, D.I. and Bolognesi, M. (2004)
Structural characterization of the nonameric assembly of an Archaeal
alpha-t-fucosidase by synchrotron small angle X-ray scattering.
Biochem. Biophys. Res. Commun., 320, 176–182.
26. Cobucci-Ponzano, B., Trincone, A., Giordano, A., Rossi, M. and
Moracci, M. (2003) Identification of the catalytic nucleophile of the
family 29 alpha-t-fucosidase from Sulfolobus solfataricus via chemical
rescue of an inactive mutant. Biochemistry, 42, 9525–9531.
27. Cobucci-Ponzano, B., Mazzone, M., Rossi, M. and Moracci, M. (2005)
Probing the catalytically essential residues of the alpha-t-fucosidase
from the hyperthermophilic archaeon Sulfolobus solfataricus.
Biochemistry, 44, 6331–6342.
28. Condo, J., Ciammaruconi, A., Benelli, D., Ruggero, D. and Londei, P.
(1999) Cis-acting signals controlling translational initiation in the
thermophilic archaeon Sulfolobus solfataricus. Mol. Microbiol., 34,
377–384.
29. Bradford, M.M. (1976) A rapid and sensitive method for the
quantitation of microgram quantities of protein utilizing the principle of
protein-dye binding. Anal. Biochem., 72, 248–254.
30. Lübben, M. and Schäfer, G. (1989) Chemiosmotic energy conversion of
the archaeabacterial thermooacidophile Sulfolobus acidocaldarius:
oxidative phosphorylation and the presence of an F0-related
N,N′-dicyclohexylcarbodiimide-binding proteolipid. J. Bacteriol.,
171, 6106–6116.
31. Moracci, M., La Volpe, A., Pulitzer, J.F., Rossi, M. and Ciaramella, M.
(1992) Expression of the thermostable beta-galactosidase gene from the
archaeaetium Sulfolobus solfataricus in Saccharomyces cerevisiae
and characterization of a new inducible promoter for heterologous
expression. J. Bacteriol., 174, 873–882.
32. Tschirihashi, Z. and Brown, P.O. (1992) Sequence requirements for
efficient translational frameshifting in the Escherichia coli dnaA gene
and the role of an unstable interaction between tRNA(Lys) and an AAG
lysine codon. Genes Dev., 6, 511–519.
33. Mølbech, N., Atkins, J.F. and Neuhard, J. (1991) Ribosomal –1
frameshifting during decoding of Bacillus subtilis cdd occurs at the
sequence CGA AGA. J. Bacteriol., 181, 2930–2937.
34. Baranov, P.V., Gesteland, R.F. and Atkins, J.F. (2002) Release factor
2 frameshifting sites in different bacteria. EMBO Rep., 4, 373–377.
35. Plant, E.P. and Dinman, J.D. (2006) Comparative study of the effects of
the methanogen methyltransferase. Science, 396, 1462–1466.
36. Baranov, P.V., Gesteland, R.F. and Atkins, J.F. (2004) P-site tRNA is
required for efficient translational frameshifting in the SARS coronavirus.
mRNA. Nature, 431, 273–275.
37. Herr, A.J., Wills, N.M., Nelson, C.C., Gesteland, R.F. and Atkins, J.F.
(2002) Release factor 2 requires high-efficiency frameshift.
38. Sang, D. and Kang, H. (2003) Prokaryotic and eukaryotic translational
machineries respond differently to the frameshifting RNA signal from
plant or animal virus. Virus Res., 92, 165–170.
39. Baranov, P.V., Gesteland, R.F. and Atkins, J.F. (2004) P-site tRNA is
required for efficient translational frameshifting in the SARS coronavirus.
mRNA. Nature, 431, 273–275.
40. Perrodou, E., Deshayes, C., Muller, J., Schaeffer, C., Van Dorsselaer, A.,
Ripp, R., Poch, O., Reynaert, J.M. and Lecompte, O. (2006) ICDS database:
interrupted CoDing sequences in prokaryotic genomes. Nucleic Acids
Res., 34, 338–343.