The EMBO Journal vol.4 no.1 pp.223–229, 1985

Two translational initiation sites in the *infB* gene are used to express initiation factor IF2α and IF2β in *Escherichia coli*

Jacqueline A.Plumbridge, Françoise Deville[1], Christine Sacerdot, Hans Uffe Petersen[2], Yves Cenatiempo[1], Alain Cozzone[1], Marianne Grunberg-Manago and John W.B.Hershey[3]

Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, [1]Laboratoire de Biologie Moléculaire, Université Claude Bernard Lyon 1, 43 Bd. du 11 Novembre 1918, 69621 Villeurbanne, France, [2]Department of Chemistry, Aarhus Universitet, Langelandsgade 140, 8000 Aarhus, Denmark, and [3]Department of Biological Chemistry, School of Medicine, University of California, Davis, CA 95616, USA

Communicated by M.Grunberg-Manago

The gene *infB* codes for the two forms of translational initiation factor IF2: IF2α (97 300 daltons) and IF2β (79 700 daltons). To determine whether the two forms differ at their N terminus, purified IF2α and IF2β were subjected to 11 or more steps of Edman degradation. The N-terminal amino acid sequences are completely different, but match perfectly the DNA sequences at the beginning of the *infB* open reading frame and an in-phase region 471 bp downstream. A fusion was constructed between the proximal half of the *infB* gene and the *lacZ* gene lacking the region coding for the first eight amino acids. The fused gene expresses two products of 170 000 and 150 000 daltons, corresponding to the fused proteins IF2α-β-galactosidase and IF2β-β-galactosidase, which confirms *in vivo* that the IF2 forms differ at their N terminus. A deletion of the 5'-non-translated region of the fused gene, including the Shine/Dalgarno ribosomal binding site, results in the expression of IF2β-β-galactosidase but not IF2α-β-galactosidase. This strongly suggests that IF2β results from independent translation rather than from a precise proteolytic cleavage of IF2α. Further evidence for initiation of protein synthesis at the putative IF2α and IF2β start sites was sought by using an *in vitro* dipeptide synthesis assay. A DNA fragment containing the entire *infB* gene was cloned into three plasmid vectors and the resulting recombinant DNAs were used as templates in assays containing fMet-tRNA and various labelled aminoacyl-tRNAs. Synthesis of both formylmethionyl-threonine (corresponding to the IF2α start sequence) and formylmethionyl-serine (corresponding to the IF2β start sequence) was detected in a ratio of 6:1. These results are consistent with independent initiation events at the IF2α and IF2β sites. The weaker expression at the IF2β start site *in vitro* is also consistent with the fact that IF2α is more abundant than IF2β in intact cells. Thus all of the experimental evidence suggests that the *infB* gene uses two translational initiation sites to express IF2α and IF2β.

Key words: infB gene/initiation factors IF2α and IF2β/*E. coli*

Introduction

Initiation of protein synthesis in *Escherichia coli* is promoted by three proteins called initiation factors IF1, IF2 and IF3. IF2 catalyzes the binding of fMet-tRNA to ribosomes and possesses GTPase activity, IF3 prevents the association of ribosomal subunits and is involved in the binding of mRNA to ribosomes, and IF1 appears to assist in the action of the other factors (for recent reviews, see Grunberg-Manago *et al.*, 1978; Maitra *et al.*, 1982). IF2 is unusual in that it may be isolated from cells as two forms which differ in mol. wt., IF2α (mol. wt. 97 300) and IF2β (mol. wt. 79 700), but which are closely related immunologically and by tryptic peptide mapping (Howe and Hershey, 1982; Fakunding *et al.*, 1972). Although IF2 is sensitive to proteolysis in crude lysates, immunoblot analysis of cells rapidly lysed in SDS buffer consistently revealed IF2α and IF2β in a ratio of ~2:1 (Howe and Hershey, 1982). Moreover two forms of IF2 have been found in every species of Enterobacteria examined (Howe and Hershey, 1984). These results indicate that the two forms are actually present in bacteria and that IF2β is not merely the result of proteolytic cleavage of IF2α during isolation. The existence of two IF2 forms suggests that the two factors may differ in activity and thereby contribute to the control of gene expression. IF2α, but not IF2β, catalyzes the synthesis of β-galactosidase (Eskin *et al.*, 1978) and binds to RNA-Sepharose (Domogatckii *et al.*, 1979), whereas both factors are active in promoting R17 bacteriophage coat protein synthesis from the viral RNA (A.Taguchi, J.G.Howe and J.W.B.Hershey, in preparation).

The gene for IF2, called *infB* has been isolated (Plumbridge *et al.*, 1982). It is located at 69 min on the *E. coli* chromosome and is part of a complex operon of at least five genes. The promoter proximal gene is *metY* coding for a minor form of $tRNA_f^{Met}$, followed by a gene for an unidentified 15-K protein, *nusA* coding for a protein involved in transcription termination, *infB*, and a gene for another unidentified 15-K protein (Ishii *et al.*, 1984; Plumbridge and Springer, 1983; Kurihara and Nakamura, 1983). The *infB* gene subcloned as a 3-kb fragment expresses both IF2α and IF2β (Plumbridge and Springer, 1983), indicating that both IF2 forms are expressed from the same gene. The DNA sequence of this 3-kb fragment shows an open reading frame of 2667 bases preceded by a strong translational initiation site (Sacerdot *et al.*, 1984). This size is consistent with the mol. wt. of IF2α. We were interested to know what is the origin of the lower mol. wt. (IF2β) form of the protein. Initial experiments designed to demonstrate a precursor-product relationship between IF2α and IF2β were completely negative. Exponentially growing *E. coli* cells were pulsed labelled with [^{35}S]methionine and then chased with excess non-radioactive methionine for four generations. Samples analyzed by immunoprecipitation with anti-IF2 showed consistently the same ratio of radioactively labelled IF2α to IF2β throughout the labelling and chase period. Similarly a pulse-chase experiment on plasmid-encoded proteins labelled in a maxicell system showed no hint of a change in the ratio of IF2α to IF2β. Such a change should be expected if IF2β were a proteolyzed version of IF2α. We report here a variety of data which show that IF2α and IF2β differ at their N termini and strongly indicate that IF2α and IF2β are independently synthesized from the same gene.

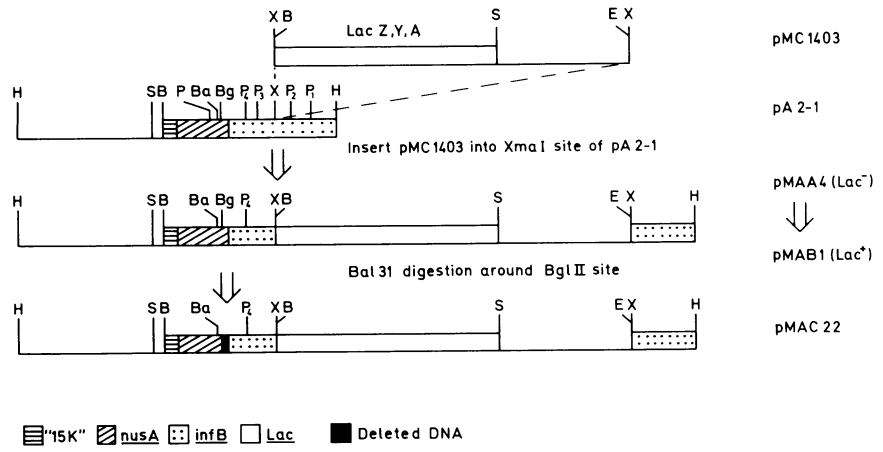

Fig. 1. Construction of *infB-lacZ* fusion plasmids. The single line indicates plasmid vector sequences. The double lines show *E. coli* DNA. The various genes are indicated. Restriction sites are H = *Hind*III; B = *Bam*HI; Bg = *Bgl*II; X = *Xma*I; S = *Sal*I; P = *Pst*I; E = *Eco*RI; pMAB1 and pMAA4 have the same restriction pattern, pMAB1 is a Lac$^+$ derivative of pMAA4 selected as described in Materials and methods.

Results

The two forms of IF2 differ at their N terminus

To determine whether IF2α and IF2β differ at their N or C terminus, a hybrid *infB-lacZ* gene was constructed. The proximal half of the *infB* gene was fused to the *lacZ* gene lacking the region coding for the first eight amino acids, as shown in Figure 1 and described in detail in Materials and methods. This plasmid was Lac$^-$ because the *infB* and *lacZ* coding sequences are not in phase. Spontaneous Lac$^+$ derivatives were selected and analyzed in maxicells. All showed the synthesis of two high mol. wt. protein bands of ~170 and 150 kd, e.g., pMAB1 (Figure 2, lane 1). The difference in mol. wt. between the two protein bands is comparable with that between IF2α and IF2β, and the size of the larger band is consistent with the addition of ~50 kd of protein from IF2 (up to the *Xma*I site) (Sacerdot *et al.*, 1984) to 116 kd of β-galactosidase. Both protein bands are precipitated by antibodies against IF2 and against β-galactosidase (data not shown). Therefore, the two forms of the fusion protein seem to correspond to an IF2β-β-galactosidase and an IF2α-β-galactosidase hybrid protein. This is strong evidence that the IF2α and β proteins differ at the N terminus since a difference at the C terminus would be expected to result in a single fusion protein.

N-terminal amino acid sequences of IF2α and IF2β

Independent evidence that IF2α and IF2β differ at the N terminus was obtained by determining the N-terminal amino acid sequences of the two proteins. Highly purified preparations of the two factors were sequenced by Edman degradation as described in Materials and methods. The sequence for IF2α is:

?-?-Asp-Val-Thr-Ile-Lys-Thr-Leu-Ala-Ala

that for IF2β is:

Ser-Asn-Gln-Gln-Asp-Asp-Met-Thr-Lys-Asn-Ala-Gln-Ala-Gln-Lys-Ala

It is obvious that the two proteins possess entirely different N termini. Both amino acid sequences correspond exactly with sequences derived from the DNA sequence of the *infB* gene (Sacerdot *et al.*, 1984). For IF2α, the first two steps of the Edman degradation gave ambiguous results. The DNA sequence predicts Met and Thr for these two amino acids. The AUG codon for Met is preceded by an excellent Shine/Dalgarno sequence (AAGGA) 6–10 bp upstream, strongly suggesting

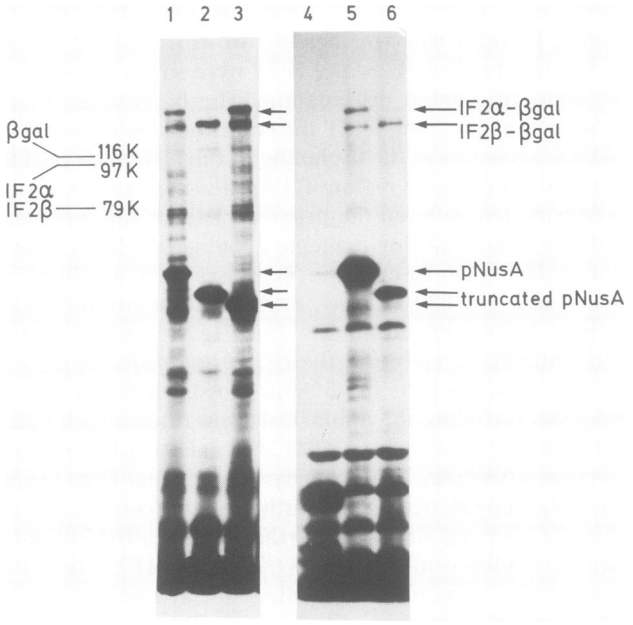

Fig. 2. Synthesis of IF2α- and IF2β-β-galactosidase fusion proteins. Proteins labelled with [^{35}S]methionine were separated on an 8% polyacrylamide SDS gel and located by fluorography. **Lanes 1–3:** maxicell analysis of plasmid-encoded proteins. **Lane 1**, pMAB1, **lane 2**, pMAC22, **lane 3**, pMAC4. **Lanes 4–6:** analysi of phage-encoded proteins in u.v.-irradiated and infected cells. **Lane 4**, λSEW (parental vector), **lane 5**, λMAB1, **lane 6**, λpMAC22. The position of migration of IF2α, IF2β and β-galactosidase are indicated.

that this AUG is the translational start site for IF2α. Initiation at this site would produce a protein of 890 amino acids with a mass of 97 300, in good agreement with the mol. wt. reported for IF2α.

The N-terminal amino acid for IF2β, Ser, is preceded in the DNA sequence by a GUG codon which could act as an initiation site. The putative initiator site lies 471 bases downstream from the initiator site for IF2α in the same reading frame and would produce a protein of 79 700 daltons. The GUG codon is preceded by a purine-rich sequence GGAAAAAGA (base pairs −13 to −5) which is a rather weak ribosomal binding site (Stormo *et al.*, 1982). The amino acid sequence results show that if synthesis is initiated at the GUG site, the fMet is removed from the mature protein.

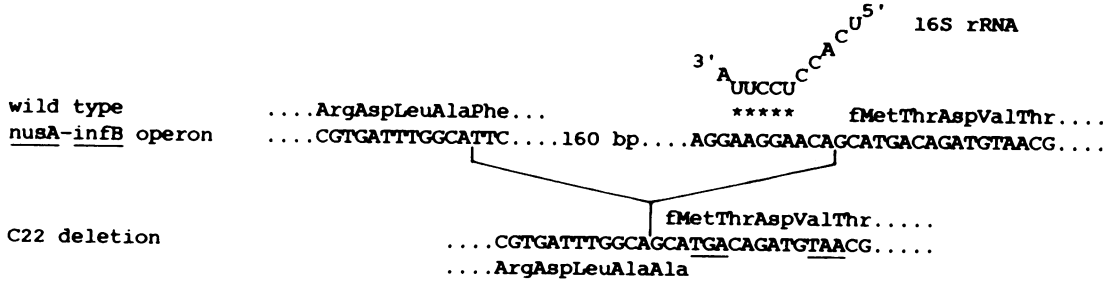

Fig. 3. Sequence around the IF2α initiation site. The DNA sequence around the IF2 initation codon is given and the possible base pairing interactions with 16S rRNA shown. The new sequence caused by the $\triangle$C22 deletion is shown in the bottom part of the figure together with the translated sequence of the *nusA* reading frame. In-phase translational termination sites are underlined. The CA bases at positions -4 and -3 before the fMet AUG codon of IF2 in the $\triangle$C22 strain can be derived either from the *nusA* or *infB* sequence.

IF2β is the result of a second translational initiation site within infB

Two simple models can be proposed to explain the synthesis of IF2β. Either IF2α is precisely cleaved to a lower mol. wt. form, IF2β; or IF2β is synthesized by initiation at its own translational initiation site within the coding sequence of IF2α. A precedent for the former is the *in vitro* conversion of newly synthesized IF3α into IF3β (Lestienne *et al.*, 1982). However, in this case the proteolyzed form is an *in vitro* artifact not found *in vivo*. To investigate the origin of IF2β, deletions were made near the IF2α initiator region. If the two forms of IF2 are the result of independent translational starts, then it should be possible to remove the IF2α start while leaving that for IF2β intact. On the other hand, if IF2β is a proteolyzed product of IF2α, then eliminating the IF2α initiation site must also eliminate IF2β synthesis. Deletions were created by nuclease *Bal*31 digestions around the unique *Bgl*II site of pMAB1 carrying the *infB-lacZ* protein fusion, as described in Materials and methods. The DNA of plasmids with reduced expression of β-galactosidase activity was analyzed with restriction enzymes. A subset with deletions of up to 1 kb around the *Bgl*II site were examined in maxicells. Amongst these, the plasmid pMAC22 (Figure 2, lane 2) no longer causes the synthesis of the IF2α fusion protein, but the IF2β fusion protein is still strongly expressed. pMAC22 also causes the synthesis of a protein somewhat smaller than the wild-type NusA protein synthesized by pMAB1 (Figure 2, lane 1). This is presumed to be a truncated NusA protein. The size of the deletion carried by pMAC22 is ~160 bp. For comparison, Figure 2, lane 3, shows the proteins synthesized from a plasmid, pMAC4, which also carries a deletion of ~160 bp. This plasmid still causes the synthesis of both forms of the IF2-β-galactosidase fusion but the NusA truncated protein is smaller. An explanation for these observations is that in pMAC22 the deletion removes more DNA to the right of the *Bgl*II site, removing the IF2α initiation site and leaving a larger pNusA fragment than in pMAC4 where the deletion leaves both IF2 initiation sites intact but creates a smaller pNusA fragment.

The exact location of the deletion in pMAC22 was determined by sequencing the convenient *Bal*I-*Pst*I$_4$ fragment of pMAB1 and pMAC22 cloned into M13 mp9. The DNA sequence of this region is known (Ishii *et al.*, 1984; Sacerdot *et al.*, 1984). The pMAC22 deletion starts 38 bp before the *Bgl*II site and finishes 136 bp afterwards which places it just in front of the IF2α ATG initiation codon. There is an ambiguity as to the origin of the CA bases at positions -3 and -4 of $\triangle$C22 which can be derived from either *nusA* or *infB* DNA. The deletion removes the Shine/Dalgarno sequence and replaces it with the sequence shown in Figure 3 which does not show the characteristics of a ribosome binding site (Stormo *et al.*, 1982). The *nusA* coding sequence is now out-of-phase with *infB* and presumably terminates at the TGA codon or TAA codon as shown in Figure 3. The mass of pNusA removed is ~5000 daltons, in agreement with the putative pNusA fragment migration observed in gels (Figure 2, lane 2). The *Sal*I fragment of pMAB1 and pMAC22 was cloned into the *Sal*I site of λ bacteriophage and the resultant phages analyzed in the u.v.-irradiated cell system (Plumbridge and Springer, 1983). Once again λMAB1 directed the synthesis of both forms of the fusion protein whereas λMAC22 only caused the synthesis of the IF2β-β-galactosidase protein (Figure 2, lanes 6 and 7).

Table I. Immunoblotting with anti-IF2 of lysates from strains carrying pMAE4 and pMAE$\triangle$C22

Plasmid	IF2α-βlac	IF2β-βlac	IF2α	IF2β
pMAE4 wt operon	10 000	1300	3070	1000
pMAE$\triangle$C22 IF2 deletion	380	1200	3100	1070

Immunoblotting was carried out as described (Howe and Hershey, 1981). The numbers given are the c.p.m. of ^{125}I-labelled protein A bound to each of the four protein bands detected with anti-IF2. Numbers are corrected to 10 μg protein of cellular extract. IF2α and IF2β arise from expression of the *infB* gene in the *E. coli* chromosome, and serve as internal controls.

The *nusA-infB* fragment of DNA used in these constructions is missing the principle promoter of the operon which is situated 850 bp in front of *nusA* (Ishii *et al.*, 1984). The *infB-lacZ* fusion has been reconstructed on a larger fragment carrying the whole operon (pMAE4) (manuscript in preparation). The $\triangle$C22 has been transferred to this plasmid giving pMAE$\triangle$C22. Strains carrying either pMAE4 or pMAE$\triangle$C22 were analyzed by quantitative immunoblotting with anti-IF2. By this technique, a very small amount of the IF2α-β-galactosidase protein was detected with pMAE$\triangle$C22. Much larger amounts of the IF2β-β-galactosidase fusion were detected. The amount of IF2α fusion protein from pMAE$\triangle$C22 was ~5% of that from the pMAE4 plasmid (Table I). This gives a quantitative indication of the change in efficiency of the expression of IF2α in the two cases.

The infB gene directs the in vitro synthesis of two initiating dipeptides

In order to test whether IF2α and IF2β synthesis originates at independent translational initiation sites, we employed the *in vitro* dipeptide assay system of Weissbach and co-workers (Cenatiempo *et al.*, 1982a; Robakis *et al.*, 1981). The coupled transcription-translation system is dependent on exogenous

Table II. Plasmids used for dipeptide synthesis

Vector plasmid	Antibiotic resistance	Recombinant plasmid	Antibiotic resistance
pBR322	Amp, Tc	pB16-1	Amp
pACYC184	Cm, Tc	pA3-1	Cm
pRLS100	Cm, Tc	pFY16-1	Cm

All recombinant plasmids carry the *Bgl*II-*Hind*III segment of pA2-1 containing the *infB* gene inserted within the Tc region of their respective vectors.

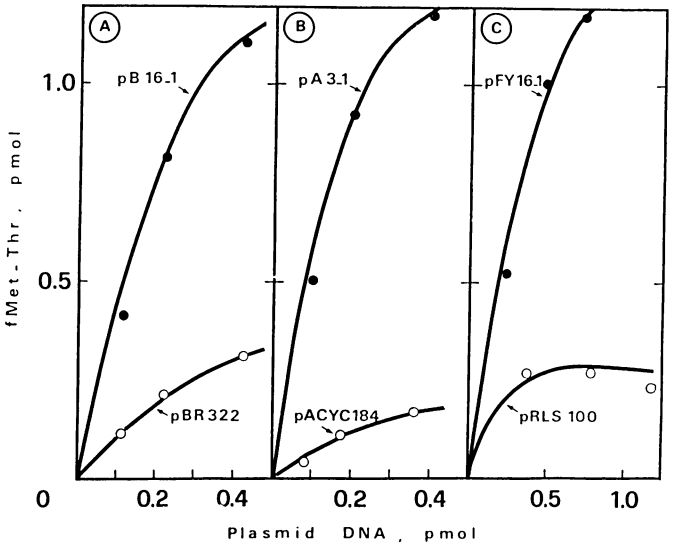

Fig. 4. fMet-Thr synthesis direcdted by various plasmids. The amount of fMet-Thr was determined using [^{3}H]Thr-tRNA$_4^{Thr}$. Incubation conditions are described in Materials and methods.

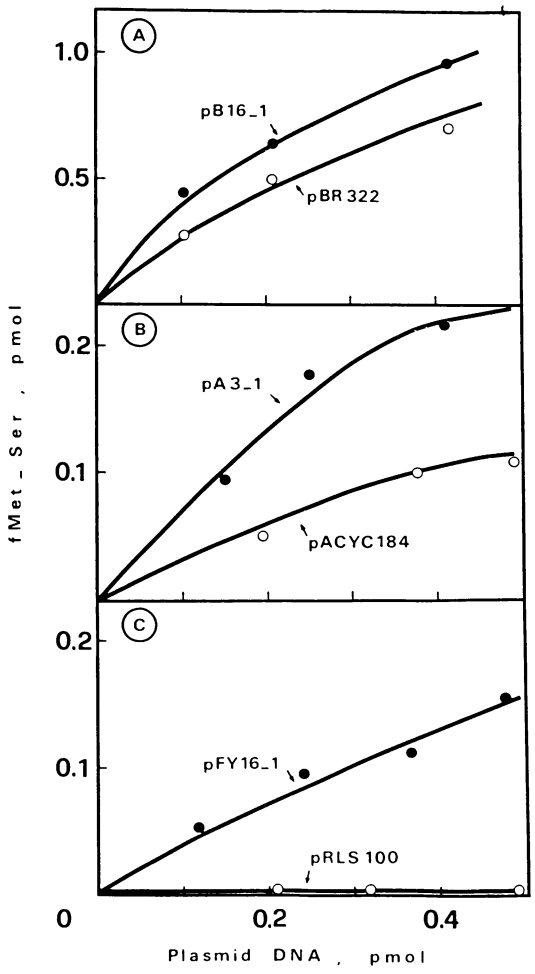

Fig. 5. fMet-Ser synthesis directed by various plasmids in the presence of [^{3}H]Ser-tRNA$_3^{Ser}$. Incubation conditions are described in Materials and methods.

DNA and synthesizes formylmethionyl dipeptides from highly purified aminoacyl-tRNA derivatives. The *in vitro* synthesis of dipeptides faithfully mimics *in vivo* gene expression. The *Bgl*II to *Hind*III fragment of pA2-1 contains DNA coding for the C terminus of pNusA, the whole of IF2 and the N terminus of an unidentified 15-kd protein (Sacerdot *et al.*, 1984). Therefore, the fragment carries the putative initiation sites of IF2α which codes for fMet-Thr (AUG-ACA), IF2β coding for fMet-Ser (GUG-AGC), as well as the 15-kd protein coding for fMet-Ala (AUG-GCG). The same fragment has been cloned in three different plasmid vectors (Table II). Initiation at the IF2α site was tested with purified tRNA$_4^{Thr}$ (anticodon UGU) as described in Materials and methods, and the results with the three plasmid DNAs are shown in Figure 4. Some synthesis of fMet-Thr is observed with the control plasmid DNAs not carrying the *infB* insert (pBR322, pACYC184, pRLS100), whereas 4- to 6-fold greater synthesis is seen when the recombinant *infB* plasmids (pB16-1, pA3-1, pFY16-1) are used. This indicates that the *infB* insert directs the synthesis of fMet-Thr, presumably from the putative IF2α initiator site (see Discussion).

To test for initiation at the IF2β site, tRNA$_3^{Ser}$ (anticodon GUC) was used and the results are shown in Figure 5. The control plasmid pBR322 yields a relatively high amount of fMet-Ser and only a small increase in the formation of the dipeptide is detected when pB16-1 is used as template. This result is not surprising since fMet-Ser is the initial dipeptide of β-lactamase, whose gene is present on both plasmids. Although it does not contain the β-lactamase gene, pACYC184 also directs the synthesis of a significant amount of fMet-Ser. Nevertheless, a 2-fold increase in dipeptide formation is observed when pA3-1 is used as template. In contrast, pRLS100 is totally inactive in fMet-Ser formation, whereas pFY16-1 directs the synthesis of a significant amount of fMet-Ser. This indicates that a second initiation site in the *infB* DNA insert is recognized by the assay system, albeit a somewhat weaker site. If one assumes that the amounts of fMet-Thr and fMet-Ser formed under our experimental conditions are directly comparable, then the former dipeptide is generally synthesized to a greater extent than the latter. For example, when pFY16-1 is used as template, a fMet-Thr/fMet-Ser ratio close to six can be calculated (Figures 4 and 5).

To demonstrate the specificity of the synthesis of the two dipeptides, a number of assays were carried out by omitting the isoacceptor species used in the complete system or replacing them by other species or unfractionated tRNA. The results in Table III show first that dipeptide synthesis is completely dependent on the presence of the initiator tRNA (tRNA$_f^{Met}$). Replacing tRNA$_4^{Thr}$ (anticodon = UGU) by tRNA$_3^{Thr}$ (anticodon = GGU) does not adversely affect fMet-Thr synthesis. One explanation for the lack of dependency on tRNA^{Thr} isoacceptor species could be that a misreading by a 'two out of three' mechanism is observed in this particular case (Mitra *et al.*, 1977). On the other hand, fMet-

Table III. tRNA dependency of dipeptide synthesis

Initiator fMet-tRNA$_f^{Met}$	Second aminoacyl-tRNA	Dipeptide Type	Dipeptide Amount, pmol
+	Thr-tRNA$_4^{Thr}$	fMet-Thr	1.1
−	Thr-tRNA$_4^{Thr}$		0
+	Thr-tRNA$_3^{Thr}$		1.1
+	Unfract. Thr-tRNA		1.1
+	Ser-tRNA$_3^{Ser}$	fMet-Ser	0.2
−	Ser-tRNA$_3^{Ser}$		0
+	Ser-tRNA$_1^{Ser}$		0
+	Unfract. Ser-tRNA		0.2
+	Ala-tRNA$_3^{Ala}$	fMet-Ala	0
+	Unfract. Leu-tRNA	fMet-Leu	0
+	Unfract. Asn-tRNA	fMet-Asn	0

The recombinant plasmid pFY16-1 (1 pmol) was used to study the synthesis of five different types of dipeptide, as listed. Most assays were carried out in the presence (+) of the initiator fMet-tRNA$_f^{Met}$, and a few others in its absence (−). The second aminoacyl-tRNA used for dipeptide formation was either a pure preparation of one particular charged tRNA or a mixture of all tRNA species, as indicated. Incubation conditions were those described in Materials and methods. Expressed values are corrected for the background dipeptide synthesis directed by the control vector plasmid pRLS100.

Ser synthesis is dependent on the specific tRNA isoaccepting species: if tRNA$_1^{Ser}$ (anticodon = UGA) is substituted for tRNA$_3^{Ser}$ (anticodon = GCU), no dipeptide formation occurs. Interestingly, unfractionated *E. coli* tRNAs aminoacylated with Thr or Ser are as active in dipeptide formation as pure tRNA$_4^{Thr}$ or tRNA$_3^{Ser}$ species, respectively. In contrast, Table III shows that tRNA$_3^{Ala}$ (anticodon = GGC) or unfractionated tRNA charged with Leu or Asn are inactive in the presence of pFY16-1.

Discussion

The results reported here strongly indicate that IF2α and IF2β are synthesized from two independent translational initiation sites both *in vitro* and *in vivo*. Analyses of *infB-lacZ* gene fusion products and the determination of the N-terminal amino acid sequences clearly establish that the two proteins differ at their N terminus. Deletion of the Shine/Dalgarno ribosomal binding site for IF2α results in greatly decreased synthesis of IF2α but has no effect on IF2β synthesis. This observation supports the hypothesis of independent intiation sites and argues against the view that IF2β is a degradation product of IF2α. The presence of two distinct translational initiation sites in the *infB* region is also supported by the *in vitro* synthesis of fMet-Thr and fMet-Ser dipeptides. There is a little doubt that the IF2α site is active, given its AUG initiator codon and strong Shine/Dalgarno sequence. However, the putative IF2β site involves a GUG initiator codon and a rather weak Shine/Dalgarno sequence, and therefore is likely less active than the IF2α site. A weaker IF2β site is consistent with lower IF2β levels in cells and 6-fold less efficient fMet-Ser synthesis *in vitro*. However, it cannot be excluded that the fMet-Ser dipeptide is coded, at least in part, by a different region of the *infB* gene. The sequence of the *Bgl*II-*Hind*III fragment was examined in all three reading frames for Ser codons of the type AGX preceded by an ATG or GTG. In addition to the IF2β initiation site, three other sequences in the *Bgl*II-*Hind*III fragment (Sacerdot *et al.*, 1984) could in theory code for fMet-Ser: GTG-AGC (residues 668–673); ATG.AGC (residues 1273–1278); and ATG.AGC (residues 2485–2490). In the first and third instances, possible Shine/Dalgarno sequences are weak or non-existent. However, residues 1273–1278 are preceded by a strong ribosomal binding site, GAGG (residues 1264–1267), and thus could score in the dipeptide assay. It is clear, however, that this in-phase site is not expressed *in vivo* since no IF2-related protein of ~56 kd is detected in cell lysates.

The present data do not give any information about the mRNA species which code for the two forms. The major promoter for the operon has been located before the *metY* gene nearly 2 kb before *infB* (Ishii *et al.*, 1984). Other, presumably minor, promoters have been located within the operon (Plumbridge and Springer, 1983; Y.Nakamura, personal communication). It is not known, for example, whether IF2α is made from a long transcript and IF2β from a second transcript. However, preliminary S1 mapping experiments failed to reveal mRNAs which begin between the two putative IF2 initiation sites and thus contain the IF2β site but not the IF2α site (results not shown). Positioning a λpL promoter before *nusA* equally stimulates the synthesis of pNusA, IF2α and IF2β, suggesting that the mRNA transcript from this promoter is used to translate all three proteins. Some insight into how both IF2α and IF2β may be expessed from the same mRNA is gained by an analysis of the codon usage in *infB* (Sacerdot *et al.*, 1984). The region of *infB* corresponding to IF2β contains codons used frequently, whereas the region upstream from the IF2β start site, which is unique to IF2α, has numerous rarer codons and is expected to be less efficiently translated. The pausing of ribosomes at such rarer codons might be necessary for other ribosomes to initate at the IF2β start site.

Two in-phase translational initiation sites in a gene, giving two related protein products of differing mol. wt., is unusual but not without precedent in bacteria. Two proteins, 78 kd and 69 kd, are synthesized from the *cheA* locus which is required for chemotactic behavior; analysis of nonsense and missense mutations show that the two proteins are translated in the same phase and differ at their N termini (Smith and Parkinson, 1980). In addition, multiple translational reinitiations have been observed after an amber codon in *lacI* (Files *et al.*, 1974). Several examples of two in-phase proteins being synthesized from the same mRNA at different initiation sites are known in bacteriophages: the λ gene coding for proteins *C* and *Nu3* involved in prohead formation (Shaw and Murialdo, 1980); gene 4 of T7, the DNA primase (Dunn and Studier, 1981); and the two forms of the gene A protein in φX174 and G4 (Linney and Hayashi, 1973; Godson *et al.*, 1978).

A reasonable rationale for multiple initiation sites in bacteriophages is to maximize use of DNA in systems where DNA coding capacity is limited. However, this cannot be the reason for genes of the *E. coli* chromosome. For the case of the λ*C* and *Nu3* genes, Shaw and Murialdo (1980) propose that there is a functional interaction between pC and pNu3 which is facilitated by their common structural domains. Smith and Parkinson (1980) consider the same explanation for the *cheA* products, but also suggest a second explanation: the common C terminus would allow them to carry out a similar function whereas the different N termini would give a specificity to their interactions. No evidence exists for an interaction between IF2 molecules and so the latter hypothesis, a discriminatory role for the two forms of IF2, seems more attractive. It has been reported that IF2β does not sup-

port the *in vitro* synthesis of β-galactosidase (Eskin *et al.*, 1978) whereas both IF2α and IF2β support R17 coat protein synthesis (A.Taguchi, J.G.Howe and J.W.B.Hershey, in preparation). It is interesting to compare these results with the observation that nitrofuran antibiotics specifically inhibit the initiation of certain genes like β-galactosidase whereas phage encoded proteins are relatively insensitive to the antibiotic (Herrlich and Schweiger, 1976).

Materials and methods

Materials

Unfractionated *E. coli* tRNA and purified $tRNA_f^{Met}$ were purchased from Boehringer (Mannheim, FRG), the other purified tRNA isoacceptor species were from Subriden RNA (Rolling Bay, WA) or kindly supplied by N.Brot (Roche Institute of Molecular Biology, Nutley, NJ). L-[^{3}H]amino acids (17–55 Ci/mmol) were from CEA (Saclay, France) or from Amersham Center (Amersham, UK). A 0.25 M DEAE salt eluate (Kung *et al.*, 1975) was used as a source of enzymes to acylate tRNA species and to formylate L-methionyl-$tRNA_f^{Met}$ (Dickerman *et al.*, 1967). N-Formyl-L-methionine and N-formyl-L-methionyl-L-alanine were purchased from Serva (Heidelberg, FRG). N-Formyl-L-methionyl-L-serine and N-formyl-L-methionyl-L-threonine were synthesized as already described (Robakis *et al.*, 1981). Pre-coated t.l.c. plates (silica gel 60, 250 μm) were obtained from Merck (Darmstadt, FRG). Dowex 50WX-8 was purchased from Fluka (Buchs, Switzerland).

Initiation factors IF1, IF2α, IF2β and IF3 were prepared according to Hershey *et al.* (1977), and were >95% pure. Anti-IF2 was obtained from rabbits as described (Howe *et al.*, 1978). EF-Tu was a generous gift of R.Leberman and B.Antonsson (E.M.B.L., Grenoble, France). RNA polymerase was purified as previously reported (Burgess and Jendrisak, 1975). Restriction enzymes were from commercial sources and used as recommended by the manufacturers or as described (Maniatis *et al.*, 1982).

Standard techniques

General genetic and cloning techniques have been described previously (Miller, 1982; Maniatis *et al.*, 1982). Antibiotics were employed at the following concentrations: ampicillin (amp) 100 μg/ml in plates and 500 μg/ml in liquid cultures; chloramphenicol (cm) 25 μg/ml; tetracycline (tc) 10 μg/ml.

Construction of a fusion between infB and lacZ

The cloramphenicol-resistant plasmid, pA2-1, carries the genes *nusA* and *infB* (Plumbridge and Springer, 1983). pMC1403 is an ampicillin-resistant plasmid carrying the *lacZ* gene starting from the eighth amino acid codon onward, thus missing all the *lac* transcription and translation signals (Casadaban *et al.*, 1980). Both plasmids were digested with *Xma*I [unique sites in both plasmids (Figure 1)], mixed, ligated and used to transform a Δ*lac* strain. Amp- and cm-resistant plasmids were selected on plates containing 5-bromo-4-chloro-3-indolyl-β-D-galacto-pyranoside (Xgal). After 48 h a few colonies were very pale blue. These were purified, plasmid DNA analyzed and the plasmid pMAA4 (Figure 1) identified. The lack of a strong blue (LacZ$^+$) response to Xgal from this plasmid implies that the fusion between *infB* and *lacZ* at the *Xma*I site was not in-phase. Comparison of the DNA sequence of *infB* (Sacerdot *et al.*, 1984) with that at the beginning of *lacZ* shows this to be true. Strongly Lac$^+$ derivatives of pMAA4 were selected by plating 1 ml of a saturated culture (resuspended in 10^{-2} M $MgSO_4$) on minimal lactose plates containing ampicillin and chloramphenicol and the auxotrophic amino acids. Twelve colonies were picked and purified several times on minimal lactose plates and also on MacConkey lactose plates both containing ampicillin and chloramphenicol to eliminate any Lac$^-$ segregants. Plasmid DNA was analyzed and one clone, pMAB1, was chosen since its restriction pattern was identical to that of pMAA4. pMAB1 still carries both the junction sites of *Xma*I and *Bam*HI with no appreciable deletion of DNA. We presume a small rearrangement of DNA has occurred either in *infB* or *lacZ* near the junction which puts *infB* and *lacZ* into the same reading phase.

Construction of deletions

*Bal*31 digestions around the unique *Bgl*II site of pMAB1 were carried out for varying times as described (Maniatis *et al.*, 1982). After ligation, the DNA was used to transform a Δ*lac*, recA strain and the resultant plasmids analyzed on LB Xgal and MacConkey lactose plates containing amp and cm. Plasmid DNA was prepared from strains which were less red than pMAB1 on lactose but which were still blue on Xgal plates and examined with a variety of restriction enzymes known to cut near the *Bgl*II site. Those plasmids with deletions near the beginning of *infB* were examined by the maxicell system as described (Plumbridge *et al.*, 1982). To sequence the IF2α deletion, the fragment *Bal*I-*Pst*I_4 of pMAB1 and pMAC22 was cloned into M13 mp9 between the *Sma*I and *Pst*I sites. Dideoxy sequencing was carried out as described (Sanger *et al.*, 1980).

Construction of pFY16-1

The recombinant plasmid pB16-1 (Plumbridge and Springer, 1983) which expresses *infB* carries the 3.1-kb *Bgl*II-*Hind*III fragment of pA2-1 cloned into the *Bam*HI-*Hind*III site of pBR322. A *Sal*I-*Hind*III fragment (carrying the *infB* gene between *Bgl*II-*Hind*III plus a small amount of Tc^R gene) was recloned into pRLS100 selecting for cm resistance and screening for amp- and tc-sensitive clones. pRLS100 is a pBR325 derivative where the beginning of the β-lactamase gene has been deleted (kindly provided by N.Brot, Roche Institute of Molecular Biology, Nutley, NJ, USA).

DNA-directed dipeptide synthesis

The incubation conditions for *in vitro* synthesis of dipeptides were slightly modified from Cenatiempo *et al.* (1982b); Robakis *et al.* (1983). Briefly, the incubation mixture contained in a total volume of 35 μl: 30 mM Tris-acetate (pH 7.5). 10 mM sodium dimethylglutarate (pH 6.0); 35 mM ammonium acetate; 2 mM DTT; 9–11 mM magnesium acetate; 2.9 mM ATP, 0.7 mM CTP, GTP and UTP; 28 mM phosphoenolpyruvate; 0.5 μg of pyruvate kinase; 39 mM potassium acetate; 0.8 mM spermidine; 4% polyethyleneglycol 6000; 0.3 μg of IF1; 0.5 μg of IF2; 0.6 μg of IF3 (or 15 μg of a ribosomal salt-wash containing all three factors); 0.4 μg of EF-Tu; 10 μg of RNA polymerase; 0.6 A_{260} unit of NH_4Cl-washed *E. coli* 70S ribosomes; 10 pmol of fMet-$tRNA_f^{Met}$ and 10 pmol of the second aminoacyl-tRNA, labelled with a [^{3}H]amino acid. The reaction was initiated by adding 0.1–1.0 pmol of plasmid DNA. The reaction was stopped with 2.5 μl of 1 M NaOH and the mixture was incubated for an additional 10 min at 37°C to hydrolyze any peptidyl-tRNA. The standard assay for dipeptide synthesis was either extraction into ethylacetate or chromatography on a minicolumn of Dowex 50WX-8 (H^+ form) as described previously (Robakis *et al.*, 1981; Weissbach *et al.*, 1984). Alternatively, separation of dipeptides by t.l.c. was used, and migration was compared with synthetic dipeptide standards.

N-terminal amino acid sequence determination

Approximately 10 nmol each of IF2α and IF2β were precipitated with 10% trichloroacetic acid. The precipitates were washed twice with acetone and dissolved in 70 μl formic acid. The N-terminal amino acid sequences were determined by step-wise Edman degradation in a Beckman 890 C sequencer using a 0.25 M Quadrol in the Beckman protein program 12 2974 (Skorstensgaard *et al.*, 1982).

Acknowledgements

We thank T.E.Petersen and S.Magnusson for the N-terminal amino acid sequencing, M.Springer for helpful discussions and Dorothy Derania and Sylvie Laveran for expert typing. This work was supported by grants from the American Cancer Society (NP-70) (to J.H.), The Danish Natural Science Research Council and The Niels Bohr Fund (to H.U.P.), and from the CNRS (Grant No 18), MRI [No 82 V 1289, No 5072 (to Y.C.)], INSERM (No 823.008 and 831.013), Fondation pour la Recherche Médicale and E.I.du Pont (to M.G.-M.) and from the A.T.P. CNRS-NSF (to J.H. and M.G.-M.).

References

Burgess,R.R. and Jendrisak,J.J. (1975) *Biochemistry (Wash.)*, **14**, 4634-4638.
Casadaban,M.J., Chou,J. and Cohen,S.N. (1980) *J. Bacteriol.*, **143**, 971-980.
Cenatiempo,Y., Robakis,N., Meza Basso,L., Brot,N., Weissbach,H. and Reid,B. (1982a) *Proc. Natl. Acad. Sci. USA*, **79**, 1466-1468.
Cenatiempo,Y., Robakis,N., Reid,B., Weissbach,H. and Brot,N. (1982b) *Arch. Biochem. Biophys.*, **218**, 572-578.
Dickerman,H.W., Steers,E., Reidfield,B. and Weissbach,H. (1967) *J. Biol. Chem.*, **242**, 1522-1525.
Domogatckii,C.P., Vlassik,T.H. and Beziepkina,T.A. (1979) *Proc. Acad. Sci. USSR*, **248**, 240-243.
Dunn,J.J. and Studier,F.W. (1981) *J. Mol. Biol.*, **148**, 303-330.
Eskin,B., Treadwell,B., Redfield,B., Spears,C., Kung,H. and Weissbach,H. (1978) *Arch. Biochem. Biophys.*, **189**, 531-534.
Fakunding,J.F., Traugh,J.A., Traut,R.R. and Hershey,J.W.B. (1972) *J. Biol. Chem.*, **247**, 6365-6367.
Files,J.G., Weber,K. and Miller,J.H. (1974) *Proc. Natl. Acad. Sci. USA*, **71**, 667-670.
Godson,G.N., Barrel,B.G., Staden,R. and Fiddes,J.C. (1978) *Nature*, **276**, 236-247.
Grunberg-Manago,M., Buckingham,R.H., Cooperman,B.S. and Hershey, J.W.B. (1978) *Symp. Soc. Gen. Microbiol.*, **28**, 27-110.
Herrlich,P. and Schweiger,M. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 3386-3390.
Hershey,J.W.B., Yanov,J., Johnston,K. and Fakunding,J.L. (1977) *Arch. Biochem. Biophys.*, **182**, 626-638.
Howe,J.G. and Hershey,J.W.B. (1981) *J. Biol. Chem.*, **256**, 12836-12839.
Howe,J.G. and Hershey,J.W.B. (1982) *Arch. Biochem. Biophys.*, **214**, 446-451.

Howe,J.G. and Hershey,J.W.B. (1984) *Arch. Microbiol.*, in press.
Howe,J.G., Yanov,J., Meyer,L., Johnston,K. and Hershey,J.W.B. (1978) *Arch. Biochem. Biophys.*, **191**, 813-820.
Ishii,S., Kuroti,K. and Imamoto,F. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 409-413.
Kung,H.F., Speaks,C. and Weissbach,H. (1975) *J. Biol. Chem.*, **250**, 1556-1562.
Kurihara,T. and Nakamura,Y. (1983) *Mol. Gen. Gent.*, **190**, 189-195.
Lestienne,P., Dondon,J., Plumbridge,J.A., Howe,J.G., Mayaux,J.F., Springer,M., Blanquet,S., Hershey,J.W.B. and Grunberg-Manago,M. (1982) *Eur. J. Biochem.*, **123**, 483-488.
Linney,E. and Hayashi,M. (1973) *Nature New Biol.*, **245**, 6-8.
Maitra,N., Stringer,E.A. and Chandhuri,A. (1982) *Annu. Rev. Biochem.*, **51**, 869-900.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.
Miller,J.H. (1972) *Experiments in Molecular Genetics*, published by Cold Spring Harbor Laboratory Press, NY.
Mitra,S.K., Lustig,F., Akesson,B., Lagerkvist,U. and Strid,L. (1977) *J. Biol. Chem.*, **252**, 471-478.
Plumbridge,J.A. and Springer,M. (1983) *J. Mol. Biol.*, **167**, 227-243.
Plumbridge,J.A., Howe,J.G., Springer,M., Touati-Schwartz,D., Hershey, J.W.B. and Grunberg-Manago,M. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 5033-5037.
Robakis,N., Meza-Basso,L., Brot,N. and Weissbach,H. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4261-4264.
Robakis,N., Cenatiempo,Y., Meza-Basso,L., Brot,N. and Weissbach,H. (1983) *Methods Enzymol.*, **101**, 690-706.
Sacerdot,C., Dessen,P., Plumbridge,J.A., Hershey,J.W.B. and Grunberg-Manago,M. (1984) *Proc. Natl. Acad. Sci. USA*, in press.
Sanger,F., Coulson,A.R., Barrel,B.G., Smith,A.J.H. and Roe,B.A. (1980) *J. Mol. Biol.*, **143**, 161-178.
Shaw,J.E. and Murialdo,H. (1980) *Nature*, **283**, 30-35.
Skorstengaard,K., Thøgersen,H.C., Vibe-Pedersen,K., Petersen,T.E. and Magnusson,S. (1982) *Eur. J. Biochem.*, **128**, 605-623.
Smith,R.A. and Parkinson,J.S. (1980) *Proc. Natl. Acad. Sci. USA*, **77**, 5370-5374.
Stormo,G.D., Schneider,T.D. and Gold,L.M. (1982) *Nucleic Acids Res.*, **10**, 2971-2976.
Weissbach,H., Robakis,N., Cenatiempo,Y. and Brot,N. (1983) *Biotechniques*, **2**, 16-22.

Received on 23 October 1984