Volume 15 Number 13 1987 Nucleic Acids Research

Cloning and mapping of *infA*, the gene for protein synthesis initiation factor IF1

John F.Sands, Helen S.Cummings, Christine Sacerdot[1], Liliane Dondon[1], Marianne Grunberg-Manago[1] and John W.B.Hershey

Department of Biological Chemistry, School of Medicine, University of California, Davis, CA 95616, USA and [1]Institut de Biologie Physico-Chimique, 13 rue Pierre et Marie Curie, 75005 Paris, France

Received March 30, 1987; Accepted June 8, 1987 Accession no. Y00373

ABSTRACT

The gene for translation initiation factor IF1, *infA*, has been identified by using two synthetic oligonucleotides to screen a Charon 30 library of *Escherichia coli* DNA. A recombinant lambda phage, C1921, was purified from a plaque positive for both probes. A 2.8 kb *Bgl*II fragment and a 2.0 kb *Hind*III fragment isolated from C1921 were subcloned into the *Bam*HI and *Hind*III sites of pBR322 to yield pTB7 and pTH2 respectively. Synthesis of IF1 in maxicells transformed with pTB7 or pTH2 indicates the presence of *infA* in both inserts. This was confirmed by DNA sequencing: a region was found that codes for a 8,119 dalton protein with an amino acid sequence corresponding to IF1. The chromosomal location of *infA* was determined by mapping the closely linked beta-lactamase gene (Amp^r) in pTB7 and pTH2. pTB7 and pTH2 were transformed into *polA* Hfr hosts, and integration of the plasmid by homologous recombination near *infA* was selected on the basis of ampicillin resistance. The site of integration was confirmed by Southern blot analysis of restriction nuclease digested wild type and transformed genomic DNA. The Amp^r marker (and therefore *infA*) was mapped to about 20 minutes by Hfr interrupted matings and P1 transduction experiments. The structure and regulation of the *infA* operon currently are being investigated.

INTRODUCTION

Initiation of protein synthesis in bacteria is promoted by three initiation factors (6): IF1, IF2 and IF3. IF2 occurs in two molecular weight forms, IF2α (97,300 Da) and IF2β (79,700 Da). It is involved in the binding of formyl-methionyl-tRNA to 30S ribosomal subunits, and possesses a ribosome-dependent GTPase activity. IF3 (20,530 Da) functions as a ribosomal subunit anti-association factor and is required for translation of natural mRNAs. The mechanism of action of IF1 (8,119 Da) is less clear. It enhances IF2 binding to 30S subunits and stimulates *in vitro* the binding of fMet-tRNA to 70S ribosomes or the synthesis of β-galactosidase in a DNA-coupled transcription-translation system. Whereas some structural and kinetic information is available concerning how initiation factors bind to the 30S ribosomal subunit, detailed knowledge of how these proteins act at the molecular level is lacking.

The cloning of initiation factor genes has been undertaken in order to better study the function of these proteins and to elucidate how their genes are regulated. Analysis of initiation factor cellular levels indicates that the three factors are present in equimolar concentrations and that their levels increase as a function of growth rate in parallel with

ribosome levels (7). The gene for IF3, *infC*, was cloned by complementation of a mutant defective in IF3 activity *in vitro* (26). *InfC* maps at 38 minutes on the *E. coli* genome and is tightly linked to *thrS*, the gene for threonyl-tRNA synthetase, and *rplT*, the gene for ribosomal protein L20 (27). *InfC* has been sequenced (21), and exhibits the unusual initiator codon, AUU. The gene for IF2, *infB*, was cloned by screening for IF2 overproduction by immunoblotting lysates derived from clones carrying large *E. coli* genomic fragments in a cosmid vector (16). The gene maps at 68.5 minutes and is expressed in a operon that begins with *metY*, a minor form of the initiator $tRNA^{Met}_{f}$ and includes *nusA* coding for a transcriptional termination factor (8, 16, 17). The gene has been sequenced (22) and homologies with EF-Tu and EF-G were found, especially in the domain thought to be involved with GTP binding. Both molecular weight forms of IF2 are encoded by the same gene, apparently by independent translational initiation events on the same mRNA (15).

We are especially interested in cloning the gene for IF1, *infA*, in order to determine whether or not this factor is required for cell growth, a point not proven by *in vitro* assay procedures. We also wish to study the regulation of *infA* in relation to the other factor genes. No mutant forms of IF1 have been reported, but the amino acid sequence of the factor has been determined (18). We report here the cloning of *infA* by screening an *E. coli* genomic library with oligonucleotides based on the amino acid sequence, the sequencing of the structural gene and its mapping to 20 minutes on the *E. coli* genome.

MATERIALS AND METHODS

Bacterial strains, plasmids and phages.

The *E. coli* strains and plasmids used in this work are listed in Table 1. Standard bacteriological procedures were employed as described elsewhere (14). The Charon 30 library was constructed from a partial *Sau*3A digest of *E. coli* genomic DNA (19) and was kindly provided by F. R. Blattner (Genetics Dept. U. Wisconsin). A stock of the bacteriophage $P1^{vir}$ was obtained from R. Merker (Dept. Biological Chemistry, U. California, Davis). Plasmid pTB7 was constructed by ligating the purified 2.8 kb *Bgl*II fragment containing *infA* into the *Bam*HI site of pBR322. Plasmid pTH2 was constructed by ligating the purified 2.0 kb *Hind*III fragment containing *infA* into the *Hind*III site of pBR322.

Recombinant DNA techniques.

Restriction enzymes, polynucleotide kinase, *E. coli* DNA polymerase I, Klenow fragment and T4 DNA ligase were obtained from commercial suppliers and used according to the supplier's specifications. Radioactive nucleotides were obtained from Amersham Corp., Arlington Heights, IL. Transformation of cell lines with plasmid DNA, gel electrophoresis, maxicell assays, restriction fragment isolation from preparative polyacrylamide gels, chromosomal DNA isolations, large scale preparations of plasmid DNA prepared by cesium chloride-ethidium bromide equilibrium centrifugation and rapid miniscreens of plasmid DNA

Table 1. Bacterial strains, phage and plasmids.

Strain	Relevant characteristics	Source or Reference
Y1090	*lon supF lacI*$^+$	Young & Davis, 1983
BW360	Tet^r *polA*	Silver & Wickner, 1983
CSH57	*leu purE trp his argG ilv met thi strA*	Miller, 1972
KL16	Hfr (see Fig. 5)	CGSC
JFS1	KL16 Tet^r *polA*	This work
JFS2	KL16 Amp^r Tet^r *polA*	This work
Hfr3000	Hfr (see Fig. 5)	CGSC
JFS3	Hfr3000 Tet^r *polA*	This work
JFS4	Hfr3000 Amp^r Tet^r *polA*	This work
AB1321	*proA aroA his thi*	CGSC
AT3143	F$^-$ *proC purE pdxC pyrF his ilv met rpsL*	CGSC
JF703	*proC ompF his purE ilv met rpsL*	CGSC
KL282	*serC relA*	CGSC
LCB274	*thr leuB rpsL thi zbj-274*::Tn9	CGSC
RE103	*proA trp his cmlA strA*	CGSC
JM103	(*lac pro*) *hsdR supE*/F' *traD proA proB lacI*^{Q} *lacZ*M15	Messing
CSR603	*recA uvrA*	CGSC
Phage or plasmid		
C1921	Charon 30 isolate containing *infA*	This work
pBR322	Amp^r Tet^r	Bolivar *et al.*, 1977
pTB7	Amp^r *infA*$^+$	This work
pTH2	Amp^r *infA*$^+$	This work
mp10	M13 *lac*	Messing, 1983

CGSC designates Coli Genetic Stock Center, Yale University, New Haven, Connecticut

were performed as described (20). DNA sequencing of the M13mp10 recombinant clones was by the dideoxynucleotide chain termination method (12).

Screening the Charon 30 library.

Two mixed tetradeca-deoxyribonucleotides (Table 2) were chemically synthesized: oligomer 1 was provided by Beckman Instruments, Inc. (Fullerton, CA) and oligomer 2 was provided by Nicole Beauchemin and Jean Paquette (Dept. Biochimie, U. Montreal, Canada). Each probe was 5' end-labeled with polynucleotide kinase and [γ-^{32}P]ATP (sp. act. 5000 Ci/mmol) and purified by electrophoresis in polyacrylamide gels. The Charon 30 phages were mixed with *E. coli* strain Y1090 in appropriate dilutions to give about 200 plaques per plate. Phage from 550 plaques were patched onto a lawn of Y1090 bacteria (138 plaques per plate) and allowed to grow for 5 hours. A nitrocellulose filter (Schleicher and Schuell BA85, 82mm circles) wetted in water was overlaid on the plaques for 10 minutes and then treated to fix the phage DNA on the filter (24). The labeled oligonucleotides (10^6 cpm/ml) were hybridized to the fixed phage DNA at 30°C or 37°C (see Table 2) according to Wallace *et al.* (28). Positive phages were identified by exposure of the filters to Kodak X-Omat film for 1 to 3 days and were purified as described (24). Patching of phage in a grid pattern gave larger plaques and a

Table 2. Oligonucleotides and hybridization temperatures used to select *infA*-containing Charon phage.

Oligonucleotide	DNA Sequence	Hybridization Temperature
Number 1	ATTGAAATGCAAGG C G G A	37°C - 40°C
Number 2	AAAGAAGACAATAT C G T C	30°C

more readily identified autoradiographic signal, thus allowing an easier identification of positive phages.

Chromosomal mapping by Hfr conjugation and P1 cotransduction.

The construction of two ampicillin-resistant Hfr *polA* strains was performed as described (25). The Hfr strains KL16 (transfers chromosomal DNA counterclockwise beginnning at 62 minutes) and Hfr3000 (transfers chromosomal DNA clockwise beginning at 97 minutes) were transduced to Tet^r and *polA* with P1 transducing phages grown on *E. coli* BW360, to yield strains JFS1 and JFS3, respectively. JFS1 was transformed with pTH2 and JFS3 with pTB7, and ampicillin- and tetracycline-resistant colonies were isolated to yield strains JFS2 and JFS4, respectively. Interrupted matings of JFS2 and JFS4 with *E. coli* CSH57 were performed (14) and the following markers were scored: leu^+, trp^+, and Amp^r. Streptomycin was used for counter selection of the donor. The time of first appearance of the markers was used to estimate the location of the Amp^r gene in the chromosome. About 5 x 10^9 Hfr cells were combined with about 1 x 10^9 F^- cells in 5 ml medium and approximately 6 x 10^7 cells (50 µl) were plated in order to obtain enough recombinants to score the time of entry. The conjugation frequency when the Hfr strain is *polA* seems to be reduced by a factor of 10^2 to 10^3 over what is normally observed when the Hfr strain is $polA^+$.

P1 cotransductions were performed as described elsewhere (24). The donor and recipient strains and the markers used in the cotransductions are listed in Table 3. Glucose minimal plates supplemented with the appropriate nutrient requirements were used to score the auxotrophic markers. YT plates containing 20-50 µg/ml ampicillin were used to score for Amp^r. YT plates containing 5, 8 or 25 µg/ml chloramphenicol were used to score the *ompF* (23), *cmlA* and *zbj-274*::Tn*9* markers respectively. The colonies were counted after 2-5 days in order to ensure that all possible transductants were included.

RESULTS

Cloning *infA*.

The strategy for cloning *infA* was to screen a Charon 30 library of genomic *E. coli* DNA with oligonucleotides deduced from the IF1 amino acid sequence. Examination of the

Table 3. Transductional mapping of the infA-Amp^r locus.

Cross No.	Donor	Recipient	Selected Marker	No. of Transductants With Selected Marker	Unselected Marker	No. of Transductants With Unselected Marker	Cotransduction Frequency
1	JFS2	AB1321	aroA$^+$	84	Amp^r	74	88%
2	JFS4	AB1321	aroA$^+$	293	Amp^r	224	76%
3	JFS2	AT3143	pdxC$^+$	45	Amp^r	42	92%
4	JFS4	AT3143	pdxC$^+$	285	Amp^r	242	85%
5	JFS2	KL282	serC$^+$	91	Amp^r	79	87%
6	JFS4	RE103	Amp^r	212	cmlA	51	24%
7	JFS2	LCB274	Amp^r	130	Cm^r(Tn9)	87	67%
8	JFS4	LCB274	Amp^r	246	Cm^r(Tn9)	109	44%
9	JFS4	JF703	Amp^r	256	ompF	30	12%

IF1 sequence reported by Gualerzi and co-workers (18) led to the identification of two adjacent sequences near the N-terminus from which we deduced 14-mers with minimal ambiguities. As shown in Table 2, oligomer 1 has 12-fold ambiguity and 5 to 8 potential G/C base pairs, whereas oligomer 2 has 16-fold ambiguity and from 2 to 6 potential G/C base pairs. Each of the two oligomers was synthesized as a mixed probe which was then radiolabeled and used to screen the Charon 30 library, as described in detail in Materials and Methods. From 550 screened plaques, two hybridized to oligomer 1 only, one hybridized to oligomer 2 only, and one plaque hybridized to both oligomers (results not shown). Phage were purified from the four positive plaques, and phage DNA was isolated. The four DNAs were cleaved with either *Bgl*II or *Hind*III and examined by Southern blotting the cleaved DNA to nitrocellulose and probing with the two ^{32}P-labeled oligomer probes. Both hybridized to a 2.8 kb *Bgl*II fragment and to a 2.0 kb *Hind*III fragment from the DNA of the phage (C1921) positive with both probes (results not shown). The other DNAs gave fragments of different sizes which hybridized to one or the other, but not both probes. Therefore the C1921 phage which hybridized to both probes was selected for further characterization.

The 2.8 kb *Bgl*II fragment and the 2.0 kb *Hind*III fragment which hybridized to both probes were purified from restriction nuclease digested phage C1921 DNA. Each fragment was subcloned into pBR322 as described in Materials and Methods, yielding recombinant plasmids pTB7 and pTH2, respectively. The plasmids were transformed into *E. coli* strain CSR603 and expression of their genes was analyzed by the "maxicell" assay. As shown in Figure 1, both plasmids express a single low molecular weight protein which comigrates with IF1 at 8,000 daltons. This result is consistent with the view that pTB7 and pTH2 carry the complete structural gene for *infA*.

Sequence of *infA*.

Confirmation that *infA* was cloned was obtained by DNA sequencing. First, a restriction nuclease map of the inserts in pTB7 and pTH2 was generated (Figure 2). The two inserts overlap extensively; together they represent about 3 kb of *E. coli* DNA. The 2.8 kb *Bgl*II fragment from C1921 was purified, digested with either *Hae*III or *Hpa*II, and fragments were shot-gun cloned into M13mp10. White M13 plaques were selected, phage grown, and

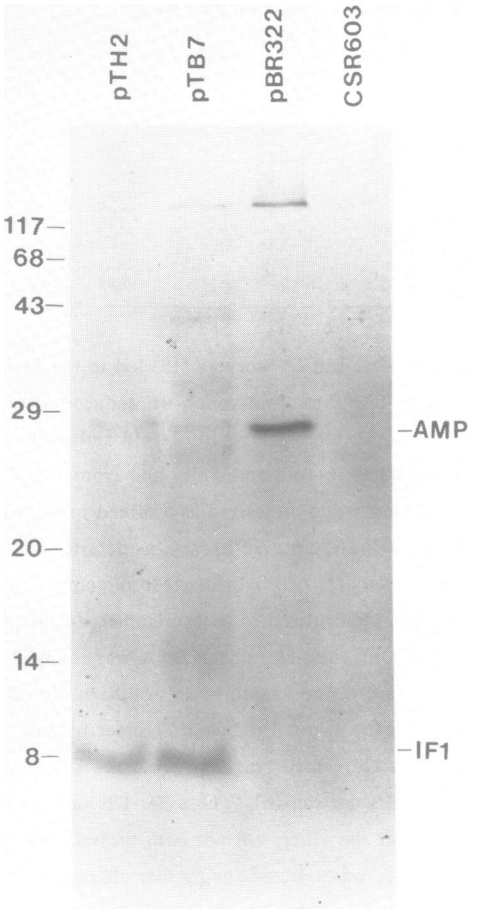

Figure 1. Synthesis of IF1 with pTB7 and pTH2. CSR603 was transformed with pTB7 and pTH2 and Amp^r strains were isolated. The cells were UV irradiated and labeled with [^{35}S]methionine as described for the maxicell assay (20). The cells were lysed in a buffer containing β-mercaptoethanol and sodium dodecyl sulfate, heated to 65^{o}C and electrophoresed in a 15% polyacrylamide gel (10). The dried gel was exposed to Kodak X-Omat film for 2 days; a photograph of the autoradiogram is shown. Migration positions for purified IF1 and the predominate product of the β-lactamase (Amp^r) gene are shown on the right. Molecular weight markers (kDa) are shown on the left.

their DNA inserts were sequenced. One of the *Hae*III inserts in M13 hybridized to the oligomer probes. This 190 nucleotide insert contained an open reading frame coding for 63 amino acids which match perfectly the IF1 sequence from Lys-2 to Gly-64 (18). This fact confirms that pTB7 contains the gene for IF1. The sequence of the entire *infA* coding region, as well as 38 bp upstream from the AUG initiator codon and 71 bp downstream from the

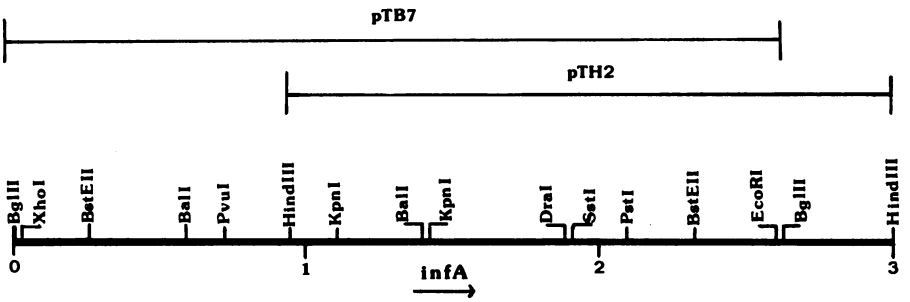

Figure 2. Restriction enzyme map of the DNA region containing *infA*. The restriction enzyme sites were determined by digesting pTB7 or pTH2 with single or double digests of various restriction enzymes. The coding region for *infA* was positioned in the middle of the 3.0 kb region by the *Bal*I and *Kpn*I restriction enzyme sites found in the 5' region of the DNA sequence of *infA* (see Figure 3). The arrow under *infA* indicates the direction of transcription of *infA*. The scale is in kilobase pairs.

UGA termination codon, was determined for both strands and is reported in Figure 3. The extended sequence was determined by sequencing other inserts which overlapped with the *infA* region. The translational signals and putative transcriptional signals are discussed later. Identification of *Bal*I and *Kpn*I restriction sites in the sequence allows us to locate and orient the *infA* sequence (Figure 3) within the 3 kb chromosomal region shown in Figure 2.

InfA maps to 20 minutes on the *E. coli* genome.

Since no mutants in *infA* were available, mapping of the gene on the *E. coli* genome

```
                                                  BALI
CCGGTTCAAATTACGGTAGTGATACCCCAGAGGATTAG ATG GCC AAA GAA GAC AAT ATT GAA ATG
                                        MET ALA LYS GLU ASP ASN ILE GLU MET
     KPNI
CAA GGT ACC GTT CTT GAA ACG TTG CCT AAT ACC ATG TTC CGC GTA GAG TTA GAA AAC
GLN GLY THR VAL LEU GLU THR LEU PRO ASN THR MET PHE ARG VAL GLU LEU GLU ASN

GGT CAC GTG GTT ACT GCA CAC ATC TCC GGT AAA ATG CGC AAA AAC TAC ATC CGC ATC
GLY HIS VAL VAL THR ALA HIS ILE SER GLY LYS MET ARG LYS ASN TYR ILE ARG ILE

CTG ACG GGC GAC AAA GTG ACT GTT GAA CTG ACC CCG TAC GAC CTG AGC AAA GGC CGC
LEU THR GLY ASP LYS VAL THR VAL GLU LEU THR PRO TYR ASP LEU SER LYS GLY ARG

ATT GTC TTC CGT AGT CGC TGA TTGTTTTACCGCCTGATGGGCGAAGAGAAAGAACGAGTAAAAGGTCG
ILE VAL PHE ARG SER ARG STOP

GTTTAACCGGCCTTTTTATTTTAT
```

Figure 3. The DNA sequence surrounding *infA* and the corresponding amino acid sequence. The DNA sequence of *infA* was determined as described in the text; the antisense strand is shown. The derived amino acid sequence corresponds to the previously published sequence of IF1 (18). The *Bal*I and *Kpn*I restriction enzyme sites used to position *infA* on the larger 3.0 kb *Bgl*II-*Hind*III region (Figure 2) are indicated. The Shine/Dalgarno site before *infA* and the putative transcriptional terminator following *infA* are underlined.

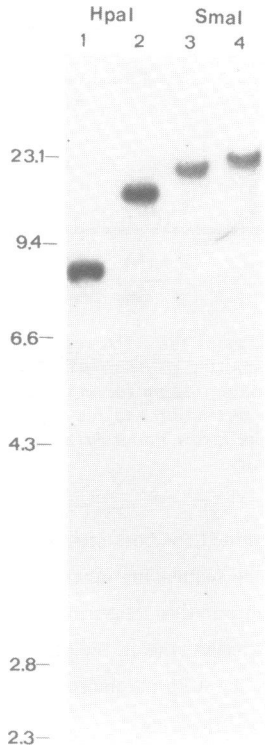

Figure 4. Southern blot analyses of chromosomal DNA. Analyses of strains with (JFS4) and without (JFS3) plasmid integration were performed as described (11). The purified DNA was restricted with *HpaI* and *SmaI* which do not cleave the plasmid containing *infA*. The restricted DNA was electrophoresed in a 0.8% agarose gel and then Southern blotted to nitrocellulose. The transferred DNA was hybridized with the 2.8 kb *BglII* fragment carried in pTB7 nick translated with [α-^{32}P]dATP. A photograph of the autoradiogram is shown, with molecular weight markers on the left (lambda digested with *HindIII* and the 2.8 kb *BglII* fragment from C1921). Odd numbered lanes contain DNA from JFS3; even numbered lanes, from JFS4.

was accomplished by a method involving the insertion and mapping of an easily scored marker closely linked to *infA* (25). Two Hfr strains were transduced to *polA*, then transformed with an *infA*-containing plasmid, and ampicillin-resistant transformants JFS2 and JFS4 were selected (see Materials and Methods). Since the plasmids cannot replicate autogenously in a *polA* background, the presence of the Amp^r gene must be due to integration of the plasmid into the bacterial chromosome by homologous recombination. The only region of homology between the chromosome and the recombinant plasmid is the cloned insert. Therefore integration is expected to occur in or near *infA*, thereby placing the plasmid's Amp^r gene near to *infA* in the chromosome. Confirmation that plasmid insertion occurred in the *infA* region was obtained by analyzing restriction nuclease digests of DNA purified from the transformed

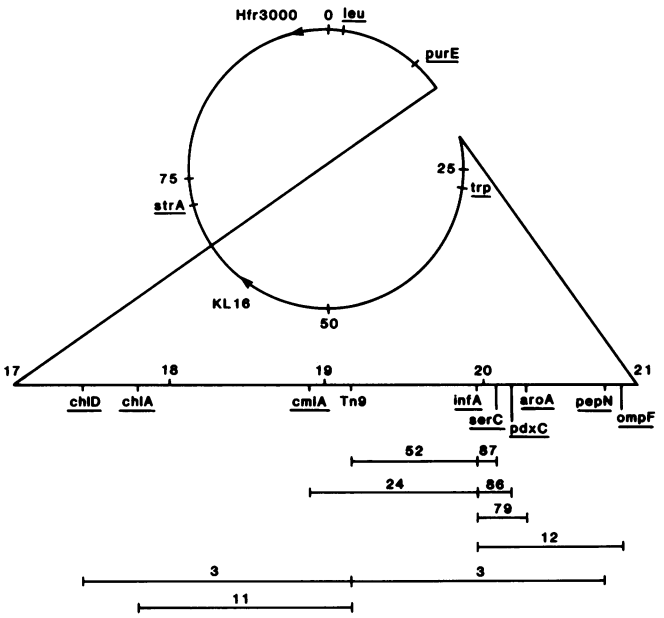

Figure 5. Hfr and P1 mapping of *infA*. The chromosomal map position of *infA* was determined by Hfr interrupted matings and P1 cotransductions. Two Hfr strains were used to map *infA* as described in the text. The figure shows the map positions at which the F factors of Hfr3000 and KL16 are integrated. The figure also shows an expanded region of the *E. coli* map from 17 to 21 minutes, with a number of marker genes used in the Hfr conjugations and in the P1 cotransduction experiments described in the text. The numbers above the horizontal lines at the bottom of the figure represent cotransduction frequencies for the two genes linked by each line (see text and Table 3). The Tn9-*chlD*, Tn9-*chlA*, and Tn9-*pepN* percentages are from data provided by Barbara Bachmann of the *E. coli* Genetics Stock Center (Yale University).

(Amp^r) Hfr strain, JFS4, and the non-transformed (Amp^s) Hfr strain, JFS3 (Figure 4). Digestions with *Hpa*I or *Sma*I, which do not cut pTB7, show a single band for both strains when probed with the labeled 2.8 kb *Bgl*II DNA fragment. The bands from the Amp^r strain are larger than those from the Amp^s strain, and differ in size by about 7 kb, the size of the putative pTB7 insert. A single band with homology in the *infA* region indicates that insertion of pTB7 occurred at or near *infA*. If insertion occurred elsewhere in the chromosome, two bands would be detected, one corresponding to chromosomal *infA*, the other to *infA* within the inserted plasmid.

Hfr strain JFS2 and JFS4 were first used to roughly map the Amp^r gene by interrupted matings with an F^- strain. Both Hfr strains gave times of entry for the Amp^r marker that placed the Amp^r gene near 20 minutes (2). The map position was confirmed and refined by P1 cotransduction experiments as described in Table 3 and Figure 5. P1 phage were grown on strains JFS2 and JFS4, then used to transduce strains carrying markers near 20 minutes.

Transductants were usually screened first for the marker gene, then scored for the presence of Amp^r. Cotransduction frequencies of about 85% were obtained with *serC*, *pdxC*, and *aroA*, which lie clustered at about 20.1 to 20.3 minutes. Cotransduction frequencies of 52% with Tn9 (near 19.2 minutes), 24% with *cmlA* (18.8 minutes) and 12% with *ompF* (20.8 minutes) indicate that the Amp^r gene (and therefore *infA*) maps at about 20 minutes.

DISCUSSION

We report here the cloning of *infA*, the last of the initiaiton factor genes to be identified in *E. coli*. Identification of *infA* is based on three kinds of evidence: 1) putative clones carrying *infA* hybridize to DNA oligomers whose mixed sequences are based on the IF1 amino acid sequence; 2) maxicell assays with cells transformed by the *infA*-carrying plasmids pTB7 and pTH2 express a low molecular weight protein that co-migrates with purified IF1; and 3) the same plasmids contain a sequence of DNA with an open reading frame coding for an amino acid sequence that exactly matches the IF1 sequence (18).

The structural part of the IF1 gene codes for 72 amino acids and begins with an AUG initiator codon. The N-formyl-methionyl residue is presumably removed post-translationally, since the N-terminal amino acid of the native protein (71 amino acids) is alanine (18). The initiator AUG is preceded by a moderately strong Shine/Dalgarno sequence, GAGG, although its placement relative to AUG is somewhat close, being separated by only 5 nucleotides. The termination codon, UAG, next to the initiator AUG is uncommon, but probably is not statistically significant (4). An alanine codon adjacent and downstream from the AUG occurs frequently in *E. coli*, especially in genes that are highly expressed (4). A moderately strong initiation signal for IF1 suggests that the mRNA may be reasonably efficiently translated. The fairly intense IF1 band in maxicell analyses (Figure 1) also is consistent with this view. The translational initiation signal of *infA* is comparable to that for *infB* and most other proteins, but contrasts with the signal for *infC*, which employs the unusual codon AUU and a very weak Shine/Dalgarno signal.

Codon usage in *infA* resembles that of other highly expressed genes such as those coding for ribosomal proteins and elongation factors. The coefficient of expressivity (*e*), the average of the ratios of the use frequency of a codon relative to the frequency of the most used codon (5, 9), is 0.716 for *infA*. For *infB*, $e = 0.835$ and for *infC*, $e = 0.642$. Thus *infA*'s coefficient is comparable to that for the other initiation factors, but is slightly lower than those of the genes for elongation factors (*tufA* = 0.904; *tufB* = 0.892; *fus* = 0.877). The IF1 structural gene terminates with a UGA codon.

We have not yet experimentally addressed the problem of how the *infA* mRNA is transcribed. There appears to be a promoter within the ca. 400 bp region between the *Hind*III site upstream of *infA* and *infA* itself (Figure 2), since IF1 is expressed in maxicells containing pTH2. The 2.0 kb *Hind*III insert of pTH2 interrupts the Tet^r promoter of pBR322 and the

orientation of *infA* is in the same direction as the Tet^r gene. Therefore expression of IF1 is not expected to result from plasmid promoters but rather from transcriptional signals in the 400 bp of *E. coli* DNA before *infA*. Whether or not there is a promoter immediately before the coding region for IF1 must be established by further experimentation. Examination of the DNA sequence downstream of the structural gene for *infA* allows us to identify a putative transcriptional termination signal (see Figure 3). About 35 bp downstream from the *infA* termination codon begins a strong palindromic sequence followed by an A-T rich sequence. This sequence might function as a rho-independent terminator signal (1). Preliminary evidence has been generated from DNA sequencing that *infA* is directly followed by a tRNA gene (unpublished results). Should termination of transcription not occur at the putative terminator site identified in Figure 3, then processing of the *infA* transcript should occur at the tRNA gene. It remains to be shown where the *infA* transcript(s) begins and ends. However, it seems possible that IF1 is translated from a monocistronic mRNA.

The mapping of *infA* to about 20 minutes indicates that none of the initiation factor genes is located close to another. With *infB* at 68.5 minutes and *infC* at 38 minutes, the three genes are widely dispersed on the *E. coli* chromosome. This is perhaps unexpected since all three factors are present in about stoichiometric amounts in cells, and their levels are coordinately regulated (7). Since *infB* and *infC* map close to other genes involved in transcription and translation, we ask if *infA* lies close to any such genes. Except for the tRNA gene just downstream from *infA*, we have not yet been able to identify any other genes near to *infA*. Nor has the orientation of *infA* and its restriction map (Figure 2) yet been determined relative to the chromosomal map. From the genomic map (2), we note that *serS*, which encodes seryl-tRNA synthetase, lies near 20 minutes. The gene for ribosomal protein S1, *rpsA*, maps at about 20.5 minutes, and therefore seems to lie too far away to be linked transcriptionally. Work is in progress to better define the DNA sequences near *infA* and the transcripts that code for IF1.

The cloning of *infA* makes possible investigations of its expression and requirement for cell viability. It also allows us to construct overproducing strains so that large quantities of IF1 can be prepared for biophysical studies. Sequence information and gene fusions should reveal how *infA* is regulated and may shed some light on how the translational components are coordinately regulated in the cell. The cloned gene also provides a substrate for doing site-directed mutagenesis to genetically define the function of the protein. All of these problems currently are being investigated.

ACKNOWLEDGEMENTS

This work was supported by: grant NP-70 from the American Cancer Society (to J.H.); grant INT8312982 from the National Science Foundation (to J.H.); and grant A.T.P.C.N.R.S.-N.S.F. "Etats-Unis 1984" from the Centre National de Recherche Scientifique (to M.G.-M.).

J.F.S. was also supported by a National Institutes of Health Training Grant in Genetics, GM07467. H.S.C. was supported by a National Institutes of Health INRSA fellowship, GM10914.

REFERENCES

1. Adhya, S. and Gottesman, M. (1978) *Ann. Rev. Biochem.* 47, 967-996.
2. Bachmann, B. J. (1983) *Microbiol. Rev.* 47, 180-230.
3. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. and Falkow, S. (1977) *Gene* 2, 95-113.
4. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. and Stormo, G. (1981) *Ann. Rev. Biochem.* 35, 365-403.
5. Grosjean, H. and Fiers, W. (1982) *Gene* 18, 199-209.
6. Hershey, J. W. B. (1980) In Prescott, D. M. and Goldstein, L. (eds.), Cell Biology: A comprehensive treatise, Vol. 4, Gene expression: Translation and the behavior of proteins. Academic Press, NY, pp. 1-68.
7. Howe, J. G. and Hershey, J. W. B. (1983) *J. Biol. Chem.* 258, 1954-1959.
8. Ishii, S., Kuroki, K. and Imamoto, F. (1984) *Proc. Natl. Acad. Sci. USA* 81, 409-413.
9. Itemura, T. and Ozeki, H. (1982) *Cold Spring Harbor Symp. Quant. Biol.* 47, 1087-1097.
10. Laemmli, U. K. (1970) *Nature* 227, 680-685.
11. Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982) Molecular Cloning. A Laboratory Manual. Cold Spring Harbor Laboratory Press, NY.
12. Messing, J. (1983) *Methods Enzymol.* 101, 20-78.
13. Messing. J., Crea, R. and Seeburg, P. H. (1981) *Nucl. Acids Res.* 9, 309-321.
14. Miller, J. H. (1972) Experiments in Molecular Genetics, Cold Spring Harbor Laboratory Press, NY.
15. Plumbridge, J. A., Deville, F., Sacerdot, C., Petersen, H. U., Cenatiempo, Y., Cozzone, A., Grunberg-Manago, M. and Hershey, J. W. B. (1985) *EMBO J.* 4, 223-229.
16. 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.
17. Plumbridge, J. A. and Springer, M. (1983) *J. Mol. Biol.* 167, 227-243.
18. Pon, C. L., Wittmann-Liebold, B. and Gualerzi, C. (1979) *FEBS Lett.* 101, 157-160.
19. Rimm, D. L., Horness, D., Kucera, J. and Blattner, F. R. (1980) *Gene* 12, 301-309.
20. Rodriguez, R. L. and Tait, R. C. (1983) Recombinant DNA Techniques: An Introduction, Benjamin/Cummings Inc., Menlo Park, CA.
21. Sacerdot, C., Fayat, G., Dessen, P., Springer, M., Plumbridge, J. A., Grunberg-Manago, M. and Blanquet, S. (1982) *EMBO J.* 1, 311-315.
22. Sacerdot, C., Dessen, P., Hershey, J. W. B., Plumbridge, J. A. and Grunberg-Manago, M. (1984) *Proc. Natl. Acad. Sci. USA* 81, 7787-7791.
23. Sato, T. and Yura, T. (1979) *J. Bacteriol.* 139, 468-477.
24. Silhavy, T. J., Berman, M. L. and Enquist, L. W. (1984) Experiments With Gene Fusions, Cold Spring Harbor Laboratory Press, NY.
25. Silver, P. and Wickner, W. (1983) *J. Bacteriol.* 154, 569-572.
26. Springer, M., Graffe, M. and Grunberg-Manago, M. (1977) *Mol. Gen. Genet.* 151, 17-26.
27. Springer, M., Graffe, M. and Grunberg-Manago, M. (1979) *Mol. Gen. Genet.* 169, 337-343.
28. Wallace, R. B. Shaffer, J., Murphy, R. F., Bonner, J., Hirose, T. and Itakura, K. (1979) *Nucl. Acids Res.* 6, 3543-3557.
29. Young, R. A. and Davis, R. W. (1983) *Science* 222, 778-782.