Expression of the human interferon-γ cDNA in yeast

Rik Derynck, Arjun Singh and David V.Goeddel

Department of Molecular Biology, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080, USA

Received 8 December 1982; Revised and Accepted 22 February 1983

ABSTRACT

Several plasmids which direct the expression of human interferon-γ (IFN-γ) cDNA in the yeast Saccharomyces cerevisiae under the transcriptional control of the 3-phosphoglycerate kinase (PGK) promoter have been constructed. Up to $25x10^6$ units per liter culture of biologically active IFN-γ are synthesized in yeast transformed with these expression plasmids. The amount of IFN-γ mRNA in yeast transformed with the IFN-γ expression plasmids is much lower than the amount of PGK mRNA present when the PGK gene and promoter are inserted on a similar plasmid. The major and three minor sites of transcription initiation have been localized. The major starting point of the IFN-γ mRNA is at 4 basepairs upstream of the PGK mRNA start. All transcription starts map at an ApG as does PGK mRNA. The yeast IFN-γ mRNA terminates in a specific segment of the 3' untranslated region of the IFN-γ cDNA. The polyadenylation site at about 182-184 nucleotides 3' of the IFN-γ stop codon is preceded by a sequence which is very similar to a sequence proposed to be involved in transcription termination and polyadenylation (49).

INTRODUCTION

The yeast Saccharomyces cerevisiae has received much attention in recent years as a model system for genetic and biochemical studies of lower eukaryotes. These investigations have been greatly stimulated and facilitated by recombinant DNA techniques and by the establishment of a procedure for the transformation of yeast with plasmids (1,2). This methodology has contributed towards the understanding and characterization of the cellular processes and regulatory elements in yeast. Heterologous gene expression in yeast has been achieved not only for some bacterial genes (3), but also for several genes of higher eukaryotes (4-7).

We have previously described the high level expression in yeast of human IFN-α1 (leukocyte interferon D; ref. 6), using the endogenous promoter for the alcohol dehydrogenase I (ADH1) gene (8). This same promoter has recently also been used for the expression of the coding sequence for the Hepatitis B surface antigen (7).

We report here the expression in yeast of the coding sequence for mature

human interferon γ (IFN-γ) or "immune" interferon. We have isolated and characterized the cDNA for this protein and have expressed it in E. coli and monkey cells (9). The cDNA codes for 166 amino acids, of which the amino-terminal segment of 20 amino acids represents the signal peptide, involved in the secretion from induced lymphocytes. The DNA sequence for mature IFN-γ is expressed in yeast using the endogenous promoter for the 3-phosphoglycerate kinase (PGK) gene. The gene for this yeast glycolytic enzyme and its flanking regions have recently been characterized, and the transcription initiation and termination sites have been determined (10). The sequence of the PGK promoter region and an approximate localization of the transcriptional start have also been reported by Dobson et al. (11).

MATERIALS AND METHODS

Enzymatic reaction conditions and transformations

Restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase, exonuclease BAL 31 and bacterial alkaline phosphatase, were purchased from Bethesda Research Laboratories or New England Biolabs and used as recommended by the manufacturers.

The 5' protruding ends of DNA restriction fragments were "filled in" using E. coli DNA polymerase I (Klenow fragment; Bethesda Research Laboratories). These reactions were performed in 6mM Tris-HCl pH 7.5, 6mM $MgCl_2$, 7mM β-mercaptoethanol, 50mM NaCl and 2μM of each deoxynucleotide triphosphate. Incubations were for 30 min. at room temperature using 5 units enzyme per pmole restriction fragment. Single stranded ends of restriction fragments were removed by digestion with S1 nuclease (Miles) using 200 units per pmole for 1 hr at 14°C in 25mM sodium acetate pH 4.5, 0.3M NaCl, 1mM $ZnCl_2$. After electrophoresis in 5 percent polyacrylamide gels, specific restriction fragments were isolated by electroelution.

Plasmid DNA was prepared from 5 ml bacterial cultures using the procedure of Birnboim and Doly (12). Large scale plasmid DNA preparations from 500 ml cultures were performed by a simplified cleared lysate method (13,14).

Bacterial transformations in E. coli 294 (15) were done as described (16). Transformation of the yeast S. cerevisiae strain 20B-12 (α trp1 pep4-3) (17) was according to Hinnen et al. (1). The plasmid stability in yeast grown in selective medium was determined by comparing the plating efficiency of the yeast under selective (YNB+CAA: 0.67 percent Difco yeast nitrogen base, 10 mg/l adenine, 10 mg/l uracil, 0.5 percent Difco casamino

acids, 2 percent glucose) and nonselective (YEPD: 1 percent yeast extract, 2 percent peptone, 2 percent glucose) conditions.

Construction of pPGK-γ4

Plasmid pYE-HBSAg-5 (Hitzeman, R.A., unpublished), which contains the coding sequence for the hepatitis B surface antigen under transcription control of the PGK promoter, was cleaved with EcoRI and SalI and the 6.8 kbp fragment was isolated. This fragment contains the pBR322 origin of replication and the β-lactamase gene (18) as well as the yeast replication origin ars1 (19), preceded by an incomplete TRP1 gene. Plasmid pIFN-γtrp48 (9) was digested with HincII, cleaved with EcoRI and the 938 basepair IFN-γ cDNA fragment was isolated. A third plasmid, pFRM31 (10), was digested with EcoRI, "filled in" with Klenow fragment and cut with SalI. The 1.1 kbp fragment, containing the 3' untranslated end of the PGK gene and the necessary signals for transcription termination in yeast, was isolated. Ligation of these three restriction fragments from three different plasmids resulted in the construction of plasmid pPGK-γo (not shown). This plasmid cannot, however, be used for transformation of yeast because of its non-functional TRP1 selection marker. pPGK-γo was therefore cleaved with PvuI. The fragment containing the ars1 replication origin and the defective TRP1 gene was replaced with the PvuI fragment of pFRM31 (10) which contains the ars1 sequence and the functional, intact TRP1 gene (19). The resulting plasmid pPGK-γ4 is shown in Fig. 1.

Construction of pPGK-γ4Δ3

Plasmid pIFN-γtrp48 (9) was digested with BglI and the mixture of restriction fragments was treated with exonuclease BAL 31. About 200 to 300 basepairs were removed from the ends of the restriction fragments. In this way we removed most of the 3' untranslated end of the IFN-γ cDNA, since the BglI recognition site is situated 257 nucleotides 3' of the stop codon. The ends were "filled in" with E. coli DNA polymerase I (Klenow fragment) and the mixture of DNA fragments was digested with HindIII. The fragments containing the IFN-γ coding sequence were then ligated into pUC9 (20), which had been opened with SalI, "filled in" and cut with HindIII. Plasmid pUC9 contains an array of unique restriction sites. A resulting plasmid, pUCγBR1 (not shown), contains the coding sequence for mature IFN-γ with an XbaI and EcoRI site immediately preceding the start codon, as was the case in the parent plasmid pIFN-γtrp48. The stop codon is followed by about 50 basepairs of the 3' untranslated region and unique BamHI, SmaI and EcoRI sites. The IFN-γ cDNA fragment was then excised from pUCγBR1 with EcoRI and BamHI. It

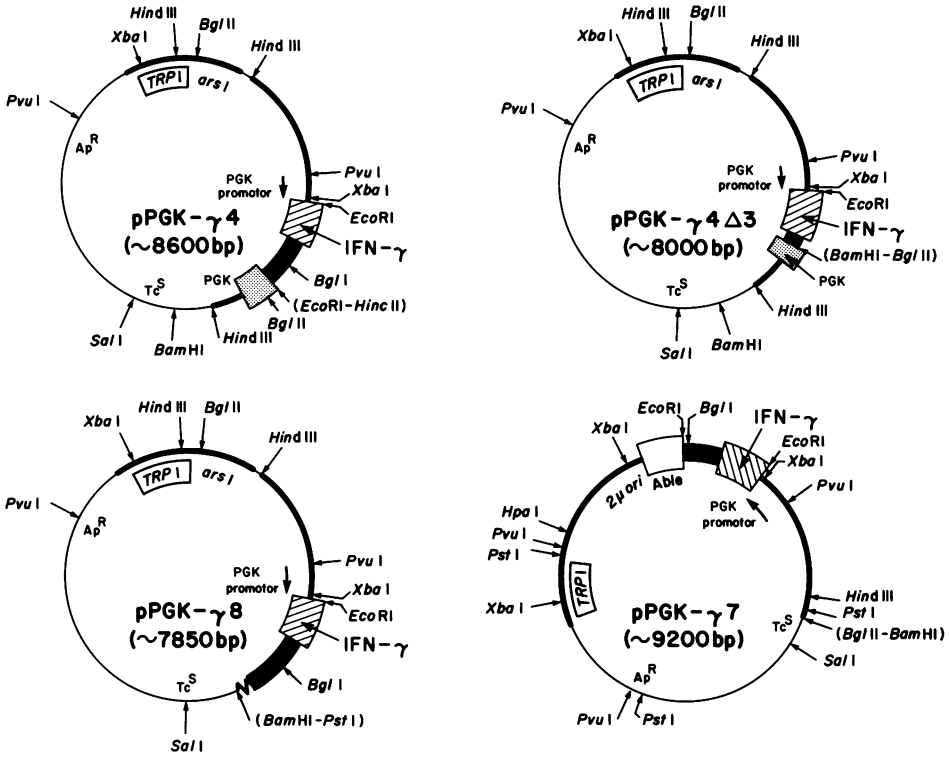

Fig. 1. Schematic representation of the plasmids for expression of IFN-γ in yeast. The thin lines show the pBR322 sequences, while the yeast DNA is represented by thick lines. The boxed regions are the total or partial coding sequences for IFN-γ (hatched), PGK (dotted) and the "Able" protein (open). The 3' untranslated region of IFN-γ is shown as a black boxed region. The arrow indicates the direction of transcription from the PGK promoter. Ap^R, ampicillin-resistance; Tc^S, tetracycline-sensitivity; ori, origin of replication.

was ligated to the large SalI-EcoRI fragment of the above described plasmid pPGK-γ4 containing the replication origins and selection marker genes and the BglII-SalI fragment of pPGK-γ4, which encompasses the 3' untranslated end of the PGK gene. This resulted in the construction of pPGK-γ4Δ3 (Fig. 1).

Construction of pPGK-γ8

Plasmid pPGK-γ4 (Fig. 1) was opened with BamHI and the protruding ends were "filled in". This plasmid was then digested with EcoRI and the large fragment was isolated. The EcoRI-BglI fragment, containing the IFN-γ cDNA sequence, was also prepared from plasmid pPGK-γ4. Plasmid p69 (9) was

digested with PstI and treated with S1-nuclease, followed by digestion with BglI. The resulting small fragment containing the terminal segment of the IFN-γ cDNA 3' untranslated region was isolated. Ligation of these three fragments resulted in the construction of pPGK-γ8 (Fig. 1).

Construction of pPGK-γ7

The major part of the IFN-γ cDNA sequence was isolated from pIFN-γtrp48 (9) as an EcoRI-BglI fragment. Also, the 410 bp BamHI-BglI fragment was prepared from plasmid pSVγSau (21). This fragment contains the 26 bp BglI-Sau3A fragment of the IFN-γ cDNA 3' untranslated region with the Sau3A site converted to an EcoRI site, followed by the 375 bp EcoRI-BamHI fragment of pBR322 (18). Both fragments were ligated and inserted into pBR322, which had been cleaved with EcoRI and BamHI. The plasmid thus obtained, pBRγR1 (not shown), has an EcoRI site just in front of the start codon for mature IFN-γ and at the original Sau3A site in the IFN-γ 3' untranslated region. The IFN-γ cDNA, thus flanked by EcoRI recognition sites, was then inserted into the EcoRI-opened and dephosphorylated expression vector YEp1PT (22) in the orientation of transcription from the PGK promoter. The resulting plasmid pPGK-γ7 is shown in Fig. 1.

Determination of interferon activity in yeast extracts.

Yeast cells were grown to the desired cell density at 30°C in the trp^+ selection medium YNB+CAA (0.67 percent Difco yeast nitrogen base, 10 mg/l adenine, 10 mg/l uracil, 0.5 percent Difco casamino acids, 2 percent glucose). Cells from 5 ml cultures were pelleted, washed with PBS (20mM NaH_2PO_4 pH 7.4, 150mM NaCl), 0.1 percent bovine serum albumin (BSA), resuspended in 0.5 ml PBS, 0.1 percent BSA and disrupted by vigorous shaking for 2 min. at 4°C with 0.3 ml of glass beads (0.45-0.5 mm, Braun Melsungen AG).

In other experiments, the yeast cells were spheroplasted with zymolyase (23). The spheroplasts were washed with PBS and lysed by hypotonic treatment with 0.1X PBS, 0.01 percent BSA for 30 min. on ice. The extracts were assayed for interferon activity as described (9).

Preparation of mRNA

The transformed yeast S. cerevisiae 20B-12 was grown in 100 ml YNB+CAA at 30°C to an A_{660} of 4.0 to 5.0. The cells were spun down, washed with water, resuspended in 5 ml 0.1M NaCl, 50mM Tris-HCl pH 7.5, 10mM EDTA, 0.5 percent SDS, and mechanically disrupted with siliconized, sterile glass beads in the presence of an equal volume phenol-chloroform-isoamyl alcohol (50:50:1). The aqueous phase was repeatedly re-extracted with phenol-

chloroform-isoamyl alcohol. Polyadenylated RNA was isolated by chromatography over oligo-dT cellulose (24). Any contaminating plasmid DNA was removed by precipitation of the RNA in 2M LiCl for 4 hrs. on ice.

Northern analysis of mRNA

RNA samples were electrophoresed in 6 percent formaldehyde-1.2 percent agarose gels (25), transferred to nitrocellulose and hybridized as described (26). Restriction fragments ^{32}P-labelled with the calf thymus DNA primer method (27) were used as hybridization probes. After autoradiographic exposure, the hybridized DNA was removed from the nitrocellulose filter by boiling in water, allowing rehybridization of the same filter.

Determination of the 5' and 3' ends of mRNA

Specifically primed cDNA extension (14) was used to determine the 5' ends of IFN-γ mRNA from S. cerevisiae 20B-12/pPGK-γ7. The 5' labelled synthetic deoxyoligonucleotide dCATCTGAATGACC, complementary to nucleotides 229 to 241 of the IFN-γ cDNA (9), was used as cDNA primer. The discrete cDNA products were separated in a 7M urea-7 percent polyacrylamide gel and, after electroelution, sized by electrophoresis in a 7M urea-8 percent polyacrylamide gel next to a sequencing ladder, obtained by "supercoil" dideoxy-sequencing (14,28,29,30) of pPGK-γ7 using the same 13-mer, and next to an unrelated sequencing ladder.

The polyadenylation site of yeast IFN-γ mRNA was determined by the S1-nuclease mapping procedure (31) using the conditions of Mantei et al. (32). The 630 basepair long TaqI-PstI fragment from pIFN-γtrp48 (9), encompassing the C-terminal part of the IFN-γ coding sequence and the total 3' untranslated region was 3' labelled with DNA-polymerase I (Klenow fragment) in the presence of α^{32}P-dCTP (33). Hybridization was performed with 10μg of polyadenylated RNA from S. cerevisiae 20B-12/pPGK-γ7. S1-nuclease treatment was at 37°C for 30 min. using 50, 100 or 200 units of enzyme. The reaction products were analyzed on a 7M urea-6 percent polyacrylamide gel using ^{32}P-labelled HpaII-fragments of pBR322 (34) as size markers.

RESULTS

The IFN-γ expression plasmids

Four different shuttle vectors, designed to direct the synthesis of mature human IFN-γ in yeast, were constructed (Fig. 1). These plasmids contain the origin of replication of pBR322 (18) and thus can autonomously replicate in E. coli. Transformed bacteria can be selected for ampicillin

resistance due to the presence of the β-lactamase gene, also from pBR322. In this way the newly constructed plasmids can be easily prepared and characterized from E. coli cultures. The yeast TRP1 gene, originally situated on an EcoRI-PstI fragment (19) is incorporated in these plasmids for use as a selection marker in yeast. Its presence on a plasmid renders a transformed trp^- yeast strain, like S. cerevisiae 20B-12, prototrophic for tryptophan. The plasmids pPGK-γ4, -γ4Δ3 and -γ8 replicate in yeast due to the presence of the chromosomal ars1 replication origin (19), originating from chromosome IV. Comparison of the plating efficiency under selective and nonselective conditions showed that these plasmids are retained in intact form in only 25 to 30 percent of the cells of S. cerevisiae 20B-12, grown in selective medium. A similar stability was previously observed with the ars1 containing plasmids for expression of IFN-α1 in yeast (6). Plasmid pPGK-γ7, in contrast, carries the origin of replication of the 2μ circle (35), which results in a much higher plasmid stability of 90-95 percent.

The expression of the mature IFN-γ sequence is under the transcriptional control of the PGK promoter. The 10 basepair segment immediately preceding the translational start of the natural PGK gene had previously been replaced by a synthetic combined XbaI-EcoRI recognition site (5'dTCTAGAATTC 3') (36). The introduction of these sites allows the versatile insertion of a coding sequence of a heterologous gene. For expression in E. coli an EcoRI recognition site, followed by the ATG initiation codon, was present in front of the coding sequence for mature IFN-γ (9). Therefore, the IFN-γ sequence could be inserted via the EcoRI site under the control of the PGK promoter.

The DNA sequence for mature IFN-γ, which codes for 147 amino acids including the initiating methionine, is followed by the complete 3' untranslated region of the IFN-γ cDNA in the plasmid pPGK-γ8. Plasmids pPGK-γ4, -γ7 and -γ4Δ3 retain a segment of the 3' untranslated region, proximal to the IFN-γ stop codon, of 495, 287 and 50 bp respectively (Figs. 1 and 5). The introduced IFN-γ gene should be followed by the sequences required for transcription termination and polyadenylation of yeast mRNA to allow transcription into a biologically functional mRNA. For this purpose, a restriction fragment containing the C-terminal part of the PGK coding sequence and its downstream region has been introduced behind the IFN-γ sequence in plasmids pPGK-γ4 and -γ4Δ3. A functionally similar restriction fragment, containing the transcription termination signal of the "Able" gene of the 2μ circle (35) is incorporated in pPGK-γ7. Plasmid pPGK-γ8 does not contain a yeast transcription terminator fragment.

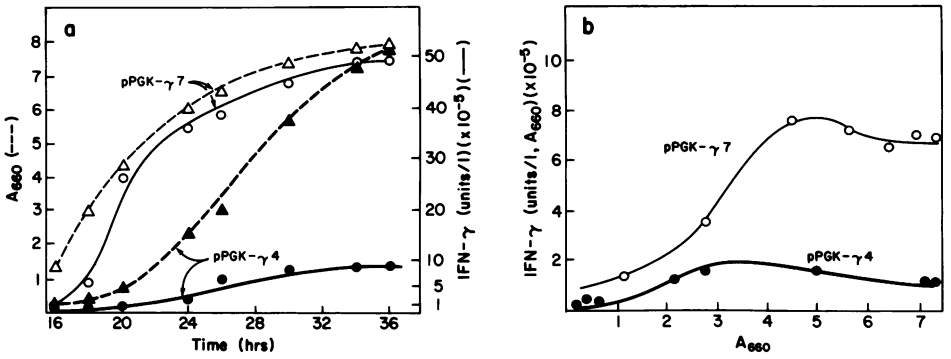

Fig. 2. Time course of IFN-γ production in S. cerevisiae 20B-12 containing either pPGK-γ4 (thick lines) or pPGK-γ7 (thin lines). The cultures were inoculated at an A_{660} of 0.01 and grown in YNB+CAA. Duplicate 10 ml samples were removed every 2 to 4 hrs, starting at 17 hrs after inoculation and extracts were made using glass beads. Each extract was assayed in duplicate for interferon activity, both at a tenfold dilution and undiluted, and the mean interferon titer was calculated. (a) Kinetics of interferon production and growth curve. The full lines represent the interferon titers, while the dashed lines show the cell density. The triangles show the mean A_{660} values at individual time points for pPGK-γ4 (Δ) and pPGK-γ7 (Δ), while the circles indicate the corresponding mean interferon titers: pPGK-γ4 (O) and pPGK-γ7 (O). (b) Kinetics of interferon production per A_{660} of yeast cells and per liter versus increasing cell density. The calculated mean values are shown for pPGK-γ4 (O) and pPGK-γ7 (O).

Synthesis of IFN-γ in yeast

Yeast transformed with the various expression plasmids were grown to an A_{660} of 4.0 in the medium YNB+CAA. After mechanical disruption of the cells with glass beads $8x10^5$ units of antiviral activity per liter of culture were found for pPGK-γ4 and -γ4Δ3, while about $3x10^5$ units/l were synthesized in the case of pPGK-γ8 and $5x10^6$ units/l in the case of pPGK-γ7. The kinetics of IFN-γ production in yeast is shown in Fig. 2. Only a minimal expression level is detected at low cell density while the IFN-γ activity per cell peaks at an A_{660} of 4 to 5.

The level of IFN-γ activity is clearly dependent upon the extraction procedure. The much milder method of lysing spheroplasts (23) by osmotic shock, resulted in the release from the cells of a 3 to 5 fold higher interferon level than after mechanical disruption of the yeast with glass beads. However, since this milder method is more difficult and time-consuming we used the glass bead extraction method for routine analysis and for the kinetics determinations. Electrophoretical analysis of

Plasmid	IFN-γ titer per liter culture (A_{660} = 4.0) Disruption with glass beads	Spheroplasting and osmotic shock
pPGK-γ4	$8x10^5$ units/l	$2.5-3.5x10^6$ units/l
pPGK-γ4Δ3	$8x10^5$ units/l	N.D.
pPGK-γ8	$3x10^5$ units/l	N.D.
pPGK-γ7	$5x10^6$ units/l	$16-25x10^6$ units/l

Table I: Determination of IFN-γ levels in the yeast S. cerevisiae 20B-12 transformed with the indicated plasmids. The extraction methods have been described in the "Materials and Methods" section. In the case of disruption with glass beads, the average titer from 10 individual experiments is given. The values for the spheroplasting, followed by osmotic shock, are based on four different experiments. N.D. = not determined.

^{35}S-methionine labelled yeast proteins in SDS-polyacrylamide gels, however, did not allow unambiguous detection of an IFN-γ specific protein band (not shown).

"Northern" hybridization analysis of the IFN-γ mRNA

Polyadenylated RNA was prepared from the yeast cultures at cell densities (A_{660} = 4) giving maximal IFN-γ synthesis. The mRNA was analyzed by electrophoresis in formaldehyde agarose gels (25), followed by "Northern" hybridization (26) with a probe for either the IFN-γ cDNA or the yeast transcription termination region. An EcoRI-PstI fragment of pIFN-γtrp48 (9), which encompasses the mature IFN-γ coding region and the complete 3' untranslated region, was used as IFN-γ cDNA probe. The EcoRI-XbaI fragment of YEp1PT (22), the parent plasmid of pPGK-γ7 (Fig. 1), containing the end of the coding sequences and the transcription termination region of the "Able" gene (35), was used as probe for the "Able" terminator. Similarly, the BglII-HindIII fragment from pPGK-γ4 (Fig. 1) containing the downstream sequences of the PGK gene (10), was used as probe for the PGK terminator.

Figure 3a shows the results of Northern hybridization of IFN-γ mRNAs from the transformed yeast strains with the radioactive IFN-γ cDNA probe. The size of this mRNA is 720 to 750 nucleotides for plasmids pPGK-γ4, -γ8 and -γ7 and 850 bases in the case of pPGK-γ4Δ3. In the former cases, no hybridization of this mRNA species with the PGK terminator probe (Fig. 3b), nor with the "Able" terminator probe (not shown) was observed. In the case of plasmid pPGK-γ4Δ3, the IFN-γ mRNA hybridized with the PGK terminator (Fig. 3b). The hybridization of this same mRNA with both the IFN-γ cDNA probe and the PGK terminator probe was confirmed by using the same filter.

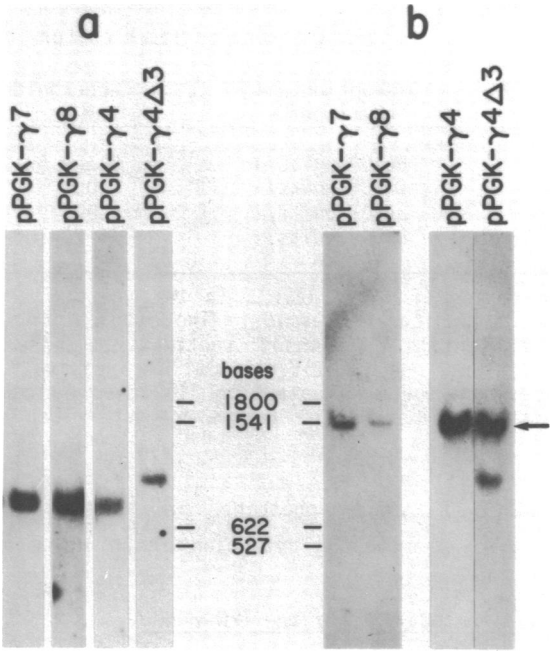

Fig. 3. Northern analysis of polyadenylated RNA from yeast 20B-12 containing the indicated plasmids after electrophoresis on formaldehyde - 1.2 percent agarose gels. The intensities of the visualized mRNA bands do not reflect differences in abundance of the mRNAs due to variations in the amounts of mRNA per lane and in autoradiographic exposure time of the lanes. (a) Hybridization with the IFN-γ cDNA probe. (b) Hybridization with the PGK-terminator probe. The arrow indicates the endogenous PGK mRNA, transcribed from the chromosomal gene. The indicated size markers are yeast 18S ribosomal RNA (1800 bases, ref. 50), E. coli 16S ribosomal RNA (1540 bases, ref. 53) and the denatured 620 and 520 bp HpaII fragments of pBR322.

The results of hybridizations with the probes for the IFN-γ cDNA and the terminator regions, and the size determinations of the IFN-γ mRNAs suggest that the transcription terminates in the 3' untranslated region of the IFN-γ cDNA in plasmids pPGK-γ4, -γ7 and -γ8. The yeast transcription terminator, which is inserted downstream of the IFN-γ cDNA in the case of pPGK-γ4, -γ7 and -γ8, is then functionally not needed. However, in the case of pPGK-γ4Δ3, most of the IFN-γ cDNA 3' untranslated region is removed and the transcription terminates in the PGK terminator region. The size of this IFN-γ mRNA is compatible with transcriptional termination at the same position as for the intact PGK gene, i.e., at about 75 to 82 nucleotides behind the PGK stop codon (10).

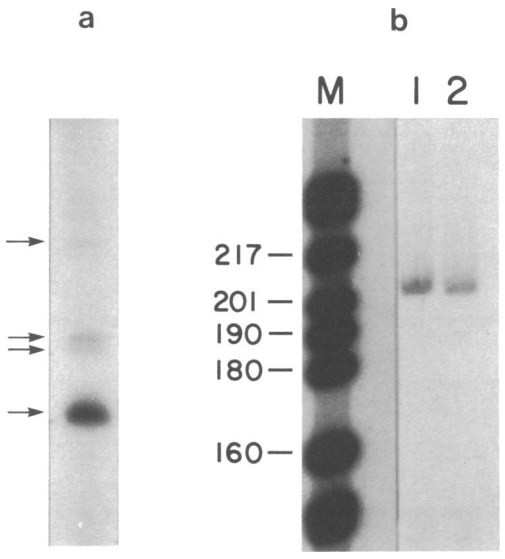

Fig. 4. Determination of the 5' and 3' termini of the yeast IFN-γ mRNA. (a) The transcription initiation points were determined by specifically primed cDNA synthesis. The discrete products, shown by arrows, were separated in a 7M urea-7 percent polyacrylamide gel. (b) S1-nuclease mapping of the polyadenylation site of the yeast IFN-γ mRNA. The RNA-DNA hybrids were treated for 30 min. with 50 units (lane 1) or 100 units (lane 2) S1-nuclease and then sized on 7M urea-6 percent polyacrylamide gel. The lane M shows the ^{32}P-labelled HpaII fragments of pBR322 with their sizes indicated on the left.

It has previously been shown that insertion of the PGK gene and its promoter on a plasmid results in a much higher level of PGK mRNA than with a single chromosomal PGK gene, probably reflecting the copy number of the plasmid (10 and unpublished data). In the case of the IFN-γ mRNA expression plasmids, however, the level of mRNA is much lower than expected. This is clearly illustrated for plasmid pPGK-γ4Δ3 (Fig. 3b). In this case, both the IFN-γ mRNA and the PGK mRNA hybridized with the PGK terminator probe. It is striking that the level of IFN-γ mRNA is lower than the PGK mRNA transcribed from the single chromosomal gene. The amount of IFN-γ mRNA is higher in the case of pPGK-γ7 than for pPGK-γ4, -γ4Δ3 and -γ8, probably due to the higher stability of this 2μ origin based plasmid (data not shown). A higher copy number of the latter plasmid might also contribute to this difference.

Mapping of the 5' and 3' end of the yeast IFN-γ mRNA

Yeast transformed with plasmids pPGK-γ4, -γ8 and -γ7 synthesize an IFN-γ mRNA of the same size (Fig. 3a). The transcriptional initiation and

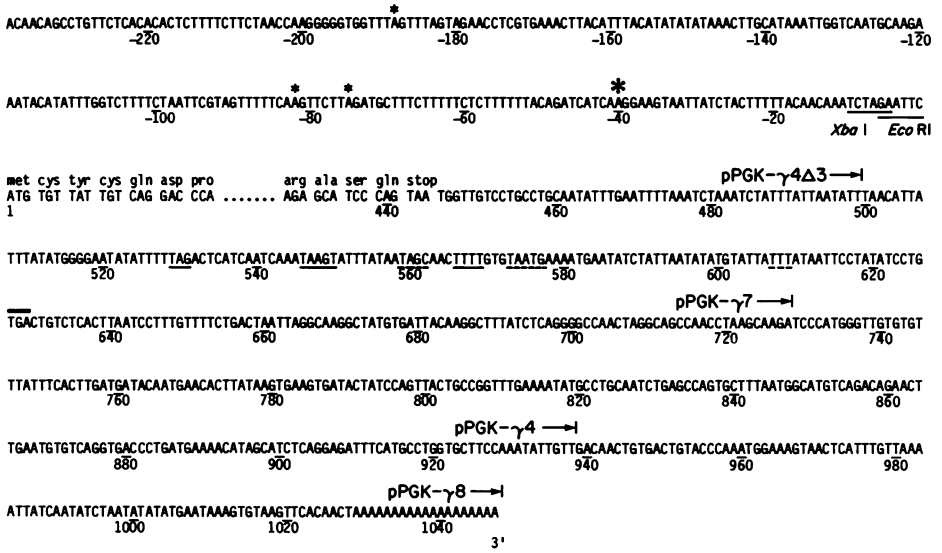

Fig. 5. Nucleotide sequence of the PGK promoter segment ligated through a synthetic XbaI-EcoRI site to the coding sequence for mature IFN-γ, which is in turn followed by the IFN-γ 3' untranslated region. The length of the segment of this 3' untranslated sequence, contained in the different plasmids is indicated. The position indicated for pPGK-γ4Δ3 is approximate (± 5 bp). The asterisks in the 5' flanking sequence show the initiation points of the three minor (small asterisks) and the major (large asterisk) transcripts. The underlined residues (full or dashed lines) in the 3' untranslated end indicate the sequences homologous to the tripartite consensus sequence for transcription termination (49). The polyadenylation site is indicated by the overlined nucleotides.

polyadenylation sites were therefore determined only in the case of plasmid pPGK-γ7, which yielded the most IFN-γ mRNA.

A cDNA extension reaction (14) was performed on the mRNA from S. cerevisiae 20B-12/pPGK-γ7 using an IFN-γ specific primer complementary to the mRNA and corresponding to nucleotides 229 to 241 in the original IFN-γ cDNA (9). A major product, accounting for 98 percent of the transcripts, and three minor transcripts were observed after gel electrophoresis (Fig. 4a). These were mapped on a denaturing polyacrylamide gel next to an unrelated sequencing ladder and a sequencing ladder obtained by the dideoxy method (14,29,30) on plasmid pPGK-γ7 using the same primer (not shown). The major transcript starts at the ApG at position -40, probably at the A, while the minor transcripts start with the ApG at -75, -82 and -187 (Fig. 5).

The 3' polyadenylation site of the IFN-γ mRNA from S. cerevisiae 20B-12/pPGK-γ7 was determined using the S1-nuclease procedure (31), after hybridization of the mRNA with a TaqI-PstI restriction fragment from pIFN-γtrp48 (9), covering the complete 3' untranslated region. Analysis of the reaction products on a denaturing gel (Fig. 4b) allowed the mapping of the polyadenylation site at 182-184 nucleotides 3' of the IFN-γ stop codon (Fig. 5).

DISCUSSION

The introduction of the coding sequence for mature human IFN-γ under the transcriptional control of the PGK promoter results in the synthesis of biologically active IFN-γ molecules in yeast. The levels obtained with zymolyase treatment, followed by osmotic shock, are 3 to 5 times higher than after disruption of the cells with glass beads and reach up to 25×10^6 units/l at an A_{660} of 4.0 in the case of S. cerevisiae 20B-12/pPGK-γ7. It seems likely that the use of other extraction procedures could result in the detection of still higher interferon levels. Alternatively, the use of other yeast strains as hosts for the plasmids described here might allow higher expression levels. For example, S. cerevisiae GM3C-2 (37) has consistently resulted in IFN-γ titers one and a half times higher than strain 20B-12 (22).

The synthesis of IFN-γ is dependent on the cell density as shown in Fig. 2. The difference in growth curve of S. cerevisiae 20B-12/pPGK-γ4 and -γ7 probably reflects the difference in stability of the plasmids. A 25 to 30 percent stability is observed for pPGK-γ4 and other plasmids carrying the ars1 replication origin, while the use of the 2μ origin, as in plasmid pPGK-γ7, results in the presence of the intact plasmid in about 95 percent of the cells. The growth curve of strain 20B-12/pPGK-γ7 is indeed very similar to the growth of untransformed S. cerevisiae 20B-12 in trp- supplemented medium (data not shown). The higher IFN-γ expression level and the increased relative amount of IFN-γ mRNA observed with plasmid pPGK-γ7 as compared with pPGK-γ4 or -γ4Δ3 might, to a large extent, be due to this higher plasmid stability. In addition, the presence of the 2μ origin in pPGK-γ7 might result in a higher copy number when compared with the other ars1-containing plasmids. However, it is not clear what causes the difference in expression between pPGK-γ4 and -γ8, since both have the same ars1 origin. It is possible that the presence and position of some sequences on the plasmids exerts an influence on the expression level. This

might also be a contributing factor to the difference in the IFN-γ levels between pPGK-γ4 and -γ7.

The observed expression level of the IFN-γ cDNA is markedly lower than that of the yeast PGK gene and also lower than for several other human cDNAs, e.g. IFN-α1 (10) using the same promoter. There are several possible explanations for this difference. In addition to the instability of IFN-γ during the extraction procedure, IFN-γ might be very labile in the cells. This latter possibility is hard to investigate since no IFN-γ band could be clearly visualized after SDS-polyacrylamide gel electrophoresis. The translation of IFN-γ mRNA might be inefficient due to strong differences in codon usage between human and yeast genes. Bennetzen and Hall (38) observed an extreme codon bias in highly expressed yeast genes like alcoholdehydrogenase I (ADH1, ref. 8), glyceraldehyde-3-phosphate dehydrogenase (39) and enolase (40). We have found a similar codon preference in the PGK gene (10). A strong correlation exists between the relative abundance of tRNA species in yeast and the codon preference in most yeast genes (41). The codon usage of human IFN-γ is in this respect very unfavorable. There are in the human IFN-γ sequence several codons that are not present in the yeast genes for ADH1, glyceraldehyde-3-phosphate dehydrogenase, enolase and PGK, e.g. certain codons for serine, leucine, isoleucine and threonine.

Furthermore, initiation of translation of the IFN-γ mRNA might be less efficient than for other mRNAs. Most yeast genes and all of the highly expressed yeast genes for glycolytic enzymes characterized so far (8,10,39,40) have pyrimidines in the second and third position downstream of the start codon, in striking contrast with the higher eukaryotes (42). In the case of the IFN-γ gene, there is a G residue at position 2. However, it is not known how differences at this position affect the translation efficiency *in vivo*, although the importance in a cell-free wheat germ system has been demonstrated (42). Initiation of translation is certainly influenced by changes in the sequence immediately preceding the initiation codon. Based on *in vitro* studies, using synthetic oligonucleotides, Kozak (42) has shown the importance of the sequence just in front of the start codon for efficient ribosomal binding. We have also determined the *in vivo* influence of the nucleotide at position -3 before the start codon on the expression of a heterologous gene in yeast (R.D. and D.V.G., unpublished results). Any or all of these factors could be responsible in part for the low expression level of the IFN-γ cDNA, when compared to the PGK gene with its own promoter. However, it seems unlikely that the difference between

the expression of IFN-α1 cDNA, using either the ADH 1 (6) or the PGK (10,22) promoter, and IFN-γ cDNA is due solely to these factors. As for the IFN-γ cDNA, the codon choice for IFN-α1 is also unfavorable when compared with the codon bias in highly expressed yeast genes (38). Also, the sequence just in front of the initiation codon and the second and third nucleotide behind the ATG is the same for both IFN-α1 and IFN-γ.

The difference in expression level between the PGK gene and the IFN-γ cDNA can be attributed in part to a difference in specific mRNA abundance. While the insertion of the PGK gene and promoter on a plasmid results in a much higher level of PGK mRNA than in the case of a single chromosomal gene (10), we find a very low level of IFN-γ mRNA as illustrated for plasmid -γ4Δ3 (Fig. 3b). Even considering the 25 to 30 percent stability of this plasmid, the amount of IFN-γ mRNA in a plasmid harboring yeast cell is still no higher than the PGK mRNA derived from the single chromosomal PGK gene. The higher IFN-γ mRNA level observed in the yeast, transformed with plasmid pPGK-γ7, is probably mainly due to the higher plasmid stability or to a difference in plasmid copy number. The low level of mRNA, transcribed from a heterologous cDNA, is not restricted to the case of IFN-γ cDNA, since a similar phenomenon has been observed with the expression of the IFN-α1 cDNA using the PGK promoter (C. Chen and R. Hitzeman, unpublished results). It is not known if this low mRNA level is due to instability of the mRNA or to a decreased efficiency of transcription and this requires further investigation.

The trancripts of the IFN-γ cDNA have been mapped and all start at an ApG (Fig. 4). The presence of multiple initiation sites for transcription has been found for several yeast genes such as ADH1 (8), CYC1 (37) and TRP5 (43). This striking feature of transcription initiation at an ApG has also been observed for the ADH1 (8) and the PGK (10) genes, but does not seem to be a general phenomenon in yeast. We previously localized the initiation of transcription of the PGK gene at 36 nucleotides upstream of the start codon (10). It is remarkable that the IFN-γ mRNA initiates at 4 nucleotides upstream from the ApG initiation point for the PGK mRNA. Furthermore, minor transcription starts, as in the case of the IFN-γ cDNA, have not been detected for the PGK gene. The start of PGK mRNA and the major transcriptional start for the IFN-γ cDNA are both localized within the region determined for the PGK mRNA start by Dobson et al. (11). It should be stressed that what is considered here as initiation of transcription can also be the result of RNA processing, although no indications for this have

been reported.

For the insertion of a heterologous gene under the control of the PGK promoter, the sequence ...TATAAAAACA... immediately preceding the PGK initiation codon (10) has been replaced by an XbaI-EcoRI recognition site ...TCTAGAATTC... of the same length (Fig. 5, ref. 36). It is possible that the differences in transcriptional initiation between the IFN-γ mRNA and the PGK mRNA are due to this replacement, since no other sequences upstream of the PGK gene have been altered. This would imply that this sequence plays a role in the localization of the initiation site for transcription by RNA polymerase II. It is not clear whether the resemblance of this TATAAAAACA sequence with the Goldberg-Hogness box or TATAAA consensus sequence (44) in higher eukaryotes or eukaryotic viruses is meaningful. The latter consensus sequence is generally found 25 to 30 nucleotides upstream from the mRNA start and seems to be important for proper transcription initiation (for review, ref. 45 and 46). Although a TATAAA-like sequence is present in front of transcribed sequences in yeast, the distance from the transcriptional start is longer and more variable than in higher eukaryotes (8,37,43,47). A TATAAA sequence is also found 145 bp upstream from the PGK initiation codon (10, Fig. 5). No unambiguous evidence for the functional importance of this sequence in yeast has as yet been reported. On the other hand, it is also possible that the insertion of the heterologous gene itself causes the difference in initiation of transcription. The above-mentioned fact that the level of IFN-γ mRNA transcribed from the PGK promoter is much lower than the PGK mRNA, derived from the gene inserted in a similar plasmid (10 and unpublished results), could be related to this difference in transcriptional initiation with the PGK mRNA.

"Northern" analyses of the IFN-γ mRNA from plasmids pPGK-γ4, -γ7 and -γ8 (Fig. 3a) and also from other similar plasmids for IFN-γ expression (unpublished results) reveal a mRNA of 720 to 750 nucleotides long. Considering the initiation of the major transcript at 40 nucleotides upstream of the start codon, this result indicates that the termination of transcription takes place in the 287 bp long segment of the IFN-γ 3' untranslated region, which is retained in these plasmids. The polyadenylation site of this mRNA has been localized at about 182-184 nucleotides behind the stop codon (Fig. 5). However, when in the case of pPGK-γ4Δ3 only the 50 bp of the 3' untranslated region proximal to the stop codon are present, transcription continues into the PGK terminator sequence. This demonstrates that a segment of the 3' untranslated region of

the IFN-γ cDNA can act as an efficient transcription terminator in yeast. In contrast, analysis of the IFN-α1 mRNAs, transcribed from the PGK promoter, has shown that the absence of a specific terminator sequence results in a heterogeneous population of long transcripts (C. Chen and R. Hitzeman, unpublished results).

The size of the IFN-γ mRNAs, as determined by "Northern" analysis, is compatible with the mapping data of the 5' and 3' ends. On the basis of these results a length of about 664 nucleotides is predicted for the major IFN-γ mRNA in the case of pPGK-γ4, -γ7 and -γ8, in agreement with the observed mRNA size of 720 to 750 nucleotides, since polyadenylation adds about 50 to 100 bases to the transcript (48). In the case of pPGK-γ4Δ3 a transcript length of 735-742 bases (without polyadenylation) is expected since the PGK transcription terminates at about 75-82 nucleotides behind the PGK stop codon (10). The observed length of 850 nucleotides for the polyadenylated IFN-γ mRNA transcribed from this plasmid is in agreement with this calculated value.

Based on transcription termination studies of the CYC1 gene and on observed sequence homologies of 3' flanking gene sequences, Zaret and Sherman (49) have recently proposed a consensus sequence, involved in transcription termination and polyadenylation for most yeast genes. This sequence ...TAG...TA(T)GT...(AT rich)...TTT... or a variant of it precedes the polyadenylation site of most yeast mRNAs at variable distance. Similar sequences are found at positions 85 to 126 and 116 to 166 downstream from the IFN-γ stop codon (Fig. 5). The fact that these sequences are located in the segment of the 3' untranslated region of the IFN-γ cDNA which causes termination of transcription and their close proximity to the polyadenylation site is striking and suggests a role in the transcription termination and subsequent polyadenylation. This sequence could be responsible for the homogeneity of the IFN-γ mRNA population in S. cerevisiae 20B-12/pPGK-γ4, -γ7 and -γ8, in contrast to the IFN-α1 case and its removal would result in the continued transcription observed in pPGK-γ4Δ3.

In contrast to the well documented studies on the regulation of transcription in higher eukaryotes (for review see refs. 45 and 46), limited data are available on the identification of specific control elements in yeast and have as yet not revealed a clear similarity with the regulation of transcription in mammalian cells (37,50,51). The expression of the coding sequence of human mature IFN-γ can serve as a model system to investigate transcription and translation in yeast.

Acknowledgments

We thank Dr. R. Hitzeman for numerous discussions and advice. We are grateful to J. Lugovoy and F. Hagie for advice and help during the early stages of this work, to S. Norton for technical support and to Dr. H. Oppermann and L. Yelverton for advice. We acknowledge L. May for performing interferon assays and M. Vasser for oligonucleotide synthesis. We thank J. Arch for excellent editing assistance.

Genentech Contribution No. 156

REFERENCES

1. Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Proc. Natl. Acad. Sci. USA 75, 1929-1933.
2. Beggs, J.D. (1978) Nature 275, 104-109.
3. Hollenberg, C.P. (1982) Current Topics in Microbiology and Immunology 96, 119-142.
4. Henikoff, S., Tatchell, K., Hall, B.D. and Nasmyth, K.A. (1981) Nature 289, 33-37.
5. Mercereau-Puijalon, O., Lacroute, F. and Kourilsky, P. (1980) Gene 11, 163-167.
6. Hitzeman, R.A., Hagie, F.E., Levine, H.L., Goeddel, D.V., Ammerer, G. and Hall, B.D. (1981) Nature 293, 717-722.
7. Valenzuela, P., Medina, A., Rutter, W.J., Ammerer, G. and Hall, B.D. (1982) Nature 298, 347-350.
8. Bennetzen, J.L. and Hall, B.D. (1982) J. Biol. Chem. 257, 3018-3025.
9. Gray, P.W., Leung, D.W., Pennica, D., Yelverton, E., Najarian, R., Simonsen, C.C., Derynck, R., Sherwood, P.J., Wallace, D.M., Berger, S.L., Levinson, A.D. and Goeddel, D.V. (1982) Nature 295, 503-509.
10. Hitzeman, R.A., Hagie, F.E., Hayflick, J., Chen, C.Y., Seeburg, P.H. and Derynck, R. (1982) Nucl. Acids Res. 10, 7791-7808.
11. Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J., Kingsman, S.M., Perkins, R.E., Conroy, S.C., Dunbar, B. and Fothergill, L.A. (1982) Nucl. Acids Res. 10, 2625-2637.
12. Birnboim, H.C. and Doly, J. (1979) Nucl. Acids Res. 7, 1513-1523.
13. Clewell, D.B. (1972) J. Bacteriol. 110, 667-676.
14. Derynck, R., Leung, D.W., Gray, P.W. and Goeddel, D.V. (1982) Nucl. Acids Res. 10, 3605-3615.
15. Backman, K., Ptashne, M. and Gilbert, W. (1976) Proc. Natl. Acad. Sci. USA 73, 4174-4178.
16. Hershfield, V., Boyer, H.W., Yanofsky, C., Lovett, M.A. and Helinski, D.R. (1974) Proc. Natl. Acad. Sci. USA 71, 3455-3459.
17. Jones, E. (1976) Genetics 85, 23-33.
18. 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.
19. Tschumper, G. and Carbon, J. (1980) Gene 10, 157-166.
20. Vieira, J. and Messing, J. (1982) Gene 19, 259-268.
21. Simonsen, C.C., Shepard, H.M., Gray, P.W., Leung, D.W., Pennica, D., Yelverton, E., Derynck, R., Sherwood, P.J., Levinson, A.D. and Goeddel, D.V. (1982) in "The Interferons" (eds. T.C. Merigan, R.M. Friedman and C.F. Fox), 1-14.
22. Hitzeman, R.A., Leung, D.W., Perry, L.J., Kohr, W.J., Levine, H.L. and Goeddel, D.V. (1983) Science 219, 620-625.

23. Kuo, S.-C. and Yamamoto, S. (1975) Meth. Cell. Biol. 11, 169-183.
24. Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. USA 69, 1408-1412.
25. Dobner, P.R., Kawasaki, E.S., Yu, L.Y. and Bancroft, F.C. (1981) Proc. Natl. Acad. Sci. USA 78, 2230-2234.
26. Thomas, P.S. (1980) Proc. Natl. Acad. Sci. USA 77, 5201-5205.
27. Taylor, J.M., Illmensee, R. and Summers, S. (1976) Biochim. Biophys. Acta 442, 324-330.
28. Smith, M., Leung, D.W., Gillam, S., Astell, C.R., Montgomery, D.L. and Hall, B.D. (1979) Cell 6, 753-761.
29. Smith, A.J.H. (1980) Meth. Enzym. 65, 560-580.
30. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucl. Acids Res. 9, 309-321.
31. Berk, A.J. and Sharp, P.A. (1978) Proc. Natl. Acad. Sci. USA 75, 1274-1278.
32. Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S. and Weissmann, C. (1980) Gene 10, 1-10.
33. Fiddes, J.C. and Goodman, H.M. (1980) Nature 286, 684-687.
34. Sutcliffe, J.G. (1978) Nucl. Acids Res. 5, 2721-2728.
35. Hartley, J.L. and Donelson, J.E. (1980) Nature 286, 860-865.
36. Hitzeman, R.A., Leung, D.W., Perry, L.J., Kohr, W.J., Hagie, F.E., Chen, C.Y., Lugovoy, J.M., Singh, A., Levine, H.L., Wetzel, R. and Goeddel, D.V. (1982) in Proceedings of the Berkeley Workshop on Recent Advances in Yeast Molecular Biology: Recombinant DNA, 173-190.
37. Faye, G., Leung, D.W., Tatchell, K., Hall, B.D. and Smith, M. (1981) Proc. Natl. Acad. Sci. USA 78, 2258-2262.
38. Bennetzen, J.L. and Hall, B.D. (1982) J. Biol. Chem. 257, 3026-3031.
39. Holland, J.P. and Holland, M.J. (1980) J. Biol. Chem. 255, 2596-2605.
40. Holland, M.J., Holland, J.P., Thill, G.P. and Jackson, R.A. (1981) J. Biol. Chem. 256, 1385-1395.
41. Ikemura, T. (1982) J. Mol. Biol. 158, 573-597.
42. Kozak, M. (1981) Nucl. Acids Res. 9, 5233-5252.
43. Zalkin, H. and Yanofsky, C. (1982) J. Biol. Chem. 257, 1491-1500.
44. Goldberg, M. (1979) Ph.D. thesis, Stanford University.
45. Breathnach, R. and Chambon, P. (1981) Ann. Rev. Biochem. 50, 349-383.
46. Darnell, J.E. Jr. (1982) Nature 297, 365-371.
47. Gallwitz, D., Perrin, F. and Seidel, R. (1981) Nucl. Acids Res. 9, 6339-6350.
48. McLaughlin, C.S., Warner, J.R., Edmonds, M., Nakazato, H. and Vaughan, M.H. (1973) J. Biol. Chem. 248, 1466-1471.
49. Zaret, K.S. and Sherman, F. (1982) Cell 28, 563-573.
50. Beggs, J.D., van den Berg, J., van Ooyen, A. and Weissmann, C. (1980) Nature 283, 835-840.
51. Struhl, K. and Davis, R.W. (1981) J. Mol. Biol. 152, 535-552.
52. Warner, J.R. (1982) in The Molecular Biology of the Yeast Saccharomyces cerevisiae (J.N. Strathern, E.W. Jones and J.R. Broach, eds.) vol. 2 (Cold Spring Harbor Laboratories, New York) in press.
53. Brosius, J., Palmer, M.L., Kennedy, P.J. and Noller, H.F. (1978) Proc. Natl. Acad. Sci. USA 75, 4801-4805.