Volume 12 Number 17 1984 Nucleic Acids Research

Cloning with tandem gene systems for high level gene expression

Nancy Lee*, J.Cozzitorto, N.Wainwright and D.Testa

Department of Molecular Genetics, Interferon Sciences, Inc., 783 Jersey Avenue, New Brunswick, NJ 08901, USA

Received 27 July 1984; Accepted 13 August 1984

ABSTRACT

A method has been devised for increasing the copy number of a gene (or genes) cloned into a plasmid while minimizing the size of the plasmid. If n copies of a transcriptional unit are cloned, including the promoter, coding region and terminator, the size of the plasmid will increase by n times the total size of the unit. However, if we borrow the concept of polycistronic operon and sandwich n structural genes, each with its own ribosome binding-site, between a promoter and a transcription terminator, there will be a space saving equivalent to n-1 promoters and n-1 transcription terminators. We have constructed plasmids in which an E. coli lipoprotein promoter is followed by 1 to 4 human leukocyte interferon genes and a transcription terminator. The applications of this method in genetic engineering are discussed.

INTRODUCTION

One of the purposes of genetic engineering is to increase the production of desired gene products in a cell. Gene expression can be enhanced at the levels of transcription and translation by gene manipulation. Moreover, for genes located in plasmids, it has been found that the amount of a gene product produced by a cell can also be raised by either increasing the copy number of the plasmid harboring the gene, or by increasing the copy number of the promoter-gene complex in the plasmid (1-2).

Theoretically, the higher the copy number of a cloned gene a plasmid contains, the higher will be the gene dosage effect. However, there are certain constraints. Temperature-regulated runaway plasmids achieve high plasmid copy number, but, the cells usually die as a result of overdosage and/or disruption of essential cellular functions. Moreover, as plasmid size increases, transformation efficiency of many cell types becomes

reduced, and larger sized plasmids tend to have low copy numbers (3-4). In particular, it has been found for many cell types that transformation efficiency decreases once the size of a plasmid increases above 15Kb (5). Accordingly, for these cell types, plasmid length must be carefully controlled.

Until now, hybrid plasmids and other genetically engineered hybrid nucleic acid sequences constructed for production of a particular gene product have contained at least the following: a promoter region; a ribosomal-binding region;a translation initiation region; a coding region for the gene of interest; a translational termination region; and a transcriptional termination region.

The presence of a strong promoter and a fine tuned ribosome-binding region within this kind of construction will result in high levels of gene expression. However, if this scheme is used in cloning multiple copies of a gene so as to obtain gene dosage effect, for every structural gene cloned, a promoter region and a transcriptional termination region also have to be cloned. The presence of multiple copies of these regions contributes significantly to the overall size of the hybrid plasmid. This leads to the problems of poor transformation efficiency and low copy number described above for larger sized plasmids.

Many bacterial operons are polycistronic, that is, several different genes are transcribed as a single message from their operons. Some of the examples are the lactose operon which contains 3 linked genes (lac Z, lac Y, and lac A) and the tryptophan operon which contains 5 linked genes (trp E, trp D, trp C, trp B, and trp A) (6). In an operon, the messenger RNA is initiated at the promoter-operator region. Within the transcript, coding regions are separated by intercistronic regions of varying lengths. This concept can be borrowed to construct plasmids containing "homopolycistronic" (HPC) sequences for high-level gene expression. In this construction, a promoter (with or without an operator) is followed by more than one copy of the coding region.

Each gene is preceded by a suitable ribosome-binding region, with only one 5' untranslated region and one transcription

terminator needed. Therefore, economy of promoter and end regions is achieved. This is to be distinguished from a more conventional method of cloning multiple whole transcriptional units, which we refer to as "polymonocistronic" (PMC) configuration.

In this paper, a human leukocyte interferon gene (LeIFN) was used to construct a hybrid plasmid containing homopolycistronic sequences of up to four interferon genes. Hybrid plasmids containing PMC sequences were also constructed, and the interferon activities expressed in E. coli cells transformed with these plasmids were compared.

MATERIALS AND METHODS

Strains and Plasmids

E. coli K-12 strains JA221 (hsd M+ hsd R- recA Len Lac Y trp) and JA221 (hsd M+ hsd R- recA Len Lac Y trp/F'lac I^q lac Z+ lac Y+ pro B+) were used as transformation hosts.

The cloning vehicles pIN-II-A3, pIN-I-A3, pKEN030 were kindly supplied by Dr. M. Inouye. Construction of these vehicles is as described by Nakamura and Inouye (7). pCGS261 (provided by Collaborative Research, Inc.) contains a leukocyte interferon gene (LeIFN-B', or, as it is also known, IFN alpha 8') on a 1.1 Kb HindIII fragment (see Figure 1). In this construction, 69 base pairs coding for the signal sequence of the preinterferon protein has been removed and replaced by an ATG translation initiation codon as well as ClaI and HindIII linkers. The nucleotide sequence of the coding region of the LeIFN-B (IFN alpha-8) reported by Goeddel et. al. (8) is identical to this leukocyte interferon gene, except for a single base deletion and, twelve bases away, a single base insertion. This causes the sequence coding for the 99th to 102nd amino acid residues to be changed from Val-Leu-Cys-Asp to Ser-Cys-Val-Met (9).

Plasmids were isolated by the method described by Tanaka & Weisblum (10).

All restriction enzymes, T_4 DNA ligase, XhoI linker d(CCTCGAGG), and the DNA polymerase I Klenow fragment were obtained from New England Bio Labs. Bacterial alkaline phosphatase was obtained from Worthington Diagnostic Systems,

Inc., Freehold, New Jersey. Nuclease S_1 was obtained from Boehringer-Mannheim, Indianapolis, Indiana.

Preparation of Total Lysate

One ml of E. coli JA221 cells containing the plasmids to be assayed were grown in L-Broth containing 50 ug/ml ampicillin. For inducible systems, cells were grown to an optical density of 0.4 at 600 nm and isopropyl-β-D-thiogalactoside (IPTG, Sigma) was added to a final concentration of 2mM. After further incubation of the culture at 37°C for 1 hour, cells were sedimented and resuspended in 5M guanidine chloride containing 0.5mM phenylmethanesulfonyl fluoride (PMSF, Sigma) and kept on ice for 10 minutes with vigorous vortexing. The extent of breakage was monitored by light microscopy. The extracts were centrifuged 10,000 rpm for 10 minutes and supernatants diluted for assay. For constitutive systems, cells were harvested at optical densities within 0.5 to 1.0 at 600 nm and lysates prepared as described.

Measurement of Antiviral Activity

Antiviral assays were performed using cytopathic effect inhibition (CPE) of Vesicular Stomatitis Virus on HEp-2 cells (11). All interferon titers were determined relative to the NIH alpha (human) leukocyte/Sendai interferon standard #GA-23-902-530.

Gel Electrophoresis

Agarose gel electrophoresis was performed as described by Sharp, et. al. (12). Fragments were isolated from the gel by electroelution onto a Whatman NA-45 paper strip. The DNA was was eluted by NET buffer (1M NaCl, 0.1mM EDTA, 20mM Tris HCl, pH 8). Recoveries were found to be greater than 90% with no contamination of the final product by agarose.

RESULTS

Preparation of Homopolycistronic (HPC) Hybrid Plasmids

Following the procedure shown in Fig. 1, a 1.1 Kb HindIII fragment containing the IFN-B' (IFN alpha-8') gene from pCGS261 (see Materials and Methods) was cloned into the HindIII site of the pIN-II-A3 cloning vehicle (7). Expression in this system is under the control of an E. coli outer membrane lipoprotein (lpp)

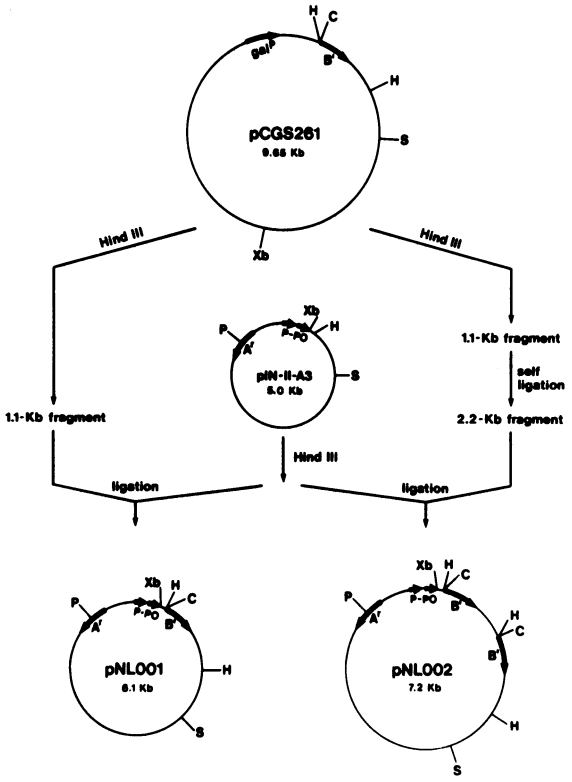

Fig. 1 Cloning either one or two copies of an IFN-B' containing HindIII fragment (1.1 Kb) into the inducible expression vehicle pIN-II-A3. pNL001 contains one HindIII fragment, while pNL002 contains two HindIII fragments in tandem. Abbreviations used: galP = yeast galactose promoter, P-PO = lipoprotein promoter + lac promoter-operator, A^r = ampicillin resistance, B' = interferon B' gene, Xb = XbaI, H = HindIII, C = ClaI, S = SalI, P = PstI.

promoter and a lactose (lac) promoter-operator. One of the resulting hybrid plasmids contained one copy of the LeIFN-B' (IFN alpha-8') gene and was designated pNL001.

A 2.2 Kb DNA fragment, as is also shown in Fig. 1, had two 1.1 Kb HindIII fragments linked in the same orientation, and was cloned into the HindIII site of pIN-II-A3. This hybrid plasmid was designated pNL002. Restriction enzyme mapping confirmed the reading direction of the LeIFN-B' genes in pNL001 and pNL002 to be the same as that of the lipoprotein promoter and lactose

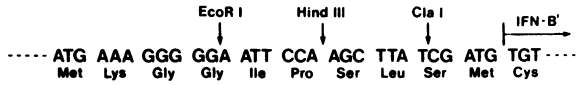

Fig. 2 N-terminus of the coding region immediately 3' to the lipoprotein promoter or lactose promoter-operator of the hybrid interferon gene produced by the processes of Figs. 1 and 4-7.

promoter-operator.

Fig. 2 shows the N-terminus of the coding region immediately after the lactose promoter-operator expected for pNL001 and pNL002. As can be seen in that figure, the interferon coding region is preceded by sequence coding for 10 amino acid residues which are introduced by pIN-II-A3, the HindIII and ClaI restriction sites, and an ATG codon The Shine-Delgarno sequence is 7 bases 5' to the first ATG. For pNL002, as shown in Fig. 3, the translation termination site (TGA) of the first LeIFN-B' is separated from the translation initiation site (ATG) of the second LeIFN-B' by 635 bases. There is a Shine-Delgarno sequence 12 bases 5' to the ATG of the second LeIFN-B'.

Due to the lac promoter-operator element they contain, pNL001 and pNL002 are both inducible systems. In order to obtain constitutive systems which do not need an inducer for expression, pNL011 was constructed from pIN-I-A3 and pCGS261. pIN-I-A3 is exactly the same as pIN-II-A3 except it lacks the lac promoter-operator (7). Following the procedure shown in Fig. 1, pNL011 was formed from pIN-II-A3 and pCGS261. As with pNL001, pNL011 contains only a single copy of LeIFN-B'.

To obtain a constitutive system having two copies of LeIFN-B', a 3.1 Kb fragment was obtained from SalI digestion and ClaI partial digestion of pNL002 and ligated by T4 DNA ligase to

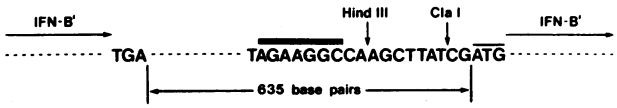

Fig. 3 Intercistronic region between two IFN-B' genes tandemly arranged in accordance with the processes of Figs. 1 and 4-5. Eleven base pair 5' to the second IFN-B' gene is a ribosome-binding site marked by the heavy bar. The initiation codon for the second IFN-B' gene is marked by a light bar.

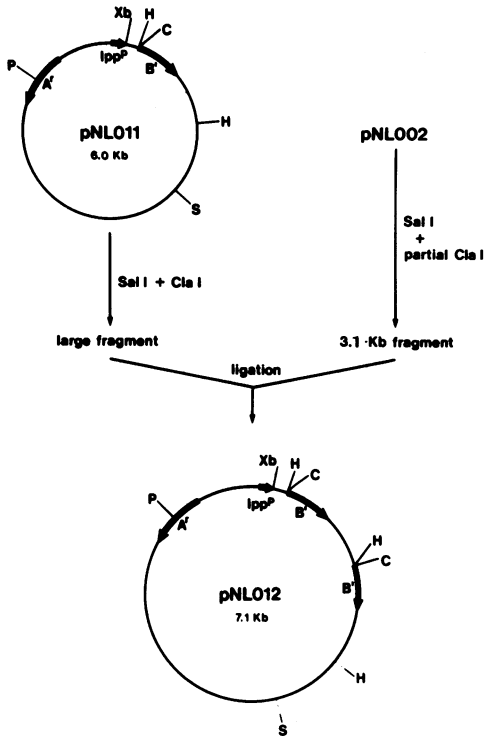

Fig. 4 Construction of hybrid plasmid pNL012 which contains two linked IFN-B' genes. pNL011 is similar to pNL001 except that the lac promoter operator region is removed so that transcription therefore becomes constitutive. Abbreviations used: $1pp^P$ = lipoprotein promoter; B', Xb, H, C, S, P, and A^r same as in Fig. 1.

a 4.0 Kb fragment from SalI and ClaI digestion of pNL011. The resulting plasmid was designated pNL012 (see Figure 4). Analyzing by restriction enzyme digestion, the two LeIFN-B' genes were found to be in the same orientation as the lpp promoter.

Using pNL012 as starting material, plasmids containing tricistronic and tetracistronic LeIFN-B' genes were constructed as shown in Figure 5. After partial digestion with ClaI, 7.1 Kb DNA fragments which were pNL012 linearized by a single cut at either one of the two possible ClaI sites were isolated and ligated by T4 DNA ligase with 1.1 Kb interferon-containing ClaI fragment obtained from pNL012. The resulting plasmid can have

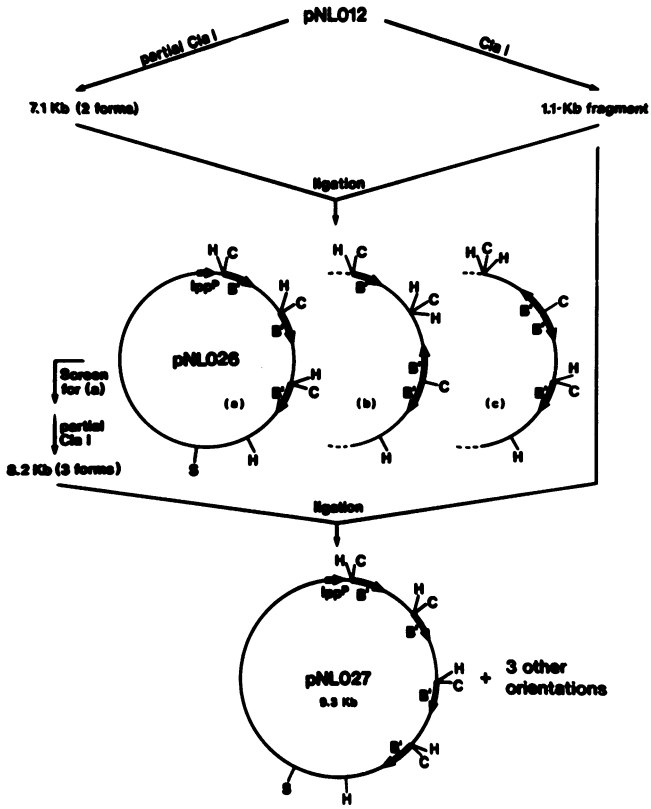

Fig. 5 Construction of hybrid plasmids pNL026 and pNL027 which contain three linked IFN-B' genes and four linked IFN-B' genes, respectively. The abbreviations used are the same as in Fig. 4.

three possible structures. After transformation of E. coli JA221 cells, plasmids were isolated, purified, and restriction pattern analyzed. The plasmids containing three and four interferon genes in the same orientation were designated pNL026 and pNL027, respectively.

Preparation of Polymonocistronic (PMC) Hybrid Plasmids

For construction of a PMC hybrid plasmid, it was necessary to have convenient restriction sites 5' to the promoter and 3' to the transcription termination site. The nature of the restriction sites were chosen so that they would facilitate the formation of tandem repeats, control the relative orientation of

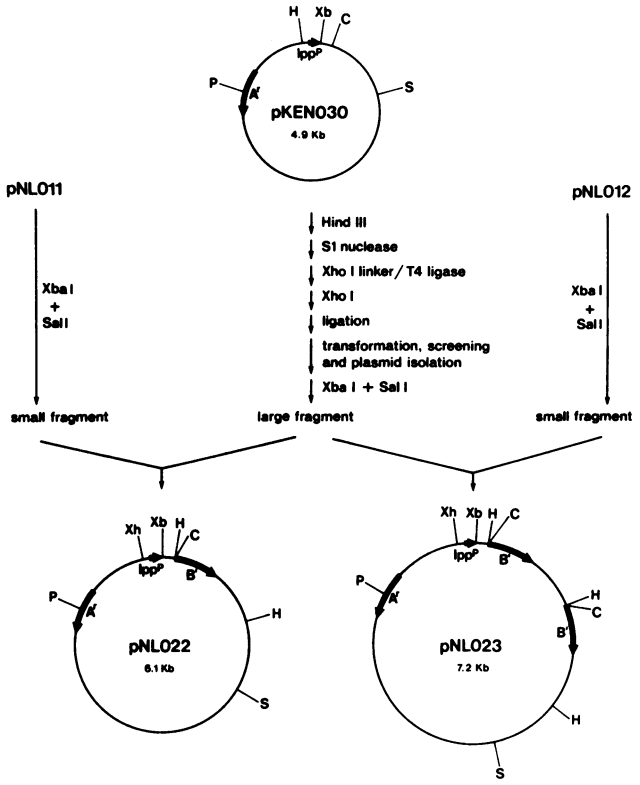

Fig. 6 Construction of hybrid plasmid pNL022. pNL022 is similar to pNL011 except that a XhoI site is located 5' to the lipoprotein promoter. Abbreviations used: Xh = XhoI; remainder as in Fig. 4.

the tandem repeats, and be suitable for checking the size of the tandem formed.

For this purpose, pNL022 was constructed in accordance with the procedures shown in Figure 6. This plasmid has a XhoI site 5' to the lipoprotein promoter and a SalI site 3' to the transcriptional terminator. pKEN030 (7) which has a HindIII site 5' to the lipoprotein was used as the starting material. pKEN030 was linearized by a HindIII restriction enzyme, cut and made blunt-ended by S1 nuclease. XhoI linkers were then attached by blunt-end ligation, cohesive ends were formed by XhoI cleavage, and the DNA was recircularized by T4 DNA ligase.

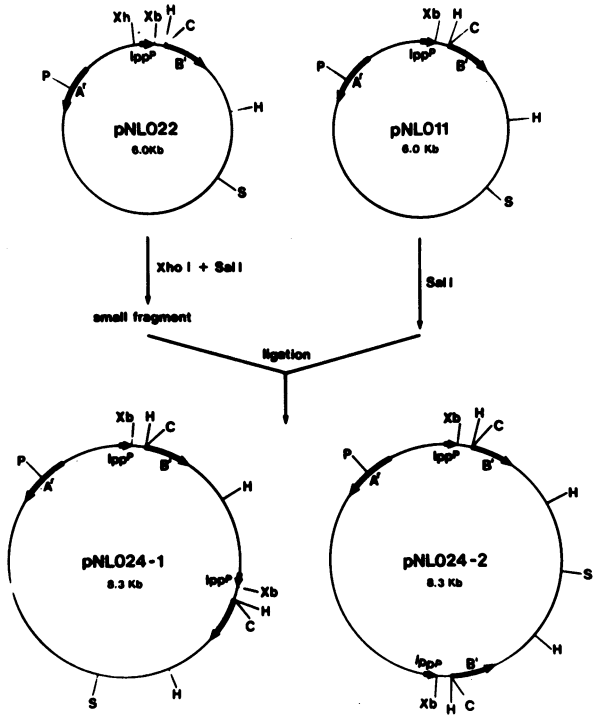

Fig. 7 Construction of hybrid plasmids pNL024-1 and pNL024-2. These plasmids each contain two promoter-gene units. Abbreviations as in Fig. 6.

After transformation and screening, the large fragment obtained was ligated to the small XbaI/SalI fragment from pNL011 to form pNL022. Other than the additional XhoI site, pNL022 is identical with pNL011.

To form the desired PMC hybrid plasmid, pNL022 was cut by XhoI and SalI and the resulting small fragment was ligated to pNL011 previously linearized by SalI (see Figure 7). The two possible hybrid plasmids with opposite orientation were designated pNL024-1 and pNL024-2. The joining of a SalI cohesive end with another SalI cohesive end recreated a SalI site. The joining of SalI and XhoI cohesive ends results in a sequence that cannot be cut by either enzyme. The SalI site of each plasmid is located at different positions. Hence, pNL024-1 was easily distinguished from pNL024-2 by cutting with XbaI and

evaluating the sizes of the resultant restriction fragments.

It was found that the expression of interferon was affected by the presence of an 8 base pairs XhoI linker upstream of the lipoprotein promoter (see below). So that PMC hybrid plasmid pNL024-1 could be compared with an equivalent homopolycistronic hybrid plasmid containing a XhoI site upstream of the lipoprotein promoter, another HPC hybrid plasmid, designated pNL023, was constructed. The pNL023 plasmid was prepared in the same way as pNL022 (see Figrue 6), except that pNL012 instead of pNL011 was used as starting material. This plasmid contains two interferon coding regions downstream of a single lipoprotein promoter.

Interferon Produced by Tandem Hybrid Plasmids

The interferon activities expressed by E. coli cells (JA221 strain) harboring various plasmids were determined by cytopathic effect inhibition (CPE) of Vesicular Stomatitis Virus on HEp-2 cells (see Table 1).

Within the HPC series pNL011 (monocistron), pNL012 (dicistron), pNL026 (tricistron), and pNL027 (tetracistron), the ratio of relative interferon activities were 1.0:2.3:3.3:4.5 (see Table 1). Within the pNL022 (monocistron) and pNL023 (dicistron) series, the ratio of relative interferon activities were 1.0:2.0. Although this ratio is similar to that of pNL011 and pNL012, the interferon titers of the pNL022 and pNL023 series were found to be consistently three to four fold higher. This difference may be due to the presence of the 8 bp XhoI linker in the pNL022 and pNL023 plasmids.

A PMC plasmid (dicistronic), pNL024-1, was found to have a level of expression about the same as the comparable HPC plasmid, pNL023.

Preparation of an Interferon Titer Produced by a Promoterless Polycistronic Plasmid

Possibly there exist sequences that function as a promoter and a transcription terminator within the 638 base pairs intercistronic region. These plasmids might actually contain tandem transcriptional units rather than a HPC system. The interferon titer reported can be explained as well if this is the case. In order to investigate this possibility, a plasmid containing a

TABLE 1
INTERFERON ACTIVITIES IN TANDEM CONSTRUCTIONS

(A)

	Interferon Titers U x 10^4/ml/OD (±S.D.***)	Relative Ratio
ISE011* (pNL011) (m**=1)	2.1 (± 0.8)	1.0
ISE012* (pNL012) homopolycistronic (m**=2)	5.0 (± 1.6)	2.3
ISE026* (pNL026) homopolycistronic (m**=3)	6.9 (± 2.7)	3.3
ISE027* (pNL027) homopolycistronic (m**=4)	9.5 (± 3.2)	4.5

(B)

	Interferon Titers U x 10^4/ml/OD (±S.D.***)	Relative Ratio
ISE022* (pNL022) (m**=1)	8.0 (± 4.6)	1.0
ISE023* (pNL023) homopolycistronic (m**=2)	16 (± 6.7)	2.0
ISE024* (pNL024-1) polymonocistronic (m**=2)	14.7 (± 6.2)	1.9

* In ISE strains, E. coli JA221 is used as host cell.
** m=number of cistrons in homo-polycistronic and polymonocistronic systems
*** Standard Deviation

dicistronic sequence was constructed with no promoter 5' to the first interferon gene. If there is a functional promoter within the intercistronic region, there will be expression from the second, but not from the first interferon gene. In its absence no interferon will be produced.

A 2 Kb ClaI-SalI fragment containing two interferon genes, but no lipoprotein promoter, was isolated after partial ClaI and SalI digestion of pNL023. This fragment was ligated to the large ClaI-SalI fragment of pBR322 and the resulting plasmids

were used to transform E. coli strain JA221. Plasmids were isolated from ampicillin resistant and tetracycline sensitive colonies and analyzed by restriction enzyme digestion. Positive clones were grown and assayed. The interferon titers were found to be insignificant and indistinguishable from that of the pBR322 control. This result demonstrates the absence of a functional promoter in the intercistronic region.

In Vivo Stability of Tandem Hybrid Plasmids

The stability of tandem hybrid plasmids in rec A host was studied. E. coli containing pNLO22, pNLO23, pNLO26, and pNLO27 were cultured in broth under ampicillin selection for more than a hundred generations. For each strain, six individual clones were picked and the sizes of the plasmids were determined by agarose gel electrophoresis. The plasmid sizes as well as the gel patterns resulting from Pst I and Sal I digestion of these plasmids were identical with that of the original plasmid.

DISCUSSION

A series of tandem hybrid plasmids containing HPC as well as PMC sequences have been constructed. The HPC plasmids were stably maintained in a rec A host for more than a hundred generations. We have shown that it is unlikely that functional promoters exist in the intercistronic regions. The absence of a promoter also implies the absence of a transcriptional terminator. If terminators existed in the absence of internal promoters, then only the first gene would be translated. In such a situation, all homopolycistronic systems would give similar interferon titers. However, this is not the case.

The interferon titers for cell extracts harboring various HPC and PMC sequences were found to be directly proportional to the number of interferon genes incorporated in the sequences. The activity values, especially for those plasmids containing more than one interferon gene, compare favorably with those reported by Yelverton et. al. ([13]). It should be noted that the interferons expressed from the first coding region 3' to the lipoprotein promoter contain 9 extra amino acids at the amino terminus. Elimination of this additional sequence may produce even higher titers.

The level of interferon expression is about the same for comparable HPC and PMC plasmids, such as pNL023 and pNL024-1, which contained the same number of interferon genes per plasmid. This suggests that the efficiency of gene expression for both tandem gene arragnements is very similar.

These results demonstrate that cloning in HPC systems, one has advantages of maximizing the copy number of the cloned gene in a plasmid to achieve gene dosage effect. Plasmid size is minimized and high plasmid copy number is maintained while transformation efficiency and gene expression are not compromised. The optimal copy number of genes to be incorporated into a plasmid to obtain maximum expression can be established experimentally.

The space saving that can be achieved by adopting the HPC tandem gene format is illustrated by using pNL027. In this case, the sizes of the DNA sequences containing the lipoprotein promoter, the interferon gene, and the lipoprotein transcription terminator are 0.23 Kb, 1.1 Kb, and 1.0 Kb, respectively. The length of the HPC sequence is (0.23 + 4(1.1) + 1.0 Kb) = 5.63 Kb. When the PMC format is used, the length would be 4(0.23 + 1.1 + 1.0) Kb = 9.32 Kb. Therefore, a space saving of almost 40% is achieved.

Even greater packing density than that described in the foregoing examples can be achieved by shortening the intercistronic region. In naturally occurring polycistronic operons, the intercistronic regions typically range from a few to several hundred base pairs ([14-16]). The intercistronic region for the plasmids of the foregoing examples is a 638 base pairs sequence as shown in Figure 3. Accordingly, more than 600 base pairs of intercistronic region can be deleted.

It is interesting to note that the seemingly similar level of expression of each component gene in the HPC systems we have constructed is in contrast to many naturally occurring polycistronic operons where the levels of expression of non-identical genes are markedly different ([17, 18]). This may be due to the fact that in the HPC systems, all the component genes and intercistronic regions are identical. Hence, the various factors affecting transcription, translation and the degradation of the

products of transcription and translation will affect each component to the same extent. It may be useful to use the HPC tandem system as a model to study the mechanisms involved in the regulation of polycistronic operons. This can be achieved by making systematic insertions, deletions, or replacements within the HPC operon.

It can be envisaged that one application may be in the design and construction of polycistronic sequences which contain more than one type of gene, heteropolycistronic (HEPC). In view of the results which show that a gene product is produced in an amount directly proportional to the number of genes in the HPC sequence, the expression levels of each type of gene within the HEPC arrangement can be predetermined by carefully controlling their copy number in the arrangement. For this analysis, we have constructed HEPC plasmids containing equal number of two different genes, and equimolar expression was observed (unpublished results). Further characterizations of the various HPC, PMC, and HEPC systems are in progress.

ACKNOWLEDGEMENTS

We thank L. Sendelbach, S. Goshorn, and J. Enright for performing interferon assays and T. Pilecki for preparation of the manuscript.

*To whom reprint requests should be addressed

Abbreviations used in this manuscript: LeIFN, Leukocyte interferon; HPC, homopolycistronic; PMC, polymonocistronic; HEPC, heteropolycistronic.

REFERENCES

1. Uhlin, B., Molin, S., Gustafsson, P. and Nordstrom, K. (1979) Gene 6, 91-106.
2. Lee, N., Nakamura, K. and Inouye, M. (1981) J. Bacteriol 146, 861-866.
3. Kushner, S.R. (1978) in Genetic Engineering, Boyer, H.W. and Nicosia, S. (eds.), pp. 17-23, Elsevier/North Holland Biomedical Press, Amsterdam.
4. Broda, P. (1979) in Plasmids, W. H. Freeman and Co., Ltd., San Francisco.
5. Rodriguez, R. and Tait, R. (1983) in Recombinant DNA Techniques, p. 4, Addison Wesley, Reading, Massachusetts.
6. Miller, J.H. and Reznikoff, W. S. Eds. (1978) The Operon, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
7. Nakamura, K. and Inouye, M. (1982) EMBO J. 1, 771-775.

8. Bowden, D.W., Mao, J., Gill, T., Hsiao, K., Lillquist, J., Testa, D. and Vovis, G. (1984) Gene 27 87-99.
9. Goeddel, D. Leung, D., Dull, T., Gross, M., Lawn, R., McCandliss, R., Seeburg, P., Ullrich, A., Yelverton, E., and Gray, P. (1981) Nature 290, 20-26.
10. Tanaka, T. and Weisblum, B. (1975) J. Bacteriol. 121, 354-362.
11. Chang, T., Testa, D., Kung, P., Perry, L., Dreskin, H.J. and Goldstein, G. J. Immunology (1982) 128, 585-589.
12. Sharp, P.A., Snyder, B., and Sambrode, J. (1973) Biochemistry 12, 3055-3063.
13. Yelverton, E., Leung, D., Weck, P., Gray, P. and Goeddel, D. (1981) Nucleic Acid Res. 9, 731-741.
14. Christie, G.E. and Platt, T. (1980) J. Mol. Biol. 142, 519-530.
15. Selker, E. and Yanofsky, C. (1979) J. Mol. Biol. 130, 135-143.
16. Barry, G., Squires, C. and Squires, C.L. (1980) Proc. Natl. Acad. Sci. USA 77, 3331-3335.
17. Zabin, I. and Fowler, A. (1978) in The Operon, Miller, J. and Reznikoff, W. Eds. pp. 89-121, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
18. Higgins, C., Ames, G., Barnes, W., Clement, J. and Hofuung, M. (1982) Nature 298, 760-762.