The EMBO Journal vol.5 no.11 pp.2987–2994, 1986

# Promoters of *Escherichia coli*: a hierarchy of *in vivo* strength indicates alternate structures

**Ulrich Deuschle[1], Wolfgang Kammerer[1], Reiner Gentz and Hermann Bujard[1]**

Central Research Units, F.Hoffmann-La Roche & Co. AG, CH-4002 Basel, Switzerland, and [1]Zentrum für Molekulare Biologie, Universität Heidelberg, PO Box 10 62 49, 6900 Heidelberg, FRG

Communicated by H.Bujard

**The strength *in vivo* of 14 promoters was determined in a system which permits the quantitation of RNA synthesis with high accuracy. Up to 75-fold differences in promoter strength were measured and the most efficient signals are promoters from coliphages T7 and T5. Their activity approaches the strength of fully induced promoters of the rRNA operons which may be close to the functional optimum of a single sequence. By contrast, a synthetic 'consensus promoter' belongs to the less efficient signals. Our data show that optimal promoter function can be achieved by alternate structures and strongly suggest that information outside of the 'classical' promoter region contributes to promoter activity.**

*Key words: E. coli* promoters/*in vivo* strength/alternate structures

## Introduction

Despite a wealth of information on the structure of *Escherichia coli* promoters, the topography of RNA polymerase/promoter complexes and on the processes governing the onset of specific transcription (for review, see Rosenberg and Court, 1979), our understanding of structure/function relationships of *E. coli* promoters still remains unsatisfactory. The steadily increasing number of elucidated promoter sequences has allowed the extensive study of structural homologies and refined 'consensus sequences' were proposed (for review, see Hawley and McClure, 1983). So far, however, such model sequences were of little predictive value, whenever functional parameters of an individual promoter were to be derived from structural information alone. Quantitative information on the function of defined promoter sequences is obviously required for a better understanding of how the complex functional programme of a promoter is stored in a DNA sequence.

Here we describe an experimental system for the accurate determination of promoter strength *in vivo* and *in vitro*. Promoter activities are measured by monitoring RNA synthesis in relation to an internal standard. Thus, the results are independent of translational effects and of gene dosage. Differences in mRNA half-life can be taken into account. Fourteen promoters were characterized. The up to 75-fold differences in promoter strength found *in vivo* yielded a functional hierarchy of sequences which permits novel correlations between structural and functional parameters. Evidence for a hypothesis derived from these and other data (Gentz and Bujard, 1985) is also presented in the accompanying publication (Kammerer *et al.*, 1986), which shows that sequence elements located in the transcribed region can significantly contribute to promoter strength.

## Results

### *Experimental strategy*

To determine accurately the efficiency of transcriptional signals we have developed a system in which the RNA produced under the control of such signals is quantified in relation to an internal standard. The system outlined in Figure 1 consists of: (i) plasmid pDS2 carrying the β-lactamase (*bla*) transcriptional unit as an internal standard as well as indicator regions (the sequences for dihydrofolate reductase of the mouse, *dhfr*, and of the chloramphenicol acetyltransferase, *cat*) which can be brought under the control of transcriptional signals in various ways, and (ii) three coliphage M13 derivatives containing sequences complementary to coding regions of the internal standard (*bla*) or the indicator regions (*dhfr, cat*), respectively. Radioactively labelled RNA from these regions is hybridized in solution against the single-stranded (ss) DNA probes of the three M13 recombinants. Adsorption of the hybrids to nitrocellulose via their ssDNA portion allows a rapid quantitation of the hybridized RNA. Analysis of the S1-resistant material of such hybrids permits the characterization of the transcription products.

Promoters of interest are integrated in front of the *dhfr/cat* indicator region and the amount of RNA specific for *dhfr* and/or *cat* sequences is compared with the transcripts produced in the *bla* region. Consequently, all promoters are characterized with respect to the β-lactamase promoter ($P_{bla}$) and promoter strengths are given in '$P_{bla}$-units'.

Three more aspects were considered. Firstly, intense transcription interferes with plasmid replication (Stüber and Bujard, 1982; Bujard *et al.*, 1985). Insertion of an efficient terminator at site 1 or 2 of pDS2 (Figure 1) is therefore essential for the stable integration of strong promoters. Secondly, fragments carrying the promoter of interest may also initiate transcription in the opposite direction, i.e. into the *bla* region, thereby increasing the level of expression of the internal standard. Such activities can be picked up by analysing the S1-resistant RNA/DNA hybrids (Figure 4). The integration of an additional terminator between $P_{bla}$ and the promoter insertion site (site 3 in Figure 1) resolves this problem. Finally, the half-life times of *in vivo* RNA can vary significantly and are often influenced by the structure of the 5′ region of the molecule (Yamamoto and Immamoto, 1975; Belasco *et al.*, 1986). Since different promoters yield RNAs with differing 5′ sequences it is necessary also to monitor the stability of the various *in vivo* transcripts.

### *Integration of promoters into the pDS2 system*

The promoter-carrying fragments summarized in Table I were integrated into the polylinker of pDS2/$t_0$1 (Figure 1). Deleterious overproduction of DHFR or CAT protein was avoided by fusing the *dhfr* sequence to the promoter fragment out-of-frame with potentially efficient ribosomal binding sites and by reducing transcription into the *cat* gene at site 1 (Figure 1) with terminator $t_0$ of phage lambda. Plasmids containing strong promoters inserted in the proper orientation were selected by plating the trans-

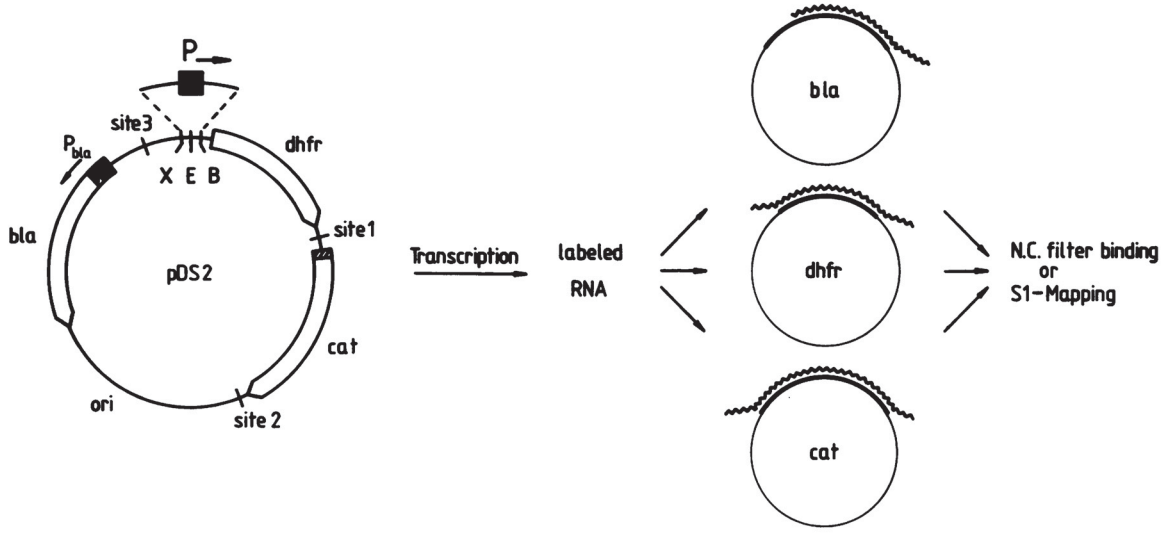


**Fig. 1.** Experimental principle. Promoters integrated into the polylinker site of pDS2 (a derivative of pDS1, Stüber and Bujard, 1982) control the transcription of two indicator sequences, *dhfr* and *cat*. Sites where terminators can be inserted allow the prevention of clockwise readthrough transcription into the replication region (sites 1, 2) as well as safeguarding the *bla* region (site 3) which is used as an internal transcriptional standard unit. The directions in which transcription is initated by $P_{bla}$ and by the integrated promoters (P ■) are indicated by arrows. Translational initiation signals are denoted as ▨. X, E and B designate *Xho*I, *Eco*RI and *Bam*HI cleavage sites respectively. Labelled RNA transcribed from these plasmids is analysed by hybridization against three M13-derived ssDNA probes which carry a portion of the *bla* gene, the coding sequence of DHFR or of CAT, respectively. Analysis of the hybrids by nitrocellulose adsorption or electrophoresis of the S1-resistant material permits a quantitative and a qualitative characterization of the three RNA species. Transcripts originating within the cloned promoter fragment and directed towards the *bla* region are identified by the M13mp9 *bla* probe which extends up to the *Eco*RI site. In such cases two species of S1-resistant hybrids are obtained (Figure 4). Insertion of a terminator at site 3 prevents interference of such transcription with the *bla* standard unit.

**Table I.** Promoter carrying fragments integrated in pDS1/$t_0$1

| Promoter | Fragment size (in bp) | Integration sites | Transcript length up to the coding sequence of DHF |
|---|---|---|---|
| $P_{H207}$ | 212 | *Eco*RI | 157 |
| $P_{D/E20}$ | 471 | *Eco*RI | 132 |
| $P_{N25}$ | 254 | *Xho*I | 106 |
| $P_{G25}$ | 233 | *Eco*RI | 81 |
| $P_{J5}$ | 220 | *Eco*RI | 106 |
| $P_{A1}$ | 330 | *Eco*RI/*Bam*HI | 96 |
| $P_{A2}$ | 143 | *Eco*RI/*Bam*HI | 76 |
| $P_{A3}$ | 154 | *Eco*RI/*Bam*HI | 64 |
| $P_L$ | 278 | *Eco*RI | 135 |
| $P_{lac}$ | 129 | *Xho*I | 92 |
| $P_{lacUV5}$ | 91 | *Eco*RI/*Bam*HI | 47 |
| $P_{tacI}$ | 285 | *Eco*RI | 63 |
| $P_{con}$ | 61 | *Eco*RI/*Bam*HI | 23 |

Thirteen promoters were integrated into pDS1 derivatives using the restriction sites indicated. Of these promoters, nine are coliphage promoters, namely $P_{H207}$, $P_{N25}$, $P_{J5}$, $P_{D/E20}$, and $P_{G25}$ from phage T5; $P_{A1}$, $P_{A2}$ and $P_{A3}$ from phage T7 and $P_L$ from phage lambda. $P_{lac}$ and $P_{lacUV5}$ are promoters of the *E. coli lac* operon; $P_{tacI}$ is a trp/lacUV5 promoter hybrid and $P_{con}$ is a promoter synthesized according to a consensus sequence. The exact position of a promoter can be derived from the distance between the first transcribed nucleotide and the A of the initiation codon of the *dhfr* sequence.

formed bacteria onto plates containing 30 μg/ml chloramphenicol. Due to the leakiness of terminator $t_0$, resistance of this magnitude is indicative of strong promoters.

### *Phage M13 derivatives for quantifying RNA by liquid hybridization*

Hybridization probes for the *bla*- and the *dhfr*-specific RNA were obtained by integrating the 752-bp *Pst*I – *Eco*RI fragment of pBR322 and the 672-bp *Bam*HI – *Hin*dIII fragment of pDS1 into the properly cleaved M13mp9. As *cat*-specific probe the 865-bp *Hin*dIII – *Xba*I fragment of pDS1 was integrated into the *Hin*dIII – *Hin*cII cleaved M13mp8 after the ends of the *Xba*I cleavage site had been filled in. The 752-bp fragment contains part of the *bla* region (570 bp) and a stretch of 200 bp upstream of the *bla* promoter. This region is convenient for identifying transcripts which originate within the cloned promoter fragment and are directed towards the *bla* region (Figure 4).

**Table II.** Half-life of mRNAs: influence of the 5′ region

| Promoter (in pDS1/to1) | Half-life of transcripts (min) | | |
|---|---|---|---|
| | *dhfr* | *cat* | *bla* ($P_{bla}$ directed) |
| $P_{H207}$ | 0.6 ± 0.1 | 0.7 ± 0.1 | 1.6 ± 0.4 |
| $P_{J5}$ | 1.2 ± 0.1 | 1.4 ± 0.2 | 1.8 ± 0.3 |
| $P_{lacUV\ 5}$ | 1.4 ± 0.3 | 1.5 ± 0.2 | 1.4 ± 0.4 |

RNA was labelled *in vivo* with [$^3$H]uridine for 1 min before rifampicin was added to the culture (200 μg/ml) and samples were withdrawn at defined times (0.5, 1.0, 2.0, 3.0, 5 and 10 min). RNA was isolated and quantified by hybridization against the proper M13mp9 derivatives. The errors given are maximal deviations from the mean of four independent measurements. Whereas the *dhfr*- and *cat*-specific RNA is synthesized under the control of the promoters indicated, the *bla*-specific transcripts are always initiated by $P_{bla}$ (rightmost column).

### *Examination of RNA stability in vivo*

Half-life times of most mRNAs in *E. coli* are ~1–3 min (Mangiarotti and Schlessinger, 1967). Therefore labelling periods of 5 min and longer yield values describing essentially the steady-state whereas short labelling periods are indicative for newly synthesized RNA and reflect the rate of RNA synthesis. Since *dhfr*-specific RNAs initiated at different promoters have different 5′ regions, their stability *in vivo* cannot be predicted. Therefore RNA was labelled *in vivo* for 5 and 1 min, respectively and the ratio of *bla*- to *dhfr*-specific RNA was determined. Except for

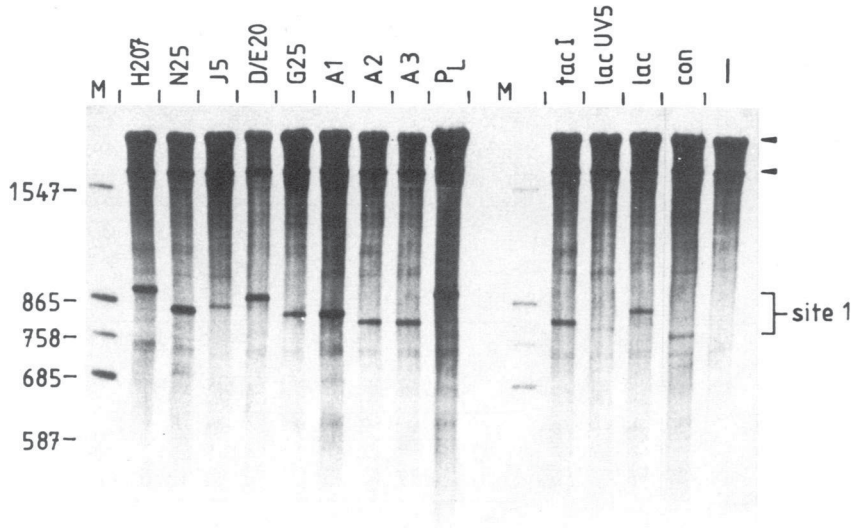


**Fig. 2.** Direct visualization of *in vivo* RNA synthesized under the control of various promoters. Aliquots of $^3$H-labelled total *in vivo* RNA were separated in 4% polyacrylamide/8 M urea gels (20 V/cm; 1 × TBE). The major RNA species seen around position 865 nucleotides are transcripts initiated by the promoters indicated and terminated at site 1 by $t_o$. The differences in abundance of these transcripts correlate with the promoter strengths summarized in Table III. The size differences of the transcripts are due to the differing locations of the promoters within the cloned fragment. The position of the rRNAs are indicated (◄), the rightmost lane contains RNA prepared from a promoter-less plasmid construct. M denotes size markers.

promoter $P_{H207}$ differences in the relative abundance of the two RNAs was 10% or less (data not shown). In the case of $P_{H207}$, however, the apparent promoter strength increased from 20 to 30 $P_{bla}$-units upon reduction of the labelling time from 5 to 1 min. The half-life time of the *bla*-, *dhfr*- and *cat*-specific RNA was therefore determined with plasmids containing the promoters $P_{H207}$, $P_{J5}$ and $P_{lacUV5}$, respectively. Whereas the stability of the transcripts initiated by promoters $P_{J5}$ and $P_{lacUV5}$ were found to be indistinguishable from that of the *bla* transcript, the RNA originating from $P_{H207}$ is distinctly less stable exhibiting a half-life time of only 36 s (Table II).

*Determination of in vivo promoter strength*

Various promoters integrated into pDS1 or pDS2 derivatives were transferred into either *E. coli* C600 or M15. For each promoter three independent cultures were grown up and labelled RNA was extracted. Each RNA preparation was directly analysed by electrophoresis (Figure 2) and aliquots were hybridized to the respective ssDNA probes, M13mp9 *bla*, M13mp9 *dhfr*, M13mp8 *cat*. The RNA/DNA hybrids were then analysed by nitrocellulose adsorption as well as by electrophoresis of the S1-resistant material. Results of a typical experiment in which four different promoters were examined are shown in Figures 3 and 4. The reproducibility of the nitrocellulose adsorption method is demonstrated in Figure 3 where a single RNA preparation has been analysed in duplicate for each promoter. For low background values it is essential that the RNA preparation is virtually free of cellular DNA. The variation in the absolute amount of *bla*-specific RNA is a reflection of gene dosage due to changes in plasmid copy number (Stüber and Bujard, 1982). The analyses of the same RNA/DNA hybrids after nuclease S1 digestion are shown in Figure 4. Although this latter procedure permits only a rough estimate of the relative abundance of the different RNA species it is an important control for the number of promoters within a cloned fragment. Thus, the fragment carrying promoter $P_{H207}$ directs transcription also towards the *bla* region, as revealed by a second RNA species of ~800 nucleotides in length (Figure 4). The efficiency of this second promoter, which is ~60% of $P_{bla}$, was taken into account when calculating the strength of $P_{H207}$. A similar transcript was observed for the fragment containing $P_{A1}$ of phage T7 and, though in much lower abundance, for $P_{tacI}$. These promoters were therefore analysed in a plasmid containing terminator T1 at site 3 (Figure 1). As expected, the apparent promoter strength was increased for $P_{H207}$ and $P_{A1}$.

To examine the reliability of $P_{bla}$ as an internal standard we replaced this promoter by $P_{lacUV5}$ and determined the hierarchy of four promoters with respect to this standard. As seen in Table III there is full agreement between the data derived from both sets of experiments. Furthermore as can be seen in Figure 3, the amount of $P_{bla}$-specific RNA is not influenced by the

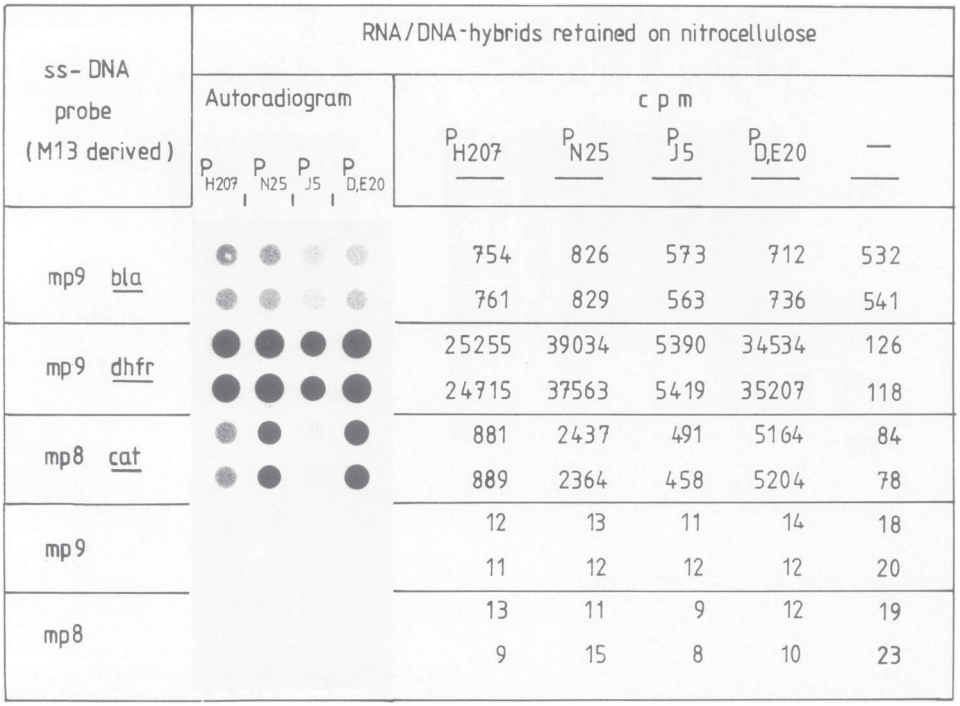


**Fig. 3.** Quantitation of *in vivo* RNA. RNA labelled *in vivo* with [$^3$H]uridine was hybridized to single-stranded M13-derived DNA probes which contain *bla*-, *dhfr*- and *cat*-specific sequences, respectively (mp9/bla, mp9/dhfr, mp8/cat). The RNA/DNA hybrids were adsorbed to nitrocellulose in a 'minifold' filtration unit. The left part of the figure displays an autoradiogram of a nitrocellulose sheet in which RNA synthesized from promoters $P_{N25}$, $P_{J5}$, $P_{D/E20}$ and $P_{H207}$ integrated in pDS1/to1 was analysed in duplicate. The leftmost column shows the DNA probes used; mp9 and mp8 are control DNAs without any plasmid-specified sequence. After dissection of the nitrocellulose filter, the radioactivity of the individual probes was determined (right part of the figure). The difference in *dhfr*- and *cat*-specific RNA reflects the efficiency of terminator $t_0$ whereas the difference in *dhfr*- and *bla*-specific RNA reflects the relative efficiency of promoters. The variation in absolute counts of the *bla*-specific RNA is due to differences in plasmid copy number which changes significantly during cellular growth (Stüber and Bujard, 1982).

simultaneous presence of a strong promoter in the plasmid.

In summary, the relative strengths of 14 promoters were determined as exemplified for the four promoters in Figures 3 and 4 and the results are summarized in Table III. Wherever necessary, corrections for additional counter-clockwise transcription (into the bla region) and mRNA half-life time were performed. These promoter strengths correlate well with the abundance of *dhfr*-specific transcripts in the electrophoretic pattern of total *in vivo* RNA as depicted in Figure 2.

### *Comparison of rRNA synthesis with $P_{N25}$-directed transcription in vivo*

Up to 23% of the newly synthesized RNA in *E. coli* can be *dhfr*-specific if a strong promoter is carried by pDS2. It was therefore of interest to estimate the amount of rRNA synthesized under the same conditions and to compare the efficiency of our cloned promoters with promoters of the rRNA operon. In analyses as depicted in Figure 5, the mass ratio between rRNA and *dhfr*-specific RNA initiated by $P_{N25}$ was found to be 4:1. On a molar basis this corresponds to two *dhfr*-specific transcripts per rRNA transcript. At the generation time prevailing during our experiments (1.5 per h) an *E. coli* cell contains 2.5 genomes (Sarmientos and Cashel, 1983) and consequently 17 – 18 rRNA operons. Due to the copy number of pDS2 (Stüber and Bujard, 1982) the $P_{N25}$ – *dhfr* transcription unit is present ~80 times in such cells. Thus, ribosomal promoters initiate transcription under these conditions ~2.5 times more efficiently than $P_{N25}$.

**Table III.** The strength of cloned promoters *in vivo*

| Promoter | Relative strength | |
|---|---|---|
| | $P_{bla}$-units | $P_{lacUV5}$-units |
| $P_{H207}$ | 55 ± 4[a] | |
| $P_{D/E20}$ | 56 ± 8 | |
| $P_{N25}$ | 30 ± 5 | |
| $P_{G25}$ | 19 ± 2 | |
| $P_{J5}$ | 9 ± 1 | |
| $P_{A1}$ | 76 ± 9 | |
| $P_{A2}$ | 20 ± 4 | |
| $P_{A3}$ | 22 ± 3 | |
| $P_L$ | 37 ± 7[b] | 12 ± 0.6[b] |
| $P_{lac}$ | 5.7 ± 0.5 | 2.1 ± 0.2 |
| $P_{lacUV5}$ | 3.3 ± 0.3 | 1.4 ± 0.2 |
| $P_{tacI}$ | 17 ± 2 | 5.2 ± 0.4 |
| $P_{con}$ | 4 ± 0.2 | |
| $P_{bla}$ | 1 | |

[a]Value corrected for short half-life of RNA.
[b]$P_L$ was examined in *E. coli* C600 since it is suppressed in our DZ291-strain. The relative promoter strengths were calculated from hybridization experiments as exemplified in Figure 4. Depending on whether $P_{bla}$ or $P_{lacUV5}$ were used to initiate transcription of the *bla* region, promoter strengths were calculated in $P_{bla}$ or $P_{lacUV5}$ units, respectively.

## Discussion

The goal of this study was to accurately measure the *in vivo* strength of a group of well defined promoter sequences and to attempt an interpretation of sequence data based on functional information. A prerequisite for this goal was a reliable procedure for quantifying promoter strength. With the construction of plasmid pDS2 and three phage M13 derivatives (Figure 1) a system was established which fulfilled all requirements: (i) pro-

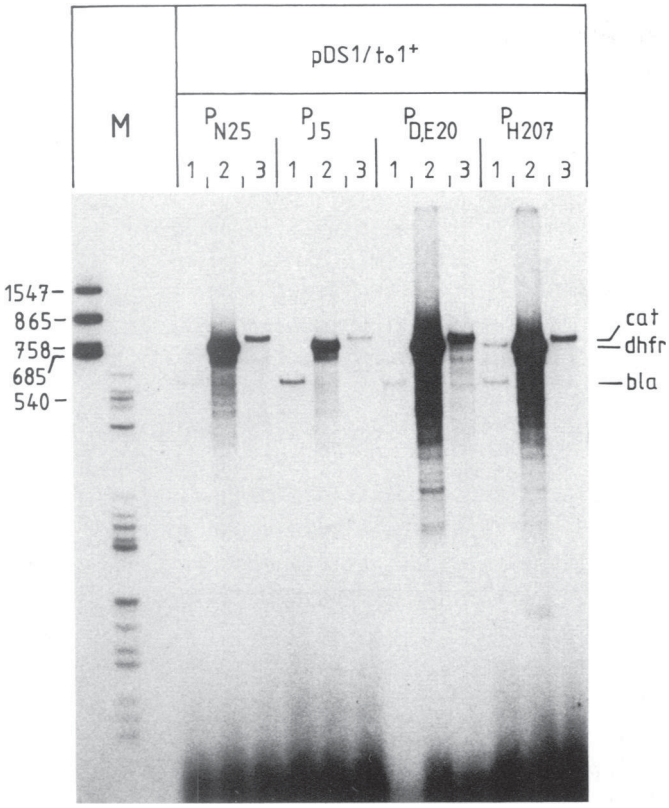


**Fig. 4.** Electrophoretic analysis of S1-resistant hybrids obtained with *in vivo* RNA. The RNA/DNA hybrids analysed are aliquots of the preparations shown in Figure 3. For each promoter ($P_{N25}$, $P_{J5}$, $P_{D/E20}$ and $P_{H207}$) the S1-resistant *bla*-, *dhfr*- and *cat*-specific RNA/DNA hybrids (**lanes 1, 2** and **3**, respectively) are seen in the autoradiogram of a polyacrylamide gel. The difference in the abundance between *bla*- and *dhfr*-specific hybrids reflects the difference in efficiency of these two promoters, whereas the effect of terminator $t_0$ can be seen by comparing *dhfr*-specific with *cat*-specific hybrids. With $P_{H207}$ two transcripts are observed in **lane 1**, of which the lower one represents RNA initiated from $P_{bla}$. The upper band around position 750 nucleotides is due to a promoter within the cloned fragment which initates transcription towards the *bla* region. The positions of *bla*-, *dhfr*- and *cat*-specific hybrids are indicated; M denotes size markers.

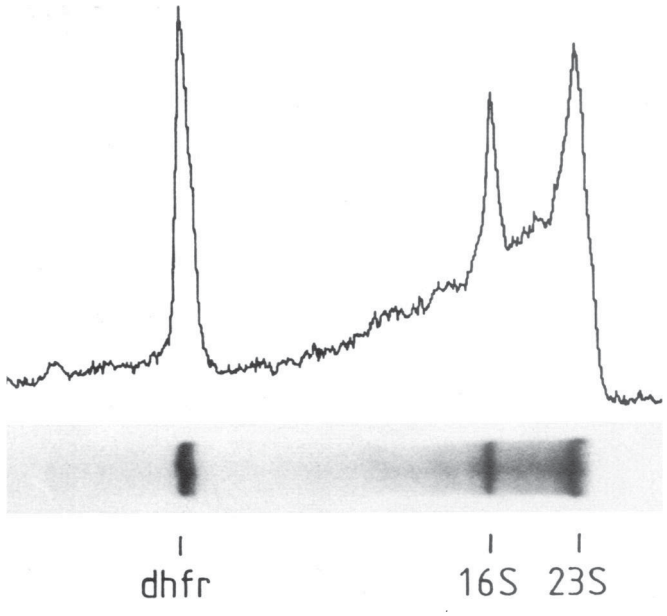


**Fig. 5.** Comparison of $P_{N25}$-directed RNA synthesis with the activity of rRNA operons. Cells harbouring a pDS2 derivative containing $P_{N25}$ and $t_0$ in site 1 were pulse labelled with [$^3$H]uridine for 1 min under standard conditions. The total RNA was extracted and separated by PAGE (4% polyacrylamide, 8 M urea). Microdensitometric evaluation of the autoradiogram revealed the ratio of rRNA to *dhfr*-specific RNA synthesized, which was 4.3:1 by mass.

moter activity is measured at the RNA level independent of any translational effects; (ii) an internal standard compensates for physiological variations such as gene dosage, etc; (iii) differences in RNA stability can easily be detected and properly corrected; (iv) the quantitation of RNA by hybridization in solution followed by adsorption to nitrocellulose is rapid and highly reproducible.

The data obtained allow for the first time arrangement of 14 promoters utilized by *E. coli* RNA polymerase into a hierarchy, based on their primary function namely, to commence RNA synthesis *in vivo*.

*Determination of promoter strength*

The accuracy and reproducibility of the hybridization procedure described here is demonstrated in Figure 3. For individual *in vivo* RNA preparations isolated from the same culture the standard deviations are <3%. Larger deviations may arise for several reasons: (i) the RNA preparation still contains DNA; (ii) *dhfr* transcripts with different 5′ and/or 3′ ends are compared; (iii) promoters are compared at different physiological states of the cell, or in different strains.

Following the protocol described, highly purified RNA is obtained yielding results as shown in Figure 3. The 3′ end of *dhfr* transcripts can be kept invariable by utilizing the same terminator at the same site (in our study $t_0$ at site 1). However, the potential influence of 5′ regions on mRNA half-life has to be examined, as demonstrated for $P_{H207}$ (Table II). Finally, we have observed that the activity of some promoters, for example $P_{J5}$, can vary significantly depending on the stage of growth and that variations of up to ±25% are found when different host cells are used. We have therefore limited our studies to two *E. coli* strains (C 600 and DZ 291) and to logarithmically growing cultures at $OD_{600} = 0.6$. By observing the above precautions relative promoter strength can be determined within one host strain to an accuracy of ±15% (Table III).

*The hierarchy of promoter strength*

The 14 promoters examined can be arranged in a hierarchy according to their activities *in vivo* under defined conditions (Table III). Using $P_{bla}$ as standard, the promoter strengths span almost two orders of magnitude. The strongest promoter identified is $P_{A1}$ which is ~76 times $P_{bla}$. It is followed by a group of phage promoters ($P_{D/E20}$, $P_{H207}$, $P_L$, $P_{N25}$) whose activity *in vivo* is between 30 and 60 units. At the lower end of the hierarchy one finds the fully induced $P_{lac}$, $P_{lacUV5}$ as well as $P_{con}$. $P_{tacI}$ takes an intermediate position; if compared with $P_{lacUV5}$, one of its parent sequences, it is five times more efficient.

Is there a way to fit this hierarchy into an absolute scale of promoter strength? In fast growing cells the operons producing rRNA belong to the most intensely transcribed regions. The rate of chain initiation per second is estimated to be ~1. This is close to the maximum which can be expected if a rate of chain elongation of 50 nucleotides/s and a space requirement of 50 bp per RNA polymerase molecule is assumed. As shown for the *rrnB* operon it is a single promoter, namely $P_1$, which is capable of directing such intense transcription (Sarmientos and Cashel, 1983). Our estimates indicate (Figure 5) that $P_{N25}$ can reach 40% of the activity of a fully induced rRNA promoter. This suggests that promoters like $P_{A1}$, $P_{H207}$ and $P_{D/E20}$ can initiate RNA synthesis with nearly maximal rates. We are presently examining

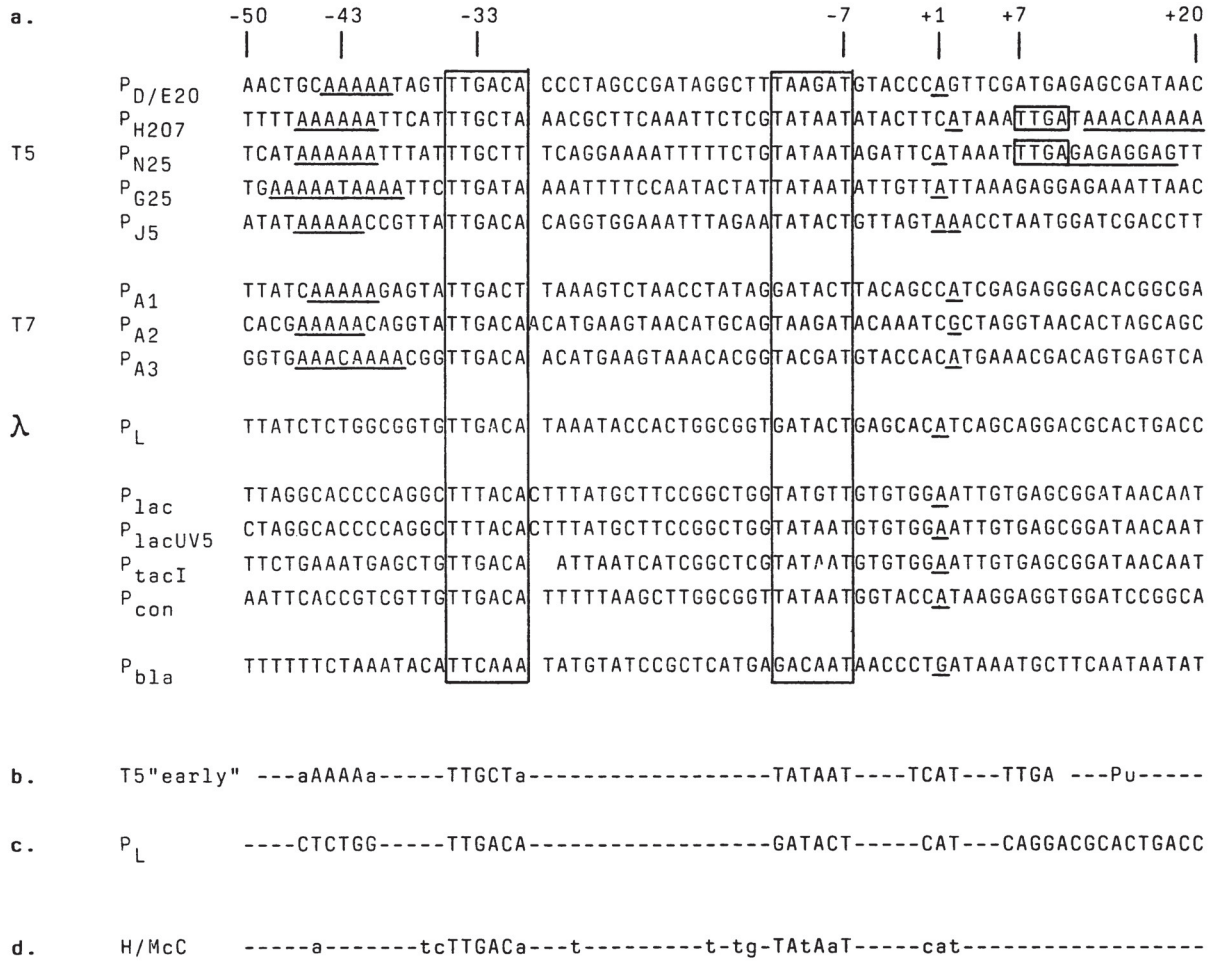


**Fig. 6.** Promoter sequences. **(a)** The sequences of the 14 promoters studied are aligned with respect to the first T of the −33 hexamer and the last T of the −10 hexamer. The conserved regions around −33 and −10 are boxed, the A clusters around −43 and the first nucleotide transcribed (+1) are underlined. The motifs typical for the downstream region of 'early' T5 promoters, TTGA around +7 and the run of purines are indicated within the sequences of $P_{H207}$ and $P_{N25}$. **(b,c)** Sequence elements of two prototype promoters. The features characterizing an 'early' T5 promoter were derived from six such signals (Gentz and Bujard, 1985). In contrast, corresponding sequence elements of $P_L$, a promoter comparable in strength to an 'early' T5 promoter, is shown in **c**. Whereas the AT content is ~80% for a typical 'early' T5 promoter, it is only 47% for $P_L$. **(d)** Essential features of a consensus sequence derived from ~150 different promoters (Hawly and McClure, 1983).

whether this upper limit can indeed be reached by a single promoter sequence like $P_{A1}$ or whether tandem arrangements of promoters, as found within the early region of phage T7 and T5, are required. In any case our data strongly suggest that some of the promoters studied here encode a program which is close to the functional optimum of a single sequence.

*Structural considerations*

The information available now allows us to examine our collection of promoter sequences with respect to functional properties. If we define a strong promoter as a sequence having at least 20 times the activity of $P_{bla}$ *in vivo*, then, with the exception of $P_{J5}$ all phage promoters of this collection are strong promoters (Table III). None of these strong promoters contains simultaneously both of the so-called 'canonical' hexamers around position −10 (TATAAT) and −33 (TTGACA), respectively (Figure 6). All phage promoters — with the exception of $P_L$ — are rich in AT and contain striking blocks of AT pairs around position −43. Promoters $P_{H207}$ and $P_{N25}$ which control 'early' genes of phage T5 show, in addition, a structure typical for this class of signals (Gentz and Bujard, 1985) namely the tetrameric sequence TTGA around +7 and a stretch of purines spanning from around +9 to +20. We have derived a consensus sequence from six 'early' T5 promoters, and the most prominent sequence elements are given in Figure 6. There are significant differences between a typical 'early' T5 promoter and a consensus sequence, as described by Hawley and McClure (1983). The most striking features are the homologies outside of the 'classical' (position +1 and −35, Figure 6) promoter region which strongly suggests contributions of these regions to promoter function. In the accompanying paper (Kammerer *et al.*, 1986) we present evidence that indeed information encoded in one of these regions can contribute significantly to promoter activity.

Most interesting is the comparison between $P_L$ from phage lambda and a typical 'early' promoter from phage T5 like $P_{H207}$. *In vivo* both promoters belong to the most efficient signals with very similar overall strength (Table III). Nevertheless, they differ clearly in their structure. (i) The AT-content (between −50 and +20) is 81% for $P_{H207}$ but only 47% for $P_L$. (ii) The most conserved hexameric sequence around −10 is 'consensus'-like for $P_{H207}$ but shows two deviations in $P_L$. Similarly, $P_{H207}$ differs from the 'consensus' around −33 exhibiting a sequence frequently found with strong T5 promoters (TTGCT) whereas $P_L$ is 'consensus'-like in this region. (iii) Around −43 $P_{H207}$, like many strong promoters, has a block of AT pairs; $P_L$, however, is GC rich (Figure 6). (iv) There are no similarities between the

two promoters downstream of +1.

Furthermore, $P_L$ and a typical 'early' T5 promoter like $P_{N25}$ differ drastically in their rate of complex formation with *E. coli* RNA polymerase as well as in their *in vitro* strength if compared under competitive conditions and the same holds when $P_{A1}$ is compared with $P_{N25}$ (von Gabain and Bujard, 1979). In both cases $P_{N25}$ readily out-competes $P_L$ and $P_{A1}$.

Complex signals like promoters encode a program for a multi-step process. Such a process can be limited at various levels (Kammerer *et al.*, 1986) and sequences optimized in different ways can result in signals with identical overall properties, i.e. promoter strength. Therefore, unlike, for example operator/repressor systems where a protein has merely to select and to bind a specific sequence, a unique consensus sequence describing the functional programs of promoters should not exist. The promoters $P_{H207}$, $P_{A1}$ and $P_L$ resemble each other closely in their overall function *in vivo*, i.e. they commence RNA synthesis with close to maximal rates. Nevertheless, they differ clearly in their sequences and in individual parameters of promoter function. Together with the results reported in the accompanying publication (Kammerer *et al.*, 1986) this supports the hypothesis proposed earlier (Bujard, 1980), namely that optimal function of complex signals like promoters can be encoded in alternate but equivalent sequences.

## Materials and methods

*Nucleic acids*

M13mp8 and M13mp9 (Messing and Vieira, 1982) were obtained from R.Cortese, EMBL Heidelberg. The pDS1 cloning system and its nomenclature has been described in detail previously (Stüber and Bujard, 1982). An earlier version of M13mp9 contained a portion of pBR322 DNA (position 2348–2718, Sutcliffe, 1979). This sequence was removed since it increased the background in hybridization experiments. Phage M13 derivatives were grown in *E. coli* 71-18 cells and isolated as described by Herrmann *et al.* (1980). Phage DNA was extracted by hot phenol (65°C) followed by chloroform/isoamylalcohol (24:1) extraction before it was ethanol precipitated and re-dissolved in hybridization buffer. Plasmids as well as RF-DNA from M13 phages were prepared as described before (Stüber and Bujard, 1982), or by alkaline extraction (Birnboim and Doly, 1979). Various DNA sequences were cloned using standard procedures (Maniatis *et al.*, 1982). *E. coli* was transformed with DNA according to Morrison (1979). All DNA constructs were verified by sequence analysis.

*Promoters.* The identification and characterization of promoter sequences of phage T5 ($P_{H207}$, $P_{D/E20}$, $P_{N25}$, $P_{G25}$, $P_{J5}$) has been reported (Gentz and Bujard, 1985). The major 'early' promoters of phage T7 ($P_{A1}$, $P_{A2}$, $P_{A3}$) were cloned as fragments delineated within the T7 genome by the nucleotide positions 258–543, 547–651 and 657–763, respectively (Dunn and Studier, 1983). Promoter $P_L$ from phage lambda was isolated as an *Hae*III fragment from plasmid pPLC28 (Remault *et al.*, 1981). $P_{tacI}$ was transferred from plasmid ptac 12 (Amman *et al.*, 1983) to the pDS1 system as a *Pst*I/*Xho*I fragment after insertion of *Xho*I linkers at the *Pvu*II cleavage site downstream of the promoter. $P_{lac}$ was recovered from pBU10 (Gentz *et al.*, 1981) as a 186-bp *Pvu*II/*Hae*III fragment. A 91-bp *Eco*RI/*Bam*HI fragment isolated from pEX110 (Weiher, 1981) carried $P_{lacUV5}$ and a 61-bp fragment obtained from V.Korobko contained a synthetically prepared 'consensus' promoter sequence (Dobrynin *et al.*, 1980) henceforth called $P_{con}$.

*Terminators.* Terminator $t_0$ from phage lambda was described before (Stüber and Bujard, 1982). T1 was isolated from the rrnB operon (Brosius *et al.*, 1981). The analysis of several terminators using the pDS system will be described elsewhere.

*Preparation of in vivo [$^3$H]RNA*

*E. coli* cells (C600 or DZ291), harbouring the proper plasmids were kept in 20% glycerol at −20°C. Cultures (10 ml Luria broth containing 100 μg/ml ampicillin) were inoculated with these stocks and grown overnight. Aliquots of 0.1 ml were then diluted into 25 ml of pre-warmed M9 medium containing 5% casamino acids, 10% Luria broth (Miller, 1972) as well as 100 μg/ml ampicillin. For experiments with promoters controlled by the *lac* repressor ($P_{lac}$, $P_{lacUV5}$ and $P_{tacI}$) glucose was replaced by glycerol and induction was achieved by IPTG (200 μg/ml). Plasmids carrying phage lambda promoter $P_L$ were transformed into a non-lysogenic C600 strain.

Cells were grown at 37°C under good aeration and at $OD_{600} = 0.6$ the cultures were divided: one aliquot of 8 ml was saved for the analysis of the plasmid and of the protein pattern; to the second aliquot of 10 ml 500 μCi of [5,6-$^3$H]uridine (Amersham, 40–60 mCi/mmol, 1 mCi/ml aqueous solution) were added and the incubation was continued for 0.5–5 min. The cultures were quickly chilled to 0°C (liquid nitrogen) and cells were collected by centrifugation and re-suspended in TES buffer (20 mM Tris–HCl, pH 8, 10 mM EDTA, 100 mM NaCl, 1% SDS). After heating to 95°C for 3 min the lysed cells were subjected to a CsCl gradient centrifugation (nitrocellulose-filtered CsCl solutions; swing out rotor; 150 000 *g* for 16 h at 20°C; Glisin *et al.*, 1974). After removal of the supernatant the 0.8 cm portion of the inverted tube containing the pellet was cut off with a heated scalpel. The RNA was then dissolved in 2 × 80 μl of TE buffer containing 0.2% SDS and was precipitated with ethanol in the presence of 0.3 M sodium acetate. The precipitates were washed (80% ethanol), dried *in vacuo* and finally dissolved in hybridization buffer. The yields were between 200 and 300 μg RNA per 10 ml culture and sp. act. of $1-3 \times 10^5$ c.p.m./μg of RNA were obtained. Electrophoretic analysis was carried out as described below.

*Hybridization of RNA to excess ssDNA*

All hybridizations were carried out at 42°C in hybridization buffer (50% formamide, 300 mM NaCl, 20 mM Tris–HCl, pH 8, 0.5 mM EDTA) in a volume of 20 μl for 2 h. In a typical experiment 10 μl of *in vivo* [$^3$H]RNA ($\sim 5 \times 10^5$ c.p.m. of labelled cellular RNA) was mixed with 10 μl of ssDNA (0.2 pmol/10 μl of M13mp9*bla*, M13mp9*dhfr*, M13mp8*cat*, M13mp8 or M13mp9), heated to 65°C for 3 min and then kept at 42°C for 2 h. These preparations were either treated with S1 nuclease or adsorbed to nitrocellulose.

*Analysis of RNA/DNA hybrids by S1 nuclease*

200 μl of S1 buffer (50 mM NaOAc, 1 mM $ZnSO_4$, 50 mM NaCl) containing 10 U S1 nuclease were added to a hybridization assay of 20 μl and the reaction mixture was incubated at 37°C for 1 h before it was extracted by an equal volume of chloroform/isoamylalcohol (24:1). The hybrids were precipitated with ethanol, dried *in vacuo* and re-dissolved in 20 μl of 20 mM Tris–HCl, pH 8, containing 50 mM NaCl and subjected to PAGE (6% polyacrylamide, 40 mM Tris–HCl, 20 mM NaOAc, 2 mM EDTA, pH 7, 8, 3 mA/cm).

*Quantitation of hybridized RNA*

Hybridization assays were diluted 10-fold with 2 × SSC and filtered through nitrocellulose (0.45 μm BA 85 filters, 'minifold'-system, Schleicher and Schüll). It is important to note that the capacity of nitrocellulose filters for ssDNA is limited. For the filters used it is ~6 pmol of M13mp8 DNA per 1 cm$^2$. The filters were washed with 2 ml of 2 × SSC, baked at 80°C for 30 min *in vacuo* and then incubated in 100 ml of 2 × SSC containing 50 μg of RNase A/ml at 42°C for 1 h. After washing (three times with 100 ml 2 × SSC, 42°C, 10 min) and drying, the radioactivity retained on the filters was monitored in universal liquid scintillator (NEN). The ratios of *dhfr*- or *cat*-specific RNAs to *bla*-specific transcripts were calculated taking in account the number of uridines within the different RNAs. The inserts of the different ssDNA probes code for 148, 169 and 215 uridines for the *bla*, the *dhfr* and the *cat* region, respectively. The strength S of a promoter in $P_{bla}$-units is then obtained by calculating $S = 0.87 \times \text{c.p.m.}_{dhfr}/\text{c.p.m.}_{bla}$.

## Acknowledgements

We would like to thank Dr D.Stüber for sequences containing terminators and for various pDS1 derivatives, and M.Brunner and R.Knaus for stimulating discussions. This work was supported in part by grants Bu 338/12-15 from the Deutsche Forschungsgemeinschaft and by the Fonds der Chemischen Industrie Deutschlands.

## References

Amann,E., Brosius,J. and Ptashne,M. (1983) *Gene*, **25**, 167–178.
Belasco,J.G., Nilsson,G., von Gabein,A. and Cohen,S.N. (1986) *Cell*, **46**, 245–251.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1517.
Brosius,J., Dull,T.J., Sleeter,D.D. and Noller,H.F. (1981) *J. Mol. Biol.*, **148**, 107–127.
Bujard,H. (1980) *Trends Biochem. Sci.*, **5**, 274–278.
Bujard,H., Stüber,D., Gentz,R., Deuschle,U. and Peschke,U. (1985) In Helinski,D.R., Cohen,S.N., Clewell,D.B., Jackson,D.A. and Hollaender,A. (eds), *Plasmids in Bacteria*. Plenum Press, NY, pp. 45–52.
Dobrynin,V.N., Korobko,V.G., Severtsova,J.V., Bystrow,N.S., Chuvpilo,S.A. and Kolosov,M.N. (1980) *Nucleic Acids Res. Symp. Ser.*, **7**, 365–376.
Dunn,J.J. and Studier,F.W. (1983) *J. Mol. Biol.*, **166**, 477–535.
Gentz,R. and Bujard,H. (1985) *J. Bacteriol.*, **164**, 70–77.
Gentz,R., Langner,A., Chang,A.C.Y., Cohen,S.N. and Bujard,H. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4936–4940.
Glisin,V.,Crkvenjakov,R. and Buys,G. (1974) *Biochemistry*, **13**, 2633–2637.
Hawley,D.K. and McClure,W.R. (1983) *Nucleic Acids Res.*, **11**, 2237–2255.
Herrmann,R., Neugebauer,K., Pirkv,E., Zentgraf,H. and Schaller,H. (1980) *Mol.*

*Gen. Genet.*, **177**, 231–242.
Kammerer,W. (1986) Thesis, University of Heidelberg, FRG.
Kammerer,W., Deuschle,U., Gentz,R. and Bujard,H. (1986) *EMBO J.*, **5**, 2995–3000.
Mangiarotti,G. and Schlessinger,D. (1967) *J. Mol. Biol.*, **29**, 395–428.
Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning. A Laboratory Manual.* Cold Spring Harbor Laboratory Press, NY.
Messing,J. and Vieira,J. (1982) *Gene*, **19**, 269–276.
Miller,J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor Laboratory Press, NY.
Morrison,D.A. (1979) *Methods Enzymol.*, **68**, 326–331.
Remault,E., Stanssens,P. and Fiers,W. (1981) *Gene*, **15**, 81–93.
Rosenberg,M. and Court,D. (1979) *Ann. Rev. Genet.*, **13**, 319–353.
Sarmientos,P. and Cashel,M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 7010–7013.
Stüber,D. and Bujard,H. (1982) *EMBO J.*, **1**, 1399–1404.
Sutcliffe,J.G. (1979) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 77–90.
von Gabain,A. and Bujard,H. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 189–193.
Weiher,H. (1981) Thesis, University of Heidelberg, FRG.
Yamamoto,T. and Immamoto,F. (1975) *J. Mol. Biol.*, **92**, 289–304.

*Received on 7 August 1986*