The EMBO Journal vol.4 no.2 pp.519–526, 1985

Translational initiation frequency of *atp* genes from *Escherichia coli*: identification of an intercistronic sequence that enhances translation

John E.G.McCarthy, Hans U.Schairer and Walter Sebald

Department of Cytogenetics, GBF – Gesellschaft für Biotechnologische Forschung mbH, Mascheroder Weg 1, D-3300 Braunschweig, FRG

Communicated by K. von Meyenburg

The c, b and δ subunit genes of the *Escherichia coli atp* operon were cloned individually in an expression vector between the *tac* fusion promoter and the *galK* gene. The relative rates of subunit synthesis directed by the cloned genes were similar *in vitro* and *in vivo* and compared favourably with the subunit stoichiometry of the assembled proton-translocating ATP synthase of *E. coli in vivo*. The rate of synthesis of subunit c was at least six times that of subunit b and 18 times that of subunit δ. Progressive shortening of the long intercistronic sequence lying upstream of the subunit c gene showed that maximal expression of this gene is dependent upon the presence of a sequence stretching >20 bp upstream of the Shine-Dalgarno site. This sequence thus acts to enhance the rate of translational initiation. The possibility that similar sequences might perform the same function in other operons of *E. coli* and bacteriophage λ is also discussed. Translation of the subunit b cistron is partially coupled to translation of the preceding subunit c cistron. In conclusion, the expression of all the *atp* operon genes could be adjusted to accommodate the subunit requirements of ATP synthase assembly primarily by means of mechanisms which control the efficiency of translational initiation and re-initiation at the respective cistron start codons.

Key words: E. coli atp operon/subunit stoichiometry/*in vitro* and *in vivo* expression/translational initiation

Introduction

The proton-translocating ATP-synthase (H^+-ATPase) of *Escherichia coli* is composed of eight different types of subunit, the complete genetic information for which lies in a single operon (designated *unc* or *atp*; see Figure 1) at ~83 minutes on the *E. coli* linkage map (Downie *et al.*, 1979; Friedl *et al.*, 1979; Futai and Kanazawa, 1983; Gay and Walker, 1981a, 1981b; Saraste *et al.*, 1981; Nielsen *et al.*, 1981). The F_1 part of the enzyme, which alone has ATP hydrolase activity, comprises five types of subunit (α, β, γ, δ and ε), whereas the membrane-integrated component F_0 comprises three types of subunit (a, b and c; see, e.g., Hoppe and Sebald, 1984). DNA sequencing (Gay and Walker, 1981b) has revealed the existence of a ninth, apparently non-essential (von Meyenburg *et al.*, 1982a) gene (designated gene 1 or *atp*I) at the beginning of the *atp* operon (see also Brusilow *et al.*, 1983).

A particularly remarkable aspect of not only the structure of the *E. coli* H^+-ATPase, but also of the mitochondrial and thylakoid H^+-ATPases, is the stoichiometry of the component subunits. The relative molar quantities of the *E. coli* subunits estimated on the basis of radioactivity incorporation studies are $\alpha_3\beta_3\gamma_1\delta_1\epsilon_1a_1b_2c_{10-15}$ (Foster and Fillingame, 1982; von Meyenburg *et al.*, 1982b; see also Lünsdorf *et al.*, 1984). How is this stoichiometry achieved in the synthesis and assembly of the *E. coli* H^+-ATPase, or indeed of other H^+-ATPases?

Several lines of evidence, including the results of S1 nuclease mapping, indicate that the *atp* operon has a single major promoter which initiates transcription 73 bp upstream from the gene 1 (*atp*I) reading frame (von Meyenburg *et al.*, 1982a; Porter *et al.*, 1983; Jones *et al.*, 1983; Nielsen *et al.*, 1984; Kanazawa *et al.*, 1981). Thus the *atp* operon is transcribed to produce a single, large mRNA species containing all of the structural gene reading frames. One hypothesis to explain the attainment of the subunit stoichiometry of the *E. coli* H^+-ATPase has accordingly been that the rates of synthesis in the respective subunits are regulated at the level of polypeptide chain elongation (see, e.g., Futai and Kanazawa, 1983). On the other hand, Brusilow *et al.* (1982) proposed that mRNA secondary structure might play an important role in suppressing translational initiation at certain sites (corresponding to the starts of proteins b, δ and γ).

The present work demonstrates that differences in the efficiency of translational initiation at different sites in the *E. coli atp* operon polycistronic message play a major role in determining the relative amounts of the different subunits that are synthesized. In particular, a specific mechanism of enhancement of translational initiation in the case of the subunit c cistron is shown to be important for the attainment of the required ratio of synthesized subunits. A preliminary report of some of the results reported in this paper has already been given (McCarthy *et al.*, 1984).

Results

Cloning and expression of the genes for H^+-ATPase subunits b, c and δ

The genes coding for subunits c, b and δ from the *atp* operon

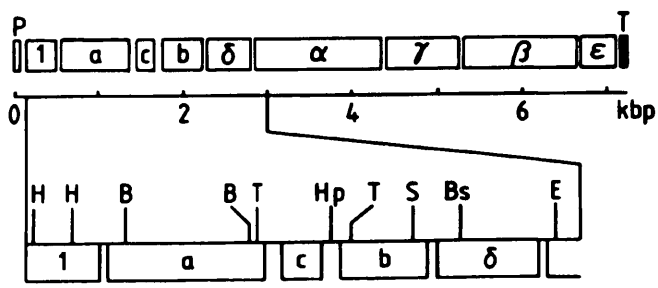

Fig. 1. The structure and a partial restriction map of the *atp* operon. The diagram shows the nine genes of the operon together with the major promoter (P) and terminator (T). Restriction endonuclease sites which were useful for the construction of the described plasmids are indicated. B, *Bam*HI; Bs, *Bst*EII; H, *Hind*III; Hp, *Hpa*I; S, *Sal*I; T, *Taq*I. Data obtained from sources given in text.

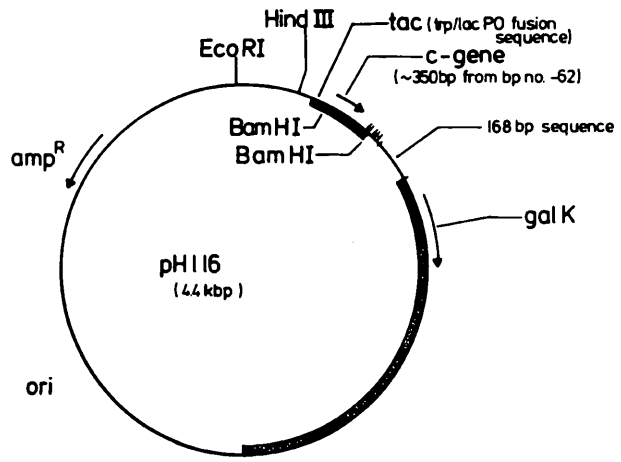

Fig. 2. The structure of a subunit c gene expression plasmid. A typical plasmid is shown (see lane 7, Figure 3). pH116(c7) was constructed by inserting a fragment of the *atp* operon bearing the subunit c gene into the *Bam*HI site of pDR540. The long arrows indicate the direction of the common transcript. The small arrows indicate the sites of five translation stop codons.

(*atp*E, *atp*F and *atp*H, respectively) were cloned using the plasmid pDR540 (Russell and Bennett, 1982; see Materials and methods). Fragments of the operon were inserted into the *Bam*HI site of the vector so that transcription of either the individual genes of the b, c and δ subunits, or the b and c subunit genes together, was initiated by the *tac* promoter and transcription continued through the *gal*K gene. A diagram of the structure of one of the recombinant plasmids bearing just the c subunit gene is shown in Figure 2. Expression of the *gal*K gene (see, e.g., McKenney *et al.*, 1981) served as a useful reference indicator of the activity of the *tac* promoter. The ^{35}S-labelled products of *in vitro* protein synthesis directed by each of the plasmids were readily identified (see Materials and methods and Figure 3). These proteins ran at precisely their expected positions on the gels. In the fluorograph depicted in Figure 3 the bands corresponding to β-lactamase and galactokinase, whose synthesis is directed by pDR540 genes, and to subunits c, b and δ, are discernible. A major objective of the present work was to quantitate the *in vitro* expression of each of the genes coding for the c, b and δ subunits; this could be most reliably achieved by determining the radioactivity of bands excised from dried gels (see Materials and methods). Pulse labelling experiments performed with whole cells (of the *atp*$^+$ strain MCG1; see Materials and methods) containing the described plasmids (Figure 4) showed that the characteristics of *in vivo* expression resembled closely the results obtained *in vitro*, whilst being overall more difficult to quantitate accurately.

Shortening the intercistronic sequence upstream of the subunit c gene reduces the synthesis of subunit c

The synthesis of the products shown in lanes 2 – 7 of Figure 3 was directed by a set of plasmids bearing the subunit c structural gene together with varying lengths of the intercistronic sequence upstream of it. The 3′ ends of the cloned DNA fragments in these plasmids were all at the *Hpa*I site between the genes encoding subunits b and c (see Figure 1); the 5′ ends are defined precisely in Figure 3. Expression of the subunit c gene *in vitro* and *in vivo* was drastically reduced when the intercistronic sequence lying upstream of the Shine-Dalgarno sequence was progressively shortened (compare Figures 3 and 4 and Table I). The expression of the subunit c gene (in Table I normalized to the expression of *gal*K) was equally high with all the plasmids bearing the complete intercistronic sequence upstream of this gene [pH184(c6), pH116(c7), pB10501(cb8) and pB10505(cb9); Figure 3; compare also Figure 4]. The ratio of subunit c:galactokinase obtained *in vitro* at a given DNA concentration with pH159(c5) and pH106(c4) was respectively 20% and 3% of that determined with pH184(c6) (Table I). The effects of shortening the intercistronic sequence upstream of the subunit c gene upon the synthesis of subunit c *in vitro* could be quantitatively confirmed *in vivo* (see, e.g., Figure 4) within the limits of accuracy of the methodology applied (see Materials and methods). The synthesis of subunit c directed by the two plasmids with the shortest inserts [pH152(c2) and pH163(c3), Figure 3] was barely distinguishable from background (and not at all discernible in Figure 3), but it was shown in both cases by complementation (see Materials and methods) that functional subunit c could be produced *in vivo*.

Relatively the same effects of shortening the intercistronic sequence upstream of the subunit c gene upon the expression of this gene *in vitro* and *in vivo* were observed when the cloned fragment of plasmids pH116(c7), pH159(c5) and pH163(c3) were inserted behind the λP_L promoter in pJLf101 (see Materials and methods). The products synthesized *in vitro* under the direction of one of these recombinant plasmids are shown in Figure 3 (lane 13).

Comparison of the expression of the different plasmid-borne genes

The relative amounts of synthesis of subunits c, b and δ and of galactokinase directed *in vitro* by the constructed plasmids were measured over a wide range of plasmid DNA concentrations. Thus in Figure 5A the relative amounts of subunits c and b and of galactokinase separated by SDS-polyacrylamide electrophoresis have been plotted as a function of the concentration of pB10501(cb8). pB10501(cb8) bears the genes of both subunits c and b together with the complete intercistronic sequence upstream of the subunit c gene and 19 bp of the subunit δ gene. The background radioactivity in the gels run with *in vitro* protein synthesis products was more significant in the case of subunit b than of subunit δ because subunit b ran close to an undefined polypeptide of the *in vitro* system. The background radioactivity in the regions of subunit c and galactokinase was negligible in relation to the incorporation into these proteins. Correcting for background radioactivity as described in Materials and methods and also for the methionine content of the respective subunits allowed the calculation of a synthesis ratio of moles subunit c: moles subunit b of between 6 and 10. The lowest ratios were generally obtained at the lower concentration of DNA, a maximum was reached at a DNA concentration in the range 30 – 70 μg/ml, and a value of ~8 was calculated at the very highest DNA concentrations (see Figure 5A and Table I).

The expression of pB10506(cb10), which is identical to pB10501(cb8) except that it lacks most of the intercistronic sequence upstream of the subunit c gene, showed different characteristics. The synthesis of subunit c directed by this plasmid was extremely low [i.e., equivalent to that obtained with pH163(c3)] compared with the expression obtained with pB10501(cb8) (compare, e.g., lanes 8 and 10, Figure 3), and

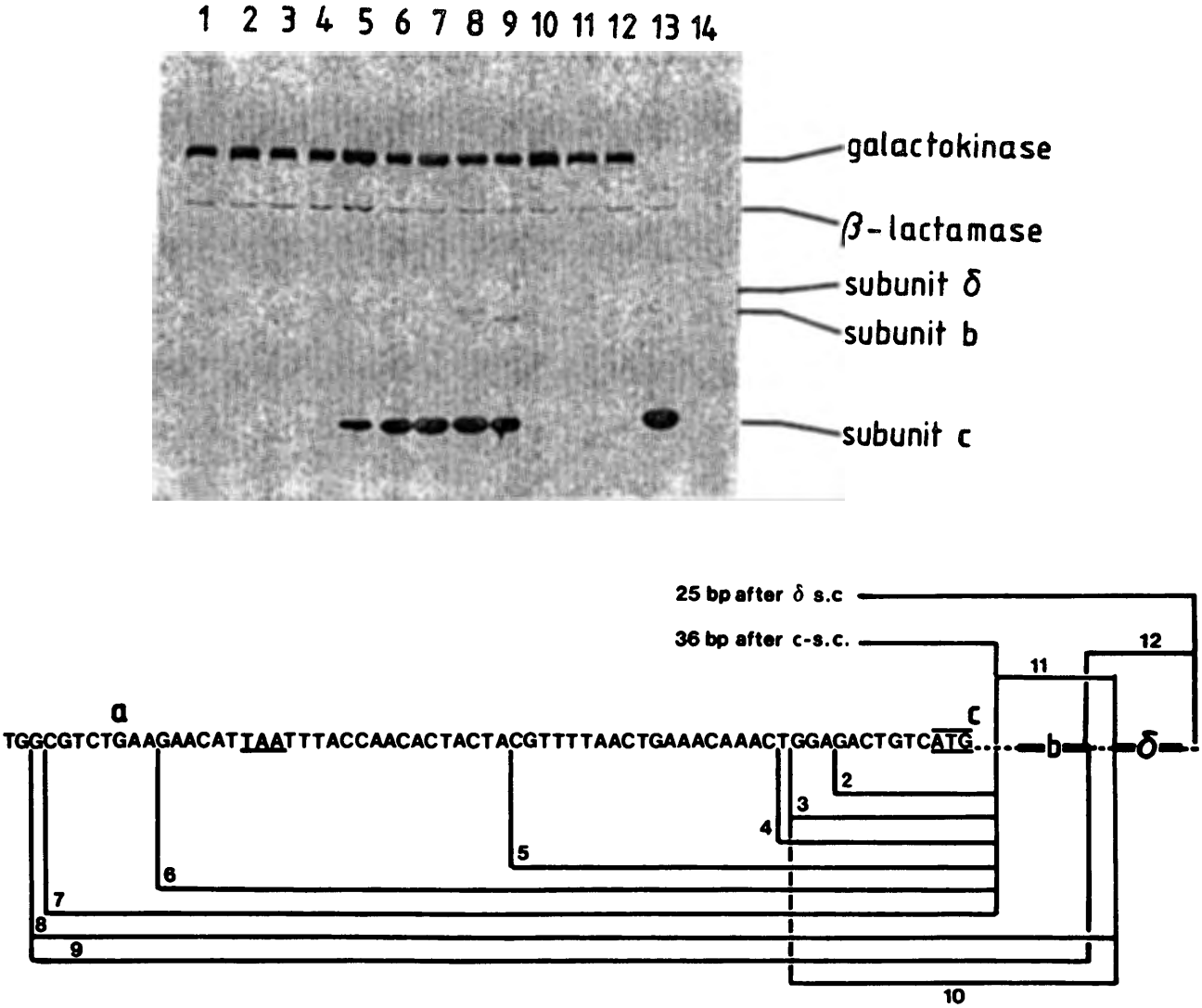

Fig. 3. *In vitro* expression of the genes of subunits c, b and δ from the *atp* operon. Fragments from the *atp* operon were cloned in the *Bam*HI site of pDR540 (see Figure 2) or the *Sal*I site of pJLf101 (see Materials and methods). The [^{35}S]L-methionine-labelled proteins synthesized in the S30 system under the direction of these plasmids were separated on a 13.5% polyacrylamide gel which was subsequently dried and fluorographed. The plasmids used (all at 56 μg/ml) were pDR540 **(lane 1)**; subunit c (*atp*E) plasmids pH152 **(lane 2)**, pH163 **(lane 3)**, pH106 **(lane 4)**, pH159 **(lane 5)**, pH184 **(lane 6)**, pH116 **(lane 7)**, pJLf116 **(lane 13)**; subunits c−b (*atp*E and *atp*F) plasmids pB10501, pB10505 and pB10506 **(lanes 8−10**, respectively); subunit b plasmid pHB6 **(lane 11)** ; subunit δ plasmid pSE1 **(lane 12)**. In the text the cloned genes and corresponding lane numbers of this figure are indicated in parentheses after the plasmid code: e.g., pB10501(cb8). A plasmid-free water blank experiment was also performed **(lane 14)**. In each case 4% of the total counts were loaded onto the gel (see Materials and methods). Also shown in the lower part of the figure is the sequence of the end of the subunit a gene followed by the whole intercistronic sequence up to and including the start codon of the subunit c gene. The lines indicate the fragments of the operon cloned into pDR540 and are numbered according to the lane order of the gel. pJLf116 contains the same insert as pH116. s.c. = stop codon.

was thus consistent with the data presented in the previous section. Moreover, the subunit b gene was also relatively more weakly expressed by pB10506(cb10). Indeed, an equally weak expression of the subunit b gene to that obtained with pB10506(cb10) was observed with a plasmid [pHB6(b11)] which bears only the subunit b gene. The fragment cloned in pHB6(b11) was derived from pB10501(cb8) and begins 22 bp upstream of the subunit b gene start codon. Comparison of the radioactive incorporation ratios of subunit b:galactokinase for pHB6(b11) and pB10506(cb10) on the one hand, and pB10501(cb8) on the other (Table I), reveals a large difference, although this is partially attributable to the apparent competitive inhibitory effect of subunit c synthesis on *gal*K expression. The number of pmol of subunit b synthesis directed per μg plasmid DNA was approximately two times less for pHB6(b11) than for pB10501(cb8) over the DNA concentration range 5 − 200 μg/ml. A similar relationship for the expression of the subunit b gene as carried by these two plasmids was observed *in vivo* (Figure 4).

The *in vitro* synthesis products of a further b and c subunit plasmid [pB10505(cb9)] are shown in Figure 3. It is remarkable that although the cloned fragment of pB10505(cb9) had lost the last 12 bp of its subunit b gene sequence, pB10505(cb9) still complemented *atp* b$^-$ mutants (see

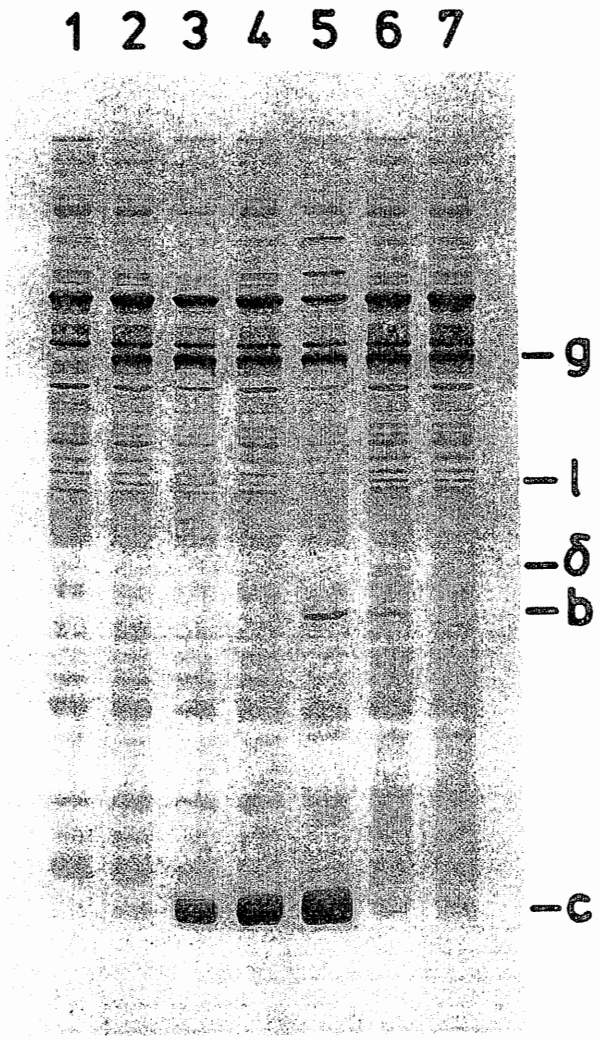

Fig. 4. *In vivo* pulse-labelling of proteins encoded by the described recombinant plasmids. Pulse-labelling was performed as described in Materials and methods using strains carrying recombinant plasmids. The host strain was MCG1, which is *atp*$^+$. A fluorograph was prepared from a 15% SDS-polyacrylamide gel which had been loaded with the sonicated and solubilized cells after pulse-labelling. Experiments were performed with MCG1 itself **(lane 1)**, or with MCG1 transformed with pH163 **(lane 2)**, pH159 **(lane 3)**, pH116 **(lane 4)**, pB10501 **(lane 5)**, pHB6 **(lane 6)** or pSE1 **(lane 7)** (compare Figure 3). The letters on the right hand side indicate the positions of galactokinase (g), β-lactamase (l), the δ subunit, b subunit and c subunit, respectively.

Materials and methods). It is unclear why the loss of the last 12 bp of the subunit b gene results in the synthesis of a polypeptide with such a noticeably increased mobility on SDS-polyacrylamide gels.

The cloned fragment of pSE1(δ12) stretches from 84 bp upstream of the subunit δ gene start codon to 25 bp after this gene's stop codon. A low incorporated radioactivity ratio (subunit δ:galactokinase from 0.091 to 0.063) was maintained over a wide range of plasmid DNA concentrations (Figure 5B and Table I). Expression of the subunit δ gene in pSE1(δ12) or of the subunit b gene in pHB6(b11) exerted no measurable effect upon expression of *galK* (relative to pDR540). Thus the ratios of subunit δ:galactokinase and subunit b:galactokinase obtained over a wide range of DNA concentrations with these two plasmids indicate, as do the specific *in vitro* yields of subunit per μg of each plasmid, that the expression efficiency is 1.5−2 times greater for the subunit b gene in pHB6(b11) than for the subunit δ gene in pSE1(δ12). Synthesis of the δ subunit directed by pSE1 (δ12) *in vivo* was so low in relation to the synthesis of the c and b subunits (and to background) as to be barely measurable (Figure 4).

Table I. Comparison of the synthesis of *E. coli* H$^+$-ATPase subunits and galactokinase in the *in vitro* system

Gene products analyzed	Plasmid[a]	Ratio of [^{35}S]-L-methionine incorporation[b]
A		
subunit c : subunit b	pB10501(cb8)	10
B		
subunit c : galactokinase	pB10501(cb8)	6.7
subunit b : galactokinase	pB10501(cb8)	0.66
subunit b : galactokinase	pHB6(b11)	0.16
C		
subunit c : galactokinase	pH184(c6)	7.6
subunit c : galactokinase	pH159(c5)	1.5
subunit c : galactokinase	pH106(c4)	0.24
subunit c : galactokinase	pH163(c3)	0.024
D		
subunit δ : galactokinase	pSE1(δ12)	0.091

[a]In each case the data given apply to a plasmid concentration of 69 μg/ml.
[b]These are molar stoichiometries calculated after correction for the content of methionines in each of the proteins [see references regarding the *atp* operon in the text and Wilson and Hogness (1969)] and also, in the cases of the b and δ subunits, for background (see Materials and methods).

Experiments were performed [using pH184(c6), pB10505 (cb9), pHB6(b11) and pSE1(δ12)] to determine whether there were significant altrations in the stoichiometry of synthesis of subunits c, b and δ *in vitro* upon reducing the temperature of incubation. It was found that at an incubation temperature of 15°C the synthesis ratio of subunit b:subunit δ remained the same as at 37°C, whilst the ratios of subunit c:subunit b and subunit c:subunit δ decreased by 40%.

The relative stabilities of the c, b and δ subunits in vitro

The possibility that selective proteolysis might play a significant role in determining the amount of each subunit that could be measured by the described methods was excluded by two types of experiment. Firstly by pulse labelling the subunits synthesized *in vitro* for 1 min and then 'chasing' with a large excess of non-radioactive L-methionine during the remaining 19 min of incubation at 37°C. Secondly by adding chloramphenicol to the *in vitro* system after 20 min of incubation as described (Materials and methods) and allowing 70 min of further incubation at 37°C. The first type of experiment revealed no significant preferential degradation of any of the subunits. The second type of experiment indicated maximum extents of degradation of 20% for the b subunit and 30% for the δ subunit.

Discussion

We have shown that the cloning of *atp* genes in expression vectors allows a detailed examination of their expression

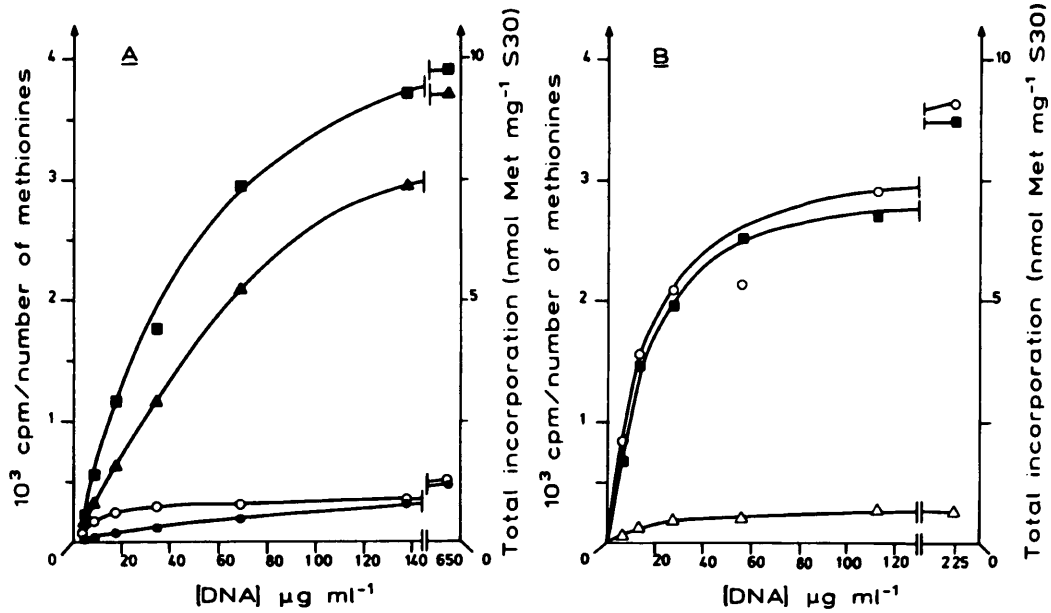

Fig. 5. Incorporation of [^{35}S]L-methionine into specific proteins in the S30 *in vitro* system. Typical plots of the incorporation into specific protein bands as a function of the concentration of added plasmid DNA [pB10501(cb8), **A**; pSE1(δ12), **B**] are shown. The radioactivity of bands corresponding to galactokinase (−○−), subunit c (−▲−), subunit b (−●−) and subunit δ (−△−) was determined and divided by the number of methionines contained in each type of protein molecule (see left hand abscissae). Thus for each of the experiments **A** and **B** the plotted values represent the relative molar amounts of each protein synthesized. Also shown are the total amounts of [^{35}S]L-methionine incorporated into protein (measured as hot tricholoroacetic acid-insoluble radioactive material corrected for the radioactivity of water blanks; −■−). Different specific activities of [^{35}S]L-methionine were used respectively in **A** and **B**.

behaviour *in vitro* and *in vivo*. The c, b and δ subunits were stable and it was therefore possible to quantify reliably the large differences in the respective synthesis rates of these polypeptides. pDR540 was especially useful for quantitative analysis of the expression of the c, b and δ subunit genes because it has a built-in reference gene (*gal*K) lying behind the cloning site which is transcribed as part of a common transcript with each of the cloned DNA sequences. Thus the measurement of the expression of *gal*K together with *atp* genes cloned in this plasmid allowed the identification of differences in the efficiency of translation of these genes. Moreover, the relative levels of expression of the cloned *atp* genes *in vitro* and *in vivo* were also generally consistent with each other.

The differences in the respective rates of synthesis of the H^+-ATPase subunits c, b and δ are determined at the level of translational initiation. This conclusion is based primarily upon the observed effects of changes in the DNA sequences upstream of the cloned *atp* genes upon their rates of translation. It should also be borne in mind that the large differences in the levels of expression *in vitro* between the c, b and δ subunit genes were observed even at the very lowest DNA concentrations. This therefore also means that the differences in synthesis rates were maintained under conditions where there cannot have been limitation of the elongation rates as the result of scarcity of factors required for polypeptide synthesis.

Most striking is the highly efficient translation of the subunit c cistron, which is dependent upon the presence of a long intercistronic sequence upstream of its start codon. The removal of the sequence >20 bp upstream of the c subunit Shine-Dalgarno sequence reduced the synthesis of this subunit by a factor of ~5 (lane 5, Figure 3 and Table I; compare also lane 3, Figure 4), and further shortening of this sequence led to much larger reductions in synthesis. The truncation of the intercistronic sequence lying before the b subunit gene (where the intercistronic sequence upstream of the c subunit had already been removed) up to 9 bp before the Shine-Dalgarno sequence by contrast, had no effect on the expression of this gene (compare lanes 10 and 11, Figure 3). The rate of translation of the subunit b gene was however affected by the presence of the intercistronic sequence lying upstream of the c subunit gene (see below). Since similar reductions in expression of the subunit c gene upon the shortening of its upstream intercistronic sequence were also observed using an entirely different expression plasmid (pJLf101), the observed effects cannot be associated with any specific local structure of either plasmid lying near the cloning site.

The codons of the genes encoding subunits c, α and β of the *E. coli* H^+-ATPase (unlike the codons of the genes encoding subunits γ, δ and ε) conform to the types commonly found in highly expressed *E. coli* genes (Futai and Kanazawa, 1983; Grantham *et al.*, 1981). These 'optimal' codons generally correspond to the most abundant iso-accepting tRNA species (Ikemura, 1981a, 1981b), and they may allow codon/anticodon interactions of intermediate energies which are postulated to be conducive to rapid turnover (Grantham *et al.*, 1981; Grosjean and Fiers, 1982). However, the data presented here contradict the proposal that codon usage directly determines the relative rates of translation of *atp* genes, and rather suggest that the above observations bear a different significance.

In Figure 6 the sequences of the translational initiation regions of the c, b and δ subunit genes are presented, together with the average (postulated as 'ideal') sequence for a translational initiation site deduced from analysis of the sequences of different genes (Scherer *et al.*, 1980; Gold *et al.*, 1981). The sequences of the three *atp* genes are generally dissimilar.

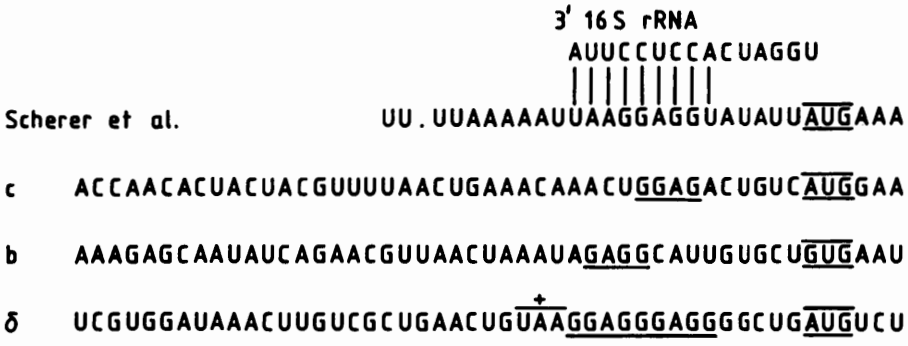

Fig. 6. Comparison of the translational initiation regions of the genes of subunits c, b and δ and of the consensus sequence of Scherer *et al.* (1980). The complete intercistronic sequence upstream of the subunit c gene is compared with the sequences upstream of the subunit b and δ genes. Only the b and c subunit genes are preceded by long intercistronic sequences. The consensus sequence of Scherer *et al.* (1980) is also presented, together with the complementary 3' end of the 16S rRNA. The translational stop codon of the b cistron is indicated in the sequence upstream of the δ cistron. The translational start codon and Shine-Dalgarno sequence of each cistron are also shown.

atpE	ACCAACACUACUACGUUUUAACUGAAACAAACUGGAGACUGUCAUGGAA
lambda D	UCUGAUGCCGUUAACGAUUUGCUGAACACACCAGUGUAAGGGAUGUUUAUGACG
lambda V	GCCGAUCUGACUUAUGUCAUUACCUAUGAAAUGUGAGGACGCUAUGCCU
rpsL	AGGACGUUUUAUUACGUGUUUACGAAGCAAAAGCUAAAACCAGGAGCUAUUUAAUGGCA
rplK	CCAACUUGAGGAAUUUAUAAUGGCUAAGAAAGUACAAGCCUAUGUCA

Fig. 7. Comparison of the translational initiation regions of the subunit c gene and of other *E. coli* and bacteriophage λ genes. The intercistronic sequence upstream of the subunit c gene (*atp*E) is compared with partially homologous translational initiation regions of two cistrons of the morphogenetic region of bacteriophage λ (Sanger *et al.*, 1982) and of two cistrons belonging to ribosomal protein operons (Post *et al.*, 1979; Post and Nomura, 1980). The two bacteriophage λ genes encode highly abundant proteins; the D major capsid protein and the V major tube protein, respectively. They are transcribed as part of the huge morphogenetic polycistronic mRNA (Ray and Pearson, 1975). The *E. coli* genes *rps*L and *rpl*K encode the ribosomal S12 and L11 subunits, respectively. The translational start codons, and in the λ sequences two stop codons, are underlined. The overall pattern common to these and other such sequences comprises a pyrimidine (U)-rich region followed by an interrupted purine (mainly A)-rich region, typically of the form GAAACAAA. The pattern does not always bear a similar relationship to the position of the translational start codon. In the case of *rpl*K the pattern may be repeated further upstream of the sequence shown in this figure.

The translational initiation region of the subunit c gene and the consensus sequence of Scherer *et al.* (1980) also have little in common apart from a high content of adenines. Local mRNA secondary structures [as initially pointed out by Brusilow *et al.* (1982)] could theoretically be formed within the sequence regions −13 bp to +24 bp of the b start codon and −12 bp to +37 bp of the δ start codon (as well as −31 bp to +12 bp of the γ start codon). Evidence has been obtained with other systems that the formation of mRNA secondary structure, which may for example sequester the Shine-Dalgarno sequence, can inhibit translational initiation (see, e.g., Hall *et al.*, 1982; Wood *et al.*, 1984; Schottel *et al.*, 1984). Brusilow *et al.* (1982) reported that the stoichiometry of synthesis of subunits of the *E. coli* H^+-ATPase *in vitro* was dramatically altered as the incubation temperature was lowered, and interpreted this as the result of changes in the stability of local mRNA secondary structures such as the ones mentioned above. It is not clear why similar experiments performed with the plasmids described here yielded much smaller, and qualitatively different, effects upon the stoichiometry of subunit synthesis than those reported by Brusilow *et al.* (1982).

It is not known to what extent the specific nature rather than purely the physical size or secondary structure of the intercistronic sequence is responsible for promoting efficient ribosomal initiation. Sequences showing striking similarity to the intercistronic sequence upstream of the subunit c gene were found in the translational initiation regions of other efficiently translated genes in *E. coli* (and in particular of ribosomal protein genes) and bacteriophage λ. Four of these regions are compared in Figure 7.

The possibility that the identified sequence pattern acts as an extra signal that enhances the rate of translational initiation, playing a particularly important role in polycistronic mRNAs, is attractive but needs to be thoroughly tested by experiment. A strong influence upon translational initiation of sequences many bases upstream from the Shine-Dalgarno sequence has also been observed in other systems (see, e.g., Kastelein *et al.*, 1983; Roberts *et al.*, 1979) where a sequence of the subunit c gene type was not present. However, direct evidence that the subunit c type sequence can act more generally to enhance translational initiation has come from the cloning of the subunit c intercistronic sequence in series with the human interleukin 2 and β-interferon genes (McCarthy *et al.*, in preparation).

The expression of the subunit b gene in the pB1050(cb) type plasmids is apparently linked to the expression of the preceding subunit c gene. This phenomenon is reminiscent of the 'translational coupling' effect identified for example in studies of the *E. coli* ribosomal protein operons (Nomura *et al.*, 1984). The parallel rise in expression of the subunit b and c genes of pB10501(cb8) with increasing DNA concentration indicates that ribosomes re-initiate at the subunit b cistron. The low frequency of re-initiation could be determined partially or

completely by the two adjacent stop codons of the subunit c cistron, the 58-bp intercistronic sequence upstream of the subunit b cistron, or the possible secondary structure in this region. The poor expression of the subunit δ gene obtained with pSE1(δ12) may also reflect the loss of some translational coupling with the subunit b cistron.

It is important to consider how well the relative efficiencies of translational initiation of the b, c and δ subunit genes correspond to the amounts of each of these subunits that are required for assembly of the *E. coli* H^+-ATPase. The synthesis of subunits c and b can be directly compared using the pB1050(cb) type plasmids, and showed a molar stoichiometry of synthesis in the range 6 – 10, which compares favourably with the estimated stoichiometry of the H^+-ATPase (Foster and Fillingame, 1982; von Meyenburg *et al.*, 1982b). The molar stoichiometry of synthesis of subunits c and δ, on the other hand, can only be estimated more indirectly. Comparison of the expression of pSE1(δ12), pHB6(b11) and the subunit c gene and pB1050(cb) type plasmids indicates that the ratio of translational initiation efficiency for subunit c:subunit δ is at least 18. By analogy to the effect of re-initiation upon expression of the subunit b gene, it might be argued that this ratio is exaggerated because the δ subunit gene in pSE1(δ12) is no longer located downstream of the b subunit gene.

Finally, the major factor underlying the synthesis of the appropriate amounts of the c, b and δ subunits of the *E. coli* H^+-ATPase has been shown to be the efficiency of translational initiation. The subunit stoichiometry of the other subunits may be achieved in the same way.

Materials and methods

Bacterial strains

MCG1 was constructed by transferring a mutated *rec*A sequence from JC10240 (Hfr Po45 *srl* 300::Tn10 recA56 *thr*300 *ilv*318 *rps*E300 *rec*A1 *thi*-1) to the *E. coli* K12 *lac* repressor overproducer strain JM103 [Δ (*lac pro*) *thi strA supE endA sbcB hsdR*$^-$ F'*tra*-D36 *proAB lacI*^{q} ZΔM15; Messing *et al.*, 1981] using the general transducing bacteriophage P1. In order to construct *atp*$^-$ (*unc*$^-$) derivatives of JM103, an *ilv*$^-$ derivative of this strain was first of all isolated after mutagenesis with N-methyl-N'-nitro-N-nitrosoguanidine and penicillin enrichment. *ilv* was used as a marker for the transfer of defective *atp* sequences from strains DG7/10 (*atp*E$^-$) and DG25/9 (*atp*F$^-$) (Schairer *et al.*, 1976) *via* P1 transduction. The *atp*$^-$ derivatives isolated in this way were rendered *rec*A$^-$ (as described) thus producing strains MCG2 (*atp*E$^-$) and MCG3 (*atp*F$^-$). The MCG strains 1 – 3 were used for transformation (Dagert and Ehrlich, 1979; Hanahan, 1983) and complementation tests involving derivatives of pDR540. Transformations with pJLf101 or derivatives thereof were performed using strains N99cI$^+$ and N4830 (Pharmacia P-L Biochemicals).

Manipulation and cloning of specific DNA fragments

The primary source of DNA from the *atp* operon was the defective specialized transducing bacteriophage λ *asn*105 (von Meyenburg *et al.*, 1978). Digestions with restriction endonucleases (performed under the manufacturer's specified conditions) together with separations and isolations using agarose and polyacrylamide gels (see, e.g., Maniatis *et al.*, 1982) yielded suitable fragments of the *atp* operon for further manipulation and cloning (see Figure 1). A *Taq*I fragment served as the basis for the construction of plasmids bearing the c subunit gene. pB1050(cb) type plasmids were constructed using shortened forms of a *Bam*HI/*Bst*EII fragment. pHB6(b11) was constructed using a part of the pB10501(cb8) insert separated after *Hpa*I digestion. The subunit δ gene was isolated and cloned as a *Sau*3A/*Eco*RI fragment. *Bal*31 exonuclease was used for progressive shortening of particular fragments at a ratio of enzyme units:μg DNA of ~1:5 at 30°C in a reaction medium described by Legerski *et al.* (1978). The DNA fragments bearing the b and c protein genes were initially cloned using the *lac* promoter-bearing plasmid pUR108-1 (U.Rüther, Cologne) and were subsequently excised and recloned using either the *tac* fusion promoter vector pDR540 (Russell and Bennett, 1982) or the expression vector pJLf101 (described below). Initial screening of transformants was performed by restriction endonuclease analysis of small-scale plasmid preparations (Birnboim and Doly, 1979; Maniatis *et al.*, 1982) and/or by colony filter hybridization using nick-translated or 5'-labelled DNA probes (Grunstein and Hogness, 1975; Jeffreys and Flavell, 1977). Preparations of plasmid DNA (Birnboim and Doly, 1979; Grosveld *et al.*, 1981) from selected clones were finally submitted to further analysis by complementation testing and DNA sequencing (Maxam and Gilbert, 1980; Sanger *et al.*, 1977). All of the cloned subunit b and subunit c gene sequences described were shown to direct the synthesis of functional subunits by complementation analysis using the constructed MCG strains 2 and 3. Complementation was not obtained when fragments were cloned in the wrong orientation. The subunits synthesized under the direction of these plasmids could also be immunoadsorbed by specific antibodies raised against the b and c subunits, respectively. 90% of the cloned δ subunit gene sequence was checked, confirming the published sequence (Nielsen *et al.*, 1981).

Construction of the expression vector pJLf101 involved insertion of the λ P_L promoter on an *Eco*RI-*Hae*III fragment (derived from pPL-Lambda, P-L Biochemicals GmbH; see Drahos and Szybalski, 1981) between the *Eco*RI and *Sal*I sites of a derivative of pBR322. pJLf101 also bears the strong fd bacteriophage transcription termination sequence (Beck *et al.*, 1978; derived from pGBu 207, see Gentz *et al.*, 1981) inserted into the *Nru*I site 320 bp downstream from the *Sal*I cloning site. The region of the original pBR322 sequence between the *Ava*I and *Pvu*II sites was deleted.

Preparation and use of the S30 fraction from E. coli

The cell-free extract was prepared from *E. coli* K12 NF1128 (*leu*$^-$ *arg*$^-$ RNase$^-$) essentially according to Zubay (Zubay, 1973; Chen and Zubay, 1983) and Pratt (1980). The assays using the S30 fraction were performed as described previously (Chen and Zubay, 1983) except that NADPH and pyridoxine were omitted. Magnesium acetate and calcium acetate were found to be optimal at final concentrations of 13 mM and 8.0 mM, respectively. The final concentration of carrier and ^{35}S-labelled L-methionine was 88 μM (up to 290 μCi/ml). Phenylmethylsulphonyl fluoride (PMSF) and *p*-aminobenzamidine were present in assays at 0.53 mM and 53 μM, respectively. Plasmid DNA (purified twice on CsCl gradients) was added at various concentrations. Protein synthesis was terminated by the addition of chloramphenicol (250 μg/ml) after 20 min incubation at 37°C or 100 min at 15°C. After a 5 min incubation at 37°C with DNase (50 μg/ml), samples were taken for analysis of incorporated [^{35}S]L-methionine and loading onto SDS-polyacrylamide gels. Samples (5 – 10 μl) taken for analysis of radioactive incorporation into protein were placed on Whatman GF/A glass fibre filters, which were washed four times (once at 90°C) with 5% trichloroacetic acid containing 1 mg/ml L-methionine and finally with ethanol and diethylether. The dependence of radioactive incorporation into protein upon time was linear during the stated incubation times. Labelled proteins synthesized in the *in vitro* system were separated on 13.5% SDS-polyacrylamide gels (with 3% stacking gels), which were subsequently stained with Coomassie Brilliant Blue R, impregnated with 2,4-diphenyloxazol (Bonner and Laskey, 1974) and dried. The positions of the radioactively labelled proteins on the gels were checked against the positions of the component proteins of standard F_1F_0 H^+-ATPase. Autoradiography was performed with Kodak X-Omat-AR or Agfa-Gevaert Curix X-ray film. Superimposition of the exposed and developed films on the dried gels allowed the precise excision of particular bands in equally sized sections for radioactivity analysis in scintillant. Control experiments in which serial dilutions of radioactively labelled proteins were loaded onto gels showed that the radioactivity measured in the gel sections was linearly related to the amount of radioactive protein in the bands. Corrections of the measured radioactive incorporation into a specific protein band for background in the gels were made by subtracting the radioactivity measured in equally sized sections cut from corresponding positions in control lanes which lacked this particular protein band. For example, for correction of the measured radioactivity in a gel section containing the subunit b band, the synthesis of which was directed by pB10501(cb8), an equally sized section was excised from the equivalent position in a lane loaded with the synthesis products of pH116(c7). pDR540 was correspondingly used as the control plasmid for pHB6(b11), pB10506(cb10) and pSE1(δ12), respectively. Identical concentrations of the experimental and control plasmids were used in the titrations.

In vivo pulse-labelling

Cells of strain MCG1 containing the described plasmids were allowed to grow in minimal medium at 37°C to $O.D._{450}$=0.5, at which point isopropyl-β-D-thiogalactoside (final concentration 2 mM) was added. After a further 20 min, 15 μCi [^{35}S]L-methionine (final concentration 15 nM) was added to 1 ml of these cells and incubation was continued for another minute. Then L-methionine (final concentration 1 mM) was added and after a further minute's incubation the cells were rapidly cooled to −20°C. After thawing, centrifuging, and washing twice with 70 mM Tris-HCl (pH 7.5), 10 mM

$MgCl_2$, 0.5 mM PMSF, 30 μg/ml *p*-aminobenzamidine, the cells were treated with lysozyme (4 mg/ml) in a solution of 25 mM Tris-HCl (pH 8.0), 50 mM glucose, 10 mM EDTA. The cells were then sonicated briefly in the first buffer containing 1 mg/ml DNase, and finally samples were taken for solubilization in SDS and loading on 15% SDS-polyacrylamide gels.

Preparation of antibodies

Antibodies against the purified subunit c were raised in rabbits (see, e.g., Wier, 1978).

Materials

Restriction endonucleases and other enzymes used for DNA manipulation and cloning were obtained from Pharmacia P-L Biochemicals, New England Biolabs, Bethesda Research Laboratories and Boehringer Mannheim (FRG). [γ-^{32}P]ATP (5000 Ci/mmol), α-^{32}P-labelled nucleotides, ^{14}C-labelled protein mixture and [^{35}S]L-methionine (1000 Ci/mmol) were from Amersham Buchler (Braunschweig, FRG). Nitrocellulose filters were from Schleicher and Schüll (Dassel, FRG).

Acknowledgements

We are grateful to Dr P.Friedl for helpful discussion and the donation of *E. coli* F_1F_0 H^+-ATPase and an antibody preparation raised against subunit b. *E. coli* strains CM845, JC10240 and NF1124 were kindly provided by Professor K.von Meyenburg, Dr W.Lindenmaier and Dr J.Collins, respectively. We also wish to thank Professor K.von Meyenburg for advice and his comments on the original manuscript. J.E.G.M. thanks the European Molecular Biology Organisation for the award of a Long Term Fellowship.

References

Beck,E., Sommer,R., Auerswald,E.A., Kurz,C., Zink,B., Osterburg,G., Schaller,H., Sugimoto,K., Sugisaki,H., Okamoto,T. and Takanani,M. (1978) *Nucleic Acids Res.*, **5**, 4495-4503.

Birnboim,A.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513-1523.

Bonner,W.M. and Laskey,R.A. (1974) *Eur. J. Biochem.*, **46**, 83-88.

Brusilow,W.S.A., Klionsky,D.J. and Simoni,R.D. (1982) *J. Bacteriol.*, **151**, 1363-1371.

Brusilow,W.S.A., Porter,A.C.G. and Simoni,R.D. (1983) *J. Bacteriol.*, **155**, 1265-1270.

Chen,H. and Zubay,G. (1982) *Methods Enzymol.*, **101**, 674-690.

Dagert,M. and Ehrlich,S.D. (1979) *Gene*, **6**, 23-28.

Downie,J.A., Gibson,F. and Cox,G.B. (1979) *Annu. Rev. Biochem.*, **48**, 103-131.

Drahos,D. and Szybalski,W. (1981) *Gene*, **16**, 261-274.

Foster,D.L. and Fillingame,R.H. (1982) *J. Biol. Chem.*, **257**, 2009-2015.

Friedl,P., Friedl,C. and Schairer,H.U. (1979) *Eur. J. Biochem.*, **100**, 175-180.

Futai,M. and Kanazawa,H. (1983) *Microbiol. Rev.*, **47**, 285-312.

Gay,N.J. and Walker,J.E. (1981a) *Nucleic Acids Res.*, **9**, 2187-2194.

Gay,N.J. and Walker,J.E. (1981b) *Nucleic Acids Res.*, **9**, 3919-3926.

Gentz,R., Langner,A., Chang,A.C.Y., Cohen,S.N. and Bujard,H. (1981) *Proc. Natl. Acad. Sci. USA*, **78**, 4936-4940.

Gold,L., Pribnow,D., Schneider,T., Shinedling,S., Singer,B.S. and Stormo, G. (1981) *Annu. Rev. Microbiol.*, **35**, 365-403.

Grantham,R., Gautier,D., Gouy,M., Jacobzone,M. and Mercier,R. (1981) *Nucleic Acids Res.*, **9**, 43-74.

Grosjean,H. and Fiers,W. (1982) *Gene*, **18**, 199-209.

Grosveld,F.G., Dahl,H.M., deBoer,E. and Flavell,R.A. (1981) *Gene*, **13**, 227-237.

Grunstein,M. and Hogness,D.S. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 3961-3965.

Hall,M.N., Gabay,J., Débarbouillé,M. and Schwartz,M. (1982) *Nature*, **295**, 616-618.

Hanahan,D. (1983) *J. Mol. Biol.*, **166**, 557-580.

Hoppe,J. and Sebald,W. (1984) *Biochim. Biophys. Acta*, **768**, 1-27.

Ikemura,T. (1981a) *J. Mol. Biol.*, **146**, 1-21.

Ikemura,T. (1981b) *J. Mol. Biol.*, **151**, 389-409.

Jeffreys,A.J. and Flavell,R.A. (1977) *Cell*, **12**, 429-439.

Jones,H.M., Brajkovich,C.M. and Gunsalus,R.P. (1983) *J. Bacteriol.*, **155**, 1279-1287.

Kanazawa,H., Mabuchi,K., Kayano,T., Noumi,T., Sekiya,T. and Futai,M. (1981) *Biochem. Biophys. Res. Commun.*, **103**, 613-620.

Kastelstein,R.A., Berkhout,B., Overbeek,G.P. and van Duin,J. (1983) *Gene*, **23**, 245-254.

Legerski,R.J., Hodnett,J.L. and Gray,H.B., Jr. (1978) *Nucleic Acids Res.*, **5**, 1445-1464.

Lünsdorf,H., Ehrig,K., Friedl,P. and Schairer,H.U. (1984) *J. Mol. Biol.*, **173**, 131-136.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*, published by Cold Spring Harbor Laboratory Press, NY.

Maxam,A.M. and Gilbert,W. (1980) *Methods Enzymol.*, **65**, 499-560.

McCarthy,J.E.G., Schairer,H.U. and Sebald,W. (1984) *Abstract 3rd EBEC Meeting*, Hannover.

McKenney,K., Shimatake,H., Court,D., Schmeissner,U., Brady,C. and Rosenburg,M. (1981) in Chirikjian,J.G. and Papas,T. (eds.), *Gene Amplification and Analysis*, Vol. II, Elsevier, North Holland, pp. 383-415.

Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 308-321.

Nielsen,J., Hansen,F.G., Hoppe,J., Friedl,P. and von Meyenburg,K (1981) *Mol. Gen. Genet.*, **184**, 33-39.

Nielsen,J., Jørgensen,B.B., von Meyenburg,K. and Hansen,F.G. (1984) *Mol. Gen. Genet.*, **193**, 64-71.

Nomura,M., Gourse,R. and Baughman,G. (1984) *Annu. Rev. Biochem.*, **53**, 75-117.

Porter,A.G.G., Brusilow,W.S.A. and Simoni,R.D. (1983) *J. Bacteriol.*, **155**, 1271-1278.

Post,L.E. and Nomura,M. (1980) *J. Biol. Chem.*, **255**, 4660-4666.

Post,L.E., Strycharz,G.D., Nomura,M., Lewis,H. and Dennis,P.P. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 1697-1701.

Pratt,C. (1980) *J. Bacteriol.*, **143**, 1265-1274.

Ray,P.N. and Pearson,M.L. (1975) *Nature*, **253**, 647-650.

Roberts,T.M., Kacich,R. and Ptashne,M. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 760-764.

Russell,D.R. and Bennett,G.N. (1982) *Gene*, **20**, 231-243.

Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 5463-5467.

Sanger,F., Coulson,A.R., Hong,G.F., Hill,D.F. and Petersen,G.B. (1982) *J. Mol. Biol.*, **162**, 729-773.

Saraste,M., Eberle,A., Gay,N.J., Runswick,M.J. and Walker,J.E. (1981) *Nucleic Acids Res.*, **9**, 5287-5296.

Schairer,H.U., Friedl,P., Schmid,B. and Vogel,G. (1976) *Eur. J. Biochem.*, **66**, 257-268.

Scherer,G.F.E., Walkinshaw,M.D., Arnott,S. and Morré,D.J. (1980) *Nucleic Acids Res.*, **8**, 3895-3907.

Schottel,J.L., Sninsky,J.J. and Cohen,S.N. (1984) *Gene*, **28**, 177-193.

von Meyenburg,K., Hansen,F.G., Nielsen,L.D. and Riise,E. (1978) *Mol. Gen. Genet.*, **160**, 287-295.

von Meyenburg,K., Jørgensen,B.B., Nielsen,J. and Hansen,F.G. (1982a) *Mol. Gen. Genet.*, **188**, 240-248.

von Meyenburg,K., Jørgensen,B.B., Nielsen,J., Hansen,F.G. and Michelson,O. (1982b) *Tokai J. Exp. Clin. Med.*, **7 Suppl.**, 23-31.

Wier,D.M. (1978) *Immunochemistry*, Vol. **1**, published by Blackwell, Oxford.

Wilson,D.B. and Hogness,D.S. (1969) *J. Biol. Chem.*, **244**, 2137-2142.

Wood,C.R., Boss,M.A., Patel,T.P. and Spencer Emtage,J. (1984) *Nucleic Acids Res.*, **12**, 3937-3950.

Zubay,G. (1973) *Annu. Rev. Genet.*, **7**, 267-287.

Received on 19 November 1984; revised on 14 December 1984