The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks

Russell Thompson*, Linda Taylor, Kevin Kelly^{1,2},
Roger Everett^{1,3} and Neil Willetts¹

Institute of Virology, University of Glasgow, Church Street, Glasgow G11 5JR, and ¹Department of Molecular Biology, University of Edinburgh, Kings Buildings, Mayfield Road, Edinburgh EH9 3JR, UK

²Present address: Department of Biochemistry, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK

³Present address: University of Glasgow, Institute of Virology, Church Street, Glasgow G11 5JR, UK

*To whom reprint request should be sent

Communicated by N.G. Murray

The DNA sequence of the F plasmid origin of conjugal DNA transfer, *oriT*, has been determined. The origin lies in an intergenic region which contains several inverted repeat sequences and a long AT-rich tract. Introduction of a nick into one of the DNA strands in the *oriT* region precedes the initiation of conjugal DNA replication, and the position of the strand-specific nicks acquired by a λ *oriT* genome upon propagation in *Flac*-carrying cells has been determined. The nicks were not uniquely positioned, rather there was a cluster of three major and up to 20 minor sites: the biological significance of this observation is not yet fully clear. Nine independent point mutations which inactivate *oriT* function have been sequenced and found to alter one or other of two nucleotide positions which lie 14 and 19 bp to one side of the rightmost (as drawn) major nick site. These key nucleotides may lie in a recognition sequence for the *oriT* endonuclease, since mutations at these sites prevent nicking at *oriT*.

Key words: conjugation/*oriT*/DNA sequence

Introduction

Conjugation is a process which allows the efficient transmission of plasmid DNA between bacterial cells. During conjugation mediated by the F plasmid of *Escherichia coli*, a unique pre-existing strand of the plasmid DNA is transferred to the recipient cell with the 5' terminus leading. Simultaneously, new complements to the 'non-transferred' and 'transferred' strands are synthesised in the donor and recipient cells, respectively (for review, see Willetts and Skurray, 1980; Willetts and Wilkins, 1984). Initiation of this conjugative DNA synthesis takes place after nicking of the transferred strand at a specific site called the origin of transfer, *oriT* (Willetts, 1972). Several F-specified proteins are required for initiation, in particular the products of the *traM*, *traY*, *traI* and *traZ* genes (Kingsman and Willetts, 1978; Everett and Willetts, 1980). Once initiated, conjugal DNA replication is largely carried out by host-encoded enzymes including DNA polymerase III and RNA polymerase (Kingsman and Willetts, 1978). The plasmid *traI* product has recently been purified and shown to be an ATP-dependent helicase (Abdel-Monem *et al.*, 1983) and may function to separate the F DNA strands during transfer.

The *oriT* site is central to the transfer process. We define it as including the nick site itself and the associated binding

site(s) for the nicking and triggering proteins. When lambda phages containing a cloned *oriT* sequence were propagated in cells containing an *Flac* plasmid, 5–10% of the progeny genomes contained a nick lying within *oriT* and, using this assay system, the *traY* and *traZ* proteins were shown to be essential for nicking (Everett and Willetts, 1980). Nicking at *oriT* occurred in the absence of mating pair formation and of conjugative DNA synthesis. It must therefore be the formation of mating pairs between donors and recipient cells which triggers DNA transfer and conjugative DNA synthesis rather than the nicking process. The mechanism of triggering is unknown though it is probably the consequence of an alteration in *tra* protein function induced by mating pair formation rather than *de novo* synthesis since all of the known F conjugation genes are constitutively expressed. The *traM* gene product is a good candidate for a protein involved in sensing a successful mating contact and triggering conjugal DNA replication since *traM* lies immediately adjacent to *oriT* and the protein is located in the inner membrane and binds to DNA in the *oriT* region (R. Thompson, D.R. Musgrave and M. Achtman, unpublished data).

The origin of transfer was located, as a site required in *cis* for transfer, within a 373-bp *BglII-HaeIII* fragment (see Figure 1) and several *oriT*[−] mutants were isolated by mutagenesis of a plasmid containing this region (Everett and Willetts, 1982). When one such mutation was tested by incorporation into a λ *oriT* phage, it was shown to prevent nicking when the phage was grown in *Flac*-containing cells (Everett and Willetts, 1982). Since it has proven impossible to isolate an F relaxation complex (Johnson *et al.*, 1981) we have used nicked λ *oriT* DNA as the closest approximation currently available to plasmid DNA nicked to initiate conjugation. Here, we present the sequence of the *BglII-HaeIII* fragment containing *oriT* and locate on this sequence the *oriT*[−] mutations and the precise position of the nicks in λ *oriT* phage. The structure and functions of the *oriT* region are discussed.

Results

Sequence of the *oriT* region

The region of F DNA that includes *oriT* lies at one end of the block of transfer genes and has been cloned as a 1.08-kb *BglII* fragment in pBE274 (Thompson and Achtman, 1978) and as a 0.53-kb *BglII-SalI* fragment in pED806 (Johnson *et al.*, 1981; see Figure 1). Both are small, multicopy plasmids suitable as sources of DNA for sequencing experiments. The sequencing strategy for the *oriT* region is shown in Figure 1 and the sequence is presented in Figure 2.

Within the *BglII-HaeIII* fragment, the only section of open reading frame in which a start codon is preceded by a Shine-Dalgarno ribosome binding site homology (Stormo *et al.*, 1982) extends leftward across the *BglII* site as indicated in Figure 2. There are 15 codons of this frame within the sequence presented here and upstream of the translation start site there are sequences homologous to the consensus *E. coli*

Fig. 1. Physical map of *EcoRI* fragment f6 of the F plasmid. Transfer genes are shown as boxes and *oriT*, the transfer origin, as an arrowhead. The *EcoRI* fragment is 8 kb in length (data from Thompson and Achtman, 1979; Willetts and Skurray, 1980). The lower section shows the restriction map and sequencing strategy of the 373-bp *BglII*-*HaeII* fragment that contains *oriT*.

Fig. 2. Nucleotide sequence of the *oriT* region. Inverted (←) and directly repeated sequence (→) are indicated together with the positions of the major nick sites in EDλ102 DNA and the C→T transitions found in *oriT*⁻ mutants. The promoters for *traM* and gene X, the ribosome binding site (SD) and the initiation codon for gene X are also shown. Nucleotide position 1 corresponds to the centre of the *BglII* site (see Figure 1). Nucleotides 321–400 and the position of *P_M* are from Thompson and Taylor (1982).

promoter sequence (Siebenlist *et al.*, 1980). The activity of this promoter has been detected *in vivo* using plasmid constructions in which it is fused to *lacZ* (Gaffney *et al.*, 1983), *galK* (P. Mullineaux and N. Willetts, unpublished data) or a tetracycline resistance gene (R. Thompson, unpublished data). The protein encoded by this gene, which we have designated gene X, has not been detected so far and it is not

known if it has a function in conjugation. At the other end of the *BglII*-*HaeII* fragment, transcription from the *traM* promoter crosses the *HaeII* site (Figure 2; Thompson and Taylor, 1982). The absence of any significant coding regions between gene X and *traM* indicates that the transfer origin lies within a 292-bp region between these two divergently transcribed genes.

The 292-bp region contains several inverted repeat sequences, one of which is particularly striking (nucleotides 228–248, 259–278, Figure 2). Secondary structures formed by these inverted repeat sequences may be responsible for the aberrant gel mobility of fragments from this region: the *BglII*-*SalI* fragment migrates on polyacrylamide gels as 600 bp although from our sequence data it is known to be 532 bp (Figure 2; Thompson and Taylor, 1982). There are two copies of a 22-bp sequence in direct repeat across the promoter for the *traM* gene and there is a closely related version of this sequence in inverted orientation between nucleotides 266 and 242. The region also includes an extensive AT-rich segment (nucleotides 148–234 are 80% A+T) a feature common to several vegetative replication origins of both plasmids and phages (Murotsu *et al.*, 1981; Shafferman and Helinski, 1983; Selzer *et al.*, 1983; Moore *et al.*, 1978) and the bacterial chromosome (Meijer *et al.*, 1979; Sugimoto *et al.*, 1979).

Location of the *oriT* nick in EDλ102

A DNA preparation containing molecules specifically nicked at *oriT* was obtained by growing a λ*oriT* phage (EDλ102) in cells containing an *Flac* plasmid (Everett and Willetts, 1980). The observation that the nick is at the expected position in the correct strand, that the presence of an *Flac* plasmid is essential for the appearance of the nicks and that λ*oriT*⁻ mutant genomes are not nicked, suggests that the nick is indeed that from which plasmid DNA transfer is initiated (Everett and Willetts, 1980, 1982).

A disadvantage of this method is that only 5–10% of the λ*oriT* genomes are nicked. However, Everett and Willetts (1980) were able to use such preparations to show that the strand interruption in the *oriT* region was probably a nick (since the complementary strand could not be cleaved by nuclease S1), that the 3' terminus was a 3'-hydroxyl or a 3'-phosphate (since it was susceptible to degradation by exonuclease III) and that the nick could not be religated using T4 DNA ligase. To locate precisely the position of the *oriT* nick we performed an experiment in which the 3' terminus formed after cleavage of nicked or un-nicked λ*oriT* DNAs at the *BglII* site nearest *oriT* was labelled with ³²P (see Figure 1 and Materials and methods). A portion of the singly end-labelled fragment was subjected to the Maxam and Gilbert sequencing reactions and a second portion was denatured and loaded alongside the sequence tracks. Figure 3 shows that the denatured sample of the nicked DNA displays an array of bands in the region between nucleotides 100 and 130. These bands are absent from the un-nicked DNA sample and presumably represent DNA chains extending from the 3' label at the *BglII* site to the position of the *oriT* nick. A similar experiment, using a 5' label at the *BglII* site, confirmed that these nicks are only present in one strand (data not shown).

Independent preparations of EDλ102 DNA showed variation in the total number of positions that were nicked ranging from eight to 21 nick sites: the preparation shown in Figure 3 is at one extreme of this spectrum. As seen in Figure 3

Fig. 3. Samples of EDλ102 DNA 3'-labelled at the *Bgl*II site run on an 8% polyacrylamide sequencing gel after denaturation or sequencing.

however, certain sites were nicked more frequently than others. Three major sites were consistently preferred in independent preparations of nicked *λoriT* phages and these are marked in Figures 2 and 5. The major nick locations also correspond to positions in the sequence tracks which have bands in all four lanes. In the course of the sequence reactions the samples had been subjected to reaction with piperidine and such alkaline conditions have been shown to break phosphoserine and phosphothreonine although not phosphotyrosine protein-DNA linkages (Tse *et al.*, 1980; Desiderio and Kelly, 1981; Tamanoi and Stillman, 1982). The co-migration of the major nicked fragments in the denatured sample with those in the sequenced sample thus indicates that there are no residual amino acids attached to the 5' nick terminus. This is important since analogies with other nicking-ligating systems such as *E. coli* topoisomerase I, DNA gyrase and the gene A

Fig. 4. Labelling the 3' ends of the *oriT* nicks. 8% polyacrylamide sequencing gel of nicked and un-nicked EDλ102 DNAs after labelling with the Klenow fragment of DNA polymerase I and a single dNTP. Markers are end-labelled *Hpa*II fragments of pBR322.

protein of ϕ X174 suggest that the 5' *oriT* terminus might be covalently linked to a protein (Depew *et al.*, 1978; Morrison and Cozzarelli, 1979; Ikeda *et al.*, 1976).

To confirm the presence of multiple nick sites, the 3'-hydroxyl end of the nick was labelled with a single deoxynucleoside triphosphate. The 1.08-kb *Bgl*II fragment from nicked and un-nicked EDλ102 DNA was purified and incubated with the Klenow fragment of DNA polymerase I in two separate reactions, one containing [α - 32 P]dGTP and the other [α - 32 P]dATP. The labelled fragments were digested with *Taq*I, denatured and analysed on an 8% polyacrylamide sequencing gel. Fragments running from the 3' end of the nick site to the *Taq*I site would be expected to be 290–320 bp long. Figure 4 shows that dGTP is incorporated into a series of bands in this size range and that dATP is incorporated into a second series of bands. This is consistent with the DNA sequence in the region in which the nicks fall since many of the 3' nick positions observed in Figure 3 would lead to incorporation of dGTP. This experiment also confirms that the nicks in EDλ102 DNA have 3' hydroxyl ends. In contrast to the ϕ X174-gene A protein system, the 3' terminus of the *oriT*

nick is not available to prime replacement strand DNA synthesis in donor cells (Kingsman and Willetts, 1978) although we have shown that it readily primes DNA synthesis by DNA polymerase I *in vitro* (data not shown). The interpretation of these observations is considered in the Discussion.

Sequence analysis of *oriT*⁻ mutations

When the 0.53-kb *Bgl*II-*Sal*I fragment containing *oriT* was cloned into pBR322 the resulting plasmid, pED806, was stable in F⁻ cells but was rapidly lost from cells containing *Flac*. The reason for this instability is unknown. However, when mutants which were stable in the presence of *Flac* were isolated they were found to be *oriT*⁻ in mobilisation assays (Everett and Willetts, 1982). We have isolated nine further mutants using other mutagens.

The *oriT* regions of nine mutants have been sequenced. All were GC→AT transitions despite the variety of mutagens used in their isolation. Two hydroxylamine-induced mutations (in pED833 and pED836; Everett and Willetts, 1982) were at position 141 of the sequence shown in Figure 2. Two further hydroxylamine-induced mutations (in pED834 and pED835) and five others with mutations induced by nitrous acid or ethylmethanesulphonate were at position 146. It should be stressed that all nine mutants were independently isolated and could not be siblings. Correlation with the residual levels of *oriT* function shows that mutations at the two sites reduce mobilisation by 200- and 9-fold, respectively (Everett and Willetts, 1982).

It has previously been shown that an *oriT* region containing a GC→AT transition at position 141 was not nicked when propagated in *Flac* cells as a λ *oriT* phage (Everett and Willetts, 1982). We have now shown that this is also the case for a transition at position 146 (data not shown). The two important nucleotides identified by these mutations lie 14 and 19 bp to the right of the rightmost major *oriT* nick described in the preceding section (see Figure 2). The presence of more than one *oriT* nick site could explain why no mutants in the GC pairs bounding these sites were obtained since mutation of one site might leave the others fully functional.

Discussion

Transfer of a unique plasmid strand in a fixed orientation has been experimentally demonstrated for F and is widely assumed to be the case for other plasmids. A site was mapped on F which was required in *cis* for DNA transfer, designated the transfer origin, and proposed as the site at which a nick would initiate single strand transfer (Willetts, 1972). It was not until the studies on *in vivo* nicking of a λ *oriT* phage genome, however, that evidence was obtained that this was indeed the case (Everett and Willetts, 1980). The map location of *oriT* has been refined by cloning progressively smaller DNA fragments using the ability to confer a high level of mobilisation onto a poorly mobilisable vectors as an assay for *oriT* function (Achtman *et al.*, 1978; Thompson and Achtman, 1978; Johnson *et al.*, 1981; Everett and Willetts, 1982); the DNA sequence presented here is of the smallest fragment yet shown to retain full *oriT* function, a 373-bp *Bgl*II-*Hae*II fragment.

Within the *oriT* region we have located the nucleotides affected in nine *oriT*⁻ point mutants. These mutations identify two positions which are critical for *oriT* function and have been shown to be essential for nicking of λ *oriT* genomes. We have also localised the nicks generated *in vivo* in a λ *oriT*

Fig. 5.(a) Potential secondary structures around the nick sites in ED λ 102 DNA. The positions of the *oriT*⁻ point mutations are also shown. (b) Sequence homology between ColE1 and F. The ColE1 nucleotide coordinates are from Bastia (1978), those of F from Figure 2.

genome. These were strand specific but did not occur at a unique site; rather the nicks were scattered over a 20-bp region with the major sites clustered at one end. Figure 5a shows the locations of the *oriT*⁻ point mutations and *oriT* nick sites relative to two potential secondary structures of this region. It is unlikely that the point mutations exert their effect by destabilising the hairpins *per se* since mutations in adjacent GC pairs which would also disrupt the hairpins have not appeared in this mutant collection. We therefore favour the hypothesis that these two nucleotides are key elements in the recognition sequence of the *oriT*-specific *traYZ* endonuclease. We have no explanation for our observation that mutations at only two positions yielded an *oriT*⁻ phenotype.

If the multiple potential *oriT* nick sites truly represent the *in vivo* situation (see below) their location to one side of the recognition sequence is reminiscent of type I restriction enzymes and the packaging signals of phages λ and P22 in which the DNA cutting and recognition sites are not congruent (Feiss *et al.*, 1983; Casjens and Huang, 1982). Once bound, the *traYZ* endonuclease would move in a 3' direction and nick at one or other of the *oriT* sites. The invariability of the *oriT* sequence in several independent clones analysed in the course of this work would then rule out models for conjugative replication and recircularisation based upon ϕ X174 RF replication, since it would be impossible to ensure that the second nicking event took place at the same position as the first. The unavailability of the 3' *oriT* nick terminus to prime replacement strand synthesis in any case shows that the two systems are different (Kingsman and Willetts, 1978). An

alternative explanation for the multiple nicks observed in ED λ 102 DNA is that they arose as a result of 'nick translation' prior to packaging of the λ genome. If a protein(s) is attached to the 5' terminus of the authentic *oriT* nick for example, this might interfere with DNA packaging and, in this case, only those molecules from which the protein had been removed would be successfully packaged. Protein removal could occur by 'nick translation', to displace the strand with the terminal protein followed by cleavage of the terminal DNA segment. The multiple nick sites would then represent different extents of 'nick translation' from an initial, possibly unique, nick lying to their right as drawn in Figure 2, perhaps nearer to the *oriT*⁻ point mutations.

One way to distinguish between these alternatives might be to transfer between RecA hosts a dimeric plasmid molecule containing two *oriT* sequences in direct repeat. This should yield some monomeric progeny molecules in which transfer was initiated at one *oriT* site and terminated at the second and these should contain duplications or deletions in the region of the nick sites. This prediction is currently being tested. A second way would be to develop an *in vitro* nicking system and re-measure the lengths of DNA molecules terminating at the *oriT* nick(s). In addition, such a system would allow confirmation that the *oriT* mutations lie in the endonuclease binding site, and enable further investigation of proteins binding to the *oriT* nick termini. Preliminary efforts in this direction have been unsuccessful (R. Everett and N. Willetts, unpublished data).

Either RNA polymerase or the primosome might initiate F conjugative DNA synthesis in donor and recipient cells (reviewed by Willetts and Wilkins, 1984). In this regard, we note the presence of a putative promoter oriented such that it could provide an RNA primer for replacement strand synthesis (-10 homology, 230-225; -35 homology, 255-250; Figure 2; Siebenlist *et al.*, 1980). We have as yet, however, no direct evidence that this promoter is active *in vivo*. Assembly of the primosome is initiated by interaction of the n' protein and a specific DNA site which includes a consensus sequence 5'GAAGCGG (Van der Ende *et al.*, 1983). The sequence 5'AAAGCGG is found at position 232-238 in the non-transferred strand of *oriT* (Figure 2) and, in common with known n' protein recognition sites it lies in the stem of a hairpin structure. There are no promoter or n' recognition sites in the 127 bp to the left of *oriT* that might initiate complementary strand synthesis in the recipient cells.

Small plasmids, such as ColE1, which do not carry complete transfer systems, nevertheless contain a transfer origin. When ColE1 is present in the same cell as a conjugative plasmid, ColE1-encoded mobility proteins ensure the interaction of the ColE1 transfer origin with the DNA transfer system of the conjugative plasmid to allow the efficient transfer of ColE1 DNA. The ColE1 transfer origin is thought to coincide with the plasmid relaxation complex nick site (Warren *et al.*, 1978) and this has been located on the ColE1 DNA sequence (Bastia, 1978). It falls within a short region with imperfect 2-fold rotational symmetry, unlike that of F. Figure 5b shows that a strongly conserved region (19/24 positions identical) lies 55 bp upstream of the ColE1 nick site and 142 bp upstream of the rightmost major λ *oriT* nick site. In both plasmids the conserved region falls within an inverted repeat region. The sequence in ColE1 is also highly conserved in the related plasmids ColK (J. Archer and D. Sherratt, personal communication) and CloDF13 (Snijders *et al.*, 1983).

Deletion of the ColE1 DNA to the right of the *HhaI* site in Figure 5b causes a 20-fold reduction in mobilisation frequency (Finnegan and Sherratt, 1982). The homology between F and ColE1 in this region may reflect a common F or host-encoded protein which interacts with both of these *oriT* sites. The sequence of the F *oriT* region is unrelated to the F *oriV* sequence, to the *cos* site of λ , to the inverted repeat sequences at the termini of the common transposable elements, to the gyrase recognition site or to the fl RF *oriV* sequence. Curiously, a sequence near the *oriT* mutations (19-159) is similar (9/11) to that flanking the ϕ X174 *oriV* nick site.

Materials and methods

Plasmids and phages

Sources of DNA fragments for sequencing were the plasmids pBE274 (Thompson and Achtman, 1978) and pED806 and its *oriT*⁻ mutants (Johnson *et al.*, 1981; Everett and Willetts, 1982). The λ tra transducing phage incorporating *oriT* was ED λ 102 (Everett and Willetts, 1980).

DNA sequencing techniques

End-labelled DNA fragments were sequenced using the chemical degradation method of Maxam and Gilbert (1977, 1980). For purine-specific cleavage, the modification described by Gray *et al.* (1978) was used. *oriT*⁻ mutants of pED806 were sequenced from the *Bgl*II site by the method of Maat and Smith, (1978). Sequence data were stored and analysed using the programs of Staden (1978) and Korn (1977).

Characterisation of the *oriT* nick termini

Nicked and un-nicked ED λ 102 DNA was prepared as described by Everett and Willetts (1980). Labelling experiments used either intact phage DNA or the 1.08-kb *Bgl*II fragment containing *oriT* which had been purified by electrophoresis through a 5% polyacrylamide gel followed by electroelution and ethanol precipitation.

3' Termini were labelled using the Klenow fragment of DNA polymerase I and appropriate single [α -³²P]deoxynucleoside triphosphates and 5' termini treated with bacterial alkaline phosphatase before labelling with polynucleotide kinase and [γ -³²P]ATP essentially as described (Maxam and Gilbert, 1980).

Acknowledgements

We are grateful to John Maule for excellent technical assistance and to Andrew Coulson and Duncan McGeoch for help with computer analysis of the sequence. This work was supported in part by MRC Programme Grant G978 88 to NSW.

References

- Abdel-Monem, M., Taucher-Scholz, G. and Klinkert, M. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 4659-4663.
- Achtman, M., Skurray, R.A., Thompson, R., Helmuth, R., Hall, S., Beutin, L. and Clark, A.J. (1978) *J. Bacteriol.*, **133**, 1383-1392.
- Bastia, D. (1978) *J. Mol. Biol.*, **124**, 601-639.
- Casjens, S. and Huang, W.M. (1982) *J. Mol. Biol.*, **157**, 287-298.
- Depew, R.E., Lin, L.F. and Wang, J.C. (1978) *J. Biol. Chem.*, **253**, 511-518.
- Desiderio, S.V. and Kelly, T.J. (1981) *J. Mol. Biol.*, **145**, 319-337.
- Everett, R. and Willetts, N. (1980) *J. Mol. Biol.*, **136**, 129-150.
- Everett, R. and Willetts, N. (1982) *EMBO J.*, **1**, 747-753.
- Feiss, M., Kobayashi, I. and Widner, W. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 955-959.
- Finnegan, J. and Sherratt, D.J. (1982) *Mol. Gen. Genet.*, **185**, 344-351.
- Gaffney, D., Skurray, R. and Willetts, N. (1983) *J. Mol. Biol.*, **168**, 103-122.
- Gray, C.P., Sommer, R., Polke, C., Beck, E. and Schaller, H. (1978) *Proc. Natl. Acad. Sci. USA*, **75**, 50-53.
- Ikeda, J., Yudelovich, A. and Hurwitz, J. (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 2669-2673.
- Johnson, D.A., Everett, R. and Willetts, N.S. (1981) *J. Mol. Biol.*, **153**, 187-202.
- Kingsman, A. and Willetts, N. (1978) *J. Mol. Biol.*, **122**, 287-300.
- Korn, L.J. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 4404-4405.
- Maat, J. and Smith, A.J.H. (1978) *Nucleic Acids Res.*, **5**, 4537-4545.
- Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560-564.
- Maxam, A.M. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499-580.
- Meijer, M., Beck, E., Hansen, F.G., Bergmans, H., Messer, W., von Meyen-

- burg, K. and Schaller, H. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 580-584.
- Moore, D.D., Denniston-Thompson, K., Kruger, K.E., Furth, M.E., Williams, B.G., Daniels, D.L. and Blattner, F.R. (1978) *Cold Spring Harbor Symp. Quant. Biol.*, **43**, 155-163.
- Morrison, A. and Cozzarelli, N.R. (1979) *Cell*, **17**, 175-184.
- Murotsu, T., Matsubara, K., Sugisaki, H. and Takanami, M. (1981) *Gene*, **15**, 257-271.
- Selzer, G., Som, T., Itoh, T. and Tomizawa, J. (1983) *Cell*, **32**, 119-129.
- Shafferman, A. and Helinski, D.R. (1983) *J. Biol. Chem.*, **258**, 4083-4090.
- Siebenlist, U., Simpson, R.B. and Gilbert, W. (1980) *Cell*, **20**, 269-281.
- Snijders, A., Van Putten, A.J., Veltkamp, E. and Nijkamp, H.J.J. (1983) in press.
- Staden, R. (1978) *Nucleic Acids Res.*, **5**, 1013-1015.
- Stormo, G.D., Schneider, T.D. and Gold, L.M. (1982) *Nucleic Acids Res.*, **10**, 2971-2996.
- Sugimoto, K., Oka, A., Sugisaki, H., Takanami, M., Nishimura, A., Yasuda, Y. and Hirota, Y. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 575-579.
- Tamanoi, F. and Stillman, B.W. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 2221-2225.
- Thompson, R. and Achtman, M. (1978) *Mol. Gen. Genet.*, **165**, 295-304.
- Thompson, R. and Achtman, M. (1979) *Mol. Gen. Genet.*, **169**, 49-57.
- Thompson, R. and Taylor, L. (1982) *Mol. Gen. Genet.*, **188**, 513-518.
- Tse, Y.-C., Kirkegaard, K. and Wang, J.C. (1980) *J. Biol. Chem.*, **255**, 5560-5565.
- Van der Ende, A., Teerstra, R., Van der Avoort, H.G.A.M. and Weisbeek, P.J. (1983) *Nucleic Acids Res.*, **11**, 4957-4975.
- Warren, W.J., Twigg, A.J. and Sherratt, D.J. (1978) *Nature*, **274**, 259-261.
- Willetts, N.S. (1972) *J. Bacteriol.*, **112**, 773-778.
- Willetts, N. and Skurray, R. (1980) *Annu. Rev. Genet.*, **14**, 41-76.
- Willetts, N. and Wilkins, B. (1984) *Microbiol. Rev.*, in press.

Received on 6 February 1984