
Rearrangement of *nif* Genes During Cyanobacterial Heterocyst Differentiation

R. Haselkorn, J. W. Golden, P. J. Lammers and M. E. Mulligan

Phil. Trans. R. Soc. Lond. B 1987 **317**, 173-181

doi: 10.1098/rstb.1987.0055

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

To subscribe to *Phil. Trans. R. Soc. Lond. B* go to: <http://rstb.royalsocietypublishing.org/subscriptions>

Rearrangement of *nif* genes during cyanobacterial heterocyst differentiation

BY R. HASELKORN, J. W. GOLDEN†, P. J. LAMMERS‡ AND M. E. MULLIGAN

Department of Molecular Genetics and Cell Biology, The University of Chicago, Illinois 60637, U.S.A.

Anabaena is a filamentous cyanobacterium that produces specialized cells, called heterocysts, at regular intervals along each filament when deprived of fixed nitrogen under aerobic conditions. Heterocysts are anaerobic factories for nitrogen fixation. In *Anabaena* vegetative-cell DNA, the *nifD* gene, encoding the α subunit of nitrogenase, is interrupted by an 11000 base pair DNA element. During the differentiation of heterocysts from vegetative cells, this 11 kilobase (kb) element is excised by site-specific recombination between short, directly repeated DNA sequences present at the ends of the element. The excision results in restoration of the *nifD* coding sequence and of the entire *nifHDK* transcription unit. A gene has been identified, within the 11 kb element, that is believed to encode the site-specific recombinase responsible for excision of the element during heterocyst differentiation. A second developmentally regulated gene arrangement has also been observed in *Anabaena*. This event occurs close to the *nifS* gene and involves a different set of repeated sequences, implying a different site-specific recombination system.

1. INTRODUCTION

Anabaena is a representative of many filamentous cyanobacteria that grow in chains of 50–100 uniform vegetative cells as long as the environment contains combined nitrogen (usually nitrate or ammonia). When the external supply of combined nitrogen is exhausted, *Anabaena* differentiates cells specialized for nitrogen fixation at regular intervals along each filament (figure 1). These cells, called heterocysts, provide the anaerobic environment necessary for nitrogen fixation (Haselkorn 1978; Wolk, 1982). The differentiation requires roughly one cell generation to complete (*ca.* 24–30 h). During that time, a new cell envelope consisting of an outer polysaccharide layer and an inner glycolipid layer is deposited outside the pre-existing cell wall; proteins involved in oxygen evolution and CO₂ fixation are destroyed by protease; and enzymes for nitrogen fixation and the production of ATP and reductant are induced.

Intercellular communication is critical for this arrangement to work. Vegetative cells continue to fix CO₂; carbohydrate is transported from them to the heterocyst to provide the substrates for the manufacture of ATP and reductant. The ultimate product of nitrogen fixation in the heterocyst is glutamine, which is exported to neighbouring vegetative cells. In spite of this extensive transport of small molecules between the two cell types, gene expression in each cell is radically different. Many genes, including those that encode the proteins of the nitrogenase complex, are expressed exclusively in heterocysts; others are turned off selectively in heterocysts. Heterocysts do not divide. In a culture fixing nitrogen, a vegetative cell midway between two heterocysts will differentiate in roughly the time required for each vegetative cell

† Present address: Department of Biology, Texas A & M University, College Station, Texas 77843, U.S.A.

‡ Present address: Department of Chemistry and the Plant Genetic Engineering Laboratory, New Mexico State University, Las Cruces, New Mexico 88003, U.S.A.

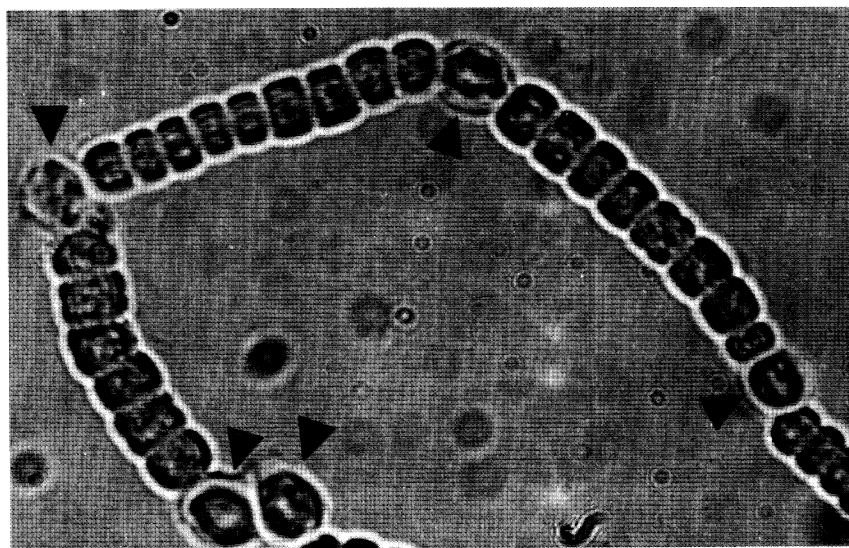


FIGURE 1. Differentiated filament of *Anabaena* 7120. The phase-contrast micrograph shows the characteristic spacing of thick-walled heterocysts (arrows). (Photograph kindly provided by R. Nagaraja.)

to divide once. This relationship between division and differentiation maintains the heterocyst spacing shown in figure 1.

We have been engaged for some time in cloning and sequencing *Anabaena* genes that are regulated differentially during heterocyst development in order to determine the molecular mechanisms involved in the control of their expression (Rice *et al.* 1982; Tumer *et al.* 1983; Nierzwicki-Bauer *et al.* 1984). In the course of these studies, we observed that DNA containing several of the genes required for nitrogen fixation is rearranged in heterocysts during their differentiation (Golden *et al.* 1985).

2. *nif* GENE ORGANIZATION

The nitrogen-fixation genes of the best-studied diazotroph, *Klebsiella pneumoniae*, are organized in eight operons clustered within 24 kb of DNA. One of these operons contains three genes, *nifH*, *nifD* and *nifK*, which encode the polypeptide components of nitrogenase reductase and the α and β subunits of nitrogenase, respectively. Other operons encode proteins required for electron transport to nitrogenase, for synthesis of the molybdenum cofactor of nitrogenase, and for other proteins involved in assembly or 'maturation' of the active enzyme complex. One of the latter is encoded by the *nifS* gene, located several genes downstream of the *nifHDK* operon (Dixon *et al.*, this symposium).

The *nif* genes of most other diazotrophs are organized in ways that are related to the *Klebsiella* arrangement. In *Azotobacter*, for example, the operon structure appears to be similar to that of *Klebsiella* but the operons are spread out, separated by stretches of non-*nif* DNA (Kennedy *et al.*, this symposium). The *nifHDK* operon, which includes the three genes encoding the polypeptide components of the nitrogenase complex, has been found to be conserved in nearly every other diazotroph. The exceptions discovered to date are *Rhizobium japonicum*, in which the gene *nifH* is some 15 kb away from the *nifDK* genes (Kaluza & Hennecke 1984) and *Anabaena*, described below.

Cloned fragments of *Klebsiella nif* DNA were used to identify and to isolate homologous fragments of *Anabaena* vegetative-cell DNA from recombinant DNA libraries (Mazur *et al.* 1980). Physical mapping of the *Anabaena nif* DNA showed it to differ from the *Klebsiella* gene organization in several ways (Rice *et al.* 1982). First, although the *Anabaena nifH* and *nifD* genes are adjacent, as in *Klebsiella*, the *nifK* gene appeared to be about 11 kb away from the *nifD* gene (figure 2). Second, a gene homologous to *nifS* is to the right of *nifH* in *Anabaena*, quite near, and not apparently neighbour to, other *nif* genes. In all, a region of over 30 kb of *Anabaena* vegetative-cell DNA was found to contain only four *nif* genes. The organization of the *nif* genes and the fact that heterocysts are terminally differentiated in *Anabaena* prompted the comparison of heterocyst and vegetative-cell DNA in the *nif* region of the chromosome.

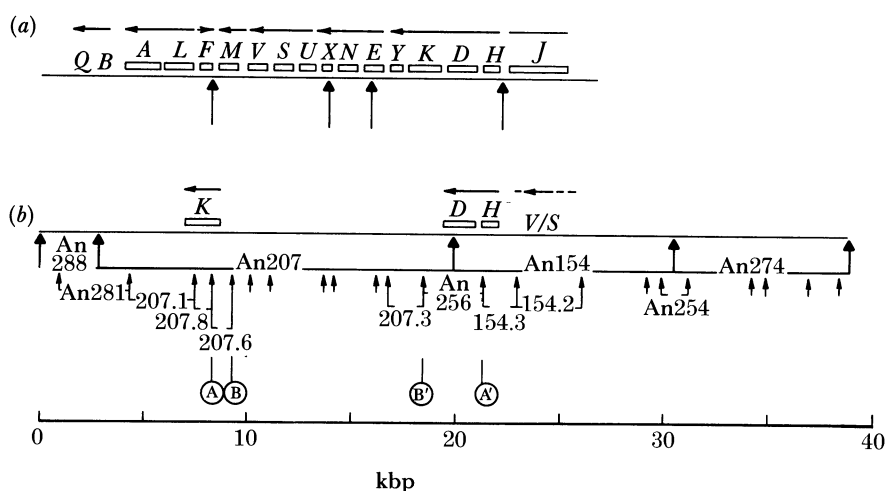


FIGURE 2. Comparison of the organization of the *nif* genes of *Klebsiella pneumoniae* (a) and *Anabaena* vegetative cells (b). Thick arrowheads indicate *Eco*RI sites, slender arrowheads *Hind*III sites. Horizontal arrows indicate direction of transcription. The *nifV/S* region of *Anabaena* was identified by hybridization; recent sequencing data indicate that the *nifS* gene is located there (M. E. Mulligan, unpublished). Numbers underneath the *Anabaena* map are clone designations. The *nifD* rearrangement involves recombination between sites located in the fragments 207.6 (A-B) and 256 (A'-B'). The *nifS* rearrangement involves one site to the right of *nifS* in 154.2 and a second distant site.

3. REARRANGEMENT OF *nifD* DURING HETEROCYST DIFFERENTIATION

The comparison of heterocyst and vegetative-cell DNA was made possible by the isolation of DNA, of relatively high molecular mass, from heterocysts. Heterocysts may be separated from vegetative cells by treating the filaments with lysozyme followed by mild sonication and differential centrifugation. The glycolipid-covered heterocyst cell walls are resistant to lysozyme digestion and physical disruption. These properties allow heterocysts to be separated from vegetative cells, but they also hamper the isolation of DNA from them. A procedure used for the preparation of RNA from *Anabaena* vegetative cells was found also to produce relatively high-molecular-mass DNA from purified heterocysts (Golden *et al.* 1985). This procedure made it possible to compare heterocyst and vegetative-cell DNA by Southern analysis, revealing differences in the sizes of the hybridizing fragments by using several of the *Anabaena nif* gene probes. The *Eco*RI fragments that hybridized with *nifK* and *nifH* probes were reduced from 17 kb and 10.5 kb, respectively, in the vegetative-cell genome (figure 2) to 6 kb and 6 kb,

respectively, in the heterocyst genome (Golden *et al.* 1985). Further Southern analysis indicated that two different DNA rearrangements occur during heterocyst differentiation, one with breakpoints near *nifK* and *nifD* and a second with breakpoints near *nifS* and elsewhere in the genome.

Both rearrangements occur at approximately the same late stage of normal heterocyst development, after morphological changes are readily apparent and at about the same time that the *nif* messages are detected by Northern analysis. Therefore the DNA rearrangements are a consequence of differentiation, not of a triggering event. The time course of the two rearrangements is illustrated in figure 3. DNA was prepared from vegetative cells, from a differentiating culture at six-hour intervals, and from purified heterocysts. In figure 3*a*, the DNA samples were digested with *Hind*III and probed with the *Hind*III fragment containing *nifD* labelled A'-B' in figure 2. This probe identifies itself in vegetative cells and in the 90% of the cells that do not differentiate. However, in the 10% of cells that do differentiate, the 2.9 kb band corresponding to A'-B' disappears, to be replaced quantitatively by two new bands at 2.1 and 1.8 kb. These new bands correspond to B-B' and A-A', respectively (see figure 2). The band B-B' is derived from an 11 kb circular molecule that is excised from the vegetative cell chromosome. The band A-A' comes from the fused *nifHDK* operon produced in the heterocyst genome as a consequence of the excision.

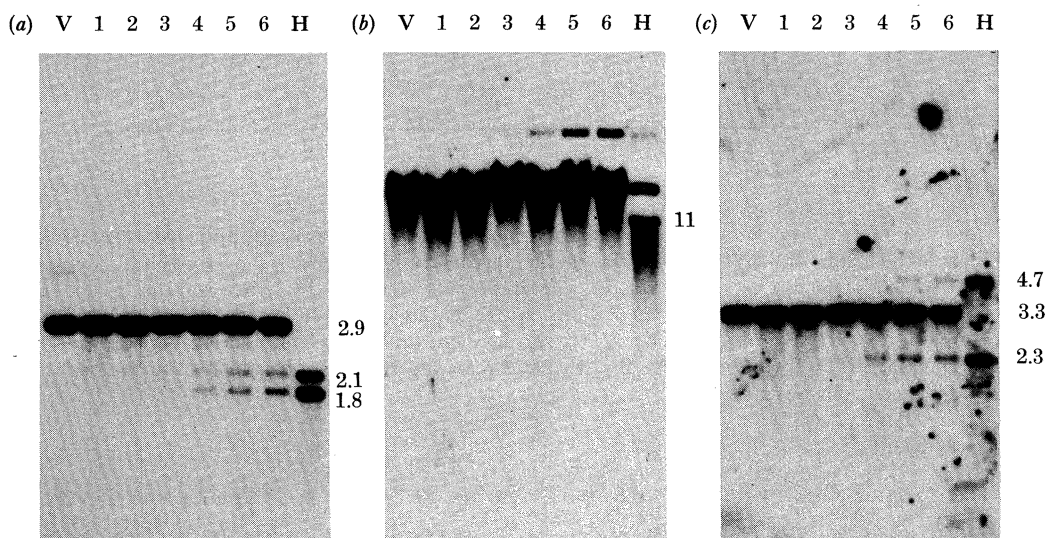


FIGURE 3. Time course of DNA rearrangement and location of DNA breakpoints. *Hind*III-digested or undigested DNA was fractionated on 0.7% agarose gels and blotted onto nitrocellulose. (*a*), *Hind*III-digested DNA probed with 256; (*b*), undigested DNA probed with 207.3; (*c*), *Hind*III-digested DNA probed with 154.2. (See figure 2 for a map of the probes.) DNA was isolated from vegetative cells (V), heterocysts (H), and at 6 h intervals during synchronous heterocyst differentiation (lanes 1-6). Lanes 1-6 contain total DNA prepared 6, 12, 18, 24, 30 and 42 h after initiation of nitrogen starvation, respectively.

Figure 3*b* shows the same DNA samples subjected to electrophoresis without prior enzyme digestion. In this case the probe is a *Hind*III fragment within the 11 kb excised region. The multiple non-chromosomal bands seen in the differentiating cells are isomeric forms of the circular 11 kb excised element (Golden *et al.* 1985). Comparison of figure 3*a* and *b* shows that the excision of the 11 kb circle and appearance of the two new *Hind*III fragments occur at the

same advanced stage of heterocyst differentiation. *nifHDK* messenger RNA is detectable only after the rearrangement has occurred. Morphological differentiation is complete by this time and the local environment in the heterocyst is likely to be anaerobic.

The 11 kb circular DNA molecule is excised from *within* the 3' end of the *nifD* gene (figures 4 and 5). The excised molecule, for convenience referred to as the 'excision', persists in heterocysts but has no known function. Recombinant libraries of heterocyst DNA were constructed and used to clone the two products of the *nifD* rearrangement: the 6 kb *Eco*RI fragment from the fused chromosome and the entire 11 kb excision. Restriction analysis of these heterocyst clones identified the regions containing the DNA breakpoints. Sequence analysis of the breakpoints showed that the *nifD* rearrangement was the result of a conservative site-specific recombination between two directly repeated 11 bp sequences (discussed below) separated by 11 kb in the vegetative-cell chromosome. The result of the recombination is the excision of the 11 kb element as a circular molecule and the fusion of the chromosome (figure 4).

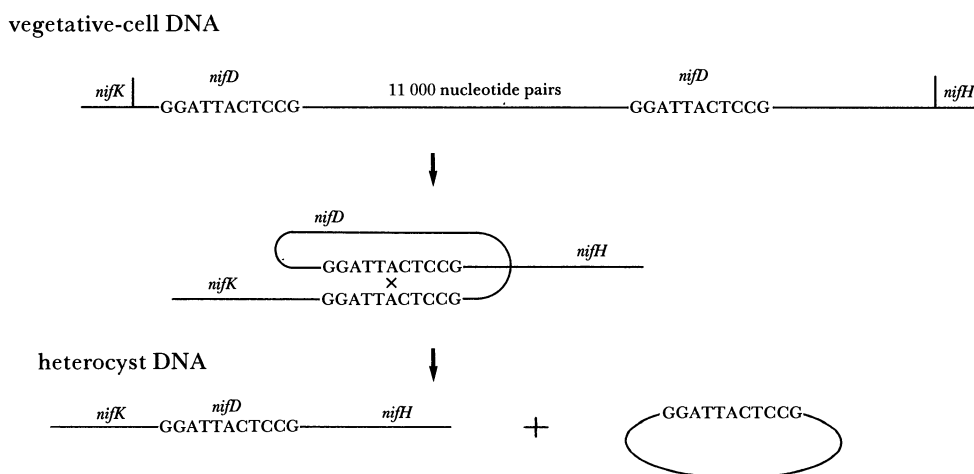


FIGURE 4. Proposed mechanism of *nifD* rearrangement. The two 11 bp directly repeated sequences that flank the excision in vegetative-cell DNA form a synapse and recombine within the repeats. The reaction products are two unlinked DNA molecules, each retaining one copy of the 11 bp repeat.

The rearrangement has two effects on the expression of the *nif* genes. One is the formation of a *nifHDK* operon. Northern analysis with *nifH*, *D*, and *K* probes identifies a common 4.7 kb RNA containing the transcript of all three genes. After rearrangement the *nifK* gene is expressed from the *nifH* promoter. It is not known whether the *nifK* gene can be expressed in the unrearranged vegetative cell genome.

The second effect of the *nifD* rearrangement is more surprising. One of the 11 bp repeats is about 330 bases 5' to the *nifK* gene. The other repeat is *within* the coding region for the *nifD* gene. The rearrangement results in the replacement of 27 amino acids encoded by the apparent vegetative cell gene with 43 amino acids encoded by DNA sequences originally 11 kb away. Comparison of the predicted carboxy-terminal amino acid sequences of two species of *Rhizobium* with the sequence derived from unrearranged vegetative cell DNA shows no homology, but there is very good homology with the rearranged heterocyst sequence. This information, along with Western blot data showing only one form of the *nifD* protein (R. Kranz and J. Golden, unpublished), indicates that only the rearranged *nifD* gene is expressed in *Anabaena*.

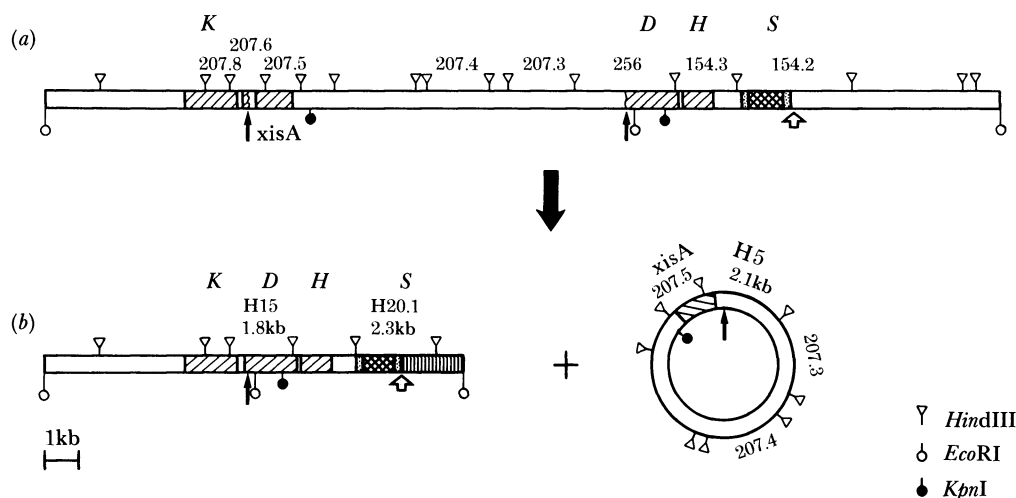


FIGURE 5. Organization of nitrogen fixation genes in *Anabaena* vegetative-cell and heterocyst DNA. (a) Two contiguous *EcoRI* fragments of vegetative-cell DNA containing the *nifK*, *nifD*, *nifH*, and *nifS* genes as well as an 11 kb element interrupting the *nifD* gene. At the left end of the element is the *xisA* gene. (b) The rearranged DNA in heterocysts. Recombination between directly-repeated 11 bp sequences at the end of the 11 kb element, indicated by solid black arrows, results in excision of the 11 kb element as a circle and restoration of the *nifD* reading frame and the *nifHDK* operon (figure 4). A second rearrangement occurs with one breakpoint just to the right of the *nifS* gene (open arrow). The *nifS* open reading frame is oriented right to left and begins very close to the breakpoint (M. E. Mulligan, unpublished). Small numbers indicate fragment sizes or the designation of cloned *HindIII* fragments.

4. REARRANGEMENT NEAR *nifS*

The second arrangement has one DNA breakpoint close to the 5' end of the *nifS* gene (figures 2 and 5). When a *HindIII* fragment, An154.2, containing the *nifS* gene is used as a probe against vegetative cell and heterocyst DNA digested with *HindIII*, it hybridizes to itself in vegetative-cell DNA but is replaced by two new fragments in heterocyst DNA (figure 3c). The DNA breakpoints involved in the *nifS* rearrangement have been cloned, two from the vegetative cell genome and two from the heterocyst genome.

All four of the *nifS* rearrangement breakpoints have been sequenced (Golden *et al.* 1987). The sequence data show that this rearrangement, like the *nifD* excision, is a conservative site-specific recombination between two repeated DNA sequences. The repeated sequences involved in the *nifD* and *nifS* rearrangements show no homology to each other, and presumably are catalysed by different enzymes (figure 6). The *nifD* 11 bp direct repeats are flanked by regions of lesser homology on both sides. The 5 bp repeated sequences found at the *nifS* breakpoints are also flanked by regions of partial homology, more extensive on one side than on the other (figure 6). Both the *nifD* and the *nifS* breakpoint sequences contain small (3 bp) inverted repeats, indicated by arrows in figure 6. Apart from these general structural features, the breakpoints bear no resemblance to the target sequences of known bacterial recombinases such as the integrase of bacteriophage lambda (Campbell 1983) or the enzyme that inverts a DNA segment in *Salmonella* (Silverman & Simon 1983).

No other rearrangements have been found in heterocyst DNA. No differences between heterocyst and vegetative-cell DNA can be detected in the ethidium bromide staining patterns of restriction digests seen on agarose gels. Other probes hybridized to Southern blots of

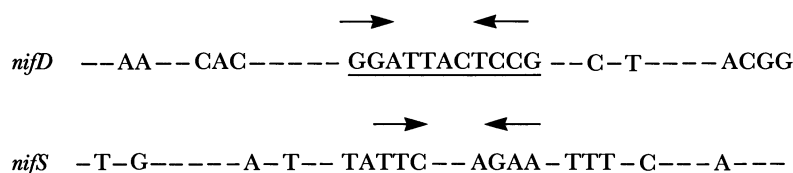


FIGURE 6. Comparison of the conserved nucleotide sequences at which the *nifD* and *nifS* elements rearrange. They share certain symmetry features (see text) but differ sufficiently to require separate site-specific recombination enzyme systems. Recombination occurs within the underlined sequences.

vegetative-cell and heterocyst DNA have shown identical banding patterns. These probes include *glnA*, encoding glutamine synthetase (Tumer *et al.* 1983), and *psbA*, encoding the Q_B protein of photosystem II (Curtis & Haselkorn 1984).

5. THE *nifD* EXCISION SITE-SPECIFIC RECOMBINASE GENE, *xisA*

The gene believed to encode the site-specific recombinase responsible for the excision of the *nifD* 11 kb element has been identified and sequenced (Lammers *et al.* 1986). It was suspected that the recombinase, which we call excisase, might be on the excision. For this reason a large-scale plasmid preparation was made of pAn207, which consists of the vector pBR322 containing the 17 kb *EcoRI* fragment of *Anabaena* DNA on which the excision is found (Rice *et al.* 1982). This plasmid preparation was found to contain, in addition to the expected 21 kb plasmid, a second plasmid 11 kb smaller. Restriction and Southern analysis of these plasmids showed that the smaller plasmid resulted from the loss of the *nifD* excision from the larger plasmid. Sequence analysis of the rearranged plasmid confirmed that the 11 kb excision had been precisely deleted from the cloned *Anabaena* DNA by recombination within the 11 bp repeated sequence.

The loss of the excision from pAn207 during propagation in *E. coli* occurred only very rarely. To monitor the excision, the excision was marked with a β -galactosidase gene *in vivo* by the insertion of MudI1734, a mini-Mu-*lac* (Castilho *et al.* 1984). This strategy also provided a method of insertional mutagenesis. A number of insertions into the excision were obtained. One of these, MX25, had the mini-Mu-*lac* inserted in the middle of the excision and was found to produce blue colonies (indicating β -galactosidase activity) with white sectors and an occasional pure white colony on agar medium containing the chromogenic indicator X-gal. When plasmid DNA was isolated from white colonies and analysed by restriction mapping, it was found to have lost the excision containing the mini-Mu-*lac*. When the excision is deleted from a plasmid it is no longer connected to an origin of replication and is eventually lost from the cells.

A different mini-Mu-*lac* insertion, MX32, was mapped inside the excision near the *nifK* 11 bp repeat. This insertion failed to give rise to white colonies. The MX32 insertion could be complemented *in trans* with a compatible plasmid containing a *HincII* fragment of *Anabaena* DNA which covered the region of the MX32 insertion (Lammers *et al.* 1986). The second plasmid restored the ability of MX32 to delete the excision containing the mini-Mu-*lac*, that is, to form white colonies on X-gal plates.

The MX32 insertion disrupted the excisase gene, *xisA*, necessary for deletion of the *nifD* element in *E. coli*. This region of the *nifD* excision was sequenced and found to contain an open reading frame of 1062 bp that codes for a 41.6 kDa polypeptide. The predicted protein is basic with a net charge of +18, consistent with expectations for a putative DNA-binding protein.

The protein encoded by *xisA* is essential for the precise deletion of the excision in *E. coli* and is very likely the site-specific recombinase for the *nifD* rearrangement in *Anabaena*.

The determination of the exact role of *xisA* in the *nifD* rearrangement requires additional experiments including chromosomal inactivation of the gene in *Anabaena* and an *in vitro* assay, for rearrangement, based on purified *xisA* gene product. The different site specificity between the *nifD* and *nifS* rearrangements suggests that they require different site-specific recombinases. At present nothing is known about the enzyme that catalyses the *nifS* rearrangement.

6. CONCLUSIONS

The rearrangement of the nitrogen-fixation genes during *Anabaena* heterocyst differentiation offers a unique example of prokaryotic gene regulation. The *nifD* rearrangement discovered in *Anabaena* probably occurs in many heterocystous cyanobacteria. Kallas *et al.* (1985) have found that the organization of the *nifHDK* genes differs between heterocystous and non-heterocystous cyanobacteria. All non-heterocystous nitrogen-fixing cyanobacteria examined to date have a contiguous *nifHDK* gene cluster, whereas all of the heterocystous strains show a separation of *nifK* from the *nifDH* genes. Rearrangement of the heterocyst genome may restrict expression of the nitrogen-fixation genes to those heterocysts that have completed development.

Although DNA rearrangement is necessary for the proper expression of the *nif* genes in *Anabaena*, our present knowledge does not indicate why this particular mechanism has been used. The deletion of the *nifD* excision from the chromosome appears to be irreversible, and therefore may contribute to the commitment of vegetative cells to terminal differentiation. Even less is known about the regulation of the *nifS* rearrangement and its effects on *nif* gene expression. The topology of the *nifS* rearrangement and the identity of the genes that are affected by the rearrangement should be known soon. It will then be important to determine how the two rearrangements are coordinated during heterocyst differentiation.

REFERENCES

- Campbell, A. 1983 Bacteriophage lambda. In *Mobile genetic elements* (ed. J. A. Shapiro), pp. 65–103. New York: Academic Press.
- Castilho, B. A., Olsson, P. & Casadaban, M. J. 1984 Plasmid insertion mutagenesis and *lac* gene fusion with mini-Mu bacteriophage transposons. *J. Bact.* **158**, 488–495.
- Curtis, S. E. & Haselkorn, R. 1984 Organization and transcription of genes coding for 32 kd thylakoid membrane proteins in *Anabaena*. *Pl. Molec. Biol.* **3**, 249–258.
- Golden, J. W., Mulligan, M. E. & Haselkorn, R. 1987 Different recombination site specificity of two developmentally regulated genome rearrangements. *Nature, Lond.* **327**, 526–529.
- Golden, J. W., Robinson, S. J. & Haselkorn, R. 1985 Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena*. *Nature, Lond.* **314**, 419–423.
- Haselkorn, R. 1978 Heterocysts. *A. Rev. Pl. Physiol.* **29**, 319–344.
- Kallas, T., Coursin, T. & Rippka, R. 1985 Different organization of *nif* genes in nonheterocystous and heterocystous cyanobacteria. *Pl. Molec. Biol.* **5**, 321–329.
- Kaluza, K. & Hennecke, H. 1984 Fine structure analysis of the *nifDK* operon encoding the α and β subunits of dinitrogenase from *Rhizobium japonicum*. *Molec. gen. Genet.* **196**, 35–42.
- Lammers, P. J., Golden, J. W. & Haselkorn, R. 1986 Identification and sequence of a gene required for a developmentally regulated DNA excision in *Anabaena*. *Cell* **44**, 905–911.
- Mazur, B. J., Rice, D. & Haselkorn, R. 1980 Identification of blue-green algal nitrogen fixation genes by using heterologous DNA hybridization probes. *Proc. Natn. Acad. Sci. U.S.A.* **77**, 186–190.
- Nierzwicki-Bauer, S. A., Curtis, S. E. & Haselkorn, R. 1984 Cotranscription of genes encoding the small and large subunits of RuP₂ carboxylase in the cyanobacterium *Anabaena* 7120. *Proc. Natn. Acad. Sci. U.S.A.* **81**, 5961–5965.

CYANOBACTERIAL *NIF* GENES

181

- Rice, D., Mazur, B. J. & Haselkorn, R. 1982 Isolation and physical mapping of nitrogen fixation genes from the cyanobacterium *Anabaena* 7120. *J. biol. Chem.* **257**, 13157–13163.
- Silverman, M. & Simon, M. 1983 Phase variation and related systems. In *Mobile genetic elements* (ed. J. A. Shapiro), pp. 537–557. New York: Academic Press.
- Tumer, N. E., Robinson, S. J. & Haselkorn, R. 1983 Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen. *Nature, Lond.* **306**, 337–342.
- Wolk, C. P. 1982 Heterocysts. In *The biology of Cyanobacteria* (ed. N. G. Carr & B. A. Whitton), pp. 359–386. Oxford: Blackwell Scientific Publications.

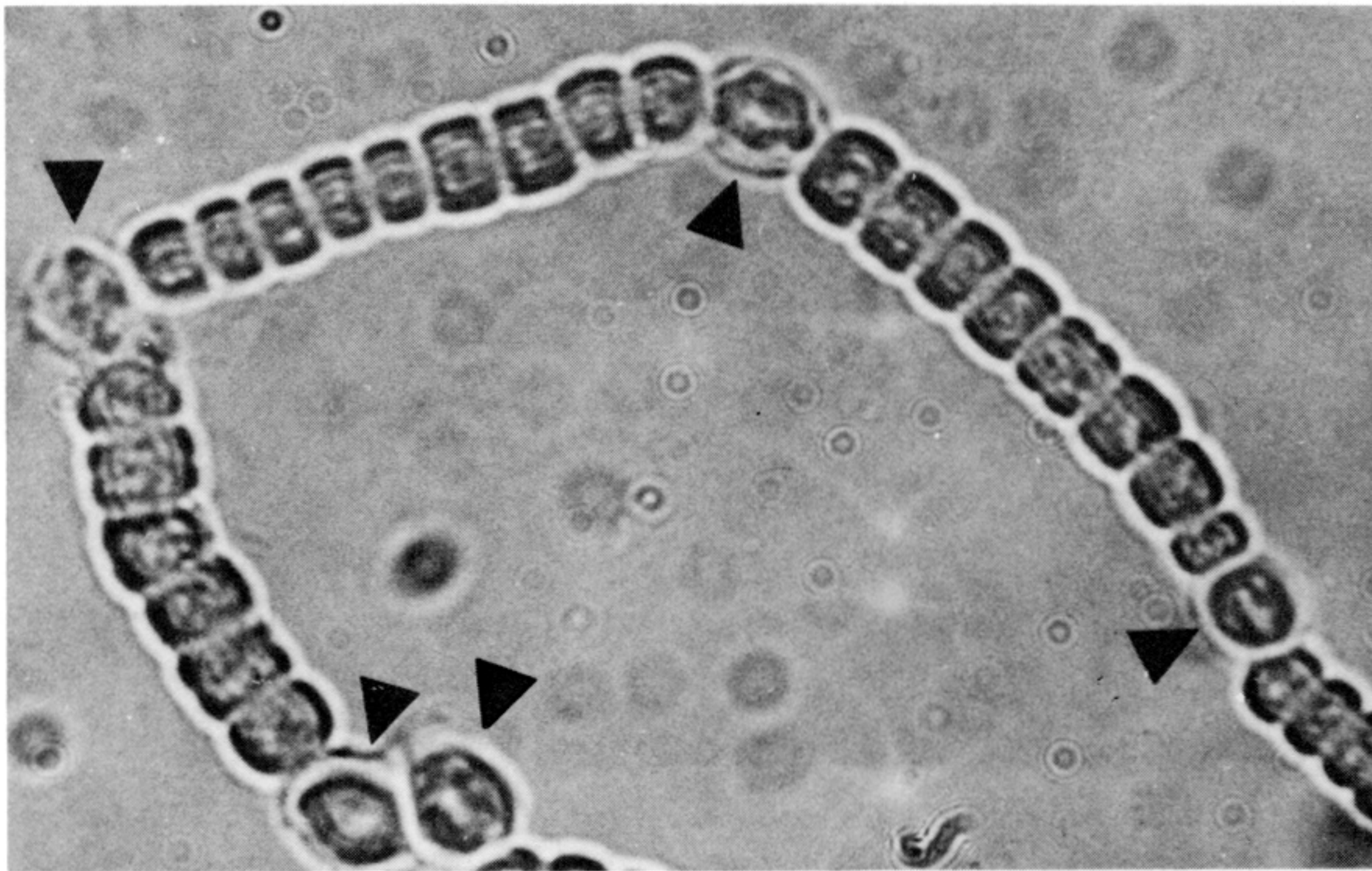


FIGURE 1. Differentiated filament of *Anabaena* 7120. The phase-contrast micrograph shows the characteristic spacing of thick-walled heterocysts (arrows). (Photograph kindly provided by R. Nagaraja.)

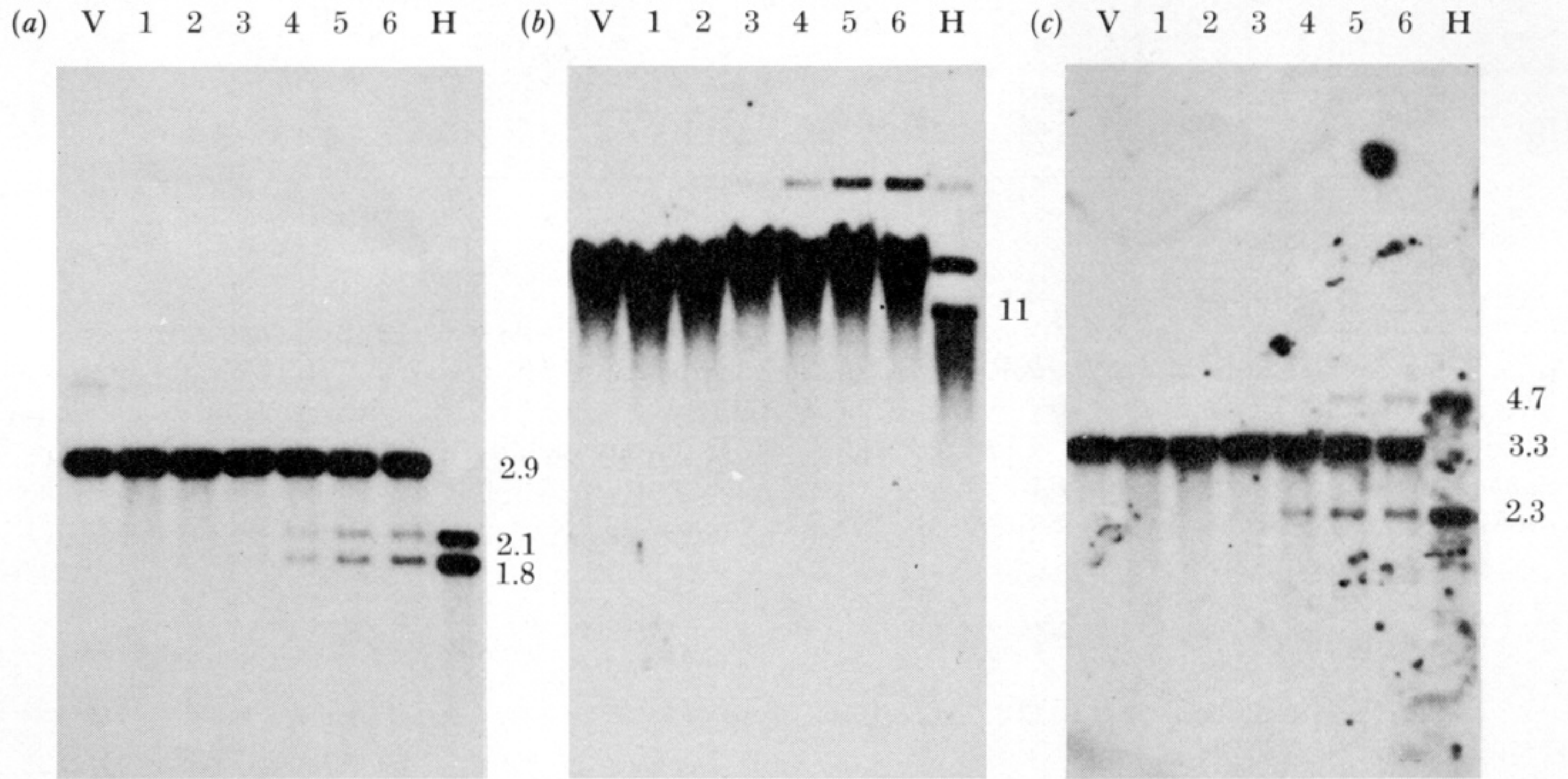


FIGURE 3. Time course of DNA rearrangement and location of DNA breakpoints. *Hind*III-digested or undigested DNA was fractionated on 0.7% agarose gels and blotted onto nitrocellulose. (a), *Hind*III-digested DNA probed with 256; (b), undigested DNA probed with 207.3; (c), *Hind*III-digested DNA probed with 154.2. (See figure 2 for a map of the probes.) DNA was isolated from vegetative cells (V), heterocysts (H), and at 6 h intervals during synchronous heterocyst differentiation (lanes 1–6). Lanes 1–6 contain total DNA prepared 6, 12, 18, 24, 30 and 42 h after initiation of nitrogen starvation, respectively.