J. Biosci., Vol. 19, Number 5, December 1994, pp 579–602. © Printed in India.

Rearrangements of nitrogen fixation (*nif*) genes in the heterocystous cyanobacteria

S K APTE* and N PRABHAVATHI
Molecular Biology and Agriculture Division, Bhabha Atomic Research Centre, Trombay, Bombay 400 085, India

MS received 9 September 1994; revised 29 October 1994

Abstract. In the vegetative cells of heterocystous cyanobacteria, such as *Anabaena,* two Operons harbouring the nitrogen fixat[i]on (*nif*) genes contain two separate intervening DNA elements resulting in the dispersion of genes and impaired gene expression. A 11 kb element disrupts the *nifD* gene in the *nifH, D-K* operon. It contains a 11 bp sequence (GGATTACTCCG) directly repeated at its ends and harbours a gene, *xisA,* which encodes a site-specific recombinase. A large 55 kb element interrupts the *fdxN* gene in the *nifB fdxN-nifS-nifU* operon. It contains two 5 bp direct repeats (TATTC) at its ends and accommodates at least one gene, *xisF,* which encodes another site-specific recombinase. During heterocyst differentiation both the discontinuities are precisely excised by two distinct site-specific recombination events. One of them is brought about by the XisA protein between the 11 bp direct repeats. The second one is caused by the XisF protein and occurs between the 5 bp direct repeats. As a consequence the 11kb and 55 kb elements are removed from the chromosome as circles and functional *nif* Operons are created. Nitrogenase proteins are then expressed from the rearranged genes in heterocysts and aerobic nitrogen fixation ensues. How these elements intruded the *nif* genes and how and why are they maintained in heterocystous cyanobacteria are exciting puzzles engaging considerable research effort currently. The unique developmental regulation of these gene rearrangements in heterocystous cyanobacteria is discussed.

Keywords. Cyanobacteria; heterocysts; *nif* genes; gene rearrangements; developmental regulation.

1. Introduction

Cyanobacteria, formerly called blue-green algae, are a fascinating group of photosynthetic bacteria many of which also fix atmospheric dinitrogen. Both academic and applied aspects of the biology of these organisms are studied extensively. They come in various shapes and sizes, some with and others without the specialized cells called heterocysts. Cyanobacteria originated three billion years ago and are supposed to have contributed significantly to the oxygenation of the primitive Earth's atmosphere (Schopf 1975). Chloroplasts of higher plants are believed to have evolved through endosymbiosis of cyanobacteria and plant cells (Walsby 1986). In sharp contrast to their "primitive antecedents" heterocystous cyanobacteria display remarkable evolutionary advances such as differentiation, pattern formation, intercellular communication, physiological division of labour among cell types, developmentally regulated gene rearrangements and gene expression and a range

*Corresponding author.

of adaptive responses of survival value resulting in their ubiquitous distribution (Apte 1992, 1993). Among the nitrogen-fixing organisms, they occupy a unique position in being the only truly photoautotrophic aerobic nitrogen fixers and are of great consequence as nitrogen biofertilizers in nature, especially in the oceans (Capone and Carpenter 1982) and in tropical rice cultivation (Venkataraman 1979). Only the heterocystous forms can fix free nitrogen in air because they differentiate heterocysts—cells designed to exclude oxygen—and therefore conducive to the activities of the oxygen-sensitive nitrogenase enzyme (Wolk 1982). Curiously, in heterocystous forms two Operons containing the nitrogen fixation genes are interrupted and have to be rearranged for their functional expression (Golden *et al* 1985). These rearrangements are the subject matter of this review.

Gene rearrangements are very common in bacteria (Berg and Howe 1989). Transposing and inverting elements are found in several bacteria and lead to gene rearrangements in a small proportion of the population. Cells containing rearranged genes in some cases acquire a surface property that allows them to evade surveillance or to respond to a new environment amicably. Gene rearrangements are also quite common among eukaryotes, the best known example being that of rearrangement of the immunoglobulin genes in human lymphocytes (Craig 1988; Blackwell and Alt 1989). Another example that has aroused considerable interest is the association of rearrangements in oncogenes with the incidence of certain types of cancers (Taub *et al* 1982). Apart from the basic interest in understanding the mechanisms involved in such genetic recombinations and their cellular physiology, gene rearrangements are becoming important in devising strategies for gene targeting and gene therapy.

Cyanobacterial gene rearrangements, although they resemble some of the site-specific recombination events found in bacteria and in higher organisms, are quite different in certain respects (Haselkorn 1992). They are the only known rearrangements within nitrogen fixation genes of a diazotroph (Golden *et al* 1985). Also, they are two of the only three known developmentally regulated gene rearrangements in bacteria (Haselkorn 1992); the third example being that of sporulation-linked rearrangements in *Bacillus subtilis* (Stragier *et al* 1989). The *Anabaena* rearrangements serve as a mechanism of gene regulation as well. Apart from being quite novel they are also of considerable applied interest since without them no nitrogen fixation is possible in heterocystous forms which are considered agriculturally important. Since their discovery in 1985 (Golden *et al* 1985) the *Anabaena nif* rearrangements have revealed several interesting facts about their *modus operandi.* We now take stock of some of the recent developments in this exciting area of developmental biology.

2. General organization and expression of nitrogen fixation (*nif*) genes in nitrogen-fixing bacteria

The organization of nitrogen fixation genes varies extensively within genomes of different nitrogen-fixing organisms. In *Klebsiella pneumoniae* the entire *nif* cluster is located between *hisD* and *shiA* on the chromosome corresponding to a map position of 43–44 min in the *Escherichia coli* genome (Bachmann 1983) which is congruent with that of *Klebsiella.* It consists of 19 *nif* genes organized in 9

transcriptional units which are closely linked, in fact "jampacked" in 24 kb of DNA (Merrick 1988). *nifK, D, H* are structural genes that encode polypeptide components of the nitrogenase complex and have been very highly conserved through bacterial evolution (Mazur *et al* 1980; Ruvkun and Ausubel 1980). *nifBQ, nifNE* and *nifV* contribute to the synthesis and integration of the active site of nitrogenase, the Fe-Mo-cofactor (Ausubel and Cannon 1980; Imperial *et al* 1984). Genes *nifM, nifS* and *nifU* are involved in the maturation of the two nitrogenase proteins (Cannon *et al* 1985a). The *nifF* and *nifJ* genes encode a flavodoxin and an oxidoreductase respectively and transfer electrons from pyruvate to nitrogenase reductase during nitrogenase catalysis (Shah *et al* 1983; Orme-Johnson 1985). The *nifL, A* operon has a regulatory function with *NifA* functioning as a transcriptional activator of all *nif* operons. Three other genes required for regulation lie outside this cluster. These are *ntrA* (Merrick 1988) encoding a nitrogen- dependent sigma factor (RpoN) (Hirschman *et al* 1985), and *ntrB* and *ntrC* whose protein products constitute a two-component (sensor-effector) signal transduction system typical of bacteria.

In *K. pneumoniae* expression of all *nif* Operons requires a NtrA-modified RNA polymerase, that binds to sequences CTGG (–26 to –23) and TTGCA (–14 to –10) (Beynon *et al* 1983), and a NifA activator that binds to a TGT-N_{10}-ACA upstream activator sequence (UAS). An integration host factor (IHF)-like protein probably binds in the intervening region and bends the DNA to bring the UAS and the NtrA- responsive -promoter sites together (Santero *et al* 1990). Interaction between RNA polymerase, NifA and NtrA proteins is then possible and transcription of *nif* genes ensues. Two environmental signals, namely absence of combined nitrogen and oxygen, trigger this transcription cascade (Cannon *et al* 1985b). The Ntr-control system senses gross changes in nitrogen levels and shuts-off/on all Ntr-controlled systems, including *nif* Operons. The *nifL, A* operon senses intermediate levels of nitrogen and of oxygen and provides a fine-tuning of *nif* expression (Merrick 1988). Together, this two-tier regulation provides a very efficient, need-based and economical mechanism of ensuring expression of Nif proteins only under conditions when they are needed, or can work.

The compact organization of *nif* genes found in *K. pneumoniae* is by no means universal. Dispersion of *nif* genes is common among *Azotobacter* and *Rhizobium* species. Scattering of the structural genes (*nifK, D, H*) is rare but occurs in certain rhizobia. Thus in *R. japonicum* and in *Parasponia rhizobium* and *nifK, D* and *nifH* occur as two independent operons separated by 15 kb DNA (Kaluza *et al* 1983; Fischer and Hennecke 1984; Weinman *et al* 1984). Discontinuities within a single *nifK, D, H* operon are, however, unique to vegetative cells of heterocystous cyanobacteria, such as *Anabaena* (Golden *et al* 1985), some of which also have a second discontinuity in another *nif* operon (see figure 1).

3. Organization of nitrogen fixation genes in cyanobacteria

The organization of nitrogen fixation (*nif*) genes in cyanobacteria differs significantly from that found in other diazotrophs. In the best studied heterocystous cyanobacterium *Anabaena* sp, strain PCC 7120 (figure 1), the major *nif* cluster comprises of several open-reading frames (ORFs), two interrupting DNA elements and just four Operons

dispersed over nearly 80 kb of DNA (Mazur *et al* 1980; Rice *et al* 1982; Golden *et al* 1985). At least 14 genes associated with nitrogen fixation have been identified in *Anabaena* sp. strain PCC 7120. These include the *nifK, D, H* operon (Mazur *et al* 1980; Rice *et al* 1982), the *nifB, fdxN, n ifS, nifU* operon (Mulligan and Haselkorn 1989), a *fdxH* gene which encodes a heterocyst-specific ferredoxim (Bohme and Haselkorn 1988), and another operon containing *orfl, orf2* which result in a Nif$^-$ phenotype, if mutated (Borthakur *et al* 1990). Another *nif* gene of unknown function and unrelated to the known *nif* genes has been located outside the major *nif* cluster by Tn*5* mutagenesis (Borthakur and Haselkorn 1989). A second copy of *nifH,* unlinked to *nifK, D* also lies outside the main cluster (Rice *et al* 1982). Two more genes *xisA* and *xisF* which reside on the interrupting DNA elements and encode for two site-specific recombinases, should also be added to this list since they are

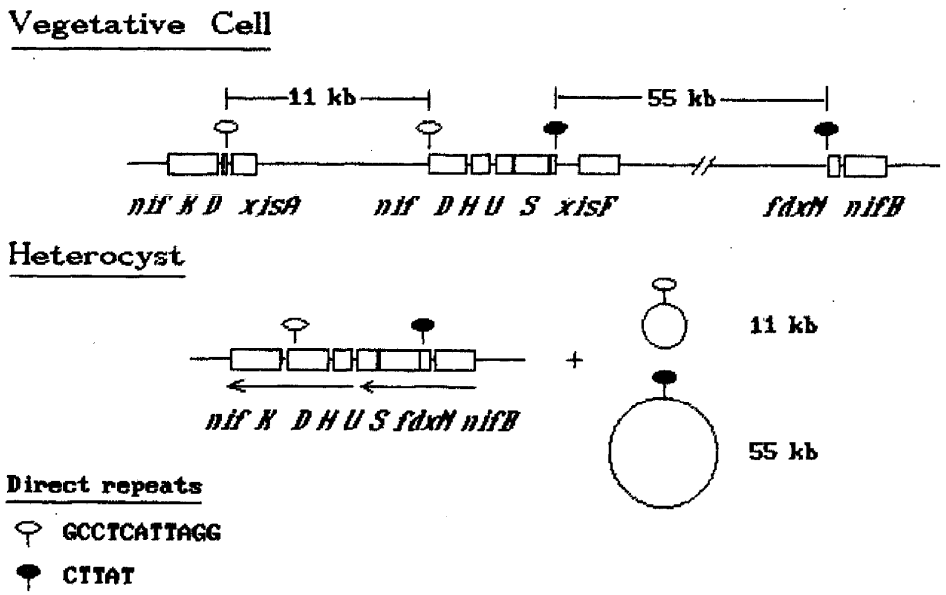

Figure 1. Rearrangement of nitrogen fixation (*nif*) genes in the heterocysts of *Anabaena* sp. strain PCC 7120. The major *nif* gene cluster consists of two Operons, the *nifH,D—K* and *nifB fdxN—nifS-nifU.* In the vegetative cells, the 11 kb *nifD* DNA dement harbours the *xisA* gene and is bordered at the ends by two 11 bp direct repeats (O). A second 55 kb element interrupting the *fdxN* gene harbours the *xisF* gene and has two 5 bp direct repeats (●) at its ends. The only other nitrogen fixation-related genes known are the *fdxH, orf*1 and *orf*2 (located about 6 kb to the left of *nifK*). The *rbcL,S* operon, that encodes the RuBP-carboxylase, lies about 10 kb to the right of *nifB.* The disrupted *nif* Operons in vegetative ells do not express. During heterocyst differentiation, the *xisA* and *xisF* express a site-specific recombinase each. These excisases cause deletion of the two elements by site-specific recombination between the respective direct repeats. In mature heterocysts, the *nif* operons become contiguous and functional and nitrogen fixation ensues. The excised 11 and 55 kb elements remain in heterocysts as unreplicated and untranscribed circles. One copy of the direct repeats goes out in each circle while the other stays back in the chromosome as part of the ORF of the fused *nifD* and *fdxN* genes, as shown. The gene rearrangements and nitrogen fixation are strictly developmentally regulated and do not occur in the vegetative cells.

absolutely essential for nitrogen fixation in *Anabaena* sp. strain PCC 7120.

In *Anabaena* sp. strain PCC 7120 the *nifD* and *fdxN* genes were found to be discontinuously organized on account of two large interruptions (Golden *et al* 1985). The interrupting DNA are excised by means of recombinases encoded by the *xisA* and *xisF* genes during heterocyst development (figure 1). The 11 kb DNA sequence has been shown to occur in a large number of heterocystous cyanobacleria such as strains of *Anabaena, Nostoc* and *Calothrix,* including their Het^- mutants (table 1). The only known exceptions are *Fischerella* (Saville *et al* 1987) and certain Symbionts *of Azolla.* Cyanobacleria isolated from symbiotic associations with *Azolla caroliniana* and *Anthoceros punctatus* showed some interesting features (Meeks *et al* 1988). A freshly isolated major symbiont from *A. caroliniana* showed a contiguous *n ifK, D, H* arrangement but when cultured with another host, *A. punctatus,* showed a discontinuous arrangement (table 1). This suggested that the major symbiont of *A. caroliniana* could not be cultured and a minor contaminant got selected during forced symbiosis with an unnatural host, *A. punctatus.* The dominant symbiont of *Azolla* could not be cultured during three previous attempts also (Tel-Or and Sandovsky 1982; Franche and Cohen-Bazire 1985; Nierzwicki-Bauer and Haselkorn 1986). Freshly isolated Symbionts from several *Azolla* fronds in Senegal have also been found to lack the 11 kb element (Franche and Cohen-Bazire 1987). In contrast, the freshly isolated as well as cultured symbiont from *A. punctatus* both showed a discontinuous *nifK-D, H* arrangement and the presence of DNA homologous to *xisA* (Meeks *et al* 1988). The 55 kb element on the other hand may have a more restricted distribution; of the four genera examined three, *Anabaena* sp. strain PCC 7120, Λ. *cylindrica* and *Nostoc* PCC MAC, possess it while one, *Anabaena variabilis* lacks it (Herrero and Wolk 1986; Meeks *et al* 1988; Haselkorn 1992). One of the salt tolerant strain in our collection, *Anabaena torulosa,* also appears to lack the 55 kb element (table 1).

The *nif* gene organization in heterocystous forms raises interesting questions about their organization in the non-heterocystous forms. A large number of non-heterocystous forms, both unicellular and filamentous, were therefore examined critically. All of them were found to lack the 11 kb element (table 1) and possessed a contiguous *nifK, D, H* operon of about 4 kb, as in every other diazotroph. The heterocystous and non-heterocystous cyanobacteria thus differ strikingly in the organization of *nifK, D, H* genes. A strange case of restriction fragment length polymorphism (a 12 kb and another 4·5 kb *ClaI* fragment both harbouring *nifK, D, H*) was reported by Barnum and Gendel (1985) in *Plectonema boryanum* and attributed to possible rearrangement of genes. We reexamined *P. boryanum* strain 594 and discounted the aforesaid results to show that (i) *nifK, D, H,* were a contiguous cluster of 4 kb which can not harbour any more genes within, (ii) no rearrangements were observed either within or in the vicinity of *nif* genes, and (iii) no homology to the 11 kb element, especially to *xisA,* could be seen in *P. boryanum,* indicating that the machinery needed for the rearrangements was also lacking (Apte and Thomas 1987). Cyanobacterial genomes contain relatively high levels of N^6-methyladenine and 5-methylcytosine and show *dcm* and *dam* methylation patterns (Padhy *et al* 1988). Also, *Cla*I is known to be sensitive to *dam* methylation (Kessler *et al* 1985). It is also known that there are *Cla*I sites in *nifS* revealed by sequence analysis but which do not cut with *Cla*I. Partial digestion with *Cla*I may, thus, explain Barnum and Gendel's (1985) results in *P. boryanum.*

Table 1. Organisation of nitrogen fixation *(nif)* genes in cyanobacteria.

Strains examined	Nitrogen fixation		*nifKDH* operon	55 kb element	Reference
	Air	Argon			
I. Non-heterocystous cyanobacteria					
Unicellular strains					
Cyanothece 7424	+	+	KDH	?	Kallas *et al* (1985)
Gloeothece 6909-1	+	+	KDH	?	Kallas *et al* (1983)
Synechococcus 7335	–	+	KDH	?	Kallas *et al* (1985)
Synechococcus 7425	–	+	KDH	?	Kallas *et al* (1985)
Filamentous strains					
'LPP' 73110	–	+	KDH	?	Kallas *et al* (1985)
Oscillatoria tenuis strain UTEX 1566	–	+	KDH	?	Saville *et al* (1987),
Plectonema boryanum	–	+	KDH	?	Apte and Thomas (1987)
Psuedanabaena 7409	–	+	KDH	?	Kallas *et al* (1985)
Psuedanabaena ATCC 29210	–	+	KDH	?	Saville *et al* (1987)
Trichodesmium NIBB 1067	+	+	KDH	?	Zehr *et al* (1991)
II. Heterocystous strains					
Free-living					
Anabaena PCC 7120	+	+	K—DH	+	Golden *et al* (1985)
Anabaena 268S10	+	?	KDH	+	Brusca *et al* (1990)
Anabaena torulosa	+	+	K—DH	–	Apte (1993)
Anabaena variabilis	+	+	K—DH	–	Herrero and Wolk (1986)
Calothrix					
7601-D (Het^-)	–	+	K—DH	?	Kallas *et al* (1983)
7601 (Het^+)	+	+	K—DH	?	Kallas *et al* (1983)
Fischerella ATCC 27929	+	+	KDH	?	Saville *et al* (1987)
Nostoc					
7121 (Het^-)	–	+	K—DH	?	Kallas *et al* (1985)
7906 (Het^+)	+	+	K—DH	?	Kallas *et al* (1985)
Symbionts					
Anthoceros symbiont					
Nostoc sp. 7801	+	+	K—DH	?	Meeks *et al* (1988)
Azolla symbiont					
Anabaena azollae					
major symbiont	+	?	KDH	?	Meeks *et al* (1988)
minor symbiont	+	–	K—DH	?	Meeks *et al* (1988)

(+) indicates presence and (-) the absence of nitrogenase activity or the 55 kb element, and (?) denotes that the feature has either not been studied or is not known.

It is noteworthy that only heterocystous forms can fix nitrogen in air and in them the *nif* operons are interrupted. The non-heterocystous forms, which have contiguous *n ifK, D, H,* fix nitrogen only under microaerobic conditions (table 1). Only a few of the non-heterocystous forms also fix nitrogen aerobically either by separating nitrogen fixation (dark) and photosynthesis (light) in the time or by excluding oxygen (e.g. bundle-forming habits of *Trichodesmium)* (Capone *et al* 1990). An exceptional case is of *Trichodesmium* sp. strain NIBB 1067, which does not form bundles and the individual trichomes fix nitrogen in air (Bergman and Carpenter 1991) in a strictly light-dependent manner (Zehr *et al* 1991). Alternate cycles of carbon and nitrogen fixation, even in continuous light, have been shown

in another non-heterocystous strain *Oscillatoria* which fixes N_2 in air (Stal and Krumbein 1985). The underlying mechanisms are currently unknown.

4. General features of *nif* gene rearrangements in *Anabaena*

Ongoing rearrangements within the *nif* region caused by an endogenous transposable element and leading to a Niff phenotype have been described for *R. meliloti* (Ruvkun *et al* 1982). Exchanges of *nif* genes between the symbiosis-related (*sym*) plasmid and the chromosome have also been noted in rhizobia. The *nif* region of organisms thus seems to be prone to gene rearrangements. Two of the *nif* operons in the vegetative cells of *Anabaena* are interrupted by large discontinuities and their removal is necessary prior to the expression of these Operons (figure 1). Although they look like introns, no RNA transcripts emanating from the 11 or 55 kb elements could be detected, indicating that they are not introns. Both these discontinuities are precisely excised during differentiation and in the mature heterocysts the two operons become contiguous. Both rearrangements are easy to detect using probes spanning the breakpoints (figures 1 and 2).

Understanding the mechanism of rearrangements required knowing the topology of the DNA involved at the breakpoints, detection and analysis of the excised DNA and characterization of the recombinases involved. The breakpoints of the 11 kb element were relatively easy to clone and sequence, and it revealed the presence of an 11 bp sequence (GGATTACTCCG) that was directly repeated at the ends of the element (figure 1) (Golden *et al* 1985). To determine the topology of the 55 kb element breakpoints, cosmid libraries of the vegetative cell DNA were constructed and used to clone the breakpoints. Elaborate chromosomal walking of these libraries with appropriate probes resulted in physically linking the left and right borders of this rearrangement and in estimating the size of the intervening DNA as about 55 kb. It also identified the topology of the DNA involved (Golden *et al* 1988). A five basepair (TATTC) sequence was found directly repeated at the ends of the element (figure 1). The direct repeats found at the ends of the two elements were, thus, not homologous (table 2). The break-points of both elements are unique and bear no resemblance to the target sequences of known bacterial recombinases such as integrase of bacteriophage lambda (Campbell 1983) or the inverting enzyme responsible for *Salmonella* flagellar phase variation (Silverman and Simon 1983).

Based on this information, it was proposed that the two rearrangements were results of two independent site-specific recombination events between the respective direct repeats (Golden *et al* 1987, 1988; Mulligan *et al* 1988; Mulligan and Haselkorn 1989). Such recombinations would result in the excision of the intervening DNA as a circle, fusion of previously distant chromosomal regions and generation of functional operons, as shown in figure 1. All of these have been experimentally shown (Haselkorn 1986, 1992). The 11 kb excison has been detected in the heterocysts as a stable non-replicating circle containing one copy of the 11 bp repeat, the other copy was approximately 130 bp from the 3'end of fused *nifD* (Golden *et al* 1985). The 55 kb element is also presumably excised as a circle but its large size has precluded its detection probably due to random degradation or shearing which occurs during DNA isolation from heterocysts. Presence of the stable 55 kb DNA

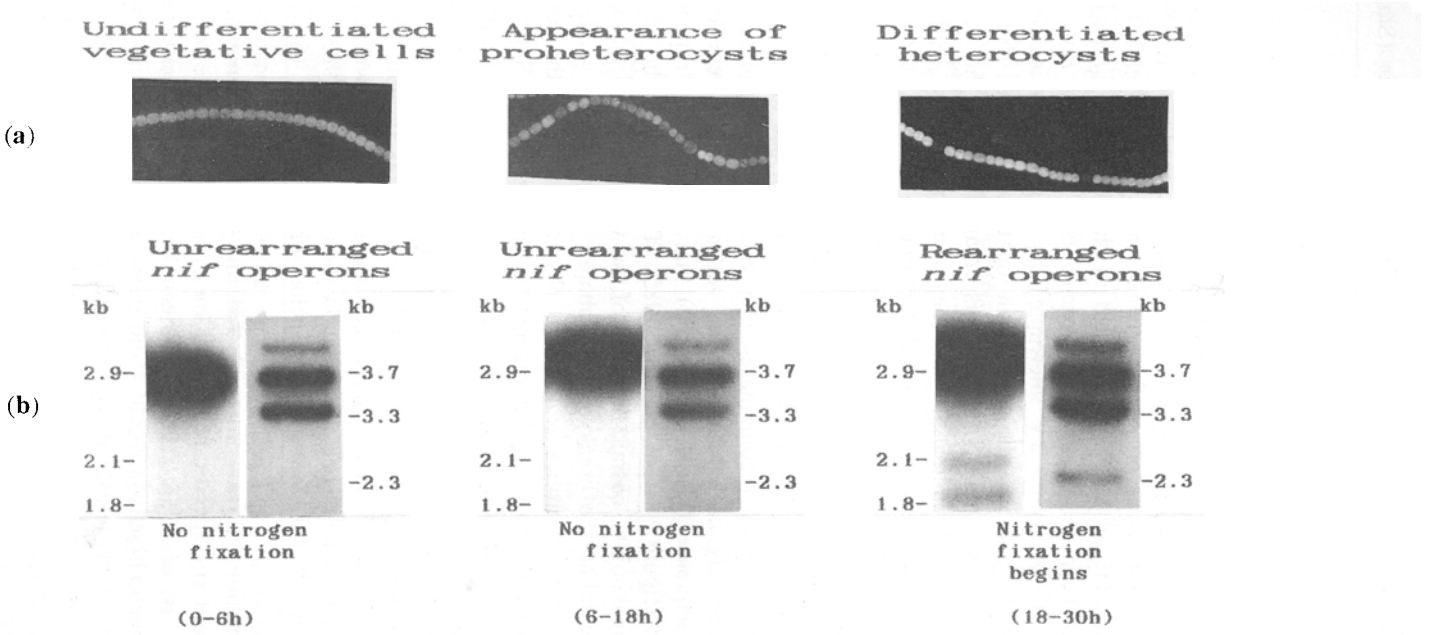

Figure 2. Chronological sequence of events leading to the expression of nitrogen fixation in *Anabaena* sp. strain PCC 7120. (**a**) Filaments were illuminated with light (520 nm) specifically absorbed by phycocyanin (pigment responsible for O_2 evolution in photosynthesis). The energy is transmitted to chlorophyll *a* and is eventually emitted as fluorescence (590 nm). Fluorescing filaments were photographed in a Carl-Zeiss fluorescence microscope. All vegetative cells fluoresce uniformly. Loss of phycocyanin is an early event (proheterocysts show partial loss of the pigment) and its complete absence in the mature heterocysts (dark gaps) helps to build an internal anaerobic milieu. (**b**) The *nif* gene rearrangements were detected using probes spanning the breakpoints of the two elements. The 11 kb element deletion is detected by probe pAn256 that spans the breakpoint near the right border of the element (figure 1) and detects a single 2·9 kb *Hind*III band in vegetative cells and two (2·1 and 1·8 kb) *Hind*III fragments in the heterocysts (Golden *et al* 1985). The 55 kb deletion is detected using the probe pANH20·1 that represents the fused *fdxN* gene in the heterocysts (figure 1). It detects a single 2·3 kb *Hind*III fragment in heterocysts and two (3·7 and 3·3 kb) *Hind*III fragments in the vegetative cells (Golden *et al* 1987). Rearrangements and nitrogen fixation occur late in development after phycocyanin has been degraded and thick walls and polar bodies have formed in the heterocysts (not shown). Under strict anaerobiosis mature heterocysts do not form, 11 kb element is not excised and filaments do not fix nitrogen.

Table 2. Developmentally regulated gene rearrangements in bacteria.

Organism	*Anabaena* sp. strain PCC 7120		*Bacillus subtilis*
1. Rearranged element	*nifD* element	*fdxN* element	Skin element
Size	11 kb	55 kb	42 kb
Interrupted gene	*nifD*	*fdxN*	*sigK*
Direct repeats	11 bp	5 bp	5 bp
Breakpoint sequence	GGATTACTCCG	TATTC	AGTAA
2. Recombinase involved	Excisase A	Excisase F	Cis A
The gene	*xisA*	*xisF*	*cisA*(*spoIVCA*)
Open reading frame	1062 bp	1545 bp	1500 bp
The protein	XisA	XisF	CisA
Amino acid residues	354	515	500
Mass (daltons)	41644	58473	57481
Homology	None	Resolvases and CisA	Resolvases and XisF
3. Initiating triggers	N deficiency, O_2	N deficiency	Nutrient deprivation
4. Cell specificity and fate of rearangements			
Unrearranged cells	Vegetative cells continue	Vegetative cells continue	Spores continue
Rearranged cells	Heterocysts are terminal cells	Heterocysts are terminal cells	Spore mother cell die
5. Consequences of rearrangements	Creation of functional *nifK, D, H* operon and N_2 fixation	Creation of functional *nifB-fdxN-nifS-nifU* operon and N_2 fixation	Expression of alternat sigma factor SigK and further deve-lopment (sporulation

in heterocyst without degradation has, however, been shown in Southern blots using cosmid probes from within the 55 kb element. Finally, transcripts corresponding to both the interrupted Operons are observed after rearrangements, establishing that the transcription read-through is possible only after deletion of the 11 and 55 kb elements (Golden *et al* 1985, 1988).

The 11 and 55 kb rearrangements are first detected in differentiating cells around 18 h, roughly at about the same time when mature heterocysts appear (Golden *et al* 1985). It seems unlikely that other secondary site-specific DNA rearrangements occur within either the 11 or the 55 kb elements. Southern blots using probes from within the two elements detected only those fragments expected from the two known rearrangements (Golden *et al* 1988). Although occurring at the same time, the two rearrangements are not obligatorily coupled. Rather they appear, for several reasons, to be unlinked events. Thus, (i) there exist cyanobacteria with only the 11 kb but not the 55 kb element (table 1), (ii) the 11 and the 5 bp direct repeats at the breakpoints of the 11 and 55 kb elements respectively are very different, (iii) manipulating conditions of differentiation, for example during anaerobiosis, one can obtain cells deleted in the 55 kb element but not in the 11 kb element (Golden *et al* 1988), and (iv) mutants of *Anabaena* sp. strain PCC 7120 have been constructed wherein the deletions of the 11 and 55 kb elements are not linked and occur independently of each other (Golden and Wiest 1988; Carrasco *et al* 1994). The two rearrangements thus seem to be distinct events mediated by relatively independent

mechanisms.

The fate and function of the excised DNA are not clear. Each of the elements seems to exist as a single circle in heterocysts and does not contain any repetitive DNA sequences (Golden *et al* 1988). The excised DNA is neither degraded (not yet shown for the 55 kb circle) nor amplified, and its function after deletion is unknown. No transcripts, corresponding to either of the elements could be easily detected in *Anabaena* sp. strain PCC 7120 (Rice *et al* 1982) and it seemed as if they did not contain any structural genes. However, as described in the sections that follow, each of the elements harbours at least one ORF that encodes a site-specific recombinase responsible for the self-excision of the element. The activities of these excisases turned out to be suicidal for the two elements.

5. The site specific excisases in *Anabaena*

The gene encoding the recombinase of the 11 kb element was discovered by chance during passage of a cloned *Anabaena* fragment in *E. coli* (Lammers *et al* 1986). The recombinant plasmid, pAn207 (containing a 17 kb EcoRI fragment from vegetative cell DNA in pBR322), carried the entire *nifK,* the 11 kb excison and some of the *nifD*. The recombinant plasmid was found to rearrange at a very low frequency, during its propagation in an *E. coli recA* strain HB 101, to form a smaller plasmid with only 6 kb insert instead of 17 kb of the original insert. To characterize the rearrangement better, a mini-Mu-Kan-*lac* was inserted by random transposition into the 11kb insert DNA of pAn207. Cells carrying pAn207 :: *Mu-Kan-lac* were Kan^r Amp^r and lac^+ while the cells wherein the 11 kb element was deleted became Kan^s Amp^r lac^-. This provided a simple and sensitive blue/white colony assay [on 5-bromo-4-chloro-3-indolyl-â-D-galactoside (X-gal) indicator plates] for studying recombination. The Mu-Kan-*lac* insertions also caused random mutations in the 11 kb element which proved very useful for mapping purposes (Lammers *et al* 1986).

One mini-Mu-Kan-*lac* insertion near the centre of the 11 kb insert DNA (called MX25) gave white colonies at a frequency of 0·3% during overnight growth of *E. coli* in Luria broth. Restriction analysis of plasmids isolated from blue and white colonies confirmed the deletion of about 11 kb DNA in the white colonies. Sequencing across the fused *nifD* junction reaffirmed site-specific recombination within same 11 bp direct repeats as in *Anabaena* (Lammers *et al* 1986). The observed frequency of 0·3% could not be enhanced by physiological manipulations such as aerobic/ anaerobic growth, nitrogen limitation or heat shock (5 min at 45°C) etc. These findings suggested that the *Anabaena* rearrangements could be precisely brought about in *E. coli,* thus making their study a lot easier. The results also established that the genetic/biochemical requirements for rearrangements were either carried by the 11 kb excison itself or could be provided by the *E. coli* host.

Another insertion (MX32) which mapped near the left border of the excision never produced any white colonies. However, the defect could be complemented in *trans* by transferring wild type DNA fragments spanning the left border region of the 11 kb excision. Functional analysis of several deletions of this region revealed a gene essential for excision (Lammers *et al* 1986). Sequence analysis of regions flanking MX32 mutation, defined an ORF of 1062 base pairs and a coding region for 354 amino acids. The ORF begins at 240 bp from the left side of the

recombination breakpoint and runs counter to the direction of *nif* gene transcription. The ORF contains two in-frame start codons 117 base pairs apart. The protein has a predicted molecular mass of 41644 daltons contains a large number of basic amino acid residues, has a positive charge of + 18 and has been called excisase A, (XisA) (Lammers *et al* 1986). The deduced amino acid sequence of *xisA* identifies a protein quite unlike any of the known recombinases. In its proximity to one of the recombination sites and its location on the excison, *xisA* resembles two other site-specific recombinases, namely the integrase of bacteriophage lambda (Campbell 1983) and the resolvase of transposons (Heffron 1983). The *xisA* inactivation could be complemented in *E. coli* by cloned *xisA* not only from *Anabaena* sp. strain PCC 7120 (Lammers *et al* 1986) but also from *A. variabilis* (Brusca *et al* 1989). Thus, the 11 kb elements from diverse *Anabaena* and *Nostoc* strains may presumably contain identical excisases and the same target sequences at the breakpoints.

The role of *xisA* in mediating the deletions of the 11 kb element was unequivocally demonstrated by site-directed mutagenesis of the gene in *Anabaena* sp. strain PCC 7120 (Golden and Wiest 1988). The cloned *xisA* was insertionally mutated *in vitro* and then used to replace the wild type *xisA* in the *Anabaena* chromosome by conjugal transfer (Wolk *et al* 1984) followed by homologous recombination. The mutant DW12–2·2 thus obtained grew normally in nitrogen supplemented media but not in combined nitrogen-free media although it made normal heterocysts. The mutant failed to rearrange the 11 kb element and did not fix nitrogen. The rearrangement of the 55 kb element was unaffected in the mutant (Golden and Wiest 1988). The *xisA* is therefore absolutely essential for nitrogen fixation in *Anabaena* sp. strain PCC 7120.

The site-specific recombinase for the 55 kb deletion and the topology of the DNA involved eluded discovery until recently. Brute-force DNA sequencing at the ends of the 55 kb element finally identified the gene encoding the recombinase excising this element (Carrasco *et al* 1994). Sequencing was undertaken based on the suspicion that like most other mobile genetic elements, the gene encoding the site-specific recombinase would be located near one end of the element. Sequencing showed a large ORF near the *nifS* proximal left border of the 55 kb element, the gene has been called *xisF*. The start codon in *xisF* was located 47 bp from the 5 bp repeat involved in site-specific recombination. The 1545 bp *xisF* ORF could encode a 515 amino acid polypeptide with a predicted molecular mass of 58473 dalton and a pI of 9·83. The sequence revealed 26% identity and 49% similarity to another developmentally regulated site-specific recombinase, CisA from *B. subtilis* over its entire length.

Insertional Inactivation of *xisF* gene *in vitro* and its transfer to *Anabaena* sp. strain PCC 7120 by conjugation resulted in site-directed mutagenesis of *xisF* of the *Anabaena* chromosome (Carrasco *et al* 1994). The resulting mutant, KSR9, lost its ability to rearrange the 55 kb element but showed normal deletion of the 11 kb element during differentiation. It also produced normal heterocysts and grew well on nitrate but not on dinitrogen. Upon complementation with cloned *xisF,* the recipient (KSR11) showed normal rearrangement of the 55 kb element. Therefore the *xisF* encoded the site-specific recombinase responsible for excising the 55 kb element. The discoveries of the two unlinked genes, *xisA* and *xisF,* each encoding a separate site-specific recombinase acting on different target sites established the

independence of the two rearrangements. The two recombination events are linked to heterocyst formation rather than to each other. Perhaps some host functions are common to both rearrangements.

6. Rearrangement of *Anabaena* elements in *E. coli*

Both the 11kb and the 55 kb elements of *Anabaena* sp. strain PCC 7120 were properly rearranged even in *recA* mutants of *E. coli,* This, as described earlier, provided the opportunity for characterization of structure-function relationships in both the excisase genes in a much more amenable host, *E, coli,* than *Anabaena.* A substrate plasmid (pAM461), containing just the left and right borders of 11 kb element flanking only a 715 bp DNA (instead of 11 kb) was constructed. The plasmid pAM46l was precisely rearranged, in *E. coli* overexpressing *xisA,* at the correct breakpoints indicating that the pAM461 contained all the sequences necessary to act (in *cis)* as substrate for the site-specific recombinase. The ability of *xisA*, provided in *trans,* to rearrange pAm461 shows that XisA protein is sufficient to cause site-specific recombination in *E. coli* (Brusca *et al* 1990).

Due to its large size the entire 55 kb element could not be cloned and so its rearrangement was difficult to demonstrate in *E. coli.* For this purpose, a rearrangement substrate plasmid pJGIA was constructed positioning a *lacZ* between the left and right break-points of the 55 kb element. *E. coli* cells harbouring this plasmid were lac^+. Upon transfer of a cloned *xisF* gene in expression vectors in *trans* to the same *E. coli* about 0·1% of the cells rearranged pJGIA DNA at the precise breakpoints and produced *lac* colonies. Partial modification of. the 5 'end of *xisF* increased this rearrangement frequency to 1% (Carrasco *et al* 1994).

Rearrangement of *Anabaena* elements occurred at rather low frequency in *E. coli* (pMX25:0·3%; pJGIA: 0·1–1·0%) as described earlier. This suggests that *E. coli* cells provide cellular conditions that are very different from that in differentiating heterocysts. Therefore the rearrangements in *E, coli* could be triggered differently and the physiological significance and relevance of their regulation in *E. coli* should be extrapolated to *Anabaena* with care.

The problem of low level expression of these genes was partly solved by overexpressing these genes by cloning them downstream of a strong promoter such as *tac.* This enhanced the rearrangement frequencies by about 10-fold, but no more. Brusca *et al* (1990) have suggested that XisA, if overexpressed, may be toxic to *E. coli* and may lead to lethality, Haselkorn (1986) has cited unpublished results of W J Buikema that a *xisA* cloned downstream of a T7 promoter in a plasmid expression vector caused 100% rearrangement of pMX25, in *trans.* How the same protein produced from two different promoters can be differentially toxic to *E. coli,* is not clear.

Studies on *xisA* and *xisF* transcription have been hampered because of the inability to detect mRNA transcripts. Attempts by Rice *et al* (1982) earlier failed to detect transcripts from any part of the 11 kb and to the left end of the 55 kb element. Thus, the transcription level may be low and mRNA may have a very short half-life. Expression from the native promoter is very weak in *E. coli* also. It is not known where the *xisA* transcript begins in *E. coli* since the transcription start point could not be established (Lammers *et al* 1986). The *xisF* transcripts also have not been

detected probably due to low level transient expression of this gene in *Anabaena* (Carrasco *et al* 1994). Reasons for low level and transient expression are beginning to be unravelled now. Transcription of *xisA* and *xisF* seems to be restricted to low levels by means of *cis*-acting negative regulatory elements located upstream of their promoters or within ORFs while the transient expression in *Anabaena* owes it to its developmental regulation (see following sections).

7. Forced gene rearrangements in *Anabaena*

Fusion of the ORF of *xisA* by cloning it downstream of a strong *glnA* promoter from *Anabaena* sp. strain PCC 7120 (pJB4) caused abundant transcription from *xisA* region in *E. coli,* as assayed by Northern analysis. But *Anabaena* containing pJB4 showed no RNA corresponding to *xisA* in either the nitrogen-supplemented or deficient media (Brusca *et al* 1990) indicating that it is not transcribed sufficiently in *Anabaena.* This suggested presence of certain negative regulatory elements(s) (NREs) upstream of *xisA* ORF blocking its expression in *Anabaena.* Several 5'deletions were then generated and cloned downstream of a strong *tac* promoter. All the recombinant plasmids carrying deletions produced easily detectable levels of *xisA* transcripts in *E. coli* (Brusca *et al* 1990). When transferred to *Anabaena* sp. strain PCC 7120, these constructs did not produce detectable levels of transcripts but showed different degrees of rearrangements of chromosome in cells carrying various deletions. Using appropriate probes, rearrangements were shown to be precisely the same which occur in wild type *Anabaena* sp. strain PCC 7120 during differentiation (Brusca *et al* 1990). The conclusion, therefore, was that the Overexpression of *xisA* was forcing rearrangements even in the undifferentiated (nitrate-growth) cultures.

One of the exconjugant strains 268 carrying a plasmid (pAM268), which contained deletions of the first start codon and extended up to 117 bp from the second start codon of *xisA,* showed maximal rearrangements. Although in a large proportion of vegetative cells DNA still remained unrearranged, it was possible, however, to sonicate strain 268 to single cells, plate them out and to grow single colonies therefrom; some of these upon screening were found to lack the 11 kb element. This led to isolation of a strain, 268S10, which completely lacked the 11 kb element and contained only rearranged contiguous *nifK,D,H* in its genome (Brusca *et al* 1990). This strain showed a normal phenotype and at least in short-term experiments showed no growth advantage over the wild type strain.

The plasmid pAM268, which resulted in the strain 268S10, wherein the upstream sequences up to -65 bp of *xisA* had been deleted, induced 50% rearrangement of the vegetative cell chromosome (Brusca *et al* 1990). The results identified a 127 bp NRE within the element—its relation to *xisA* promoter is not known because the promoter has not yet been mapped. But it lies between -65 to -192 bp region upstream of second ATG codon in *xisA* ORF. This suggests that a repressor binding to NRE may keep transcription repressed in the vegetative cells. However, the alternative possibility that the 5'region may be responsible for a post-translational regulation also exists. Thus, a deletion may cause *xisA* translation to begin at second start codon producing an active recombinase while the larger (beginning from the first start codon) XisA protein may be catalytically inactive (Brusca *et*

al 1990).

Although it has been possible to force the expression of *xisA* gene in vegetative cells of *Anabaena* sp. strain PCC 7120, the transcripts have never been detected (Brusca *et al* 1990). It is also not clear why the ratio of unrearranged to rearranged DNA remains constant. One explanation may be that excision may be reversible, except in heterocysts—but this needs to be tested. Spontaneous excision of *nifD* element has never been found. XisA and XisF activities are extremely efficient and very low levels of excisase proteins are sufficient for excision.

8. DNA-binding regulatory proteins of vegetative cells

It may be recalled that deletions upstream of *xisA* define a -192 to -130 bp region, called the negative regulatory element or NRE, which blocks the expression of *xisA* but does not prevent its expression during heterocyst differentiation. A DNA-binding factor, VF1, has been partially purified from vegetative cells of *Anabaena* sp. strain PCC 7120 by heparin-sepharose chromatography. It was found to have affinity to the *xisA* upstream region carrying the NRE by band-shift or gel retardation assays (Chastain *et al* 1990). Three independent DNA : VF1 complexes were thus identified. Increasing concentration of VF1 resulted in shift from the predominance of complex 1 (higher mobility) to complex 3 (least mobility) suggesting that three VF1 proteins bind to the DNA fragment. Estimated molecular mass of VF1 in the three complexes was approximately 45000, 80000 and 120000 daltons suggesting they were formed by the same species perhaps binding as monomer-dimer or trimer or as dimer-tetramer-hexamer. Mobility shift assays on fragments deleted for various lengths of *xisA* upstream region revealed the sequence where VF1. bound. DNAsel footprinting showed a protected region of 63 coding and 66 non-coding bp beginning 143 bp upstream of second ATG in *xisA* ORF. ExoIII digestion in this region produced deletions which allowed binding of only two or one or none of VF1 in band-shift assays (Chastain *et al* 1990). Analysis of protected DNA showed that the three binding sites were adjacent to each other—this would allow protein-protein interactions as also accommodate multimers of same protein. The actual binding sequence appeared to be ACATT and was common to all three sites. One of the (distal) ACATT repeats was found to be on the opposite strand suggesting that VF1 binds to this in an inverted orientation relative to the two *xisA* proximal sites (Chastain *et al* 1990).

The gene encoding VF1 has been cloned using a strategy that provides a genetic selection for expressed clones of genes encoding sequence-specific DNA binding proteins such as VF1 (Wei *et al* 1993). The gene encoding VF1 was named *bifA*. BifA protein produced in *E. coli* shows properties indistinguishable from those of *Anabaena* VF1 (Wei *et al* 1993). BifA and VF1 show similar elution profiles on heparin-sepharose columns and identical mobility shifts when assayed with target sequences from upstream of *xisA* and *glnA,* The 672 bp *bifA* ORF encodes a 25 kDa basic protein with a pI of 9·70 that shows 77% identity and 87% similarity to another cyanobacterial DNA-binding regulatory protein encoded by *Synechococcus ntcA* (Vega-Palas *et al* 1992). It also shows 37% identity and 74% similarity of *E. coli* Crp (Reznikoff 1992) and 27% identity and 54% similarity to *E. coli* Fnr (Shaw *et al* 1983). It, therefore, belongs to the cyclic AMP receptor protein (Crp)

family of bacterial regulators.

In *Synechococcus* sp. strain PCC 7942 (formerly *Anacystis nidulans* R2) the *ntcA* gene has been recently identified as a positive activator required for expression of a number of genes involved in nitrogen assimilation. It also belongs to the Crp family of bacterial transcriptional regulators. The mechanism activating *ntcA* is unknown (Vega-Palas *et al* 1992). The *ntcA* has been shown to be present in 9 other cyanobacteria, including strains of genera *Anabaena, Calothrix, Fischerella, Nostoc, Psuedanabaena, Synechococcus* and *Synechocystis* (Frias *et al* 1993). The *ntcA* sequence shows a DNA-binding helix-turn-helix motif and the target site (TGT-N_{10}–ACA) for its binding has been found upstream of several genes including *xisA,* The target sites of BifA also contain the same sequence. Another interesting overlap between the two proteins is that each one of them is involved in regulating genes which play a role in nitrogen control. The *ntcA* has recently been shown to play a role in heterocyst development and also during growth of *Anabaena* sp. strain PCC 7120 on nitrate (Wei *et al* 1994).

VF 1 also bound to upstream regions of *glnA, rbcL* and *nifH* genes with different affinities. The VF 1 binding to various genes followed the order *glnA* >> *xisA* > *rbcL* >> *nifH* (Chastain *et al* 1990). We may recall that *xisA* expression from *tac* promoter in vegetative cells requires deletion of VF 1 binding site upstream of *xisA*. An attractive role for VFl could therefore be to prevent transcription of *xisA* gene in vegetative cells. However, it can not play the same repressor role in *rbcL* and *glnA* which express abundantly in vegetative cells (Tumer *et al* 1983; Nierzwicki-Bauer *et al* 1994) and *nifH* which expresses in heterocysts (Wolk 1982). Regulatory proteins which act as both positive and negative regulators are known, e.g. the lambda repressor stimulates its own expression but represses expression of *cro* gene (Ptashne 1986). One plausible suggestion, therefore, is that VFl positively regulates transcription in vegetative cells at sites where it is bound as a monomer (*glnA, rbcL*) but negatively regulates transcription when bound as a multimeric complex (*xisA*) (Chastain *et al* 1990).

The target sites of both the NtcA and BifA proteins involve a sequence TGT-N_{10}-ACA which is the upstream activator sequence (UAS) where NifA (a positive activator of *nif* genes) binds. A likely role for these proteins would be that of an activator. BifA perhaps acts both as positive and negative regulator. A second DNA-binding protein factor (VF2 ?) which binds to *rbcL* and *xisA* but not to *glnA* is now known. Interestingly, its DNA binding sites overlap with those bound by VFl (Ramasubramanian *et al* 1994). Titration of the two factors (VFl and VF2) in different cell types may therefore facilitate differential gene expression.

9. Developmental regulation of *Anabaena nif* gene rearrangements

Rearrangements occur late during heterocyst differentiation at about the same time when mature heterocysts are visible, *nif* gene expression can be detected and active nitrogenase can be assayed (figure 2) (Golden *et al* 1985). Also, both rearrangements affect *nif* gene transcription; heterocyst-specific transcripts pass through the *nifD* and *fdxN* chromosomal recombination sites only after the rearrangements (Golden *et al* 1985, 1988). Excision and *nif* expression seem to occur late in heterocyst development (Golden *et al* 1985) after synthesis of the complex envelope is complete,

phycocyanin has been degraded (figure 2) and the internal anaerobic milieu has been established (Haselkorn 1989). This may be because the conditions needed for these events (N depletion and anaerobiosis) occur around the same time, i.e., late in development. Going by their late timings, these events seem to be a consequence of, rather than a trigger for, differentiation. The two rearrangements and *nif* expression are tightly coupled to heterocyst differentiation and neither can be detected in vegetative cells.

Several experiments have, however, suggested that differentiation and rearrangements may not be linked. For example the rearrangements do also occur in Het^- mutants of certain cyanobacteria that do not form identifiable heterocysts (Kallas *et al* 1985). Under certain artificial conditions they can be separated from each other and from heterocyst development. Thus, filaments induced microaerobically do not form mature heterocysts. The 55 kb deletion occurs when filaments are induced both under aerobic and microaerobic conditions, but 11 kb deletion occurs only under aerobic state (Golden *et al* 1988). The independence of differentiation from rearrangements is also suggested by mutants carrying inactivated *xisA* and *xisF* genes (Golden and Wiest 1988; Carrasco *et al* 1994). These mutants do form normal heterocysts but do not rearrange either the 11 or the 55 kb elements and do not fix nitrogen. The *nif* rearrangements have now been brought about in vegetative cells of nitrate-grown repressed cultures by overexpressing *xisA* (Brusca *et al* 1990). This raises doubts about the developmental regulation of rearrangements and their association with the developmental process appears to be coincidental. It seems likely that the rearrangements, *per se,* may be regulated by certain environmental factors.

The chronological sequence of events, namely heterocyst differentiation, *nif* gene rearrangements and expression of nitrogenase activity, and their cause-and-effect relationship, if any, have been recently investigated and these studies revealed two kinds of controls (Apte and Gowda 1992; Apte 1992, 1993). Heterocyst development was triggered by absence of combined nitrogen and presence of oxygen in the cells. However, the two *nif* rearrangements were induced by nitrogen deficiency and were differentially sensitive to oxygen status. Thus, the 55 kb rearrangement was indifferent to the presence/absence of oxygen but the 11 kb rearrangement occurred only during growth in air (Golden *et al* 1988; Apte and Gowda 1992). We have recently observed that this may be a consequence of carbon depletion since addition of CO_2 (i.e., growth in argon/CO_2) restores 11kb rearrangement in such anaerobic cultures (Prabhavathi and Apte 1994). Since the formation of mature heterocyst also shows similar requirements it is tempting to speculate that the deletion of 11 kb may be more tightly linked to the formation of heterocysts than that of the 55 kb. Although the rearrangements occur late during development after an anaerobic milieu has been set up, they are unlikely to be triggered by anaerobiosis *per se* (Golden *et al* 1988; Apte and Gowda 1992). Nitrogen fixation, on the other hand, seems to be linked to 11 kb rearrangements rather than to formation of heterocysts. The most important conclusion of these studies was that the rearrangements occurred in the absence of mature heterocysts. However, even the extremes of environments could not increase their frequency beyond that observed during normal aerobic growth (8–10% heterocysts) . In some unknown way, therefore, the maximal frequency of rearrangements appeared to be controlled by the inherent pattern of heterocyst differentiation.

The *in situ nif* expression data obtained by Elhai and Wolk (1990) help to explain some of our observations. Working with a Het$^-$ mutant strain, *Anabaena* sp. strain PCC 7118, and using video-imaging of luciferase fluorescence [from constructs carrying transcriptional fusion (P_{nifH}–*luxA, B*)] in individual cells of differentiating filaments, these workers showed that *nifH* expression was observed only in every eighth cell. This coincides with the normal heterocyst frequency in *Anabaena* and suggests that even if formation of mature heterocysts is blocked (by mutation or anaerobiosis) *nif* expression occurs only in the "prospective heterocysts" and the "built-in developmental programme" takes care of it. The same may be the case with *nif* gene rearrangements which may therefore be developmentally regulated.

It has been suspected that the developmental regulation of transcription of *nifH* and of rearrangement of the *nifK--D, H* operon are linked to each other. This was tested by Golden *et al* (1991) by separately blocking each event by appropriate genetic manipulations and determining if the other event occurred normally. When induced in combined N-free media and probed with *nifH,* wild type *Anabaena* shows three distinct transcripts of M kb (*nifH*), 2·8 kb (*nifH, D*) and 4·7 kb (*nifH, D, K*) (Golden *et al* 1985). A *xisA* mutant DW12–2·2 which forms normal heterocysts but is unable to rearrange the 11 kb element (Golden and Wiest 1988) shows only a 1·1 kb (*nifH*) transcript but none corresponding to *nifD* and *nifK* genes. The transcription, thus, can be induced from *nifH* promoter of the unrearranged operon also but stable transcripts of *nifK* and *nifD* are not produced (Golden *et al* 1991). In another experiment a cloned *nifH* contained in a recombinant plasmid, was deleted and replaced by a Sm^r-Spc^r cassette *in vitro*. Site-directed mutagenesis of *Anabaena* sp. strain PCC 7120 with this construct yielded a mutant LW1 that had *nifH* upstream region deleted. In LW1 all three *nifH* transcripts were absent but rearrangement of both the 11 and 55 kb elements occurred normally (Golden *et al* 1991). It could, thus, be shown that although 11kb rearrangement and *nifH* transcription had very similar timing, each one occurred independently of the other.

Although the mechanisms responsible for the rearrangement happen to lie on the excised elements themselves (i.e., the *cis*-acting direct repeats and the genes encoding the recombinases), the reasons for their developmental regulation are not found within the Operons. The expression of recombinases seems to be regulated by certain *trans*-acting DNA-binding proteins (*viz.,* BifA, NtcA, VF2 . . . etc.) which bind to promoters of several genes (*viz., xisA, nifH, glnA, rbcL,* etc.) which are differentially expressed in vegetative cells and heterocysts. An intricate interplay of such factors may bring this about.

10. Other developmental gene rearrangements in prokaryotes

Developmentally regulated gene rearrangements are usually found in eukaryotes, e.g., mating type switch in yeast (Haber 1983) and immunoglobulin gene rearrangements in vertebrate lymphocytes (Craig 1988; Blackwell and Alt 1989). Instances of developmentally regulated gene rearrangements in prokaryotes are quite rare and have been reviewed recently (Haselkorn 1992). Apart from the *Anabaena* rearrangements the only other developmentally regulated gene rearrangement in bacteria is the deletion of an intervening sequence from the chromosome of the

spore mother cell during sporulation of *B. subtilis* (Stragier *et al* 1989; Sato *et al* 1990; Losick and Stragier 1992) (table 2).

During *B. subtilis* differentiation, a 42 kb element named *skin* is deleted from within the *sigK* gene in the terminally differentiated mother cell (Kunkel *et al* 1990). It is a result of a site-specific recombination and requires a gene, the *spoIVCA,* located on the excised element. XisF closely resembles SpoIVCA (table 2). Amino acid sequence of the two proteins shows 49% similarity and 27% identity. The amino-terminal region of both XisF and SpoIVCA is homologous to the resolvase family of site-specific recombinases (Kunkel *et al* 1990; Sato *et al* 1990; Stark *et al* 1992) but the target sequence of SpoIVCA and XisF are quite different (Golden *et al* 1987; Stragier *et al* 1989).

Excision of lambda lysogen from *E. coli* subjected to an environmental insult can be considered a developmentally regulated gene rearrangement, except that it does not benefit the host. The main differences are that the directly repeated sequences at the end of lambda chromosome differ from those of 11 or 55 kb elements and the excised lambda element replicates and encodes several functional proteins. In sharp contrast, the rearrangements in *Anabaena* and *Bacillus* have no advantage for the excisons which bear nothing more than the genes encoding site-specific recombinases.

There are some gross similarities in the three developmental gene rearrangements observed in *B. subtilis* and *Anabaena* (table 2). They result in excision of DNA elements that interrupt coding regions of genes required for further development in *B. subtilis* or nitrogen fixation in heterocystous cyanobacteria. They are brought about by site-specific recombinases that recognize direct repeats and catalyse the excision of a non-replicating circular DNA molecule and fusion of the previously distant parts of the gene. All three occur only in one type of cell (heterocysts or spore mother cell) but not in the other (vegetative cells or forespores). Both heterocysts and spore mother cells are terminal cells providing no DNA for the next generation. *Anabaena* rearrangements are also unique in the sense that they act as mechanisms of gene regulation.

11. What are the interrupting elements and why are they maintained?

Anabaena rearrangements (as also those in *Bacillus)* may be modeled after the excision of lysogenic bacteriophage lambda (Campbell 1983). The 11 and 55 kb elements may be considered as defective prophages that have lost the ability to replicate independently and to encode any proteins, other than the site-specific recombinase. They probably are descendants of temperate cyanophages that must have evolved the ability to respond to the host's developmental regulatory mechanisms in order to co-exist with their host. Of course they excise or they would have disappeared long ago. But their excision from chromosome has come under the control of heterocyst development. Unlike most phages the excison is of no advantage to the excison but helps the host by facilitating correct, need-based expression of *nif* genes in an activity-friendly (anaerobic) atmosphere.

The machinery for site-specific recombination process is frequently encoded by episomal elements, e.g., bacteriophages, retroviruses, transposons, plasmids, etc. It is relatively straightforward to see the evolutionary advantage that is conferred on

the episome by the recombinase, e.g., integrase allows phage lambda to lysogenize, the transposase allows Mu DNA to replicate, Cre aids accurate segregation of phage P1 plasmid DNA (Sadowski 1993). In *Anabaena* the advantage seems to be to the host. The benefit to the episome, if any, is unknown at this point.

It is not clear how the 11 kb excison got inserted into *n ifD* originally or what preserves these elements in the vegetative cell DNA. BifA type of repressors may help to maintain the elements in the vegetative cells by binding to the NREs upstream of *xisA* and *xisF* keeping them repressed. Vegetative cells would then pass it on to the progeny. Heterocysts, where these elements are deleted are terminally differentiated cells that have no genetic continuity with the descendants of the filament.

The intervening elements are sometimes dispensable. For example *B. thuringiensis* does not have the interrupting element (Losick and Stragier 1992). Cyanobacterial strains lacking one or both the elements are known or have been created (table 1). Yet strains from all over the world have the same element in the same location. Why? Three suggestions have been made by Haselkorn (1992): (i) elements encode repressors that confer immunity against cyanophages, (ii) elements encode a restriction activity that protects against phages and foreign DNA, and (iii) elements encode enzyme(s) that repair DNA damage (Haselkorn 1992). For example *E. coli* carrying T*n5* has been shown to have a growth advantage due to bleomycin resistance gene that encodes a DNA repair enzyme (Blot *et al* 1991). Sequencing of the 11kb element has revealed, in addition to *xisA,* only one more ORF that could encode a cytochrome P450 like protein (Lammers *et al* 1990), but no other genes were found.

Quite mysterious is the growth advantage conferred upon the *Anabaena* host that harbours the episomal elements. A fascinating speculation is that vestiges of them may be functional in any of the ways described above. Alternatively, rearrangements may provide a function in heterocyst differentiation or nitrogen fixation that has not yet been discovered. The *nif* transcripts are not formed without rearrangements and their tight linkage to development of heterocysts ensures that the rearranged *nif* genes would always find the atmosphere conducive for their expression. Thus, once the *nif* genes have been rearranged *Anabaena* faces no further setbacks to nitrogen fixation, unlike other free-living diazotrophs. The latter have to shut off their nitrogen fixation (and sacrifice all the good work done in expressing *nif* genes and supporting activities of Nif proteins) any time when oxygen creeps into their environment. Heterocystous cyanobacteria lacking these elements or those cured of the 11 kb element show no distinct growth advantage in the short time—but a proper assessment is necessary to understand why such elements have been stably maintained by the vegetative cells of several strains from all over the world, during evolution.

12. Summing up

The rearrangements in the nitrogen fixation (*nif*) genes in heterocystous cyanobacteria were described only a decade ago. By and large the effort so far has been to characterize the gene rearrangements in terms of their location, topology of the

DNA involved at the breakpoints, the genes encoding the recombinases and their regulation and possible links between the heterocyst development and gene rearrangements. Some aspects of cyanobacterial *nif* gene rearrangements still remain somewhat enigmatic.

First of all we still do not know whether the rearrangements are indeed developmentally regulated. In many instances they do occur in the absence of mature heterocysts. There remain several unsolved puzzles regarding their association with heterocysts, *viz.,* (i) what maintains the excisons in the vegetative cells, (ii) how are the cells selected for differentiation, (iii) what are the signals or the biochemical triggers that induce rearrangements during differentiation, (iv) why under complete anaerobiosis also, the proportion of cells with rearranged *nif* gene does not exceed 8–10%, (v) how is the induction of heterocyst-specific genes confined to these cells only and (vi) why very rarely, certain heterocystous cyanobacteria do not possess these excisons?

From what is known both the rearrangements are irreversible. Do they contribute to commitment of vegetative cells to terminal differentiation into heterocysts? Fate of the excised circles is equally enigmatic. Although they do not seem to replicate or transcribe they are not degraded either. Heterocysts of *Anabaena* sp. strain PCC 7120 do not regress or de-differentiate but in another *Anabaena* sp. strain L-31 de-differentiation of heterocysts and their return from a diazotrophic to photosynthetic mode has been reported earlier (Thomas 1970, 1972). It would be of interest to know what happens to the excised DNA in such functionally de-differentiated heterocysts.

Several questions concerning the actual mechanism of rearrangements also remain unanswered. Although only XisA and XisF proteins alone have been implicated in the site-specific recombinations so far, are they enough? All recombinases bind to two inverted protein binding sites surrounding the site of cleavage and strand exchange. Some of these sites also contain binding sites for accessory site-specific DNA-binding proteins (Sadowski 1993). Are the XisA or XisF recombinases aided by such accessory proteins in *Anabaena?* If so, what are the target sites for their binding and where are they located? Are these proteins optimally present in *E. coli* or are the low levels of rearrangements in *E. coli* to be blamed on inadequate levels of any of these?

Almost unknown are the reasons for the low level expression of the *xisA* and *xisF* genes. In fact transcripts corresponding to these have never been detected in *Anabaena* sp. strain PCC 7120. Are the *xisA, xisF* transcripts unusually unstable? Are the recombinase proteins very toxic especially at higher levels? Lack of excisase-overproducing strains is hampering *in vitro* studies on the mechanisms involved in site-specific recombinations.

Finally, there is the question of preservation of such elements through evolution and the selective advantages conferred by them to the host, if any. A new *Anabaena* sp. strain 268S10 lacking the *nifD* element has been created. This or other strains lacking *nifD* element grow normally. Brusca *et al* (1990) have considered the possibility that in 268S10 the persisting plasmid pAM268 plays a role in maintaining a normal phenotype by supplying XisA protein but they have been unable to cure strain 268S10 of the plasmid pAM268 to check this possibility. Whether the 11 kb element confers some benefit to cells growing in the natural environment is, as of now, unknown.

Acknowledgements

We thank Prof. R Haselkorn, Department of Molecular Genetics and Cell Biology, University of Chicago, Chicago, USA for providing cloned *Anabaena nif* probes. The help received from Dr M Seshadri in preparation of figure 1 is greatfully acknowledged. NP thanks the University Grants Commission, New Delhi, for the award of a fellowship.

References

Apte S K 1992 Molecular Biology of Cyanobacterial Nitrogen Fixation: Recent Advances; *Indian J. Microbial.* **32** 103–106

Apte S K 1993 Cyanobacterial Nitrogen Fixation: Molecular Genetic Aspects; *Proc. Indian Natl. Sci. Acad.* **59** 367–386

Apte S K and Gowda T K S 1992 Organisation and regulation of genes involved in cyanobacterial nitrogen fixation; in *Proceedings of the DAE Symposium on Molecular Biology of Microorganisms* (Bombay: BRNS DAE) pp 1–5

Apte S K and Thomas J 1987 Nitrogen fixation genes *nifK, D, H* in the filamentous non heterocystous cyanobaclerium *Plectoneina boryanuin* do not rearrange; *J. Genet.* **66** 101–110

Ausubel F M and Cannon F C 1980 Molecular genetic analysis of *Klebsiella pneumoniae* nitrogen fixation (*nif*) genes; *Cold Spring Harbor Symp. Quant Biol.* **45** 487–497

Bachmann B J 1983 Linkage map of *E. coli* K-12, edition 7; *Microbiol. Rev.* **47** 180–230

Barnum S R and Gendel S M 1985 Organisation of *nif, D, K* in the nonheterocystous filamentous cyanobacterium *Plectonema boryanum;* in *Nitrogen fixation research progress* (eds) H J Evans, P J Bottomley and W E Newton (Dordrecht: Martinus Nijhof) p 516

Berg D E and Howe M M 1989 *Mobile DNA* (Washington DC: American Society for Microbiology)

Bergman B and Carpenter E J 1991 Nitrogenase confined to randomly distributed trichomes in the marine cyanobacterium *Trichodesmiuin thiebautii; J. Phycol.* **27** 158–165

Beynon J, Cannon M, Buchanan-Wollaston V and Cannon F C 1983 The *nif* promoters of *K. pneumoniae* have a characteristic primary structure; *Cell* **35** 665–671

Blackwell T K and Alt F E 1989 Mechanism and developmental program of immunoglobulin gene rearrangement in mammals; *Annu. Rev. Genet.* **23** 605–636

Blot M, Meyer J and Arber W 1991 Bleomycin-resistance gene derived from the transposon Tn5 confers selective advantage to *Escherichia coli* K-12; *Proc. Natl. Acad. Sci. USA* **88** 9112–9116

Bohme H and Haselkorn R 1988 Molecular cloning and nucleotide sequence analysis of the gene coding for heterocyst ferredoxin from the cyanobacterium *Anabaena PCC* 7/20; *Mol. Gen. Genet.* **214** 278–285

Borthakur D, Basche M, Buikema W J, Boithakur P B and Haselkorn 1990 Expression, nucleotide sequence and mutational analysis of two open reading frames in the *nif* gene region of *Anabaena* sp. strain PCC 7120; *Mol. Gen. Genet.* **221** 227–234

Borthakur D and Haselkorn R 1989 Tn5 mutagenesis of *Anabaena* sp. strain PCC 7120: isolation of a new mutant unable to grow without combined nitrogen; *J. Bacteriol.* **171** 5759–5761

Brusca J S, Chastain C J and Golden J W 1990 Expression of the *Anabaena* PCC 7120 *xisA* gene from a heterologus promoter results in excision of the *nifD* element; *J. Bacteriol.* **172** 3925–3931

Brusca J S, Hale M A, Carrasco C D and Golden J W 1989 Excision of an 11-kilobase pair DNA element from within the *nifD* gene in *Anabaena* variabilis heterocysts; *J. Bacteriol.* **171** 4138–4145

Campbell A 1983 Bacteriophage lambda; in *Mobile genetic elements* (*ed.*) J A Shapiro (New York: Academic Press) pp 65–103

Cannon F C, Beynon J, Buchanan-Wollaslon V, Burghoff R and Cannon M 1985a Progress in understanding organisation and expression of *nif* genes in *Klebsiella; in Nitrogen fixation research progress* (eds) H J Evans, P J Bottomley and W E Newton (Dordrecht: Holland-Nijhof) pp 453–460

Cannon M C, Hill S, Kavanagh E and Cannon F C 1985b A molecular genetic study of *nif* expression in *Klebsiella* at the level of transcription, translation and nitrogenase activity; *Mol. Gen. Genet.* **198** 198–206

Capone D G and Carpenter E J 1982 Nitrogen fixation in the marine environment; *Science* **217**

1140–1142

Capone D G, O'Neil J M, Zehr J and Carpenter E J 1990 Basis for diel variation in nitrogenase activity in the marine planktonic cyanobacterium *Trichodesmium thiebautii; Appl. Environ. Microbial.* **56** 3532–3536

Carrasco C D, Ramaswamy K S, Ramasubramanian T S and Golden J W 1994 *Anabaena xisF* gene encodes a developmentally regulated site-specific recoinbinase; *Genes Dev.* **8** 74–83

Chastain C J, Brusca J S, Ramasubramanian T S, Tai-Fen Wei and Golden J W 1990 A sequence specific DNA-binding Factor (VF1) from *Anabaena* sp. strain PCC 7120 vegetative cells binds to three adjacent sites in the *xisA* upstream region; *J. Bacteriol.* **172** 5044–5051

Craig N L 1988 The mechanism of conservative site-specific recombination; *Annu. Rev. Genet.* **22** 77–105

Elhai J and Wolk C P 1990 Developmental regulation and spatial pattern of expression of the structural genes for nitrogenase in the cyanobacterium *Anabaena; EMBO J.* **9** 3379–3388

Fischer H M and Hennecke H 1984 Linkage map of the *Rhizobium japonicum nifH* and *nifD, K* operons encoding the polypeptide components of the nitrogenase complex; *Mol. Gen. Genet.* **196** 537–540

Franche C and Cohen-Bazire G 1985 The structural *nif* genes of four symbiotic *Anabaena azollae* show a highly conserved physical arrangement; *Plant Sci.* **39** 125–131

Franche C and Cohen-Bazire G 1987 Evolutionary divergence in the *nifK, D, H* gene region among nine symbiotic *Anabaena azollae* and between *Anabaena azollae* and some free-living heterocystous cyanobacleria; *Symbiosis* **3** 159–178

Frias J E, Merida A, Herrero A, Martin-Nieto J and Flores E 1993 General distribution of the nitrogen control gene *ntcA* in cyanobacteria; *J. Bacteriol.* **175** 5710–5713

Golden J W, Carrasco C D, Mulligan M E, Schneider G J and Haselkorn R 1988 Deletion of a 55-kilobase-pair element from the chromosome during heterocyst differentiation of *Anabaena* sp. Strain PCC 7120; *J. Bacteriol.* **170** 5034–5041

Golden J W, Mulligan M E and Haselkorn R 1987 Different recombination site specificity of two developmentally regulated genome rearrangements; *Nature* (*London*) **327** 526–529

Golden J W, Robinson S J and Haselkorn R 1985 Rearrangement of nitrogen fixation genes during heterocyst differentiation in the cyanobacterium *Anabaena; Nature* (*London*) **314** 419–423

Golden J W and Wiest D R 1988 Genome rearrangements and nitrogen fixation in *Anabaena* blocked by Inactivation of *xisA* gene; *Science* **242** 1421–1423

Golden J W, Whorf L L and Wiest DR 1991 Independent regulation of *nifH, D, K* operon transcription and DNA rearrangement during heterocyst differentiation in the cyanobacterium *Anabaena* sp. Strain PCC 7120; *J. Bacteriol.* **173** 7098–7105

Haber J E 1983 Mating-type genes of *Saccharomyces cerevisiae;* in *Mobile genetic elements* (*ed.*) J A Shapiro (New York: Academic Press) pp 559–619

Haselkorn R 1986 Organisation of the genes for nitrogen fixation in photosynthetic bacteria and cyanobacteria; *Annu. Rev. Microbiol.* **40** 525–547

Haselkorn R 1992 Developmentally regulated gene rearragements in prokaryotes; *Annu. Rev. Genet.* **26** 113–130

Haselkorn R 1989 Excision of elements interrupting nitrogen fixation Operons in cyanobacteria; in *Mobile DNA* (eds) D E Berg and M M Howe (Washington DC: American Society for Microbiology) pp 735–742

Haselkorn R, Golden J W, Lammers P J and Mulligan M E 1986 Developmental rearrangement of cyanobacterial nitrogen fixation genes; *Trends Genet.* **2** 255–259

Hefron F 1983 Tn3 and its relatives; in *Mobile genetic elements* (*ed.*) J A Shapiro (New York: Academic Press) pp 223–260

Herrero A and Wolk C P 1986 Generic mapping of the chromosome of the cyanobacterium, *Anabaena variabilis; J. Biol. Chem.* **261** 7748–7754

Hirschman J, Wong P K, Sei K, Keener J and Kustu S 1985 Products of nitrogen regulatory genes *ntrA* and *ntrC* of enteric bacteria activate *glnA* transcription *in vitro:* evidence tha the *ntrA* product is a sigma factor; *Proc. Natl. Acad. Sci. USA* **82** 7525–7529

Imperial J, Ugalde R A, Shah V K and Brill W J 1984 Role of the *nifQ* gene product in the incorporation of molybdenum in to nitrogenase in *Klebsiella; J. Bacteriol.* **158** 187–194

Kallas T, Coursin T and Rippka R 1985 Different organization of *nif* genes in heterocystous and nonheterocystous cyanobacteria; *Plant Mol. Biol.* **5** 321–329

Kallas T, Rebiere M C, Rippka R and Marsac NT 1983 The structural *nif* genes of the cyanobacteria *Gloeothece* and *Calothrix* sp. share homology with those of *Anabaena* sp., but the *Gloeothece* have

different arrangement; *J. Bacteriol.* **155** 427–431

Kaluza K, Fuhrmann M, Hahn M, Regensburgen B and Hennecke H 1983 In *Rhizobium japonicum* the nitrogenase genes *nifH* and *nifD.K* are separated; *J. Bacteriol.* **155** 915–917

Kessler C, Neumaier P S and Wolf W 1985 Recognition sequences of restriction endonucleases and methylases—a review; *Gene* 33 1–102

Kunkel B, Losick R and Stragier P 1990 The *Bacillus subtilis* gene for the development transcription factor sigma K is generated by excision of a dispensable DNA element containing a sporulation recombinase gene; *Genes Dev.* **4** 525–535

Lammers P, Golden J W and Haselkorn R 1986 Identification and sequence of a gene required for a developmentally regulated DNA excision in *Anabaena; Cell* **44** 905–911

Lammers P, Mclaughlin S, Papin S, Trujillo-Provencio C and Ryncarz A J 1990 Developmental rearrangement of cyanobacterial *nif* genes—nucleotide sequence, open reading frames and cytochrome P-450 homology of *Anabaena* sp. strain PCC 7120 *nifD* element; *J. Bacteriol* **172** 6981–6990

Losick R and Stragier P 1992 Criss-cross regulation of cell-type-specific gene expression during development in *B. subtilis; Nature* (*London*) **355** 601-604

Mazur B J, Rice D and Haselkorn R 1980 Identification of blue-green algal nitrogen fixation genes by using heterologous DNA hybridization probes; *Proc. Natl. Acad. Sci. USA* **77** 186–190

Meeks J, Joseph C M and Haselkorn R 1988 Organization of the *nif* genes in cyanobacteria in symbiotic association with *Azolla* and *Anthoceros; Arch. Microbiol.* **150** 61–71

Merrick M J 1988 Organisation and regulation of nitrogen fixation genes in *Klebsiella* and *Azotobacter,* in *Nitrogen fixation: Hundred years after* (eds) H Bothe, F J deBruijn and W E Newton (Stuttgart: Gustav Fischer) pp 293–302

Mulligan M E, Buikema W J and Haselkorn R 1988 Bacterial-type ferredoxin genes in the nitrogen fixation regions of the cyanobacterium *Anabaena* sp. strain PCC 7120 and *Rhizobium meliloti; J. Bacteriol.* **170** 4406–4410

Mulligan M E and Haselkorn R 1989 Nitrogen fixation (*nif*) genes of the cyanobacterium *Anabaena* species strain PCC 7120; *J. Biol. Chem.* **264** 19200–19207

Nierzwicki-Bauer S A, Curtis S E and Haselkorn R 1984 Cotranscription of genes encoding the small and large subunits of ribulose-l,5-bisphosphate carboxylase in the cyanobacterium *Anabaena* sp. Strain PCC 7120; *Proc. Natl. Acad. Sci. USA* **81** 5961–5965

Nierzwicki-Bauer S A and Haselkorn R 1986 Differences in mRNA levels in *Anabaena* living freely or in symbiotic association with *Azolla; EMBO J.* **5** 29–35

Orme- Johnson W H 1985 Molecular basis of nitrogen fixation; *Annu. Rev. Biophys. Biophys. Chem.* **14** 419–459

Padhy R N, Hottat F G, Coene M M and Hoet PP 1988 Restriction analysis and quantitative estimation of methylated bases of filamentous and unicellular cyanobacterial DNAs; *J. Bacteriol.* **170** 1934–1939

Prabhavathi N and Apte S K 1994 Carbon status during growth regulates the rearrangement of *nif*D element in the cyanobacterium *Anabaena* PCC 7120; in *Procedings of the DAE Symposium on Stress and Adaptive Responses in Biological Systems* (Bombay: BRNS, DAE) pp 217–222

Ptashne M 1986 Λ *genetic switch; Gene control and phage lambda* (Palo Alto: Blackwell Scientific)

Ramasubramanian T S, Wei T F and Golden J W 1994 Two *Anabaena* sp. strain PCC 7120 DNA-binding factors interact with vegetative cell- and heterocyst-specific genes; J. *Bacteriol.* **176** 1214–1223

Reznikoff W S 1992 Catabolite gene activator protein activation of *lac* transcription; *J. Bacteriol.* **174** 655–658

Rice D, Mazur B J and Haselkorn R 1982 Isolation and physical mapping of nitrogen fixation genes from the cyanobacterium *Anabaena* 7120; *J. Biol. Chem.* **257** 13157–13163

Ruvkun G B and Ausubel F M 1980 Interspecies homology of nitrogen fixation genes; *Proc. Natl. Acad. Sci. USA* **77** 191–195

Ruvkun G B, Long S R, Meade H M, van den Bos R C and Ausubel F M 1982 ISRml: A *Rhizobium meliloti* insertion sequence that transposes preferentially into nitrogen fixation genes; *J. Mol. Appl. Genet.* **1** 405–418

Sadowski P D 1993 Site-specific genetic recombination: hops, flips and flops; *FASEB J.* **7** 760–767

Santero E, Hoover T and Kustu S 1990 Mechanism of transcription from *nif* promoters: involvment of IHF; in *Nitrogen fixation: Achievements and objectives* (eds) P M Gresshoff, L E Roth, G Stacey and W E Newton (New York: Chapman and Hall) pp 459–466

Sato T, Samori Y and Kobayashi Y 1990 The *cisA* cistron of *Bacillus subtilis* sporulation gene *spoIVCA* encodes a protein homologous to a site-specific recombinase; *J. Bacteriol.* **172** 1092–1098

Saville B, Straus N and Coleman J R 1987 Contiguous organization of nitrogenase genes in a heterocystous cyanobacterium; *Plant Physiol.* **85** 26–29

Schopf J W 1975 The age of microscopic life; *Endeavour* **34** 51–58

Shah V K, Stacey G and Brill W J 1983 Electron transport to nitrogenase: purification and characterisation of pyruvate:ferredoxin oxidoreductase, the n*ifJ* gene product; *J. Biol. Chem.* **258** 12064–12068

Shaw D J, Rice D W and Guest J R 1983 Homology between CAP and Fnr, a regulator of anaerobic respiration in *Escherichia coli; J, Mol. Biol,* **166** 241–247

Silverman M and Simon M 1983 Phase variation and related systems; in *Mobile genetic elements* (*ed.*) J A Shapiro (New York: Academic Press) pp 537–557

Stal L J and Krumbein W E 1985 Nitrogenase activity in a non-heterocystous cyanobacterium *Oscillatoria* sp. grown under alternating light-dark cycles; *Arch. Microbiol.* **143** 67–71

Stark W M, Boocock M R and Sherralt D J 1992 Catalysis by site-specific recombinases; *Trends Genet.* **8** 432–439

Stragier P, Kunkel B, Kroos L and Losick R 1989 Chromosomal rearrangement generating a composite gene for a developmental transcription factor; *Science* **243** 507–512

Taub R, Kirsch I, Morton C, Lenoir G, Swan D, Tronick S, Aaronson S and Leder P 1982 Translocation of the *c-myc* gene into the immunoglobulin heavy chain locus in human Burkitt's lymphoma and murine plasmocytoma cells; *Proc. Natl. Acad. Sci. USA* **79** 7837–7841

Te!-Or E and Sandovsky T 1982 The response of the nitrogen-fixing cyanobacterium *Anabaena azollae* to combined nitrogen compounds and sugar; *Isr. J. Bot.* **31** 329–336

Thomas J 1970 Absence of the pigments of photosystem II of photosynthesis in heterocysts of a blue-green alga; *Nature* (*London*) **228** 181–183

Thomas J 1972 Relationship between age of culture and occurrence of the pigments of photosystem II of photosynthesis in heterocysts of a blue-green alga; *J. Bacteriol.* **110** 92–95

Tumer N E, Robinson S J and Haselkorn R 1983 Different promoters for the *Anabaena* glutamine synthetase gene during growth using molecular or fixed nitrogen; *Nature* (*London*) **306** 337–342

Vega-Palas M A, Flores E and Herrero A 1992 NtcA, a global nitrogen regulator from the cyanobacterium *Synechococcus* that belongs to the Crp family of bacterial regulators; *Mol. Microbiol.* **6** 1853–1859

Venkataraman G S 1979 Algal inoculation in rice fields; in *Nitrogen and rice* (*ed.*) N C Brady (Las Banos, Phillippines: International Rice Research Institute) pp 311–321

Walsby A E 1986 Origins of chloroplasts; *Nature* (*London*) **320** 21

Wei T F, Ramasubramanian T S and Golden J W 1994 *Anabaena* sp. strain PCC 7120 *ntcA* gene required for growth on nitrate and heterocyst development; *J. Bacteriol.* **176** 4473–4482

Wei T F, Ramasubramanian T S, Pu F and Golden J W 1993 *Anabaena* sp. strain PCC 7120 *bifA* gene encoding a sequence-specific DNA-binding protein cloned by *in vivo* transcriptional interference selection; *J. Bacteriol,* **175** 4025–4035

Weinman J J, Fellows F F, Gresshoff P M, Shine J and Scott K F 1984 Structural analysis of the genes encoding the molybdenum- iron protein in the *Parasponia rhizobium* strain ANU289; *Nucleic Acids Res.* **12** 8329–8344

Wolk C P 1982 Heterocysts; in *The biology of cyanobacteria* (eds) N G Carr and B A Whitton (Berkeley: University of California Press) pp 359–386

Wolk C P, Vonshak A, Kehoe P and Elhai J 1984 Construction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria; *Poc. Natl Acad. Sci. USA* **81** 1561–1565

Zehr J P, Ohki K and Fujita Y 1991 Arrangement of nitrogenase structural genes in an aerobic filamentous nonheterocystous cyanobacterium; *J. Bacteriol.* **173** 7055–7058