# Chapter 1: Introduction

Cyanobacteria are an ancient and morphologically diverse group of prokaryotes. Fossil records show their existence at least 3.5 billion year ago (Schopf, 1994; 1996). They were the first organism to use the oxygenic photosynthesis, responsible for converting earth's anoxic atmosphere to oxic environment. They are also considered to be the precursors of chloroplasts. Cyanobacteria are gram negative in nature, ranging from <1 μm to >100 μm in size. They are known in free form as well as in the association with a wide range of eukaryotic partners including plants (Bergmann *et al.*, 1996, Meeks, 98, Adams 99).

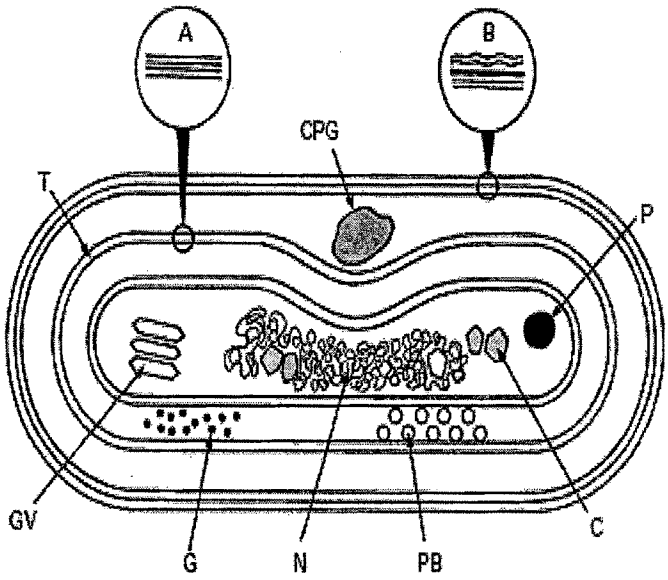



**Figure 1.1 : Schematic diagram of a thin section of a cyanobacterial cell :** C, carboxysome; CPG, cyanophycin granule; T, thylakoid; P, polyphosphate granule; N, nucleoplasmic region; G, glycogen granules; PB, phycobilisome; GV, gas vesicle. Inset (A) is an enlarged view of a thylakoid, showing the paired unit membranes. Inset (B) is an enlarged view of the cell envelope, showing the outer membrane (top), the peptidoglycan layer and the cytoplasmic membrane. *Adapted from Stanier & Cohen-Bazire (1977).*

Even though the cyanobacteria are gram negative, its cell wall is more of gram positive type (Fig. 1.1). The cell wall consists of two unit membranes, the cytoplasmic membrane and the outer membrane, separated by the electron opaque peptidoglycan layer, which is usually 1-10 nm thick (Stanier, 1988; Weckesser & Jurgens,1988; Castenholz & Waterbury, 1989;). The photosynthetic machinery of cyanobacteria is contained on, and within, the thylakoid membranes. Majority of them are facultative photoheterotrophs, capable of using light for source of energy. Cyanobacteria have been classified in to 5 classes depending on the morphology and mode of reproduction. Out of 5 classes, 4$^{th}$ class includes filamentous heterocystous morphology and mode of reproduction through fragmentation. *Anabaena* PCC 7120 belongs to this group.

*Anabaena* PCC 7120 is one of the few cyanobacteria capable of diazotrophy as well as photosynthesis simultaneously. Since the nitrogenase enzyme is extremely sensitive to even the traces of oxygen, nature has evolved different mechanisms to over come the problem. In some cases, both the processes are separated temporally, with photosynthesis occurring at daytime and nitrogen fixation occurs during night (Mishra & Tuli, 2000). However, the filamentous nature of cyanobacteria has adapted an alternative strategy, the commitment of a small number of cells to the highly specialized task of providing a suitable environment for $N_2$ fixation. These cells are known as heterocysts. Heterocyst formation occurs only when the cells are deprived of combined nitrogen source. Some other cyanobacteria like *trichodesminum* spp and *Gleotheces* shows both temporal as well as spatial separation (El-Shehawy *et al.*, 2003; Stahl, 2003). These organisms form special structures called diazocytes, where nitrogen fixation occurs. Commencing of nitrogen fixation in diazocytes show very little photosynthetic activity and *vice versa*. Activation and inhibition of these processes depends on the availability of C/N source which is sensed by the pool of glutamine.

**1.1 Heterocyst Differentiation:**

Heterocysts are highly specialized cells, required for fixation of atmospheric nitrogen under oxic condition, in the absence of combined nitrogen source. Heterocysts are terminally differentiated cells, which do not divide further, consistence with the lack of FtsZ protein in the heterocyst (Kunh *et al.*, 2000). Generally around 10% of the cells of the filament undergo heterocyst differentiation.

Heterocyst formation involves changes at genetic, morphological and biochemical level from the vegetative cell. These differences are made to make the environment suitable for nitrogen fixation, and prevent the oxygen- sensitive nitrogenase enzyme system. Morphological changes include formation of additional glycolipid and polysaccharide layer in order to prevent the diffusion of gases, degradation of photosystem II to prevent the generation of oxygen by photo synthesis, Increased rate of respiration to eliminate residual oxygen, constriction of the pore between surrounding vegetative cells to prevent the entry of gases, lack of photosynthetic $CO_2$ fixation to divert the energy and reducing equivalent towards nitrogen fixation (Walsby, 1985; Murry & Wolk, 1989; wolk *et al.*, 1994). The heterocyst formation involves the up regulation of oxidase activity as well as of uptake hydrogenase, both involved in the consumption of residual oxygen within the heterocyst (Tamagnini *et al.*, 2004). As heterocyst development progresses, the septum separating it from the adjacent vegetative cell decreases considerably in diameter, until it becomes a narrow pore channel. The number of plasma bridges (microplasmodesmata) traversing the septum also decreases three- to fivefold (Fay, 1992).

Initial signs of differentiation of heterocyst are easily observed by physiological and ultra structural changes. Immediately after the onset of heterocyst differentiation, the first major event includes degradation of large amount of proteins, especially the phycobilliproteins, which account for up to 50% of the protein. This results in the loss of fluorescence, first microscopic sign

of differentiation in the heterocyst-committed cells. Formation of mature heterocyst takes almost 24 hours after the deficiency of nitrogen fixation. Pre heterocyst- and intermediate stage differs from the vegetative cell in shape and granularity. It shows formation of additional envelop and reorganization of intracellular membrane at morphological level, and degradation of photosynthetic machinery at metabolic level (Nierzwicki-Bauer *et al.*, 1984; Sherman *et al.*, 2000). Pre heterocysts are reversible i.e. If provided the nitrogen source, can revert back to the vegetative cell. However, after the commitment phase passes, that is 8 hours onwards; it will progress for the differentiation. The exact signal for the commitment is not established yet, but the generation of anaerobic condition might be a developmental checkpoint leading to the further maturation of heterocyst (Fiedler *et al.*, 1998; Zhu *et al.*, 2001).

## 1.2 Interaction between Heterocyst and vegetative cell:

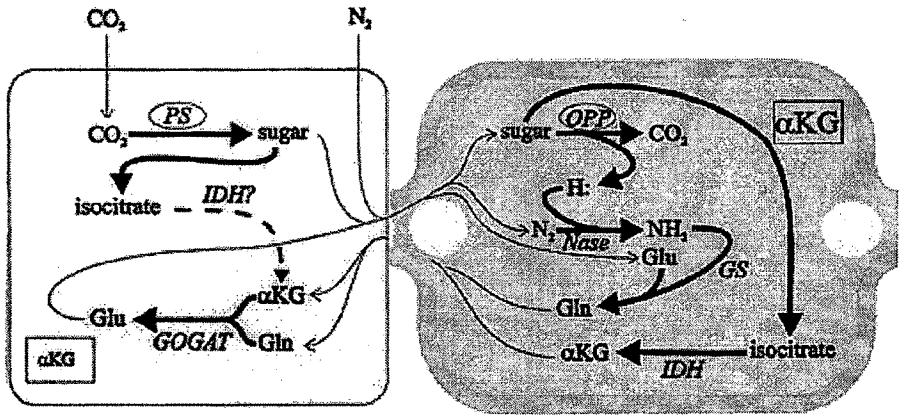
Both vegetative cells and heterocysts, in a filament are interdependent on each other for the growth. The heterocyst provides the fixed nitrogen to the entire filament. Although the definite mechanism for the transference of nitrogen source is not known, two views have been suggested. On the basis of involvement of uptake amino acid permase in the growth, has suggested a mechanism where amino acids can be exported to the periplasmic space, which is continuous with the vegetative cell and from there, taken up by amino acid permease. Another view involves transfer of nitrogen in the form of glutamine through the activity of GS. However, recently it was found that the GOGAT activity in the heterocyst seems to be low, so the glutamate required for glutamine formation, might be coming from the vegetative cell or may be synthesized in the heterocyst by pathway other than Glutamate synthase.

Since the heterocyst has lost the capacity for the $CO_2$ fixation, they are largely dependent on the vegetative cells for the source of energy and reductant required for nitrogen fixation. The vegetative cell transfers mainly sucrose to the

heterocyst, which contains enzyme alkaline invertase that converts sucrose to fructose and glucose (Figure 1.2). These hexoses are metabolized mainly by pentose phosphate pathway. Activities of the enzymes like ferrodoxin:pyruvate oxidoreductase and flavodoxin oxidoreductase provides the electrons required for the nitrogen fixation.

The morphological and biochemical changes occurring in the heterocyst formation involves differential gene expression in vegetative cell and heterocyst cell. Differentiation of heterocyst has been shown to up regulate at-least 1000 genes in the micro array data. Some genes are expressed exclusively in vegetative cells like *rbcLS*, others exclusively in heterocyst like *nifHDK* and some other gets expressed in both vegetative as well as heterocyst like *glnA*.

**Figure 1.2: Possible Metabolic interaction of vegetative cell and heterocyst cell (Meeks & Elhai, 2002)**




### 1.3 How do cyanobacteria sense Nitrogen status?

Cyanobacteria form heterocyst only in the absence of ammonia and nitrate (Wolk, 2000; Meeks & Elhai, 2002). Sensing mechanisms of nitrogen availability in cyanobacteria is different from well characterized Enterobacteraceae family (Fischer *et al.*, 1981; Merida *et al.*, 1990; 1991a; 1991b).

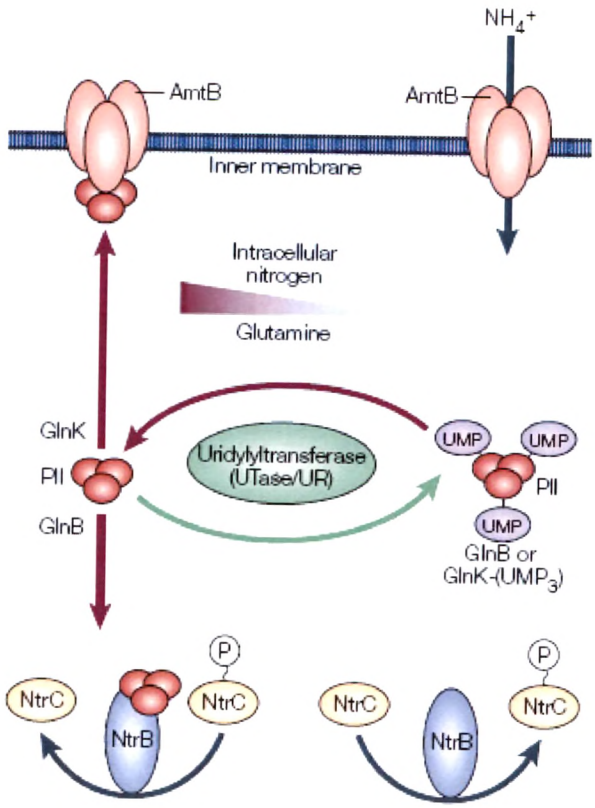



**Figure 1.3 Mechanism of nitrogen status sensing by cyanobacteria**

In *E. coli*, at high level of ammonia, Glutamate dehydrogenase (GDH) enzyme is responsible for synthesis of glutamate, while at lower concentration, Glutamine synthetase-Glutamate oxoglutarate amino transferase (GS-GOGAT) is active. GS activity is high when the intracellular glutamine levels are high. Regulation of GS activity occurs at three levels (i) allosteric regulation (ii) Protein modification (iii) gene expression (Jaggi *et al.*,1997). PII, a *glnB* gene product, is the central molecule involved in the regulation of nitrogen metabolism. $P_{II}$, a trimeric protein, is uridylated/deuridylated at 51 tyrosine residue by a single enzyme possessing uridylyltransferase (UTase)/uridyl removal (UR) activities. UTase activity is triggered by low glutamine levels whereas UR activity occurs in the presence of high levels of glutamine and ATP (Arcondéguy *et al.*, 2001). Unmodified $P_{II}$ protein confers adenylation activity to Adenylyl transferase (ATase) and deadenylation activity in $P_{II}$-UMP form, which regulates the GS

activity (figure 1.3). This modified form of PII do not interacts with the NtrB which by its kinase activity, phosphorylates of NtrC, a transcriptional activator, which in turn triggers on the transcription of the genes involved in alternate nitrogen source metabolism (Stock *et al.*, 1989).

In unicellular cyanobacteria like *Synechococcus* and *Synechocystis,* PII is involved in the nitrogen regulation however, there are some differences, instead of uridylation, it undergoes phosphorylation at serine 49 position (Forchhammer & Tandeau de Marsac 1994). PII does not appear to play a major role, since the PII mutant of *Synechococcus* retains nitrogen status dependent regulation similar to wild type. Growth of cyanobacteria on alternative nitrogen source like nitrate seems to be dependent up on another protein called NtcA, playing similar role to NtrC of *E. coli*

**1.4 Role of NtcA as sensor of nitrogen status**

NtcA is a global nitrogen regulator and binds to a consensus sequenc TGTAN$_8$TACA in which the GTN$_{10}$AC subset is essential for binding of NtcA (Jiang *et al.*, 2000). NtcA is found in wide range of cyanobacteria from unicellular to heterocystous form (Herrero *et al.*, 2001). Main function of the NtcA has been in the activation of the genes required for utilization of alternative nitrogen source. In case of *Synechococcus* sp. strain 7942, it is found to activate the operon of nitrate utilization genes including the transporter genes (Maeda *et al.*, 1998). It is also shown to up regulate *glnA, amtA* and *cynABDS* gene cluster (Luque *et al.*, 1994; Harano *et al.*, 1997; Hererro *et al.*, 2001). In Several genes of the cyanobacterium *Synechocystis* sp. strain PCC 6803 have been shown to be transcribed from N-regulated tsps which are preceded by NtcA-type promoters (Fig. 1.4). They include *amt1* (Montesinos *et al.*, 1998), *glnA, glnB* (García-Domínguez & Florencio 1997), *icd* (Muro-Pastor *et al.*, 1996), and the *rpoD2-V* gene encoding a factor affecting survival under nitrogen stress (Herrero *et al.*, 2001). In *Anabaena* PCC 7120, combined nitrogen assimilation genes or operons whose expression is NtcA dependent include the *glnA* gene and the *nir* (nitrate

assimilation) and *urt* (urea transport) operons (Frias *et al.*, 1997; Flores *et al.*, 1999; Frias *et al.*, 2000) NtcA acts as an transcriptional activator as well as repressor for genes not only associated with nitrate assimilation and nitrogen fixation but also with heterocyst differentiation.

Unicellular cyanobacterial *ntcA* mutant does not grow on alternative nitrogen source like nitrate, while the heterocystous cyanobacterial mutant fails to form heterocysts (Frias *et al.*; Wei *et al.*, 1994). NtcA is a transcriptional regulator of Catabolite Activator Protein (CAP) family. These proteins are characterized by bearing close to their C-terminus a helix-turn-helix motif for interaction with DNA, and their DNA binding and transcription activation activities are modulated through protein motifs present in their N-terminal parts (Kolb *et al.*, 1993).

### 1.5 Factors affecting NtcA activity

In cyanobacteria, 2-Oxo-glutarate pool plays a major role in the nitrogen metabolism as cyanobacteria do not contain 2-OG dehydrogenase hence the TCA cycle does not operate (Stanier *et al.*, 1977) . Thus, 2-OG serves as the carbon skeleton for the ammonium assimilation through GS-GOGAT pathway. Accumulated levels of 2-OG, both *in vivo* and *in vitro,* has been demonstrated to be the signal for the deficiency of nitrogen in cyanobacteria (Muro-pastor *et al.*, 2000; Li *et al.* 2003; Varquez-Bermudez *et al.*, 2003). 2-OG levels are higher when grown on nitrate as compared to that on ammonia. Analogues of the 2-OG, like 2,2-difluoro-pentanedioic acid (DFPA), was used to show the increase in level grown on nitrate as well as its binding with NtcA (Laurent *et al.*, 2005) . Due to the nitrogen deficiency, the increased level of 2-OG transmits the signal through its binding with NtcA. Although 2-OG binds to PII, it does not mediate the activation of NtcA gene as *Anabaena* PCC 7120 PII mutant strains show normal activation of *ntcA* mediated gene expression (Forchhammer, 2004).

Experiments have been done where second copy of *ntcA* gene was over expressed to determine its effect on heterocyst initiation. Even after over expression, the activity of NtcA, was still found under nitrogen regulation proving that NtcA activity depends on the nitrogen status of the cell, sensed by the 2-OG (Luque *et al.*, 2004; Olmedo-verd *et al.*, 2005). It has been shown that presence of 2-OG increases the binding affinity of NtcA gene to its promoter. *In vitro* studies with the NtcA and *glnA* promoter has proved that in presence of 2-OG binding affinity increased 5 fold. It has been suggested that direct interaction of 2-OG with the NtcA induces a conformational change in the NtcA, resulting in the change in its affinity towards its DNA binding sequence (Tanigawa *et al.*, 2002; Varquez-Bermudez *et al.*, 2003).

In addition to 2-OG levels, redox status of the cell has been shown to affect the NtcA binding to its promoter. Studies in *Synechocystis* PCC 6803 with various reducing agents proved that the binding of NtcA to its own promoter depends on the redox potential of the cell (Alfonso *et al.*, 2001).

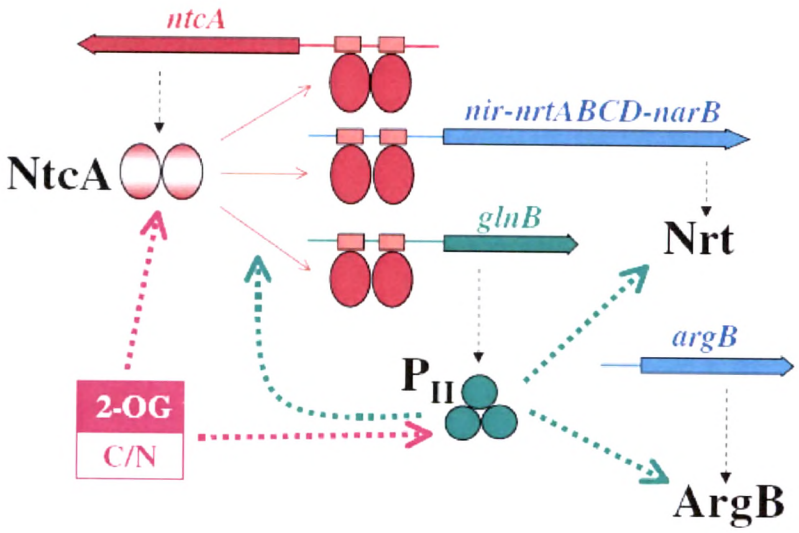



**Figure 1.4: Regulation mediated by NtcA in Cyanobacteria**

**1.6 Genes involved in Heterocyst formation**

Formation of heterocyst involves major changes at biochemical and genetic level. Microarray analysis of *Anabaena* PCC 7120 at different time points after the combined nitrogen source deficiency determined the genes associated with heterocyst differentiation (fig 5). Of the 2,380 DNA segments, 340 DNA segments, containing 1,018 ORFs, were found to be up regulated at one or more time points during the 24 hour growth without combined nitrogen, while 171 DNA segments, containing 520 ORFs were down regulated (Ehira *et al.*, 2003) These include genes involved in photosynthesis and carbon dioxide fixation, and those encoding ATP synthase, transporters of sulfate, phosphate and iron, and ribosomal proteins indicating decrease in the activity of various metabolic pathways to maintain a balance of various nutrients under the nitrogen-limited condition.

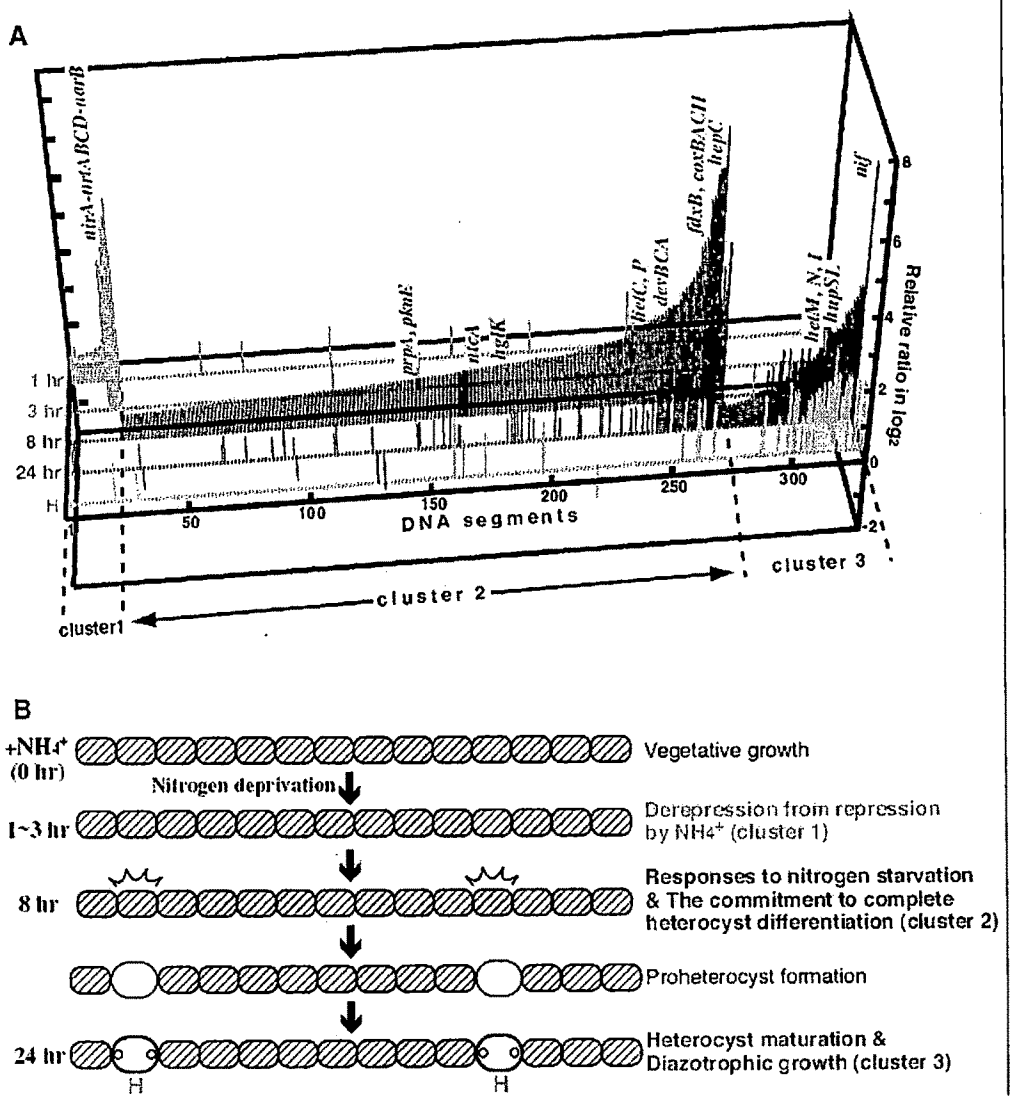



Figure 1.5 (A) Profiling of expression pattern for DNA segments upregulated by nitrogen deprivation. DNA segments upregulated at one or more time points during the 24 hr growth without combined nitrogen were used in this profiling. Each unit on the x-axis represents one DNA segment and each time point was plotted on the y-axis. 'H' on the y-axis indicates the comparison of heterocyst-enriched fraction with whole filaments at the 48th hour of nitrogen deprivation. The height of the bar is proportional to the relative ratio in the base-2 logarithm for each DNA segment at each time point. The relative ratio in the base-2 logarithm for DNA segments judged as 'no change' was assigned a value of 0.

DNA segments were classified into three clusters depending on the expression pattern, and then arranged in the ascending order of relative ratio within each cluster. (B) Schematic representation of various stages of heterocyst differentiation.

The time-course experiment showed that the response to nitrogen deprivation in *Anabaena* can be divided into several phases (Figure 1.5). The first phase involves derepression of ammonia repressed genes within 1 hr of nitrogen deprivation (cluster 1 in Figure 1.55A and '1-3 hr' in Figure 1.5 B). This mainly includes genes for transporters and metabolic enzymes required to use alternate nitrogen sources such as nitrate. Second, heterocyst differentiation is triggered by nitrogen starvation at about the eighth hour of nitrogen deprivation. Marked changes in gene expression occurred at this phase. Genes required for heterocyst differentiation and related to the function and structures of heterocysts are up regulated, perhaps in the pro heterocysts (cluster 2 in Figure 1.55A and '8 hr' in Figure 1.5 B). The eighth hour of nitrogen deprivation corresponds to the timing of the commitment to complete heterocyst differentiation which involves expression of many ORFs including *hglK*, *hetC*, and *hetP*. Finally, genes related to heterocyst function are upregulated in the heterocysts at the 24th hour of nitrogen deprivation (cluster 3 in Figure 6A and '24 hr' in Figure 6B). Fifty to 80% of ORFs whose expression was maximal at the 24th hour were preferentially expressed in heterocysts. Expressed islands were also upregulated in heterocysts. At the 24th hour of nitrogen deprivation, some ORFs were up regulated in both vegetative cells and heterocysts. These ORFs might play a role in diazotrophic growth.

Some of the genes involved in heterocyst development are discussed below.

**1.7 *ntcA*:**

*ntcA* is the first gene to be upregulated immediately after the deficiency of combined nitrogen source. Its level starts increasing and reaches maximum within 30 minutes (Golden, 2003).

### 1.8. Role of 2-OG levels in the activity of NtcA

It is known that NtcA is involved in both, activation of nitrate uptake genes as well as in heterocyst formation. At the same time, presence of nitrate inhibits heterocyst formation. In order to understand the differential regulation of NtcA on nitrate and nitrogen fixing conditions, a hypothesis has been proposed based on the accumulation of 2-OG levels. 2-OG reaches different cellular concentrations that serve as thresholds above which nitrate utilization or heterocyst formation is activated. In the presence of ammonium, 2-OG remains at a low level (Muro-Pastor *et al.*, 2001; Laurent *et al.*, 2005) but on nitrate 2-OG reaches a level sufficient for the activation of the genes involved in nitrate uptake and assimilation, but not to activate genes involved in heterocyst development. When filaments are starved of combined nitrogen, 2-OG accumulates up to the highest level (Muro-Pastor *et al.*, 2001; Laurent *et al.*, 2005) and the autoregulatory effect of *ntcA* gene expression may further amplify the nitrogen starvation signal and lead to the expression of *hetR* (Muro-Pastor *et al.*, 2002), a gene required for initiation of heterocyst differentiation (Buikema and Haselkorn, 1991).

#### 1.8.1 Expression of NtcA in *Anabaena* PCC 7120

*Anabaena* PCC 7120 *ntcA* gene has been reported to be transcribed from four N-regulated transcription start points (tsp) ( Ramasubramanium *et al.*, 1996). A tsp located 136 nucleotides upstream from the start of the coding sequence has been detected with RNA isolated from cells grown with ammonium, nitrate, or dinitrogen. One of the N-regulated transcription start points located at position 249, is preferentially used in the absence of combined nitrogen and early during heterocyst differentiation, and is active in mature heterocysts (Ramasubramanium *et al.*, 1996). Another is located at position 2180 and seems to be transiently used early in heterocyst development but not in mature heterocysts (Ramasubramanium *et al.*, 1996). The third N-regulated tsp is

located at position 2190 and appears to be used increasingly during heterocyst differentiation and also in mature heterocysts (96). Activation of the 249 and 2180 *tsps* in response to combined nitrogen deprivation is strictly dependent on NtcA, indicating that *ntcA* is also autoregulatory in *Anabaena* strain PCC 7120. NtcA has been reported (Ramasubramanium *et al.*, 1996) to bind to a region covering nucleotides 2131 to 2155, which includes a sequence ($GTAN_8AAC$) strongly similar to the consensus of NtcA-binding sites, but this sequence is located 88 bp upstream from the closer tsp (i.e., the 249 tsp). Thus, no NtcA-activated promoter could be recognized upstream from any of the three *ntcA* transcription start points in *Anabaena* strain PCC 7120. Therefore, the autoregulatory effect of NtcA in this strain might be exerted in an indirect manner or might involve a different type of NtcA-dependent promoter.

NtcA is responsible for upregulation of many genes including *hetR*, which is most important gene for heterocyst development. Expression of *hetR* gene is dependent on NtcA, even though it is indirect, since the *hetR* promoter does not contain NtcA binding sites. At the same time HetR also has been shown to positively regulate the *ntcA* expression (Muro-pastor *et al.*, 2002). Based on these facts, a model has been proposed as in figure 1.6.

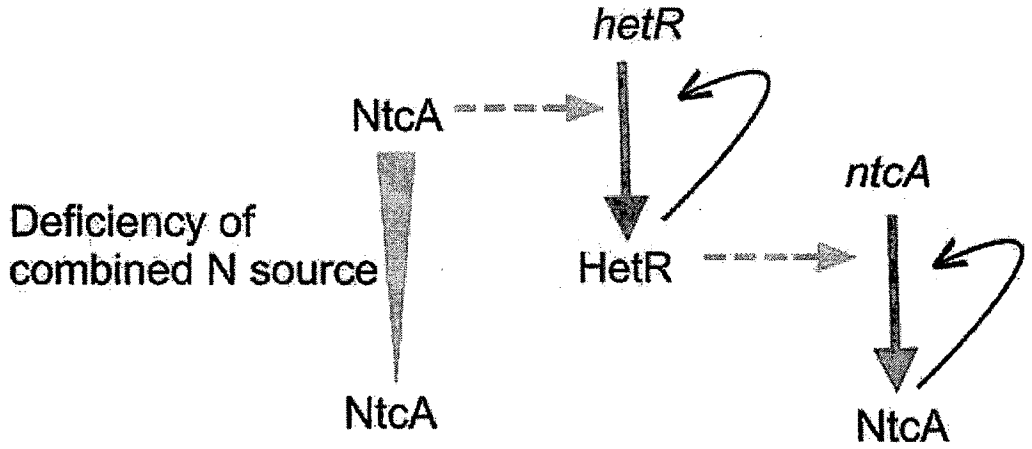



**Figure 1.6 Interdependence of NtcA and HetR for expression**

During the early stages of heterocyst initiation, residual NtcA present in the vegetative cell induces the expression of *hetR* gene. HetR protein in turn by auto regulation increases its level. This HetR up regulates the *ntcA* expression from the different tsp (-49/-180), specific for the heterocyst initiation and one for the maturation. NtcA also shows auto regulation by which its level will increase further. So both ntcA and hetR shows mutual dependence on each other for their expression (Muro-pastor *et al.*, 2002).

### 1.8.2 Role of NtcA in heterocyst differentiation

Number of genes has been shown to be positively regulated by NtcA, required for heterocyst differentiation. *hetR*, the master gene for the heterocyst formation, requires NtcA for its expression, as mutant strain of NtcA, does not show hetR expression and heterocyst formation. NtcA is also required for the activation of alternative nitrogen source utilization genes like *nir, nar, nrt,* urea utilization genes, etc. (Table 1) (Herrero *et al.*, 2004)

**Table 1.1: Genes regulated by NtcA**

| Positively regulated genes | |
|---|---|
| *glnA* | Involved in GS-GOGAT pathway |
| *nir* | Nitrate assimilation |
| *CphA1-CphB1* | Cyanophycin synthetase and cyanophycinase |
| *Cox2-cox3 operon* | Terminal respiratory oxidase |
| *urt* | Urea Transport |
| *ntcB* | Co-factor for Nitrite assimilation |
| *hepC* | Heterocyst development |
| *DevBCA* | Maturation of Heterocyst envelop |
| *petH* | Ferredoxin:$NADP^+$ reductase |
| *nifHDK* | Nitrogenase enzyme complex |
| *xisA* | Excision of *nifD* element |

| Negatively regulated genes | |
|---|---|
| *rbcLS* | Ribulose 1,5 bis phosphate carboxylase/ Oxigenase |
| *gorA* | Glutathione reductase |
| *hanA* | HU like protein |

**1.9 Relationship between 2-OG and free calcium ions**

Nitrogen starvation leads to increased concentrations of free calcium ion, which was found to be necessary for heterocyst development (Torrecilla *et al.*, 2004; Zhao *et al.*, 2005). Using the $Ca^{2+}$ binding luminescent photoprotein aequorin as a reporter, a transient, approximately threefold increase in the $Ca^{2+}$ level was detected about 1 h after combined-nitrogen step-down (Torrecilla *et al.*, 2004). The $Ca^{2+}$ signal seems to be working earlier than *hetR*, as the *hetR* mutant shows increase in the $Ca^{2+}$ concentration similar to wild type. Interestingly, the transient increase in the $Ca^{2+}$ level observed in response to combined-nitrogen removal was found to mainly result from the mobilization of internal $Ca^{2+}$ storage (Torrecilla *et al.*, 2004). This finding has been confirmed by studies on a calcium-sequestering protein, CcbP (Zhao *et al.*, 2005). The inactivation of *ccbP* leads to the formation of multiple contiguous heterocysts (Mch phenotype), whereas over expression of *ccbP* results in a low level of free $Ca^{2+}$, which is correlated with lower rates of *hetR* induction and a lack of heterocyst differentiation. A different $Ca^{2+}$ reporter, obelin, has been used to monitor the calcium levels in individual cells along the filaments (Zhao *et al.*, 2005). The $Ca^{2+}$ level was found to be about 10-fold higher in heterocysts than in vegetative cells, which correlates with the levels of expression of *ccbP* (which are high in vegetative cells and undetectable in heterocysts). CcbP itself might not be directly involved in heterocyst development but rather exert its effects by sequestering $Ca^{2+}$ under conditions of deprivation of combined nitrogen. This was supported by the finding that the heterologous calcium-binding protein calmodulin, when produced in *Anabaena* PCC 7120, also prevents heterocyst development in a similar way to CcbP. These results suggest that under nitrogen

sufficiency conditions, CcbP may bind to $Ca^{2+}$ and maintain the free $Ca^{2+}$ at a low level, and that the CcbP–$Ca^{2+}$ complex may serve as a $Ca^{2+}$ storage device. Under conditions of limitation of combined nitrogen, CcbP may be preferentially inactivated by an unknown mechanism in developing cells, which may increase the $Ca^{2+}$ levels, resulting in heterocyst differentiation. Both $Ca^{2+}$ and 2-OG levels are early signals but the relationship between the two is not known. The accumulated levels of 2-OG analogue DFPA suffice to trigger heterocyst development, even when ammonium is available; while $Ca^{2+}$ ions exert their effects only under conditions of combined-nitrogen deprivation. In the *ccbP* mutant, heterocyst differentiation did not occur when a combined-nitrogen source was present, even when the $Ca^{2+}$ ion concentrations were high. The Mch phenotype was observed in this mutant only when the combined-nitrogen source was removed (Zhao *et al.*, 2005). These data suggest that the effects of CcbP are subject to nitrogen control, which is possibly mediated by 2-OG and NtcA.

### 1.10 Role of *nrrA* in the activation of *hetR*

*nrrA* codes for nitrogen-responsive response regulator protein. Immediately after nitrogen depletion, its level starts rising and reaching at peak around 3 hours. Putative promoter region shows typical NtcA binding sites, indicating it required NtcA for the expression. Recent studies have shown that NrrA directly regulates the expression of *hetR*- the master gene for heterocyst formation. NrrA belongs to OmpR family of response regulator protein (figure 1.7)

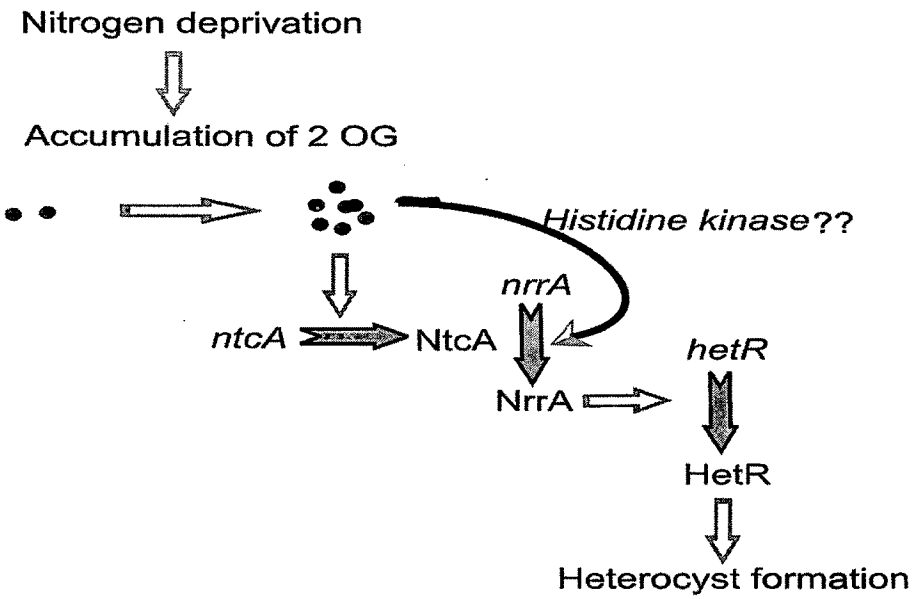



**Figure 1.7 Role of *nrrA* in *hetR* expression and heterocyst formation**

NrrA binds to the -844 to -818- a 27 base pair tsp of *hetR* gene, containing a tandem repeat sequence of CTT(A/G)AT(G/T)T. Expression of *hetR* from the tsp depends on the binding of *nrrA*, which in turn depends on NtcA. Based on the available evidences model has been proposed for the possible involvement of NrrA in heterocyst differentiation (fig ) : In the absence of combined nitrogen source, 2 OG level starts increasing which in turn enhances the DNA binding activity of NtcA. NtcA activates many genes involved in heterocyst differentiation including *nrrA*. NrrA is required for the expression of differentiation specific *hetR* transcription initiation from -728, -696 and -273 position. In the mutants of *nrrA*, heterocyst formation is delayed and over expression of *nrrA* increases the heterocyst frequency by 2-3 folds (Ehara & Ohmori, 2006).

### 1.11 HetR- the master regulator of heterocyst differentiation

HetR is the main positive regulator of heterocyst differentiation. Its expression starts after 3 hours of the nitrogen step down (Golden & Yoon, 2003). Expression of *hetR* depends on NtcA, however further expression of *ntcA* by auto regulation depends on HetR. NtcA mediates the expression of *hetR* through

NrrA. Two other proteins, PatA and HetF are also required for the up-regulation of *hetR*. Mechanism of HetF mediated activation is unknown (Wong & Meeks, 2001), while in case of PatA, which is cheY type response regulator, direct protein-protein interaction might be involved that may prevent HetR auto proteolysis or relieve the inhibition by the PatS (Liang *et al.*, 1992; Buikema and Haselkorn,2001). HetR alone seems to be enough for the heterocyst formation as over-expression of HetR from the Cu regulated promoter has resulted in the Mch phenotype and *hetR* mutants were unable to form heterocyst.

Two activities of HetR are known- (i) a serine type auto degrading protease and (ii) DNA binding activity (Zhou *et al.*, 1998; Huang *et al.*, 2004). For the DNA binding activity, it requires homodimerization involving Cys-48 residues. Side directed mutagenesis of the HetR C48A results in the loss of DNA binding activity. HetR has been shown to bind at the upstream region of *hetR*, *patS* and *hepA* thereby controlling there expression. Another HetRR223W mutant results in the loss of HetR regulation by PatS and HetN (Khudyakov and Golden, 2004).

**1.12 Pattern formation and maintenance**

In *Anabaena* PCC 7120, around 8-10% cell of the filament undergoes the differentiation and it provides one of the very few examples of the pattern formation in prokaryotes (Haselkorn, 1992). Three genes, *hetR*, *patS* and *hetN* are responsible for pattern formation and maintenance. The principle gene responsible for the pattern formation is *patS* which governs *de novo* pattern formation when filaments are induced to differentiate. A *patS* null mutant exhibits a phenotype of multiple contiguous heterocysts (Mch), an altered pattern of shortened vegetative cell intervals (Pat) on nitrogen-deficient media and abnormal differentiation in the presence of fixed nitrogen (Yoon and Golden, 1998). *patS* gene encodes a 17-amino-acid peptide, and exogenous addition of its C-terminal pentapeptide to a culture of *Anabaena* prevents differentiation. Receptor for PatS signal has been postulated to be cytoplasmic

(Wu *et al.*, 2004) and that PatS pentapeptide prevents binding of HetR to regions of the *hetR* promoter (Huang *et al.*, 2004) suggests a direct action of PatS in prevention of *hetR* autoregulation. To govern pattern formation PatS acts non-cell autonomously (Yoon and Golden, 1998). Consequently, it has been proposed that the PatS peptide diffuses away from differentiating proheterocysts along the filament to create a gradient of inhibitory signal. Expression of *patS* is constitutive in ammonia grown cultures, but disappears after the nitrogen deprivation and starts accumulating only after 6 h in heterocyst committed cells. After differentiation is complete, expression of *patS* returns to pre-induction levels, suggesting that *patS* controls *de novo* pattern formation but may not be the sole factor controlling maintenance of the pattern. The fact that a *patS*-deletion strain shifts from a Pat, Mch pattern of heterocysts to a more wild-type pattern after 72 h (Yoon and Golden, 2001) suggests that another factor(s) independent of *patS* is involved in pattern maintenance. This protein has been shown to inhibit the heterocyst formation in concentration dependent manner. Expression of *patS* depends on HetR by binding to the promoter region of *patS*. Thus, the ratio of PatS/HetR is very important for the maintenance of the pattern.

A model has been proposed where the differentiation committed cells produces the inhibition signal PatS from the mature protein, which then diffuses away in the surrounding vegetative, cell and prevents the heterocyst differentiation (figure 1.8). Signal of the patS in surrounding cells is HetR. Binding of PatS pentapeptide could inhibit the DNA binding activity of HetR thereby preventing the expression of *hetR* in the surrounding vegetative cells which in turn blocks the heterocyst formation. It is not known during which course the PatS is cleaved, converted in to pentapeptide and its diffusion to adjacent cells. Presence of PatS pentapeptide was able to inhibit the heterocyst formation however was not able to mimic the whole PatS function in PatS null mutant.

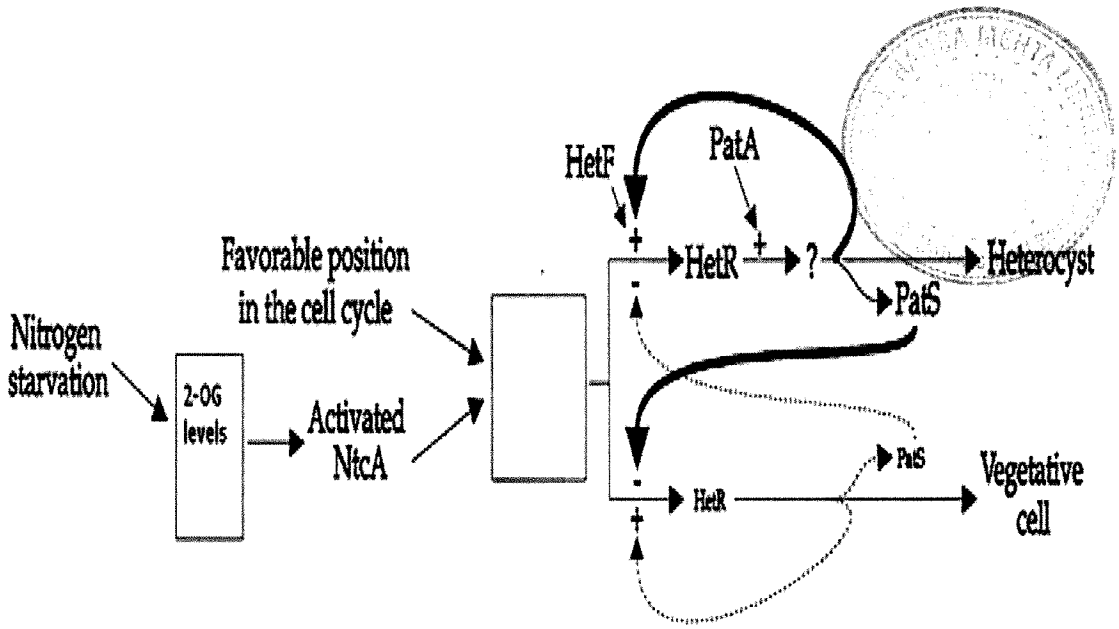



**Figure 1.8 Possible interaction of various genes for pattern formation in *Anabaeana* PCC 7120**

How the differentiating cell is immune to the PatS mediated inhibition is not fully understood. Two hypotheses have been proposed. The first hypothesis involves the ratio of PatS/HetR. In the initial stages, level of HetR is high as compared to PatS, which is expressed at low level, thereby resulting in the heterocyst differentiation (Meeks & Elhai., 2002). The second hypothesis involves the PatA mediated inhibition of PatS.

*hetN* gene is predicted to encode a ketoacyl reductase (Black and Wolk, 1994). Unlike a *patS* mutant, filaments that do not express *hetN* first develop a normal pattern of heterocysts at 24 h after induction but after 48 h shows Mch phenotype (Callahan and Buikema, 2001). *hetN* is normally not expressed until 12 h after induction (Bauer *et al* ., 1995) suggests that it does not play a role in initiation of heterocyst pattern formation but necessary for maintenance of the pattern. A low level of HetN protein is present in vegetative cells under non-inducing conditions, but accumulates exclusively to mature heterocysts (Li *et al.*, 2002) and expression of *hetN* primarily in heterocysts after induction imply that HetN is involved in production of an inhibitory signal that originates in

heterocysts and is communicated to neighboring vegetative cells. This putative HetN-dependent signal blocks heterocyst formation at points both upstream and downstream of *hetR* transcription; overexpression of *hetN* both prevents patterned expression of *hetR* and the Mch phenotype that normally results from ectopic expressionof *hetR* from an inducible promoter. These findings have led to the suggestion that HetN inhibits heterocyst formation by blocking *hetR* positive autoregulation (Callahan and Buikema, 2001). HetN inhibits the differentiation using a pathway different from the PatS and double mutants of HetN and PatS results in almost 100% cell differentiation even in presence of nitrate. *Anabaena* hetRR223W mutant showed resistance to both HetN and PatS indicating that both the HetN and PatS mediated pathways might be merging at the HetR protein. How HetN does not function within heterocyst forming cells is not known.

### 1.13 Maturation of heterocyst

#### 1.13.1 Role of *patB*

PatB contains an N-terminal domain with two putative 4Fe-4S centers and a C-terminal domain containing a DNA-binding motif. A frame shift mutation in resulting in the truncation of the C-terminal domain, showed poor diazotrophic growth and Mch phenotype (Liang *et al.*, 1993). Recent work has shown that a *patB* deletion mutant is completely defective for diazotrophic growth and that a *patB-gfp* reporter fusion shows heterocyst-specific expression at 16 hours after nitrogen step-down. Similar to the original mutant, a mutant defective in the N-terminal ferredoxin-like domain has Mch phenotype. GFP reporter fusions of *patS* and *hetR* show normal patterns of expression in a *patB* frameshift mutant, thus, indicating that PatB must function downstream of them. Earlier it has been suggested that PatB might be a redox-sensitive transcription factor that is required in the late stages of heterocyst differentiation (Jones *et al.*, 2003).

Many other genes have been identified to be required for heterocyst differentiation like *hetF* is required for localization of hetR and heterocyst

formation (Wong & Meeks, 2001), *hetC* encodes a protein that is similar to bacterial ABC protein exporters and is involved in early regulation of heterocyst differentiation (Khudyakov & Wolk,1997), *hetL* involved in early development of heterocyst (Liu & Golden, 2002), *hepA* required for polysaccharide envelope (Zhu *et al.*, 1998), *hepK/DevR* two component-regulatory system, where HepK is sensory histidine kinase while DevR is response regulator and together they regulate polysaccharide bio-synthesis, *devBCA* codes for exporter for envelop formation (Zhou & Wolk, 2003).

### 1. 14 Late events in the heterocyst maturation

Late events of the heterocyst maturation involve three site-specific recombination events followed by nitrogen fixation.

#### 1.14.1 Genomic rearrangements in *Anabaena* PCC 7120

*Anabaena* genome is interrupted by three DNA elements namely, *nifD*, *hupL* and *fdxN* element depending on the gene which they interrupt. The elements vary in size as *nifD* element is of 11.2 kb, *fdxN* element is of 54 kb and *hupL* element is of 9.4 kb (Table 2) (Ehara *et al.*, 2003). The *nifHDK* and *nifBfdxN nifSU)* operons are located next to each other while the *hupL* operon is located >700kb away on the *Anabaena* chromosome. The interrupting elements are excised out precisely during heterocyst differentiation by means of site-specific recombination events, resulting in gene rearrangement and formation of functional operons. The excised circles persist in the heterocysts with no known function (Haselkorn, 1989; 1992; Apte and Prabhavathi, 1994).

**Table 1.2 Elements of *Anabaena* PCC 7120**

| Rearranged element | *nifD* element | *fdxN* element | *hupL* element |
|---|---|---|---|
| Size (bp) | 11,278 | 59,428 | 9,419 |
| Interrupted gene | *nifD* | *fdxN* | *hupL* |
| Size of Direct repeats (bp) | 11 | 5 | 16 |
| Target Sequence | GGAGTAATCCG | TATTC | CACAGCAGTTAT AATG |

## 1.15 Presence of *nifD* element in Cyanobacteria

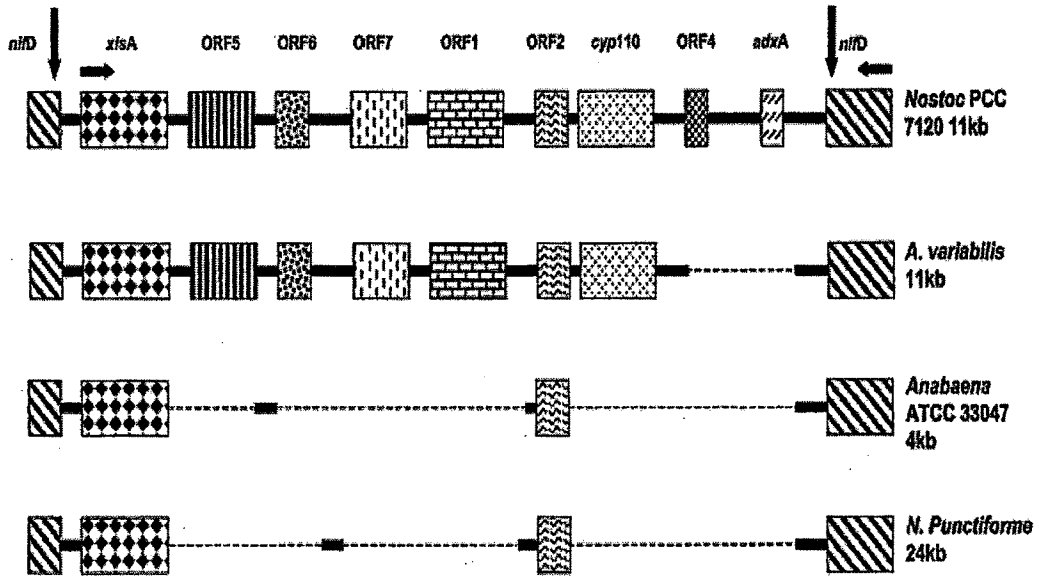



**Figure 1.9 Comparative analysis of nifD element from different cyanobacteria**

### 1.15.1 Comparative analysis of *nifD* element in different cyanobacteria

The first genetic element to be sequenced was 11.2 kb *nifD* element interrupting the *nifD* gene, which codes for α subunit of dinitrogenase, a part of part of nitrogenase enzyme complex responsible for nitrogen fixation (Mulligan & Haselkorn, 1989). The *nif* D element contains *xisA* gene at 3′ end responsible for its excision (Lammers *et al.*, 1986). It also contains several other uncharacterized ORFs and *cyp110* which codes for cytochrome 450 like protein (Hanson *et al.*, 2005). Many other cyanobacteria contains the *nifD* element however their size varies greatly from 4 – 24 kb (Figure 1.9). *Anabaena variabilis* contains 11 kb element which shows very high similarity with that of *Anabaena* PCC 7120 but the 24 kb *nifD* element from *Nostoc puctiforme* differs significantly(table 3) (Brusca *et al.*, 1989; Meeks *et al.*, 2001). 4 kb *nifD* element variant has been detected in *Anabaena* sp. Strain ATCC 33047 (Henson *et al.*, 2005). Comparison of the *nifD* elements has revealed that even though they vary in size, there are some conserved regions especially the *xisA* gene, 5′ and 3′ ends

involved in the recombination and the 11 bp target site except in *Anabaena* ATCC 33047.

**Table 1.3 Sequence similarity in % between nifD element from different cyanobacteria**

| | *A. variabilis* | ATCC 33047 | PCC7120 | *N. puncitforme* |
|---|---|---|---|---|
| (a) | The 5′ flanking region upstream of *xisA* (top) and *xisA* itself (bottom) | | | |
| *A. variabilis* | ---- | 89 | 99 | 86 |
| ATCC 33047 | 88 | --- | 90 | 81 |
| PCC7120 | 91 | 86 | --- | 86 |
| *N. puncitforme* | 83 | 84 | 83 | --- |
| (b) | The 5′ flanking region of the element (top) and ORF 2 (bottom) | | | |
| *A. Variabilis* | ---- | 90 | 88 | 74 |
| ATCC 33047 | 81 | ---- | 89 | 75 |
| PCC7120 | 93 | 81 | ---- | 73 |
| *N. puncitforme* | 80 | 83 | 80 | ---- |

Similarities are in percentage similarity in nucleotide sequences.

### 1.16 *nifD* element excision in *Anabaena* PCC 7120:

Excision of *nifD* element is carried out by XisA, coded by *xisA* gene, involving 11bp target site flanking the element (figure 1.10). This excision is required for the nitrogen fixation. XisA is absolutely required for the excision. The predicted ORF of the *nifD* element encodes a polypeptide for 480 amino acids. After excision new ORF is created by the fusion of the *nifD* ORF to *nifK* 5' flanking sequence. This new ORF codes for 497 aa, of which C-terminal 43

residues replaced 26 C-residues of vegetative cell ORF. The new ORF shows similarity to *Rhizobium* NGR 234 of *Parasponia* α subunit of nitrogenase (Golden *et al.*, 1985). Rearrangement of the *nifD* element of *Anabaena* PCC 7120 has been studied in *E. coli* with plasmids containing the entire *nifD* element (Lammers *et al.*, 1986). Insertional inactivation of the *xisA* gene of these plasmids abolishes the rearrangement and this suggested that *xisA* is necessary for the excision (Golden and Weist, 1988). Brusca *et al.* (1990 ) constructed a substrate plasmid pAM461 that contains the left and right borders of *nifD* element but does not contain an intact *Anabaena* PCC 7120 open reading frame larger than 180bp. pAM461 does not rearrange in the absence of a complementary plasmid expressing the *xisA* gene. Thus, XisA is sufficient for excision in *E. coli*. Toxic effects of XisA hindered the utility of *E. coli* to understand the mechanism of excision of *nifD* element including the nature of XisA protein.

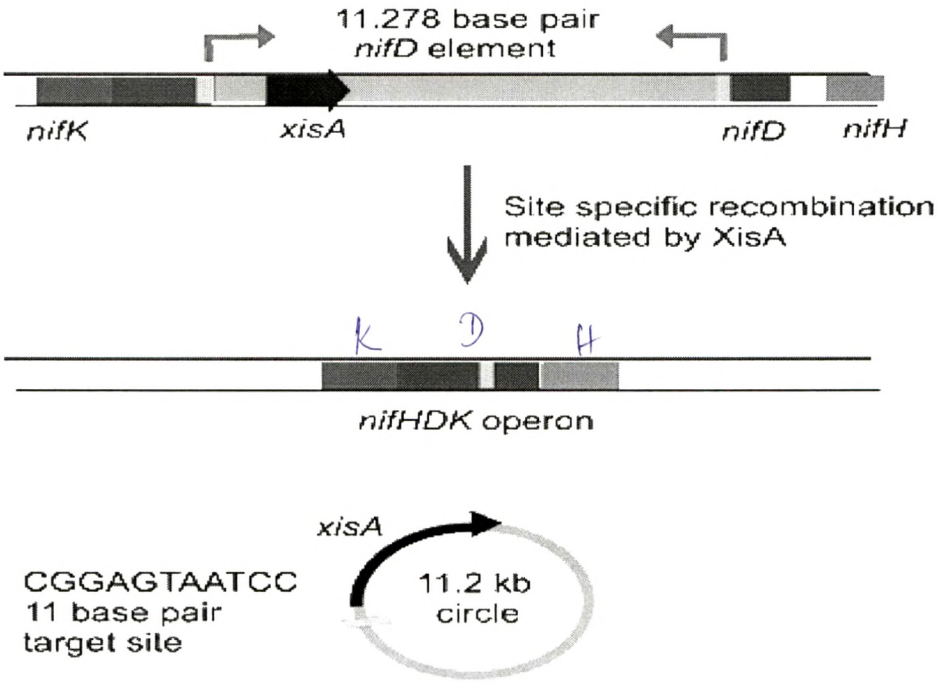



**Figure 1.10 Excision of *nifD* element**

**1. 17 *hupL* element:**

9.4 kb *hupL* element disrupts the *hupL* gene, which codes for membrane bound uptake hydrogenase. It undergoes rearrangement simultaneously with other rearrangements. DNA hybridization studies showed three strains of *Anabaena* contain *hupL* element (Carrasco & Golden, 1995; Henson *et al.*, 2005) and absent in *Anabaena variabilis* and *Nostoc punctiforme* (Meeks *et al.*, 2001). Main function of the uptake hydrogenase is in the creation/maintenance of micro-oxic condition. However, mutation or absence of this enzyme is not considered to prevent the nitrogen fixation as in case of *fdxN* element and *nifD* element, but it might decrease the efficiency of the nitrogen fixation. The predicted HupL polypeptide is homologous to the large subunit of [NiFe] uptake hydrogenase. *hupL* is expressed in a manner similar to other nitrogen-fixation genes and m-RNA was detected only during late stages of heterocyst differentiation. An ORF, named *xisC* was identified near one end of the *hupL* DNA element which encodes element's site-specific recombinase. 1.5kb ORF of XisC was identified 115 bp inside the right border of the *hupL* element (figure 1.11) (Carrasco *et al.*, 2005)

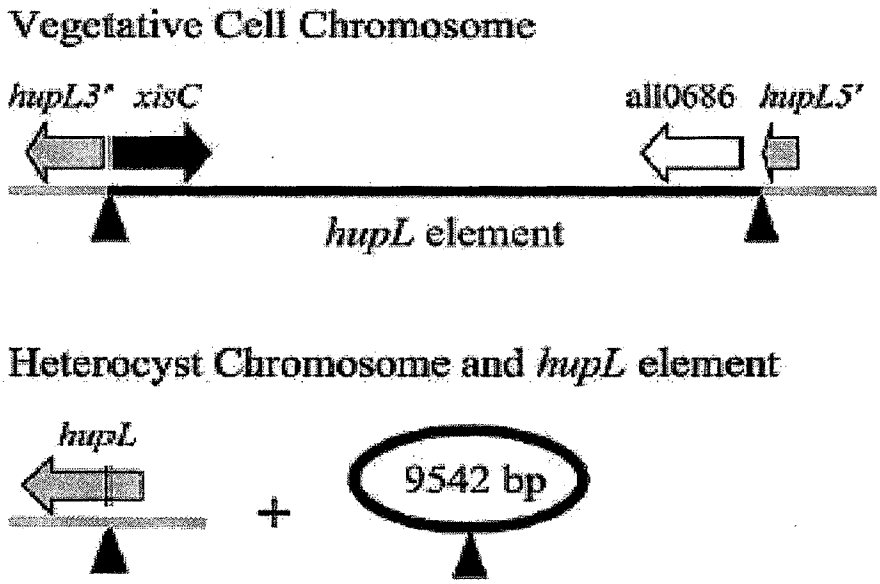



**Figure 1.11 Rearrangment of *hupL* gene**

**1. 18 *fdxN* element:**

55.2 kb *fdxN* element resides in *fdxN* gene, part of *nifB-fdxN-nifS-nifU* operon. *fdxN* element has also been detected in some other species of *Anabaena* (Figure 1.12). The presence of the 11 and 55kb elements in *Nostoc* Mac, *Anabaena cylindrica* and *Anabaena* M131are also reported (Carrasco and Golden, 1995). The Marine cyanobacterium *Anabaena* CA lacks both the elements (Carrasco and Golden, 1995). XisF is responsible for the excision of the *fdxN* element and *xisF* mutant is incapable of nitrogen fixation (Carrasco *et al.*, 1994) However XisF alone is not sufficient for recombination as it requires two other gene products i.e. *xisH* and *xisI*. Exact role of both the genes is not known, but it has been suggested that *xisF* is sufficient for the integration while the *xisHI* is required for the excision similar to integrase lambda (Ramaswamy *et al.*, 1997). The deletion of the 55 kb element results in the formation of the *fdxN* ORF and allows the expression of the downstream genes in the *nifB-fdxN- nifS-nifU* operon (Mulligan & Haselkorn, 1989). Excision of *fdxN* occurs independently form the 11kb *nifD* element excision (Carrasco *et al.*, 1994; Golden and Wiest, 1988; Golden *et al.*, 1987).

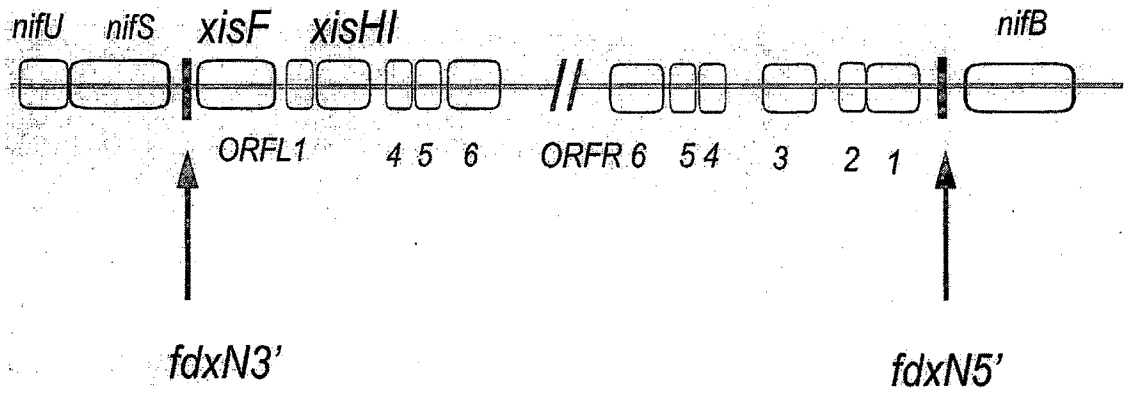



**Figure 1.12 Genetic organization of *fdxN* element of *Anabaena* PCC 7120**

Three site specific recombinases, XisA, XisC and XisF are responsible for the excision events (Table 4).

**Table 1.4 Site specific recombinases of *Anabaena***

| Recombinase | Excisase A | Excisase F | Excisase C |
|---|---|---|---|
| Gene | *xis A* | *xis F* | *xis C* |
| ORF | 1062 bp | 1545 bp | 1593 bp |
| Amino Acids | 472 | 514 | 498 |
| Estimated Mass (Da) | 41644 | 58473 | 60202 |
| Homology | **Integrase** | Resolvase | **Integrase** |

They recognize a pair of specific sequence, flanking the harboring element. XisA recognizes CGGAGTAATCC 11 base pair direct repeat sequence, XisF recognizes TATTC 5 bp sequence and XisC recognizes CACAGCAGTTATATGG 16 base pair sequence. All the three excisases are developmentally regulated *i.e.* only expressed in the heterocysts. All the recombinase are specific i.e. they do not recognize other sequence except target site. They are absolutely required for the excision as mutation in them results in the loss of excision (Golden, 1988). They are not involved in heterocyst formation as the deletion mutants of the excisase forms normal heterocyst.

XisA protein sequence that was reported earlier (Lammers *et al,* 1986) has been corrected (Nunes-Duby *et al.,* 1998). Although XisC is larger than the XisA, the amino acid sequence can be aligned along full length with XisC and their amino acid sequences are 61% similar and 43% identical and they appear to belong to Intergrase family (Nunes-Duby *et al.,* 1998). Mutation in the catalytic tyrosine of XisC results in the loss of rearrangement phenotype (Carrasco *et al.,* 2005). *Anabaena* PCC 7120 *xisF* recombinase, which is required for excision of *fdxN* element, belongs to the resolvase family and is homologous to *Bacillus subtilis* SpoIVCA site-specific recombinase.

Neither transcript nor protein of XisA, XisF and XisC has been detected so far. The only way to measure the activity has been to screen for the respective gene rearrangements. Based on these studies, the expression of these genes occurs late during heterocyst development (Haselkorn, 1992; Carrasco *et al.,* 1994). Two *Anabaena* PCC 7120 DNA binding proteins, NtcA and factor 2 has

been shown to bind the sequence upstream of *xisA* gene but the NtcA binding sites are not found in the upstream region of *xisC* (Ramasubramanian *et al.*, 1994).

### 1. 19 Site-Specific Recombinases

Site-specific recombination (SSR) involves binding, cutting, strand exchange and ligation, followed by release to complete the process. The minimum requirement for the SSR involves recombination site and recombinase for e.g. FLP and Cre recombinase, while other recombinases require additional helper proteins and DNA sequences other than the cross-over site. SSR is involved in various processes in micro-organisms (table 5)

Site-specific recombination: a sampling of enzymes and functions

| Recombinase | Biological function |
|---|---|
| Tyrosine recombinase family | |
| λ Int and many other phage integrases | Integration and excision of phage genomes |
| Int of Tn916/Tn1545 | Integration and excision: transposition of circular transposons |
| IntI | Integration and excision of gene cassettes in integrons |
| Cre | Excision: dimer reduction in phage P1 plasmids |
| XerC/D | Excision: dimer reduction in the *E. coli* chromosome as well as in many other bacterial chromosomes and some plasmids |
| TnpI of Tn4430 | Excision: resolution of cointegrates resulting from transposition of Tn4430 |
| FimB, FimE | Inversion: alternation of gene expression (fimbrial phase variation in *E. coli*) |
| Rci of R64 | Inversion of shufflon segments in plasmid R64, producing various forms of pili |
| XisA, XisC | Excision: for developmentally regulated gene activation in *Anabaena* |
| Flp | Inversion: for amplification of yeast 2-μm plasmid |
| Serine recombinase family | |
| TnpR of Tn3/γδ and related transposons | Excision: resolution of cointegrates resulting from transposition |
| Sin of *Staphylococcus aureus* | Excision: dimer reduction in staphylococcal plasmids |
| ParA of RP4 | Excision: dimer reduction in plasmid RP4 |
| Hin | Inversion: alternation of gene expression (flagellar phase variation) in *Salmonella* |
| Gin, Cin | Inversion: alternation of gene expression (tail fiber proteins) in phages Mu and P1 |
| OrfA of IS607/IS1535 | Integration and excision: transposition of the *Helicobacter pylori* element IS607 (and others) |
| Int of φC31/Bbv1/φRv1[a] | Integration and excision of *Streptomyces* and mycobacterial phages |
| TnpX of Tn4451[a] | Integration and excision: transposition of Tn4451 in *Clostridium* |
| SpoIVCA (CisA)[a] | Excision: for developmentally regulated gene activation in *Bacillus subtilis* |
| XisF[a] | Excision: for developmentally regulated gene activation in *Anabaena* |

**Table 1.5 List of various site-specific recombinase and their function**

Site-specific recombinase are divided into main two groups depending on the active site amino acid

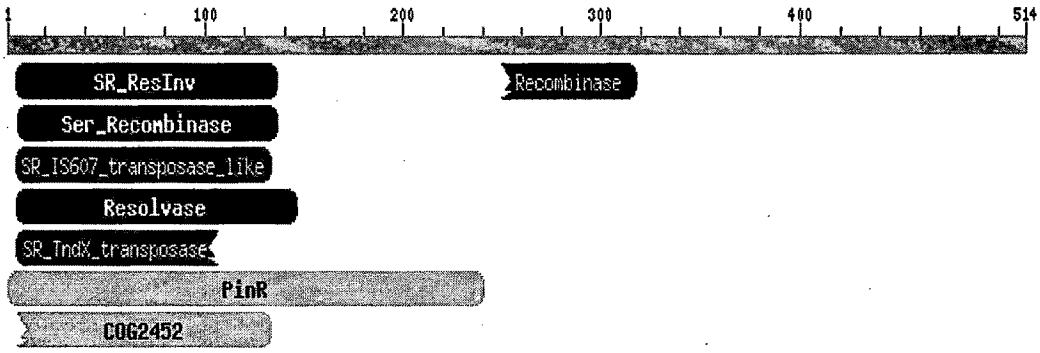
(i) Tyrosine recombinase also known as integrase family

(ii) Serine recombinase also known as resolvase family

All SSR do not require any metal for the reaction, instead the scissile phosphate is surrounded by positive charged amino acid side chain and involves the nucleophilic displacement of DNA with the protein side chain. In case of tyrosine recombinase, 5′ bridging O is displaced by the tyrosine to form a phosphotyrosyl bond to the 3′end of the broken DNA strand, while in serine recombinase 3′ O is displaced to form 5′ phosphoserine linkage. In tyrosine recombinase mediated recombination, single strand is cut at a time, followed by strand exchange, Holliday junction formation followed by second strand exchange. While serine recombinase makes double stranded cuts in all the strand at a time followed by strand exchange.

Sequence similarity of the XisA using amino acid sequence showed that XisA shows very good similarity with the integrase family members. Tyrosine integrase family members possess catalytic tetrad R-H-R-Y in the active site. Both XisA and XisC contain the R-Y-R-Y triad where second H is replaced by Y. XisF contains conserved domain of resolvase at its N-terminal (Figure 1.13).

**Figure 1. 13 Conserved domain of XisC**



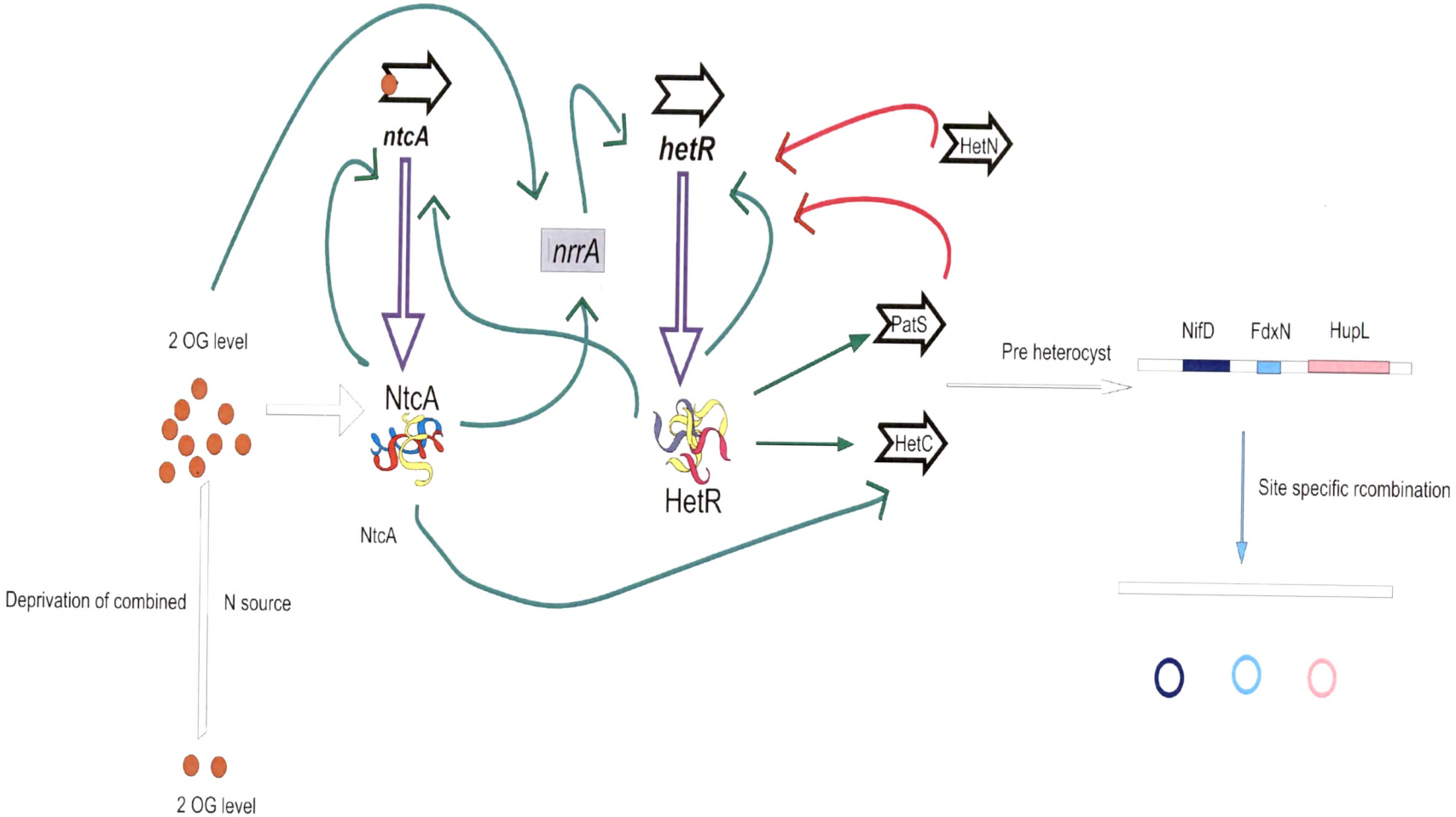



**Figure 1.14 Molecular mechanism of the Initiation of Heterocyst differentiation**

Based on all the background information, objectives of the present studies are

- To determine the role of accessory sequence(s) in the site-specific recombination mediated by XisA.
- To characterize XisA as restriction endonuclease.
- To determine the intermolecular recombination events involving *nifD* element.
- To characterize functional domains of XisA and XisC.
- To determine the regulation of XisA promoter using GFP reporter assay.