MOLECULAR GENETIC ANALYSIS OF THE *zwf* REGION OF CYANOBACTERIAL GENOMES

Haydar Karakaya

Thesis submitted in fulfilment of the requirements for the degree of PhD

Department of Biological Sciences,

University of Warwick

December 1995

Table of Contents

Summary

The oxidative pentose phosphate (OPP) cycle is the main route of carbohydrate dissimilation in cyanobacteria in the dark. In heterocysts, the OPP is a supplier of reductant to nitrogenase, thus playing an important role in nitrogen fixation. Molecular genetic analysis of glucose-6-phosphate dehydrogenase (G6PDH) and the other components of the OPP cycle have been focused on recently in an effort to understand the function and regulation of the cycle in vegetative cells and in heterocysts of cyanobacteria. This study presents data on nucleotide and amino acid sequences of three enzymes of the OPP cycle (glucose-6-phosphate dehydrogenase, transaldolase and fructose-1,6-bisphosphatase) of the filamentous, heterocystous strain *Anabaena* sp. PCC7120. Mutagenesis studies on the transaldolase gene (*tal*) of *Anabaena* sp. PCC7120 and the *opcA* gene of the unicellular strain *Synechococcus* sp. PCC7942 are also reported.

A 4,169 bp region around the *zwf* gene of *Anabaena* sp. PCC7120 was sequenced. Three genes are located in the region: *fbp, tal* and *zwf*, which encode fructose-1,6-bisphosphatase (FBPase), transaldolase and glucose-6-phosphate dehydrogenase (G6PDH), respectively. The *fbp* gene encodes a polypeptide of 349 amino acids. The product of the *tal* gene consists of 381 amino acids. The *zwf* gene encodes a protein of 509 amino acids. Four cysteine residues are present in the enzyme. Two of these cysteines (Cys-187 and Cys-445) are absolutely conserved in cyanobacterial G6PDH. This result reinforces the likelihood of their role in the regulation of enzyme activity.

Subclones carrying a range of *zwf* fragments of *Anabaena* sp. PCC7120 did not complement the *zwf* mutant *Escherichia coli* strain DF214. Western blot analysis showed that *Anabaena* sp. PCC7120 G6PDH did not express in *E. coli* DF214. Attempts aimed at production of a deletion/insertion *zwf* mutant of *Anabaena* sp. PCC7120 were not successful.

The *tal* gene of *Anabaena* sp. PCC7120 was mutated to investigate growth and survival of the mutant cells in the presence and absence of combined nitrogen. Transaldolase activity in the *tal* mutant cells was similar to that in the wild-type cells. The growth rates of the *tal* mutant cells were not significantly different from that of the wild-type cells. These results imply that more than one copy of the *tal* gene is present in the *Anabaena* sp. PCC7120 genome.

A gene, designated *opcA*, downstream from the *zwf* gene of *Synechococcus* sp. PCC7942 was mutated to investigate the possible effect of the gene on G6PDH activity. G6PDH activity in the mutant cells was reduced by 98.6%. Western blot analysis showed that the *zwf* was expressed in the *opcA* mutant cells, but the active form of the enzyme was a considerable reduced. All the results suggest that the *opcA* gene affects the assembly of G6PDH or enzyme activity, but possibly not expression of the *zwf* gene.

List of Figures

List of Tables

Acknowledgements

I should like to thank my supervisor Dr. Nick Mann for his excellent guidance and encouragement during my PhD study. I would also like to thank Professor Noel Carr for his encouragement.

I would like to thank Turkish Government and Ondokuz Mayis University for providing me a studentship for my PhD study.

Thanks are due to Dr. Dave Hodgson for introducing me conjugative plasmid RP-4 and *E. coli* strain AEE101, Julie Scanlan for plasmids pAG75 and pDB, and Dr. Shanthy Sundaram for allowing me to use the antibody.

A special thanks must go to all the members of MicroIII and Cyano lab for their co-operation and excellent friendship. In particular, I would like to thank Dr. Nigel Silman for his encouragement and his enormous patience towards my endless questions and for his generous help, acting as a trouble shooter through all stages of this study. My special thanks also go to Dr. Dave Scanlan for his help and excellent friendship.

Finally, I would like to thank all my family members and friends for their constant support and encouragement.

Declaration

I hereby declare that the work described in this thesis was conducted by myself, under the supervision of Dr. N. H. Mann, with the exception of those instances where the contribution of others has been specifically acknowledged.

None of the information contained herein has been used in any previous application for a degree.

All sources of information have been specifically acknowledged by means of references.

Haydar Karakaya

Abbreviations

6PGDH	6-phospho gluconate dehydrogenase
AMP	adenosine monophosphate
AMPS	amonium persulphate
AP	alkaline phosphatase
Ap	ampicillin
ATCC	American Type Culture Collection
bp	base pair(s)
Cm	chloramphenicol
Cys	cysteine
Da	Daltons
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E4P	erythrose-4-phosphate
EDTA	ethylenediaminotetra-acetic acid
F6P	fructose-6-phosphate
FBP	fructose-1,6-bisphosphate
fbp	fructose-1,6-bisphosphatase gene
FBPase	fructose-1,6-bisphosphatase
G6P	glucose-6-phosphate
G6PDH	glucose-6-phosphate dehydrogenase
GAP	glyceraldehyde-3-phosphate
HEPES	(N-[2-hydroxyethyl] piperazine-N'-[2-ethanesulphonic acid])
IPTG	isopropyl-β-d-thiogalactopyranoside
kb	kilobase pairs
kDa	kilo Daltons
Km	kanamycin
Nm	neomycin
OPP	oxidative pentose phosphate
ORF	open reading frame
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
PSI	photosystem I
PSII	photosystem II
psi	pounds per square inch
PCC	Pasteur Culture Collection
pgi	phosphoglucose isomerase gene
RNA	ribonucleic acid
rpm	revs per minute
RPP	reductive pentose phosphate
Rubisco	ribulose-1,5-bisphosphate carboxylase/oxygenase
RuBP	ribulose-1,5-bisphosphate
S7P	sedoheptulose-7-phosphate
SDS	sodium dodecyl sulphide
Sm	streptomycin
Spc	spectinomycin
SSC	sodium chloride-sodium citrate
SSPE	sodium chloride-sodium dihydrogen phosphate
tal	transaldolase gene
TBE	Tris-borate-EDTA
TE	Tris-EDTA
TEMED	N,N,N',N'-tetramethylethylenediamine
Tris	Tris (hydroxymethyl) methylamine
UV	ultra violet
WT	wild-type
X-GAL	5-bromo-4-chloro-3-indolyl-β-galactoside
zwf	glucose-6-phosphate dehydrogenase gene

Chapter 1

Introduction

1.1 The largest group of photosynthetic prokaryotes

Cyanobacteria are autotrophic organisms that operate an oxygen evolving photosynthesis via a plant-like photosynthetic apparatus rather than the anoxygenic bacterial photosynthesis via a bacterial type photosynthetic apparatus. Possessing this plant-like photosynthetic operation, and some other metabolic properties, cyanobacteria were considered as a group within the plant kingdom until late 1970s, named 'blue-green algae' by botanists, although they have a Gram-negative prokaryotic cellular organisation. As some of the oldest photosynthetic inhabitants of the earth, these organisms adapted successfully into almost all environments on the earth. The general properties of this group have been reviewed (see Stanier and Cohen-Bazire, 1977; Stewart, 1980; Ormerod, 1992; Whitton, 1992; Tandeau de Marsac and Houmard, 1993).

1.2 General properties of cyanobacteria

Cyanobacteria share the basic cellular features of eubacteria. However, cyanobacterial cells possess several unique and diagnostic characteristics. Through this subsection, the general characteristics of the cell and cellular organisations of cyanobacterial groups will be discussed. In addition, the recent status of cyanobacterial systematics and phylogeny will also be summarised.

1.2.1 Cellular organisation and function

Where the environmental conditions are optimal or near to the limits of growth conditions, the typical cyanobacterial cell photosynthesises using solar

radiation as an energy source harvested by phycobilin pigments and chlorophyll *a* molecules, splitting water as the electron donor. Various cell types are also represented in cyanobacterial groups, some of which are unique for cyanobacteria such as the heterocyst (for reviews see Stanier and Cohen-Bazire, 1977; Castenholz and Waterbury, 1989a; Whitton, 1992).

1.2.1.1 The cyanobacterial cell wall and other external structures

The cyanobacterial cell wall was reported to be homologous to that of other Gram-negative bacteria in respect to the basic structure and chemical composition (Stanier and Cohen-Bazire, 1977; Castenholz and Waterbury, 1989a; Gantt, 1994). It has some differences in some structural characteristics. Cyanobacterial cells are surrounded by a carbohydrate-enriched glycocalyx layer which appears as a fibrous sheath protecting the cell against desiccation and keeping the cells together in colonial and filamentous forms. However, the peptidoglycan layer of cyanobacterial cell wall is considerably thicker than that of other eubacteria. This layer provides mechanical protection and by its rigidity also determines the cell shape.

1.2.1.2 Cytoplasmic structures of cyanobacterial cell

Much of the cyanobacterial cytoplasm is occupied by thylakoid membranes which are distributed between the cytoplasmic membrane and thylakoid free centroplasm (Castenholz and Waterbury, 1989a). They are similar in most respects to the thylakoid membranes of chloroplasts, but cyanobacterial

thylakoids are non-stacked unlike those of the chloroplast. Although both cytoplasmic and thylakoid membranes share many supramolecular complexes, some e. g. photosystems are restricted to the thylakoids. Photosystem I (PSI) and photosystem II (PSII) are intrinsic complexes of thylakoid membranes, including *a* type chlorophyll molecules only (Glazer, 1987). The stromal or protoplasmic sides of thylakoids are covered by phycobilisomes which are upright hemidiscoidal structures of 20-70 nm in diameter and are attached in orderly rows. Phycobilisomes comprise the major light-harvesting complexes of most cyanobacteria and contain between 300 and 700 bilins per particle and are physically associated with one or more PSII complexes. They consist of a series of rods composed of disks of biliproteins spread out as a fan centering on the core. The core is composed of stacks of biliproteins in trimeric disks laid on their sides (Figure 1.1). The pigments of phycobilisomes are allophycocyanin in the core, phycocyanin in the rods and phycoerythrin which occurs, if present, as outer discs of the rods (Glazer, 1987).

Apart from the thylakoid membranes, there are some other common components and inclusions in the cytoplasm. Certain inclusions are always present in cyanobacterial vegetative cells: phycobilisomes, glycogen granules, polyphosphate granules, and carboxysomes (Allen, 1984). The presence or absence of these inclusions may vary from species to species, and from one environmental condition to another (Simon, 1987). Many planktonic cyanobacteria possess gas vacuoles made up of assemblages of hollow cylindrical structures. The cells with gas vacuoles select a position of optimum irradiance in a

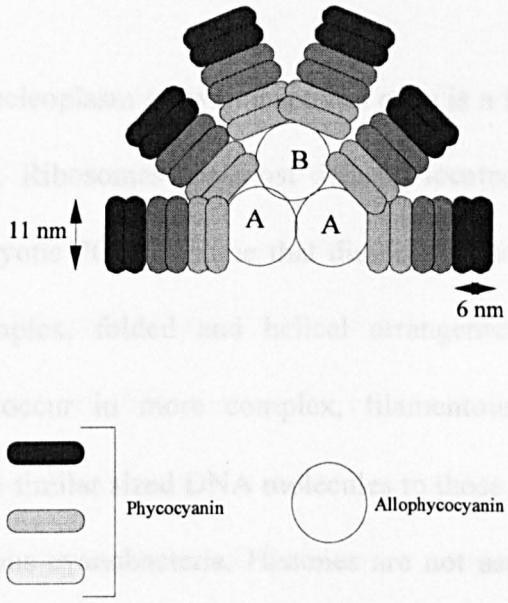


Figure 1.1. Schematic representation of a hemidiscoidal phycobilisome with a tricylindrical allophycocyanin core. Allophycocyanin disks are distributed differently in the cylinders A and in the cylinder B. Six peripheral rods are normally bound to the core. In the rods, a hexameric phycocyanin complex is always located at the rod-core linkage position. More distal complexes of peripheral rods (not shown in the figure) may be either C-phycocyanin, phycoerythrocyanin or phycoerythrin hexamer-linker polypeptide complexes depending on the organism and the environmental conditions (after Sidler, 1994).

vertically decreasing light irradiance gradient by means of regulation of gas vacuolation in response to light irradiance (Walsby, 1987).

The nucleoplasm of cyanobacterial cells is a lighter pigmented area at the centre of cell. Ribosomes are most densely located in the nucleoplasm and a typical prokaryotic 70S ribosome that dissociate into 50S and 30S. DNA fibrils are in a complex, folded and helical arrangement. Multiple copies of the chromosome occur in more complex, filamentous types. Unicellular forms, however, have similar sized DNA molecules to those known for prokaryotes other than filamentous cyanobacteria. Histones are not associated with cyanobacterial DNA. (Leach and Herdman, 1973; Castenholz and Waterbury, 1987).

1.2.2 Specialised cells of cyanobacteria

Under environmental stress, some cyanobacterial strains produce specialised cells such as heterocysts, akinetes and hormogonia. (Castenholz and Waterbury, 1989a). Akinetes are considered to be resting cells or "spores" resistant to low temperatures and to desiccation but rarely resistant to high temperature. They are differentiated from vegetative cells in having a thick wall surrounding the old one, and also possessing a granular cytoplasm accumulating cyanophycin, glycogen, lipids, and carotenoid pigments. They tolerate drying, freezing, and long term survival in anaerobic sediments; light and optimal pH and temperature are needed to germinate either through a pore at one or both ends of

the envelope or dissolving the entire envelope. (Nichols and Adams, 1982; Herdman, 1987).

Heterocysts are differentiated cells of filamentous cyanobacteria , which are specialised for the aerobic fixation of dinitrogen. Members of the Nostocales and Stigonematales produce heterocysts at intervals in the filament when the concentration of combined nitrogen has been lowered in the surrounding medium. The most conspicuous difference is the presence of a thick envelope comprising an inner laminated layer, a central homogenous layer, and an outer fibrous layer. Cell size is relatively large and they have a rounded shape. They are partially separated from neighbouring vegetative cells having a narrow waist-like connection to the vegetative cell(s). The colour is pale probably due to the diminution or absence of phycocyanin. They have a homogenous cell content lacking cytoplasmic granular inclusions. Polar nodules or polar thickenings are characteristic near the attachment to the vegetative cells. Heterocysts have been under intensive investigation because of their ability to fix nitrogen in aerobic environments and are reviewed by Fay (1972, 1992), Wolk (1982), Adams and Carr, (1981), Buikema and Haselkorn, (1993), and Wolk *et. al*, (1994).

1.3 Systematics of cyanobacterial groups

Cyanobacteria were traditionally treated as algae with a blue-green colour by phycologists for over a century. The group was defined by morphological, cytological and chemical features, most importantly their plant-like photosynthetic

properties. The history of cyanobacterial taxonomy was discussed by Stanier and Cohen-Bazire (1977), Castenholz and Waterbury (1989a), and Whitton (1992).

Under the Botanical Code, about 150 genera and well over 1000 species were described by using cyanobacterial morphological or ecological properties which are the only characters determinable in the field. Later this century, the group which used to be known as "blue-green algae" was shown to be similar to bacteria, lacking a nucleus surrounded by a membrane and possessing a cell wall structure which is similar to the Gram-negative bacterial cell wall. Eventually the group was taken as genuinely prokaryotic and named "Cyanobacteria". Despite disagreements by botanists, Stanier *et al.* (1978) proposed to place cyanobacterial nomenclature under the rules of the International Code of Nomenclature of Bacteriology rather than the Botanical Code. Today, cyanobacteria are widely accepted as a eubacterial group, although some disagreements still exist about the nomenclature. Those disagreements are essentially based on concern about confusions of the names given to same organism separately by botanists and by bacteriologists, and losing the very convenient old literature on cyanobacteria (blue-green algae).

Rippka *et al.* (1979) described 22 genera of cyanobacteria and divided them into five large sections with respect to differences in structure and development. Sections I and II include unicellular groups; the cells are either single or in colonial aggregates. The distinction between groups is the way of

reproduction. The cells in Section I reproduce either by binary fission or by budding while the cells in Section II reproduce by multiple fission giving rise to small daughter cells (baeocytes) or by both multiple fission and binary fission. Groups III, IV and V consist of filamentous strains. Reproduction in these groups is by random filament (trichome) breakage, by formation of hormogonia and sometimes by germination of akinetes (only in Sections IV and V). The trichomes of Section III organisms are composed only of vegetative cells. Under nitrogen starvation conditions the strains included in Section IV and V differentiate some of their cells into heterocysts and some strains also produce akinetes. Branching properties distinguish Section IV and V; trichomes of the former are unbranched while those of latter are branched. Castenholz and Waterbury (1989b) gave the sections of Rippka *et al.* (1979), names of ordinal rank, and sections were taken as subsections because of taking cyanobacteria as a section of oxygenic photosynthetic bacteria. Subsection I was named as Chroococcales (Waterbury and Rippka, 1989), Subsection II as Pleurocapsales (Waterbury, 1989), Subsection III as Oscillatoriales (Castenholz, 1989a), Subsection IV as Nostocales (Castenholz, 1989b) and Subsection V as Stigonematales (Castenholz, 1989c).

Giovannoni *et al.* (1988) reported phylogenetic analyses based on 16S rRNA comparisons. They provided evidence that unicellular cyanobacteria of section I and nonheterocystous, filamentous organisms of section III had multiple evolutionary origins. Thus, they concluded that taxonomic classifications based principally on morphology did not necessarily reflect phylogenetic relationships.

1.3.1 Classical and molecular approaches to cyanobacterial phylogeny

Cyanobacterial evolutional history is relatively well-documented with respect to geological and molecular evidence. Fossil records and molecular techniques have been used to build evolutional lineages. Cyanobacteria have been included into the Eubacteria, a distinct group from Archaebateria and Eukaryotes. Phylogenetic aspects of the group have been reviewed by Schopf and Walter (1982), Doolittle (1982), Woese (1987) and Wilmotte (1994).

Schopf and Walter (1982) reported that cyanobacterial fossil records were far more well-established than those of any other prokaryotic group. Cyanobacteria probably arose between 2 billion and 3 billion years ago from a photosynthetic bacterium (Brock, 1973). Cyanobacterial phylogeny has recently being studied by using molecular methods, predominantly 16S rRNA sequence comparisons (Giovannoni *et al.*, 1988; Wilmotte, 1994). Taxonomic grouping based upon molecular phylogeny pairs imperfectly with the morphological groups of Rippka *et al.* (1979). This raised questions against early fossil records attributed to cyanobacteria since interpretations of microfossil evidence is frequently based upon the assumption that morphology is phylogenetically conserved. The molecular evidence of cyanobacterial phylogenetic data suggest that cyanobacterial diversification occurred within a relatively short span of molecular evolutionary distance (Giovannoni *et al.*, 1988).

Cyanobacteria have long been hypothesised as the progenitors of eukaryotic plastids. Early reports were based on ultrastructural and biochemical similarities between chloroplasts of eukaryotes and the basic cyanobacterial cell. The hypothesis, which is known as "xenogenous" or "endosymbiont", has been reviewed by Doolittle (1982), Giovannoni *et al.* (1988) and Douglas (1994). Most of the sequencing data support the emergence of plastids from a single point with the cyanobacterial assemblage indicating a monophyletic origin. Comparisons of gene arrangements, expression and sequences from photosynthetic eubacteria and plastids have consistently indicated the cyanobacterial ancestry of plastids (Douglas, 1994).

1.4. Ecology

Cyanobacteria are the largest group of photosynthetic prokaryotes in respect of their widespread occurrence, frequent abundance and morphological diversity (for review see Stanier and Cohen-Bazire, 1977; Whitton, 1992; Tandeau de Marsac and Houmard, 1993). They are represented in most types of illuminated environment except for those with lower pH values. Cyanobacteria have been reported as the first oxygenic photosynthetic organisms of the earth creating an oxygenic atmosphere; and they are still responsible for a considerable proportion of photosynthetic oxygen evolution in the oceans. Their wide distribution reflects a large variety of diverse morphological and physiological properties such as tolerance of high temperature, high UV irradiation, desiccation, and free sulphide, and the abilities to utilise low light flux and CO_2 concentration and to fix N_2

(Whitton, 1992). Cyanobacteria play an important role in aquatic environments contributing to primary productivity (Gibson and Smith, 1982; Fog, 1982; Carr and Mann, 1994) especially as nitrogen fixers. Gas vacuoles give the ability to planktonic cyanobacterial cells to stay in an illuminated zone to utilise effective light intensity for photosynthesis (Walsby, 1987). These buoyant cyanobacteria form dense populations at the surface of lakes and reservoirs, the so called "blooms" causing deoxygenation of the habitat and also producing toxins (Codd and Poon, 1988). The temperature optimum for many cyanobacteria is higher by at least several degrees than for most eukaryotic algae. This ability gives summertime dominance of cyanobacterial populations in temperate latitudes. Some cyanobacterial species can even grow in thermal alkaline springs at temperatures of up to 74°C. They are also predominant at low and freezing temperatures such as in Antarctica where dominance may be achieved through the ability to tolerate alternating freezing and thawing or freeze-drying (Castenholz and Waterbury, 1989a). The tolerance of desiccation allows cyanobacterial populations to exist successfully in deserts and tropical regions. Some cyanobacterial strains have successfully adapted to highly saline environments and the some other strains can tolarate the water and soil which are rich in heavy metals.

Cyanobacteria occur in a wide variety of habitats successfully taking advantage of their physiological features like oxygenic photosynthesis, nitrogen fixation ability, tolerance of extreme environments etc. Their wide distribution

reflects a large variety of species with diverse morphological and physiological properties. In conjunction with their morphological and physiological diversity, their genetic diversity is also great judging from their G+C% content and genome sizes, which cover the entire range of values found for prokaryotes (Tandeau de Marsac and Houmard, 1993).

1. 5 Genetics

1.5.1 Cyanobacterial genomes

The nuclear region of cyanobacterial cells is comparable to other bacteria in lacking a nuclear membrane and compact chromosomes of the eukaryotic type (Leach and Herdman, 1973; Herdman, 1982; Tandeau de Marsac and Houmard, 1987). No histone proteins exist in the nuclear region which is consistent with the prokaryotic nature of cyanobacteria. DNA synthesis was reported to be periodic, DNA being synthesised by DNA polymerase during the light period while no DNA replication occurred in the dark (Leach and Herdman, 1973).

The genome sizes of cyanobacteria vary greatly ranging from 1.6×10^9 to 8.6×10^9 Da which far exceed the largest genome previously described in prokaryotes (Herdman *et. al*, 1979b). The genome sizes of unicellular cyanobacteria range from 1.6×10^9 to 4.2×10^9 Da, which is similar to that of other bacteria while filamentous sections have a very wide range of genome sizes from 2.60×10^9 to 8.6×10^9 Da. These figures suggested that there is a correlation

between genome size and morphological diversity, since the strains which exhibited a high degree of morphological differentiation almost invariably possessed larger genomes than those of strains which displayed more simple morphological forms. (Herdman *et. al*, 1979b)

In two subgroups (unicellular strains which reproduce in binary fission and nonheterocystous, filamentous strains), DNA base compositions are divergent ranging from 35 to 71 mol % G/C and from 40 to 67 mol % G/C, respectively. The overall range is 38 to 47 mol % G/C for all the other cyanobacteria that these include the two subgroups: heterocystous and pleurocapsalean cyanobacteria which are large, internally diverse and very different with respect to structure and development. Therefore, there is little or no correlation between base composition and structural and developmental respects (Herdman *et al.*, 1979a).

About 60% of cyanobacterial strains examined contained one or more plasmid types which have molecular sizes ranging from 1.3 kilobase pairs (kb) to 130 kb in size. The abundance of the different plasmid types varies with the metabolic states of the cells. Certain marker genes, especially any known antibiotic resistance genes from plasmids of cyanobacteria have not been reported (Tandeau de Marsac and Houmard, 1987).

1.5.2. Restriction endonucleases and modification of DNA

Restriction endonucleases (the type II restriction endonucleases) have been numerously reported in cyanobacterial strains (for review see Tandeau de Marsac and Houmard, 1987; Herdman, 1982). Strains may contain one to five different restriction endonucleases some of which are represented as isoschizomers even in very diverse strains indicating that these genes with their methylase modification counterparts could have been acquired by the transfer of plasmid or cyanophage encoded genes.

Chromosomal DNA from different filamentous and unicellular cyanobacteria is often resistant to cleavage by a number of restriction endonucleases. Use of isoschizomers clearly established that the resistance of the DNA to cleavage could either be the result of a host-controlled restriction/modification system as exists in other bacterial groups, or due to the presence of a *dam* (DNA adenine methylase) enzyme of the *Escherichia coli* type (Tandeau de Marsac and Houmard, 1987). Even if modification processes by means of methylation exist in cyanobacteria they are not sufficient to explain the lack of cleavage by the number of different restriction endonucleases. Possible explanations may be the presence of inhibitors and/or DNA binding proteins in the DNA preparations, the presence of unusual bases in the DNA other than MeC and MeA, and the absence of the specific recognition sequences in the genome of certain cyanobacteria (Lambert and Carr, 1984).

1.5.3 Transcription and translation in cyanobacteria

Cyanobacterial gene expression studies increased dramatically during the last decade, lead by the cloning and sequencing of the cyanobacterial genes. Aspects of cyanobacterial transcription has been reviewed by Leach and Herdman (1973), Tandeau de Marsac and Houmard (1987) and Curtis and Martin (1994).

The transcription apparatus of *Escherichia coli* is basically conserved in eubacteria, but cyanobacterial RNA polymerases have two subunits corresponding to the β' subunit of *E. coli*, while the equivalents of α, β and σ subunits are conserved. In cyanobacteria, sigma factor genes which are common among eubacteria have been identified showing basic characteristics of the principle sigma factors. Most cyanobacterial promoters display a conserved element at -10 from the transcription initiation site. However, for the majority of genes, an element which conforms to the *E. coli* -35 promoter element is lacking. Thus many of the cyanobacterial genes characterised are subject to regulation by activator proteins, and a number of DNA binding proteins and *cis* regulatory regions have been identified in conjunction with the control of transcription initiation (Curtis and Martin, 1994).

To analyse gene expression in cyanobacteria, reporter genes have been applied to some strains. The reporter genes were used for analysing the expression of a particular gene or regulatory region of an operon. *LacZ* (encoding β-galactosidase), *cat* (encoding chloramphenicol acetyltransferase) and *lux*

(encoding luciferase) genes have successfully employed with cyanobacteria (for reviews see Friedberg, 1988; Hunter and Mann, 1992; Thiel, 1994).

1.5.4 Molecular genetic analysis of cyanobacteria

In the last 20 years cyanobacterial molecular genetics has advanced enormously especially in the last 10 years. Gene transfer systems were described, as well as transformation for unicellular and conjugation for filamentous strains. A considerable number of cyanobacterial genes have already been cloned, and further analysed. A complete list of cyanobacterial plasmids, recombinant vectors and cloned genes were given by Houmard and Tandeau de Marsac (1987). In this section advances in molecular genetic systems for cyanobacteria are summarised.

1.5.4.1 Cyanobacterial shuttle vectors

Although one or more plasmids are resident in cyanobacteria, none of them confers convenient selectible markers (Hunter and Mann, 1992). No *E. coli* plasmid vectors or broad-host-range plasmids are capable of autonomous replication in cyanobacteria. One of the way to overcome these restrictions is construction of shuttle vectors which are capable of replication autonomously both in *E. coli* and a cyanobacterial strain. Kuhlemeier *et al.* (1981) constructed two hybrid plasmids, pUC104 and pUC105 consisting of the *E. coli* vector pACYC184 and a cyanobacterial plasmid, pUC1, derived from pUH24 which is a plasmid of *Synechococcus* sp. PCC7942. These vectors were successfully used to transform both *E. coli* and *Synechococcus* sp. PCC7942. To improve gene cloning

systems in *Synechococcus* sp. PCC7942, another hybrid plasmid, pUC303, was constructed (Kuhlemeier *et al.*, 1983). This shuttle vector contains two antibiotic resistance markers Sm^rCm^r and a unique *Eco*RI site in the *cat* gene for the negative selection of recombinant molecules and is able to transform both organisms with high frequency. To solve the problems created by the recombination between incoming plasmid and the resident plasmid pUH24, the strain *Synechococcus* sp. R-2 Spc was produced by curing of pUH24. In addition to shuttle vectors derived from pAYC184, a variety of pBR-based vectors have been developed including vectors with a polylinker site to increase versatility (Gendel *et. al*, 1983). A series of shuttle vectors were constructed for filamentous strains such as *Anabaena* and *Nostoc* strains by Wolk *et al.* (1984). These vectors (pRL family) contained the replication origins and *bom* region of pBR322 and plasmid pDU1 from *Nostoc* sp. PCC 7524 which can also replicate in *Anabaena* spp. pRL family shuttle vectors for filamentous strains confer different antibiotic selectible markers of various bacterial origin. To improve the efficiency of the vectors in cyanobacterial cells, *Ava*I and *Ava*II restriction sites (Sutcliffe and Church, 1979; Hughes and Murray, 1980; Hughes *at. al*, 1980) were deleted. A complete list of shuttle vectors were listed by Tandeau de Marsac and Houmard (1987).

1.5.4.2. Gene transfer

Three genetic transfer systems have been developed for cyanobacteria to date: transformation, conjugation and electroporation. These systems have been

reviewed by Porter (1986), Tandeau de Marsac and Houmard (1987), Hunter and Mann (1992), Saunders (1992), and Thiel (1994).

1.5.4.2.1 Conjugation

Conjugation is the most successful gene transfer method for filamentous cyanobacteria. A conjugative system to transfer DNA from *E. coli* to the filamentous cyanobacterium *Anabaena* sp. PCC7120 was developed by Wolk *et al* (1984), and the methodology was reviewed by Elhai and Wolk (1988). The conjugative apparatus is provided by an IncP conjugative plasmid RP4(RK2) which mediates DNA transfer across a broad spectrum of bacteria including many cyanobacteria. The procedure involves a triparental mating, two *E. coli* donor parents and one recipient cyanobacterial parent. One of the *E. coli* donors (donor 1) carries the broad-host-range conjugative plasmid RP4 which encodes *tra* genes for the pilus, and for an enzyme that nicks its own *ori*T (origin of transfer) and for other proteins required for conjugative plasmid to transfer itself and the ability to mobilise other plasmids into a recipient (Thomas and Smith, 1987). The other *E. coli* donor (donor 2) carries two plasmids: i) a cargo plasmid to be transferred into the cyanobacterial recipient cell and ii) a helper plasmid. The cargo plasmid is a typical shuttle vector or integrative plasmid with an *oriT* region to be nicked before mobilisation into the recipient cell. The helper plasmids contain methylase genes to protect the cargo plasmid from the restriction by host and the *mob* gene to nick the *oriT* region of the cargo plasmid. RP4 transfers itself into donor2 in which the cargo plasmid is being nicked and methylated *in trans* by the helper

plasmid. The cargo plasmid is then mobilised into the cyanobacterial recipient cell by the conjugative plasmid RP4. The precise mechanisms of bacterial conjugation were reviewed by Willetts and Skurray (1980), Simon *et al.*, (1983), and Willetts and Wilkins (1984).

Conjugation has been used successfully to transfer recombinant DNA to filamentous cyanobacteria, primarily to *Anabaena* sp. PCC7120. Gene transfer studies have been extensively carried out by conjugation in various filamentous strains such as *Anabaena variabilis*, *Nostoc* sp. ATCC29133, *Nostoc* sp. PCC 8009, and *Calothrix* sp. PCC 7601 (Thiel, 1994).

1.5.4.2.2. Transformation

Transformation is a simple way to transfer DNA to some of unicellular cyanobacteria that are members of two genera, *Synechococcus* and *Synechocystis*. Although the mechanism of transformation is poorly understood, the mechanism may share some characteristics with other transformable bacteria. *Synechococcus* sp. strains PCC 7943, PCC 6301, PCC7942, 7002, and *Synechocystis* sp. strains PCC 6803, PCC 6714 have been reported to be transformed (Grigorieva and Shestakov, 1982; Porter, 1986) and a general methodology reviewed by Porter (1988). More efforts have been focused on plasmid transformation recently. Shuttle vectors and integrative vectors, i.e. vectors that cannot replicate in cyanobacterial strain, have been transformed into unicellular strains for the purposes either to maintain shuttle vectors or to integrate a recombinant vector

into the chromosome. No transformation system for filamentous strains has been reported to date (for reviews see Thiel, 1994; Hunter and Mann, 1992; Tandeau de Marsac and Houmard, 1987; and Porter, 1986).

1.5.4.2.3 Electroporation

Recently it has been shown that electroporation is a possible route to transfer DNA into filamentous cyanobacteria. *Anabaena* sp. strain M131 has been transformed with the shuttle vector pRL6 by electroporation, and it required relatively high field strengths with short time constants (Thiel and Poo, 1989). Host restriction is a barrier to transformation by electroporation both in *Anabaena* sp. strain M131 and *Nostoc* sp. PCC 7121. This problem has been overcome by the deletion or restriction modification methylase treatment of restriction sites on recombinant vector (Thiel and Poo, 1989, and Moser *et. al* , 1993). Despite being limited to a few cyanobacterial strains at present, electroporation may be more widely applicable within the cyanobacteria (Hunter and Mann, 1992).

1.5.4.3 Mutagenesis

Classical and molecular methods have been applied for the mutagenesis of cyanobacterial genes. In this section these classical and non classical methods are summarised.

1.5.4.3.1. Chemical and UV mutagenesis

A series of chemicals and UV irradiation have used for random mutagenesis of cyanobacterial genes. Mutant characterisation is based on observation of a lost wild type phenotype. The general aspects of chemical and UV mutagenesis are reviewed by Herdman, (1982); Golden (1987), and Thiel (1994). Although both chemical and UV mutagenesis have become less popular with the development of molecular methods for mutagenesis, it has been shown that the combination of chemical and UV light mutagenesis and the complementation of mutations by wild type genes on plasmids could provide a powerful tool for the discovery and characterisation of novel genes (Thiel, 1994).

1.5.4.3.2. Nonclassical mutagenesis

Molecular genetic techniques have become more popular for mutagenesis of cyanobacterial genes or a locus in the chromosome. A few different methods are available to mutate either a cloned gene or a random array of genes (for reviews see Hunter and Mann, 1992; Saunders, 1992, and Thiel, 1994).

Targeted inactivation of a gene in chromosome by the interruption of the cloned gene with a selectible marker is the most favoured method in all cyanobacteria for which gene transfer is feasible. The method is based on recombination between a recombinant vector which contains a recombinant copy of the gene of interest. An antibiotic cassette, preferably flanked by transcription and translation termination signals, is inserted into the recombinant gene, or the

cassette replaces the gene after partial or complete deletion. After a double recombination event, the inactivated copy of the cloned gene and wild type are exchanged and thus chromosomal copy of gene is inactivated. This approach has successfully been used for mutagenesis of a variety of cyanobacterial genes. Scanlan *et al.* (1995) have characterised a *zwf* mutant of *Synechococcus* sp. PCC7942 after double recombination between a construct in pUC19 containing the *zwf* gene interrupted by an Ω fragment (Prentki and Krisch, 1984) and the wild type copy. To increase the recombination rate, some cargo plasmids such as pRL271 lacking a cyanobacterial replicon were constructed for filamentous strains (Black *et al.*, 1993). A recombinant DNA including an interrupted copy of the gene of interest is cloned into this type of vector and introduced into the filamentous cyanobacterial cells. The only route for maintenance of the recombinant vector is the integration into the chromosome at the homologous regions. These vectors also contain the *sacB* gene which maintains the positive selection of double recombinants, as well as some other antibiotic selection markers.

Apart from targeted inactivation, some other mutagenesis approaches also exist such as transposon mutagenesis and random insertional mutagenesis. Transposon Tn901 in pUH24 in *Synechococcus* sp. PCC7942 and Tn5 in *Anabaena* sp. PCC7120 have been used for mutagenesis and the subsequent cloning of novel genes. In unicellular cyanobacteria, a random insertion approach is an alternative to transposon mutagenesis. An antibiotic resistance gene is

randomly ligated into the restriction fragments of genomic DNA. The ligated DNA is transformed into cyanobacterial host. The recombinations between the ligated DNA and the chromosome create insertional or deletion-insertional mutations (Labbare *et al.*, 1989; Hunter and Mann, 1992; Thiel, 1994).

Site-directed mutagenesis has also been applied to cyanobacterial genes. A gene of interest is modified by substituting a single nucleotide or a group of nucleotides and then transferred back to the cyanobacterial host which lacks the gene of interest. The use of this technique was reviewed by Williams (1988).

Some of cyanobacteria have as many as 10-15 copies of their haploid genome (Herdman *et al.*, 1979; Labarre *et al.*, 1989). Therefore, to isolate a cyanobacterial mutant initial recombinants need to be segregated through several generations. In filamentous strains mutant segregation needs some further considerations. Even if the total segregation of mutant copies of genome occurs, a single wild type chromosome residing in a cell in the filament can confer the wild type phenotype. To overcome this problem, exconjugants can be grown through several generations and filaments can be fragmented at regular intervals until complete segregation has occurred as determined by Southern blot and phenotypic analyses.

1.6 Carbon metabolism

Cyanobacteria are autotrophs that can produce their carbon source for energy and biosynthetic constituents via photosynthetic operation in the light. The stable reserve of carbohydrate is glycogen. In the dark and in the light where necessary, these organisms are capable of generating maintenance energy during periods of darkness via dissimilation of the carbohydrate stored as glycogen. Although phototrophy is the dominant mode of carbon metabolism, some cyanobacterial strains were reported that can grow photoheterotrophically on one or more carbon compounds (Carr, 1973; Smith, 1973; Smith, 1982). In this section, photosynthesis, CO_2 fixation through the Calvin cycle and the routes of carbohydrate dissimilation are summarised.

1.6.1 Photosynthesis

Cyanobacterial photosynthesis is very similar to that of higher plants using H_2O as the electron donor and evolving O_2. The chemistry of the reaction centres of Photosystem I and Photosystem II (PSI and PSII) does not differ from that of the corresponding components of higher plant chloroplasts. However, cyanobacteria are distinct from chloroplasts in respect to the light harvesting antenna composition and structure. Higher plants utilise both chlorophyll *a* and *b* as light harvesting pigments, whereas cyanobacteria possess only chlorophyll *a*. In cyanobacteria, however, phycobiliproteins play a central role in light harvesting (Glazer, 1982). In addition to the predominant nature of plant-like photosynthesis, under anaerobic conditions, some cyanobacteria are capable of operating an

anoxygenic bacterial-type photosynthesis using H_2S as an electron donor and expelling elemental sulphur (Padan and Cohen, 1982; Stewart, 1988).

The cyanobacterial photosynthetic apparatus consists of three macromolecular complexes: PSI, PSII and phycobilisomes. Phycobilisomes serve as the primary light-harvesting antennae for PSII. Phycobilisomes are composed of a pigmented family of water-soluble proteins, the phycobilins: phycoerythrin, phycocyanin and allophycocyanin. Absorbed light energy is transferred from phycobilins to PSII or partially to PSI (for reviews see Glazer, 1987; Grossman *et. al*, 1993; Sidler, 1994). PSII is the membrane protein complex that carries out light-catalysed oxidation of water and reduction of plastoquinone. The reaction centre is composed of both intrinsic and extrinsic proteins; the prosthetic groups involved in electron transfer include chlorophyll *a*, pheophytin, quinone, tyrosine residues, and a manganese cluster (for reviews see; Sherman *et al.*, 1987; Rutherford, 1989; Barry *et al.*, 1994). PSI is also a membrane-bound complex and is composed of chlorophyll *a*, a series of inorganic and organic cofactors and a matrix of approximately eleven polypeptides to provide ligands to the photoactive components (for reviews see Golbeck, 1992, 1994).

1.6.1.1 Light reactions

Through the membrane-bound PSI and PSII, electrons are utilised by splitting H_2O at PSII and are transferred to produce NADPH and a pH gradient used for ATP synthesis. These molecules are subsequently used to fix CO_2 via the

dark reactions of photosynthesis. Light energy is absorbed primarily by phycobiliproteins at the P_{680} of PSII which are used to drive the oxidation of water and to give the excitation energy to electrons which are transferred through PSII components, and finally transferred to the Cyt b_6f complex which is a membrane bound enzyme complex (Kallas, 1994). Meanwhile, H^+ ions are also transferred by the complex from the cytoplasm to the thylakoid lumen creating a pH gradient subsequently used for synthesis of ATP by a membrane bound ATP synthase. The electrons are then transferred to cytochrome c_{553} or plastocyanin (Morand *et al.*, 1994). Plastocyanin carries an electron to PSI; the first acceptor is the P_{700} reaction centre of PSI which has already lost an electron by light energy being transferred primarily by chlorophyll *a* and occasionally by phycobiliproteins. The electrons are transferred through a series of components of PSI, and then to a soluble electron transport protein which is either ferredoxin or flavodoxin, ferredoxin: NADP oxidoreductase, and finally $NADP^+$. Cyclic electron transport also occurs at PSI. Electrons transported through PSI join the electron transport chain again possibly at Cyt b_6f complex. More details of the structure and function of the cyanobacterial light reaction centres and soluble elements were reviewed in more detail elsewhere (for reviews see Krogmann, 1973; Ho and Krogmann, 1982; Glazer, 1987; Barry *et al.*, 1994; Golbeck, 1994; Kallas, 1994; Morand *et al.*, 1994).

1.6.1.2 Dark reactions (Reductive pentose phosphate cycle)

NADPH and ATP generated by the light reactions of photosynthesis are used for the reduction of CO_2 into triose phosphates first, and then hexose phosphates, and finally stored as glycogen (for reviews see Smith, 1982; Tabita 1987, 1994). The key and unique reactions of Calvin cycle are catalysed by phosphoribulokinase and ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). The phosphoribulokinase reaction is important for regeneration of the CO_2 acceptor, RuBP (see Figure 1.2)

1.6.2 Oxidative carbon metabolism

Glycogen, which is synthesised during the light period, is degraded to CO_2 via the oxidative pentose phosphate (OPP) cycle as a predominant mode, and electrons are transferred to oxygen via a respiratory electron transport chain (Smith, 1982). However, some of the cyanobacteria can operate alternative pathways under anaerobic conditions. The unicellular cyanobacterium *Microcystis* sp. PCC 7806 was reported to be capable of carrying out the fermentation of endogenous glycogen to ethanol, acetate, CO_2, and H_2 when incubated anaerobically in the dark (Moezelaar and Stal, 1994). Some of *Oscillatoria* species carry out fermentation in the darkness under anoxic conditions where fermentation seems to be the only means of survival during dark periods (Richardson and Castenholz, 1987; Heyer *et al.*, 1989). *Nostoc* sp. strain Cc, degrades exogenous glucose as well as storage carbohydrates, using the homoacetic fermentation pathway (Margheri and Allotto, 1993). Despite the capability of some cyanobacterial strains to perform fermentation under such rare circumstances, the

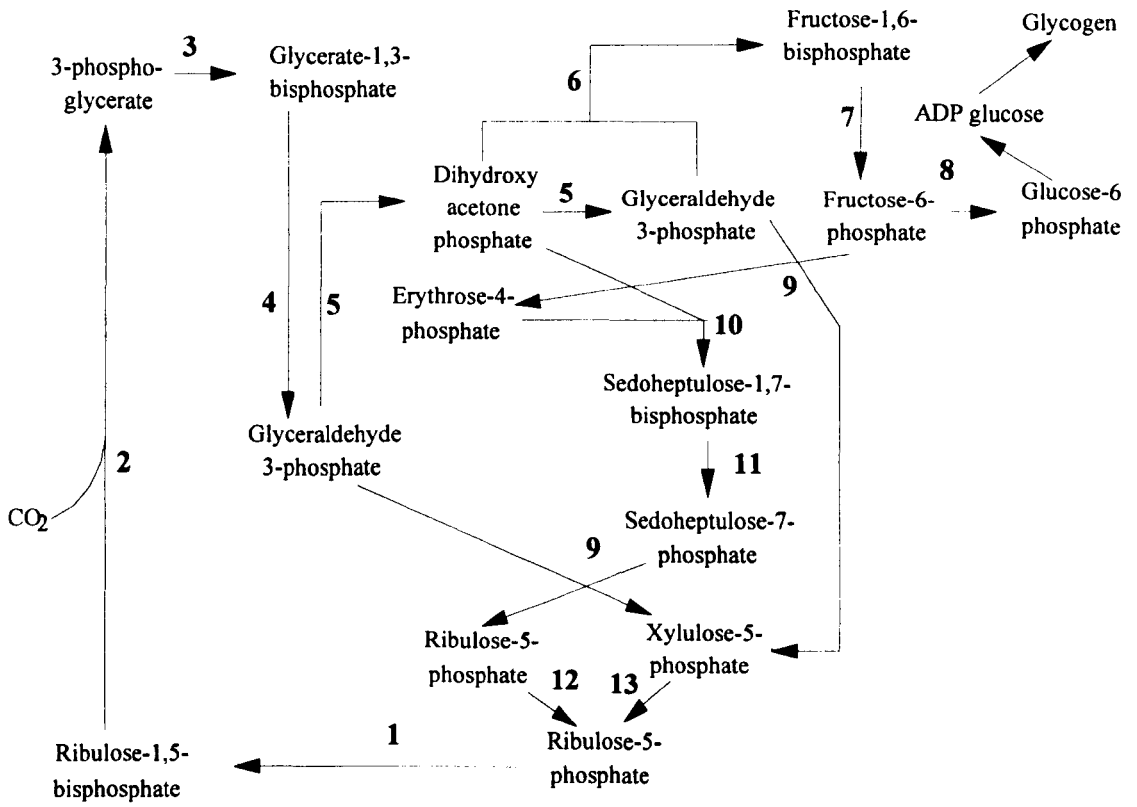


Figure 1.2. CO_2 assimilation via the reductive pentose phosphate pathway (after Smith, 1982). Numbers on the reaction lines represent the enzymes catalysing the reaction as follows: **1**, 5-Phosphoribulokinase; **2**, Rubisco; **3**, 3-Phospholycerate kinase; **4**, Glyceraldehyde 3-phosphate dehydrogenase; **5**, Triose phosphate isomerase; **6**, Fructose-1,6 bisphosphate aldolase; **7**, Fructose-1,6-bisphospate phosphatase; **8**, Hexose-6-phosphate isomerase; **9**, Glyceraldehyde transferase (transketolase); **10**, Aldolase; **11**, Sedoheptulose-1,7-bisphosphate phosphatase; **12**, Ribose-5-phosphate isomerase; and **13**, Ribose-5-phosphate 3-epimerase.

oxidative pentose phosphate cycle is the main route for carbohydrate dissimilation.

1.6.2.1 Oxidative pentose phosphate cycle

When cyanobacterial cells are in the dark, the RPP cycle is completely replaced by the oxidative pentose phosphate (OPP) cycle (Figure 1.3). Glucose-6-phosphate from the glycogen reservoir is catabolised through the cycle. During every cycle one molecule of CO_2 is released, and ribulose-5-phosphate remains in the cycle. Intermediates are incorporated via fructose-6-phosphate and then glucose-6-phosphate which enters the cycle again (Smith, 1982). Two dehydrogenases, glucose-6-phosphate dehydrogenase (G6PDH) and 6-phosphogluconate dehydrogenase (6PGDH), are unique for the OPP cycle. Reduction of two molecules of $NADP^+$ are carried out by these enzymes at reactions 1 and 3. One molecule of CO_2 is also released at reaction 3.

There is a very close interconnection between the RPP and the OPP cycles. Most of the individual reactions operate in both cycles. Two enzymes in each cycle are unique and play a key role: 5-phosphoribulokinase and Rubisco for RPP cycle, and G6PDH and 6PGDH for OPP cycle (Stanier and Cohen-Bazire, 1977). Rubisco is one of the most important targets for the regulation of CO_2 fixation. The enzyme is deactivated by accumulated metabolites especially by 6-phosphogluconate which is the first intermediate molecule of the OPP pathway. This causes an immediate cessation of CO_2 fixation (Tabita, 1994). G6PDH is the target regulation point of the OPP cycle. After a dark to light shift, the enzyme is

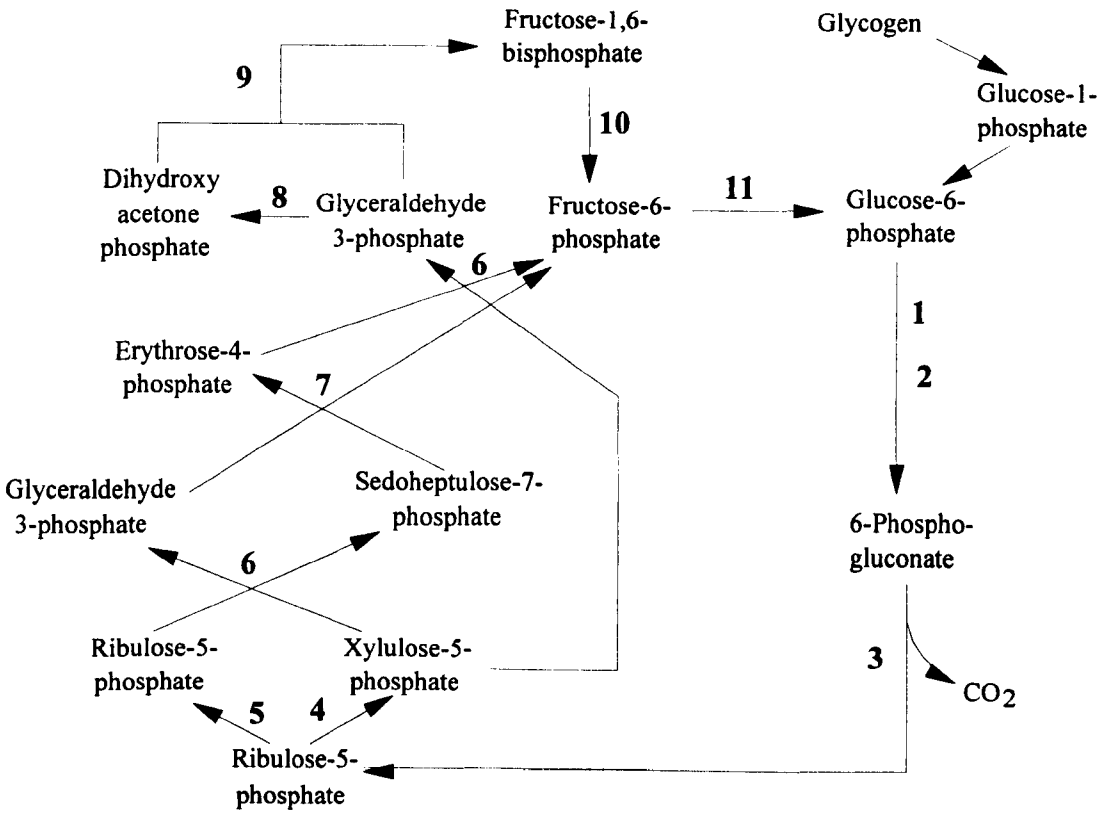


Figure 1.3. Carbohydrate catabolism via the oxidative pentose phosphate pathway (after Smith, 1982). Numbers on the reaction lines represent the enzymes catalysing the reaction as follows: **1**, Glucose-6-phosphate dehydrogenase; **2**, Phosphogluconolactonase; **3**, Phosphogluconate dehydrogenase; **4**, Ribose-5-phosphate isomerase; **5**, Ribose-5-phosphate 3-epimerase; **6**, Glyceraldehyde transferase (transketolase); **7**, Dihydroxyacetone transferase (transaldolase); **8**, Triose phosphate isomerase; **9**, Fructose-1,6 bisphosphate aldolase; **10**, Fructose-1,6-bisphospate phosphatase; and **11**, Hexose-6-phosphate isomerase.

deactivated and as a result the OPP cycle is deactivated (Smith, 1982; Copeland and Turner, 1987).

The OPP cycle has been reported to be an electron supplier to nitrogenase in heterocysts (Stanier and Cohen-Bazire, 1977; Wolk, 1987; Wolk *et. al*, 1994). Carbohydrates supplied from neighbouring vegetative cells are catabolised through the OPP cycle producing NADPH which is subsequently used by nitrogenase to reduce N_2 into NH_3 (Böhme, 1987). In the dark, ATP as a maintenance energy source is also supplied by the NADPH produced by the OPP cycle after respiratory electron transport. Aspects of respiratory electron transport in cyanobacteria were reviewed by Peschek (1987), Matthijs and Lubberding (1988), and Schmetterer (1994).

Three enzymes of the OPP pathway are the subject of this study: Glucose-6-phosphate dehydrogenase (G6PDH), fructose-1,6-bisphosphatase (FBPase) and transaldolase. G6PDH is especially focused upon, therefore, the structures and functions of these enzymes will be further summarised.

1.7. Transaldolase

Transaldolase (EC 2.2.1.2) catalyses the transfer of a C_3 unit from sedoheptulose-7-phosphate (S7P) to glyceraldehyde-3-phosphate(GAP) yielding erythrose-4-phosphate (E4P) and fructose-6-phosphate (F6P) (Figure 1.4). The enzyme is capable of carrying out the reactions in the reverse direction.

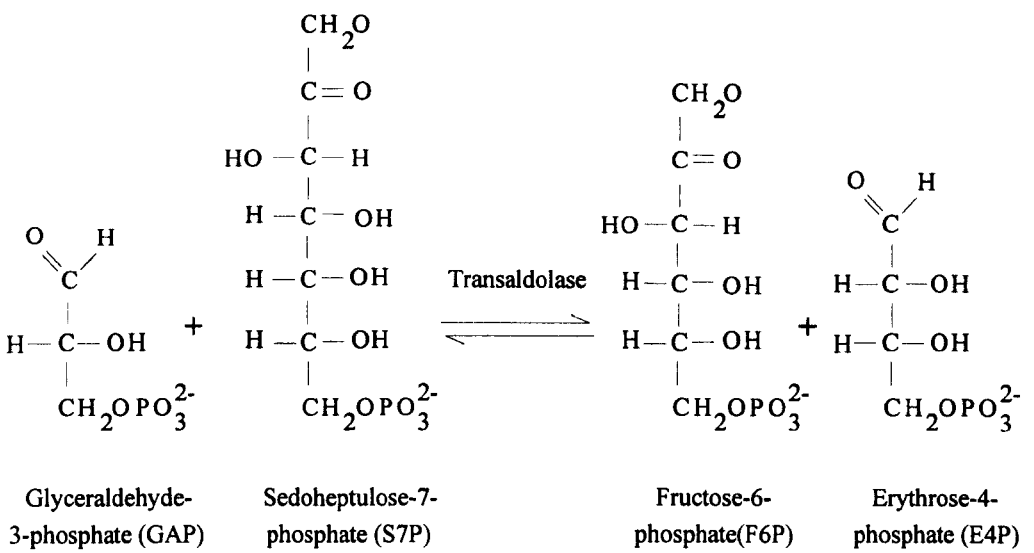


Figure 1.4. The reaction catalysed by transaldolase.

Transaldolase has a central role in the complete degradation of G6P into CO_2 via the reactions of the OPP cycle during the dark period and in heterocysts (Smith, 1982). Beyond the review by Smith (1982), there is no detailed structural or functional studies on the cyanobacterial transaldolase. Latzko and Gibbs (1969) reported transaldolase activity when the cells were evolving O_2 in *Tolypothrix tenuis* (*Calothrix* sp. PCC 7101). Recently, Summers *et. al* (1995a) have reported a transaldolase gene (*tal*) of *Nostoc* sp. ATCC29133. The characterisation of the gene was based on database comparison of the deduced amino acid sequence.

1.8. Fructose-1,6-bisphosphatase (FBPase)

Fructose-1,6-bisphosphatase (FBPase) (EC 3.1.3.11) hydrolyses fructose-1,6-bisphosphate to fructose-6-phosphate and orthophosphate (P_i) (Figure 1.5).

The enzyme participates in the photosynthetic fixation of CO_2 through the RPP cycle in photosynthetic organisms (Udvardy *et. al*, 1982). In cyanobacteria, however, FBPase was reported to be a component of the OPP cycle too (Smith, 1982). In higher plants, two isoenzymes have been reported, which are chloroplastic and cytosolic (Schnarrenberger *et al.*, 1995). Each of the isozymes is the product of a nuclear encoded gene. The cytosolic isozyme is involved in gluconeogenesis and sucrose synthesis. It is more homologous to the mammalian type than chloroplast form and inhibited by regulatory metabolites,

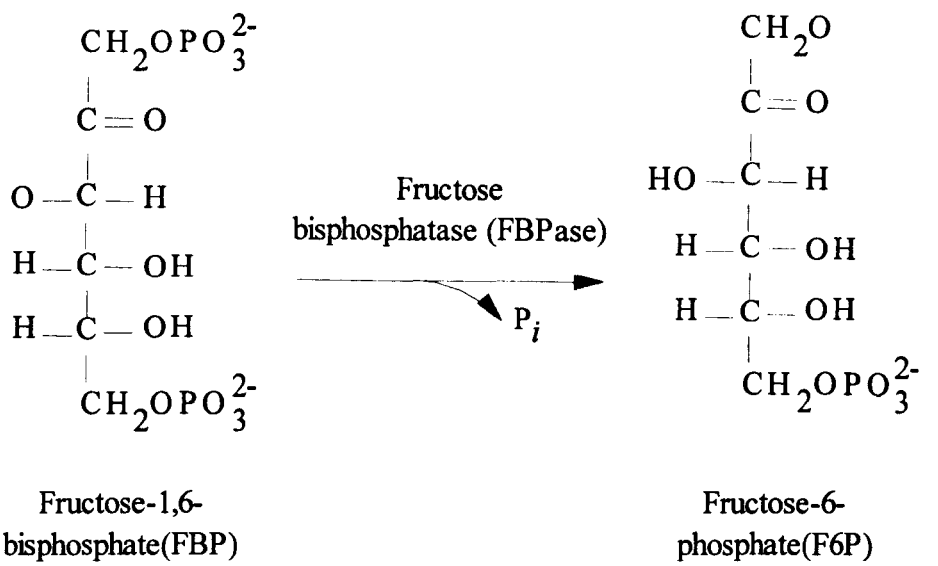


Figure 1.5. The reaction catalysed by FBPase.

such as AMP and fructose-2, 6-bisphosphate. In sugar beet (*Beta vulgaris* L.) leaves, cytosolic FBPase is also subject to light modulation in an indirect manner (Khayat *et al.*, 1993). The chloroplast isoform is present in the stroma of the chloroplasts and is one of the key regulatory components of the RPP cycle, as it is involved in the regeneration of the acceptor molecule for CO_2 fixation. The chloroplastic isoform of FBPase was reported to be a light modulated enzyme like some other RPP cycle enzymes being active in the light. Unlike the cytosolic isoform, the activation of chloroplast isoform is directly light-dependent by way of a ferredoxin/thioredoxin system and insensitive to AMP (Clancey and Gibert, 1987). Clancey and Gilbert (1987) reported that thermodynamically, the FBPase/thioredoxin f_b system was extremely sensitive to oxidation. Amino acid sequences of FBPase from various plant chloroplasts showed that a conserved region with two redox sensitive cysteine residues does not exist in the cytosolic isoforms (Marcus *et al.*, 1988; Koßmann *et al*, 1992; Carrasco *et al.*, 1994). This region is particularly involved in disulphide bond formation in the chloroplast isoform.

Udvardy *et al.* (1982) reported that cyanobacterial FBPase is a light regulated enzyme similar to chloroplast isoform of plant FBPase. Contrary to the results of Bishop (1979) which suggested a regulatory property of the enzyme intermediate between chloroplastic and cytosolic isoforms, Udvardy *et. al* (1982) reported that the properties of *Anacystis nidulans* (*Synechococcus* sp. PCC 6301) FBPase were very close to that of chloroplast FBPase. FBPase from both crude

and purified preparations is activated by the reduced form of glutathione. They concluded that the observation suggested an activation mechanism based on thiol-disulphide exchange in which mixed disulphide bonds of the enzyme were involved. In *Nostoc* sp. strain *Mac* (PCC 8003), FBPase as well as some other RPP cycle enzymes is reversibly light-activated *in vivo* by the cyanobacterial thioredoxin system (Austin *et al.*, 1992). The deduced amino acid sequence of FBPase from *Nostoc* sp. ATCC29133 did not have a conserved redox sensitive region like that of the chloroplast isoform although three cysteine residues were present (Summers *et al.*, 1995a).

1.9 Glucose-6-phosphate dehydrogenase (G6PDH)

Glucose-6-phosphate dehydrogenase (G6PDH) is the first enzyme of the OPP cycle and one of the most important regulatory enzymes of cyanobacterial carbon metabolism (Stanier and Cohen-Bazire, 1977). The enzyme plays an important role in heterocyst metabolism and is thought to be an electron donor to nitrogenase as one of the two dehydrogenases of the OPP pathway (Wolk *et. al*, 1994). The general properties of G6PDH were reviewed by Levy (1979). In this section, the general properties and the modes of regulation of G6PDH from cyanobacteria and from plants will be summarised.

1.9.1. The reaction

Glucose-6-phosphate dehydrogenase (G6PDH) catalyses the oxidation of D-glucose-6-phosphate (G6P) to D-glucono-δ-lactone-6-phosphate by $NADP^+$

(Figure 1.6) Although the reaction is thermodynamically reversible, it is rendered effectively irreversible by the rapid hydrolysis of D-glucono-δ-lactone-6-phosphate to 6-phosphogluconate. In addition to this spontaneous hydrolysis, a widely distributed lactonase ensures the irreversibility of the G6PDH-catalysed reaction under physiological conditions (Levy 1979).

1.9.2. The structure

Molecular genetic characterisations of G6PDH from *Synechococcus* sp. PCC7942 (Scanlan *et al.*, 1992) and *Nostoc* sp. ATCC29133 (Summer *et. al*, 1995a) have showed that a possible active site from some other organisms (Jeffery *et al*, 1989) has also been conserved in the amino acid sequence of glucose-6-phosphate dehydrogenase gene (*zwf*) of these cyanobacteria. The predicted molecular mass of G6PDH of *Nostoc* sp. ATCC29133 is 58,231, which is a half of the molecular weight of the hypoactive form of *Anabaena* sp. PCC7120 G6PDH as reported by Schaeffer and Stanier (1978).

G6PDHs from all organisms show an oligomeric structure; the monomer molecular mass of microbial G6PDHs are 50,000 to 60,000 while those of mammalian G6PDHs are 58,000-67,000 (for review see Levy, 1979). Schaeffer and Stanier (1978) reported that *Anabaena* sp. PCC7120 G6PDH is an hysteretic enzyme (for hysterisis of enzymes see Frieden, 1979), which undergoes relatively slow reversible transitions between different states of aggregation. Sucrose gradient centrifugation and polyacrylamide gel electrophoresis of the partially

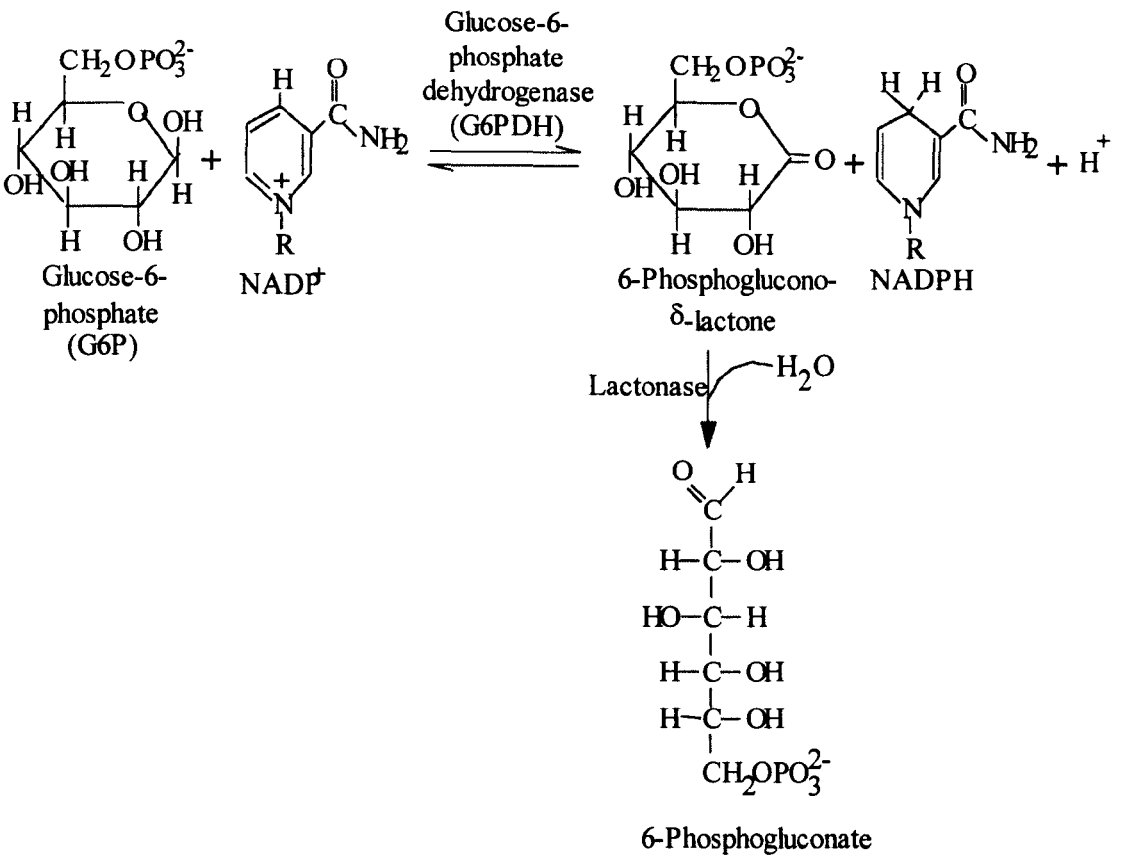


Figure 1.6. The reaction of G6PDH and lactonase. R represents 2'-phosphoadenosine-diphosphate-ribose.

purified enzyme revealed three principle forms, with approximate molecular mass of 120,000 (M_1), 240,000 (M_2) and 345,000 (M_3). The relative activities of these principle forms are $M_1<M_2<M_3$. The equilibrium favoured the more active oligomeric forms in the concentrated solutions of the enzyme. The M_1 form had a lag period in activity while M_3 had a burst. M_2 had neither a lag nor a burst in activity. Dilution in the absence of effectors shifted the equilibrium in favour of the M_1 form with a marked diminution of catalytic activity. The transition to the M_1 form was prevented by glucose-6-phosphate and by glutamine while $NADP^+$ and in crude extracts ribulose-1,5-bisphosphate tended to maintain the enzyme in the M_1 form (Schaeffer and Stanier 1978).

The G6PDH from *Anacystis nidulans* exhibited different aggregation states and the oxidation of the enzyme led to increase in the molecular weight of the enzyme (Balogh *et al.*, 1979; Cséke *et al.*, 1981; Udvardy *et al.*, 1983). Cossar *et al.* (1984) purified the enzyme from *Anabaena variabilis* and it had a molecular mass of approximately 250,000. This corresponded to the tetramer of a single polypeptide with a molecular mass of 56,200, which was determined by means of SDS-PAGE. The cytosolic isoform of G6PDH partially purified from potato (*Solanum tuberosum* L.) tubers had a molecular weight of 55,000 determined by SDS PAGE and 210,000 by gel filtration (Graeve *et. al*, 1994). Fickenscher and Scheibe (1986) found a similar molecular mass for cytoplasmic G6PDH from pea, 240,000 for native enzyme and 60,000 for the subunit. In pea chloroplasts, the enzyme has a tetrameric form with a molecular mass of 224,000, the monomer has

a molecular mass of 56,000 (Srivastava and Anderson, 1983). In conclusion, G6PDH from cyanobacteria and plants has a subunit with a molecular mass of between 55,000 and 60,000, and the native functional form is a tetramer with a molecular mass of between 210,000 and 250,000, corresponding the M2 form of *Anabaena* sp. PCC7120 G6PDH (Schaeffer and Stanier, 1978).

1.9.3. Regulation

The modes of the regulation of G6PDH have been in focus for a long time due to the importance of the enzyme in the regulation of carbon metabolism and in nitrogen fixation (Stanier and Cohen-Bazire, 1977; Buchanan, 1980; Copeland and Turner, 1987; Simpson, 1990). The dominant regulatory feature of cyanobacterial G6PDH is the light-dependence of its activity. Reduced thioredoxin, which is a product of photosynthesis, is the most powerful candidate for the light deactivation of the enzyme not only in cyanobacteria (Cseka *et al.*, 1981; Yee *et al.*, 1981; Udvardy *et al.*, 1983, 1984; Cossar *et al.*, 1984; Rowell, *et al.*, 1988) but also in plants (Buchanan, 1980; Srivastava and Anderson, 1983; Fichenscher and Scheibe, 1986; Graeve *et al.*, 1994). Some other metabolites and physiological conditions that effect the transitions between oligomeric forms may play important roles in the light modulation. Ribulose-1,5-bisphosphate (RuBP) which is the key intermediate of RPP cycle was suggested as the light modulator. Pelroy *et al.* (1972) reported that RuBP was a control compound, but Grossman and McGowan (1975) could not confirm this result. In *Anabaena* sp. PCC7120,

RuBP had no effect on the M_1 form of G6PDH activity while it acted as negative effector of transition from M_1 to M_2 (Schaeffer and Stanier, 1978).

In addition to light modulation, enzyme activity may be controlled by some other factors under certain conditions. Broedel and Wolf (1991) reported that G6PDH activity in *Synechococcus* sp. PCC7942 was growth-phase-dependent. The specific activity of G6PDH, and also that of 6-phosphogluconate dehydrogenase (6PGDH), were constant during exponential growth but it increased continuously during the transition to stationary phase. The activity stopped increasing and remained constant when the culture reached stationary phase. At the end of the transition phase, G6PDH activity was about five-fold higher than during the exponential growth. This growth-phase dependence of G6PDH activity may be due to energy starvation that occurs as the effective light intensity in the culture drops because of selfshading. This may affect the photosynthetic rate releasing the light-dependent inactivation on the enzyme and accumulation of metabolites which favour the more active oligomeric forms (Broedel and Wolf 1991). G6PDH from the mesophilic cyanobacterium *Synechococcus* sp. PCC 6307 and the thermophilic cyanobacterium *Synechococcus* sp. PCC 6716 was immediately affected by temperatures higher than the optimum and different growth temperatures did not lead to the formation of different isoenzymes or modified enzymes indicating that temperature affects the catalytic activity directly (Lubberding and Bot, 1984). It was reported that G6PDH in *Anacystis nidulans*, after a phage infection, had undergone a transition

into a hyperactive form, and the activity increased both in the light and in the dark. Apparently, infected cells lose the reducing system, which deactivates the enzyme, resulting in the oxidation of the enzyme and increasing the enzyme activity (Balogh *et al.*, 1979; Cséke *et al.*, 1981).

1.9.3.1. NADPH and ATP

NADPH and ATP are inhibitors of the enzyme. In partially purified G6PDH from *Anacystis nidulans* and *Anabaena flos-aquae*, ATP inhibited the activity competitively with NADP while NADPH inhibition was competitive with glucose-6-phosphate (Grossman and McGowan, 1975). Schaeffer and Stanier (1978) found similar results working with partially purified *Anabaena* sp. PCC7120 enzyme. NADPH inhibited the hyperactive form of the enzyme at pH 6.5, whereas ATP inhibited G6PDH in the hyperactive, normal and hypoactive states. The principle effect of both ATP and NADPH was accordingly upon catalytic function (Schaeffer and Stanier, 1978). The higher NADPH/$NADP^+$ ratios inhibited the enzyme activity in these organism. In *Anacystis nidulans*, NADPH levels dropped to 0.04 mM after a light-dark shift from 0.23 mM in the light, while $NADP^+$ remains more or less stable both in the light and in the dark, with concentrations of 0.19 mM and 0.17 mM, respectively (Ihlenfeldt and Gibson, 1975). This result reveals that the NADPH/$NADP^+$ ratio is low in the dark maintaining the enzyme in an active form.

1.9.3.2. pH

The optimal pH for catalytic activity was 7.4 for the enzyme from *Anabaena* sp. PCC7120 (Schaeffer and Stanier, 1978) and 7.2 for the enzyme from *Synechococcus* sp. PCC 6307 and *Synechococcus* sp. PCC 6716 (Lubberding and Bot, 1984). The influence of pH on G6PDH activity from *Anabaena* sp. PCC7120 is extremely complex, since it affects both catalytic activity and the equilibrium between hypoactive and normal oligomeric forms. In diluted solutions, the enzyme activity remained constant for at least 15 min at pH 6.5, whereas the activity at pH 7.9 fell to half its initial value in 10 sec. Accordingly, the M_2 form of the enzyme was favoured under mildly acid conditions (pH 6.5-6.9) while pH values above 7.0 favoured the hypoactive form. pH had a very complex effect on the activity of G6PDH as a function of G6P concentration. In the hyperactive state, the enzyme displayed Michelis-Menten kinetics at pH 6.9. At pH 8.1 the concentration of G6P necessary for half maximal activity increased greatly. The effect of alkaline conditions were even more marked if the enzyme was initially in the hypoactive state (Schaeffer and Stanier, 1978). G6PDH from *Anacystis nidulans* and *Anabaena flos-aquae* exhibited Michelis-Menten kinetics at pH 6.7 and sigmoidal (hysteretic) kinetics at pH values above 7.4 (Grossman and McGowan, 1975). In plant cells, pH changes were light induced and the increase in the pH of the stroma was about 1 (which was 7 in the dark and 8 in the light) (Buchanan, 1980). The pH of the cyanobacterial cytoplasm was lower in the dark than in the light as was observed in *Anacystis nidulans*. This difference was

maintained by light-dependent uptake of protons from the cytoplasm into the thylakoids (Falkner *et al.*, 1976).

1.9.3.3. Positive effectors, G6P and glutamine

Glutamine and G6P have both been reported to be positive effectors for the G6PDH activity in *Anabaena* sp. PCC7120 (Schaeffer and Stanier, 1978) and in *Anabaena variabilis* (Cossar *et al.*, 1984). As positive effectors they maintain the enzyme in the active state under otherwise unfavourable conditions. In *Anabaena* sp. PCC7120, both G6P and glutamine, although not a substrate of the enzyme, shared the ability to displace the equilibrium in favour of M_2 state. However, glutamine had no influence on the kinetics of the enzyme in the hyperactive and normal states (Schaeffer and Stanier, 1978). Similar results was observed in *Anabaena variabilis* (Cossar *et al.*, 1984). G6P and glutamine prevented the deactivation and promoted reactivation.

1.9.3.4. Thioredoxin and DTT

G6PDH from different sources has been reported to be under redox control (Anderson *et al.*, 1979; Cossar *et al.*, 1984; Udvardy *et al.*, 1984; Graeve *et al.*, 1994). Disulphide bond(s) may be involved in the reduction of the enzyme resulting in deactivation. The formation of disulphide bonds occurs between sulphydryl groups of cysteine residues in the enzyme (for revievs see Freedman *et al.*, 1994; Wittrup, 1995). In *Anacystis nidulans*, the powerful, non-physiological reductant dithiothreitol (DTT) decreased G6PDH activity in crude extracts (Cseke

et al., 1981). However, DTT did not affect the G6PDH activity in partially purified solutions from *Anabaena* sp. PCC7120 (Udvardy *et al.*, 1984) or from *Anabaena variabilis* (ATCC 29413) (Cossar *et al.*, 1984). When DTT was added to the partially purified enzyme in the presence of the low molecular weight (approximately 12,000) protein, thioredoxin, the activity was inhibited both *Anabaena* sp. PCC7120 and *Anabaena variabilis* (ATCC 29413) enzymes. The results implicated that thioredoxin is the *in vivo* reductant for cyanobacterial G6PDH and also for the other RPP cycle enzymes.

Thioredoxin has been widely implicated in the *in vivo* light modulation of enzyme activity in cyanobacteria and plants (for reviews see Rowell *et al.*, 1988; Gleason, 1994). In higher plants two types of thioredoxins exist, which are thioredoxin *f* and thoredoxin *m*, with two types of thioredoxin oxidoreductase enzymes, which are ferredoxin-thioredoxin oxido-reductase in chloroplasts and NADP-dependent thioredoxin oxidoreductase in non-photosynthetic plant tissues. Carbohydrate metabolism enzymes such as fructose-1,6-bisphosphatase and glucose-6-phosphate dehydrogenase are modulated by thioredoxin. Thioredoxins are ubiquitous, small redox proteins, with approximate molecular weights of 12,000 (Rowell *et al.*, 1988). Thioredoxin *m* deactivates chloroplast G6PDH whereas thioredoxin *f* activates some of the RPP cycle enzymes. Thioredoxin *m* is widely distributed in cyanobacteria (Rowell *et al.*, 1988). Another thioredoxin exists only in cyanobacteria (Alam *et al.*, 1989) and this unusual thioredoxin has a somewhat less negative redox potential (Gleason, 1994).

Cyanobacterial G6PDH is deactivated by an *m*-type thioredoxin (Gleason, 1994). At concentrations as low as 2 nM, the purified thioredoxin from *Anabaena cylindrica* deactivated partially purified *Anabaena variabilis* G6PDH indicating a potential regulatory role for this protein (Cossar *et al.*, 1984). The deactivation required reduced thioredoxin, which was reduced by DTT *in vitro*. Thioredoxin is probably reduced by ferredoxin-thioredoxin oxidoreductase *in vivo* (Yee *et al.*, 1981). The precise mechanism of deactivation may involve reduction of disulphide bond in the enzyme, or thiol-disulphide exchange in the enzyme (Anderson *et al.*, 1978).

1.9.4. Regulation in heterocysts

In filamentous heterocystous strains, the compound that donates electrons directly to nitrogenase is usually considered to be reduced ferredoxin in the light (Stewart, 1980). Ferredoxin is reduced by PSI which is the only operative photosystem in heterocysts. However, electron donation to nitrogenase in the dark and partially in the light could be from the OPP cycle. Glycolysis was also suggested to be a pathway which may donate electron to nitrogenase (Schrautemeier *et al.*, 1984). As the first enzyme of the OPP cycle, G6PDH has been widely reported to NADPH supplier in heterocysts (Stewart, 1980; Wolk, 1982; Wolk *et al.*, 1994). G6PDH from heterocysts of *Anabaena cylidrica* was reported to be 6-8 times more active than the enzyme from vegetative cells (Lex and Carr, 1974). Apte *et al.* (1978) reported similar results in terms of enzyme

activity in heterocysts. The higher activity was also confirmed for *Anabaena* sp. PCC7120, where G6PDH from heterocysts exhibited considerably higher activity than that in vegetative cells (Udvardy *et al.*, 1984). In heterocysts, as a requirement for N_2 fixation, reduced thioredoxin, ATP and NADPH are all abundant, at least in the light. These compounds favour G6PDH activity rather than inactive state resulting in a deactivation/inhibition in vegetative cells. In contrast, the enzyme from heterocysts had the same activity both in the light and in the dark at such high levels in *Anabaena cylindrica* (Apte *et al.*, 1978). Cossar *et al.* (1984) suggested two possible reason for this high activity in heterocysts in the light. The first possibility is that there may be insufficient reduced thioredoxin available in heterocysts to inactivate G6PDH in the light. In fact ATP and NADPH are rapidly used by nitrogenase preventing the accumulation of these compounds in heterocysts. The second, glutamine, which is probably the form in which fixed nitrogen is transported from heterocysts, and glucose-6-phosphate, provided from vegetative cells, accumulate in heterocysts. These two compounds, which are positive effectors of G6PDH, serve to override the deactivating effect of reduced thioredoxin.

1.9.5. Molecular biological studies on the G6PDH gene (*zwf*)

Molecular cloning and sequencing studies on the G6PDH genes (*zwf*) from various organisms have been recently reported. The *Zymomonas mobilis zwf* gene was cloned (Barnel *et al.*, 1990) by genetic complementation of a specific defect in glucose metabolism in *Escherichia coli* DF214 (Vinapol *et al*, 1975). Rowley

and Wolf (1991) cloned, sequenced and analysed the expression of the *E. coli* K-12 *zwf* gene. As the first cyanobacterial gene, the *Synechococcus* sp. PCC7942 *zwf* gene was cloned and sequenced by Scanlan *et al.* (1992) probing the chromosomal DNA with the *Z. mobilis zwf* gene. The gene consisted of 1527 nucleotides encoding a polypeptide of 524 amino acids which exhibited 41% identity with the G6PDH of *E. coli*. The enzyme has two cysteine residues which may be involved in light modulation of the enzyme activity. The *Nostoc* sp. ATCC29133 *zwf* gene was also cloned by probing with part of the *Synechococcus* sp. PCC7942 *zwf* gene (Summers *et al.*, 1995a). The deduced amino acid sequences of these two cyanobacterial G6PDH are 81% identical with each other, and the *Nostoc* sp. ATCC29133 enzyme has three cysteine residues, two of which are conserved in the *Synechococcus* sp. PCC7942 enzyme, which supports the involvement of cysteine residues in the light modulation of the enzyme. The *zwf* gene of potato (*Solanum tuberosum* L.) was sequenced on a cDNA clone (Graeve *et al.*, 1994). This gene displayed all the features of a typical eukaryotic gene. The deduced amino acid sequence was more similar to that of the other cytosolic isoforms (53.1% identity to the rat gene and 51.8% identity to yeast) rather than that of *Synechococcus* sp. PCC7942 (37% identity). A *zwf* mutant of *Synechococcus* sp. PCC7942 was produced by interposon mutagenesis and it showed no detectable G6PDH activity as assessed by both enzyme assay and immunoblot analysis (Scanlan *et al.*, 1995).

1.10. The aims of this study

This study has been carried out on two cyanobacterial strains, which are *Anabaena* sp. PCC7120 and *Synechococcus* sp. PCC7942. *Anabaena* sp. PCC7120 is a heterocystous, filamentous, freshwater strain. *Synechococcus* sp. PCC7942 is a unicellular, freshwater strain. Both of the strains, being prokaryotic organisms, are genetically amenable and are frequently used organisms for genetic analysis. The cyanobacteria are focused on not only due to their ability to operate both aerobic N_2 fixation and photosynthesis, but also as a model organism for analysing the photosynthetic properties of higher plants, and also for their potential for commercial and biotechnological applications (Kerby and Stewart, 1988; Kerby and Rowell, 1992; Richmond, 1992). One of the most important regulatory targets of cyanobacterial carbon metabolism is the G6PDH which plays an important role in transitions between light and dark carbon metabolism and in heterocysts as an electron donor to nitrogenase. Intensive studies on the regulation and light modulation of the enzyme, on one hand, were mostly based on *in vitro* analysis, and in some cases the results conflicted with each other. On the other hand, reduced thioredoxin was reported to be a light modulator of the enzyme, reducing an enzyme disulphide bond which occurred between cysteine residues in the enzyme, but there was no genetic data accumulated to judge whether the cysteine residues that were required for the formation of a disulphide bond existed in various cyanobacterial strains. As a result, it is of value to analyse the *zwf* regions of different strains to find the primary structure of the gene.

The essential aim of this study is to analyse the molecular genetic aspects of G6PDH from the cyanobacteria *Anabaena* sp. PCC7120 and *Synechococcus* sp. PCC7942. Accordingly, the aims are to sequence the *zwf* region of *Anabaena* sp. PCC7120 chromosome to find the nucleotide and amino acid sequences of the *zwf* gene. Gene arrangement in the *zwf* region of *Anabaena* sp. PCC7120 is also one of the aims. The gene organisation in this region will be compared with the other strains. If any gene exists in this region that may be involved in the regulation of G6PDH activity, these genes will be further analysed by interposon mutagenesis to characterise how the gene is involved in the regulation of G6PDH or what is the function of the gene. A deletion-insertion mutagenesis of the *zwf* gene of *Anabaena* sp. PCC7120 will be attempted. This would provide insights into the role of G6PDH both in vegetative cells and in heterocysts. In addition, such a mutant could be used for site-directed mutagenesis of cysteine residues.

Chapter 2

Materials and Methods

2.1. Bacterial strains

2.1.1. Cyanobacterial strains

The cyanoacterial strains used in this study are shown in Table 2.1.

Table 2.1 Cyanobacterial strains used in this study.

Anabaena sp. PCC7120	wild type	Rippka *et al.*, 1979
Synechococcus sp. PCC7942	wild type	Grigorieva and Shestakov, 1976
Anabaena sp. HK 28	*tal*: : pAUG401; Suc^s; Sm^r/Spc^r; Em^r	In this study
Anabaena sp. HK 29	*tal*: :Ω; Suc^r; Sm^r/Spc^r	In this study
Synechococcus sp. HK 52	*opc*A: :pSDG4; Sm^r/Spc^r; A^{pr}	In this study
Synechococcus sp. HK 55	*opc*A: :Ω; Sm^r/Spc^r	In this study

2.1.2. *Escherichia coli* strains

Escherichia. coli strains used are listed in Table 2.2.

Table 2.2. *E. coli* strains used in this study

Strain	Phenotype	Reference
TG1	K12, Δ(*lac pro*) *sup*E*thi*-1*hsd*D5 F`*tra*D36*pro*A^+B^+ *lac*I^q,ZΔM15	Ward and Howe, (1989)
MC1061	*ara*D139, Δ(*ara-leu*)7697, Δ(*lac*)X74, *gal*U^- *gal*K^-, *hsd*R^- *rps*L	Wertman *et al.*, (1986).

S17.1	RP4 2-Tc: :Mu-Km: :Tn7 Tc^r, Sm^r, Km^s, *res*$^-$, *mod*$^+$	Simon *et al.* (1983)
HB101	F^- Δ(*gpt-pro*A)62 *leu sup*E44 *ara*14 *gal*K2 *lac*Y1 Δ (*mcr*C-*mrr*) *rps*L20 (*str*)*xyl*-5 *mtl*-1 *rec*A13	Maniatis *et al.*, (1982)
AEE101	*pro met*	D. Hodgson (per. comm.)
DF214	*zwf*Δ, pgi-Mu	Vinopal *et al.*, (1975)

2.2. Plasmids used in this study

Plasmids used in this study is listed in Table 2.3.

Table 2.3. Plasmids and phage used in this study

Plasmid/Phage	Characterisation	Resistance	Reference
pUC19	Cloning vector	Ap^r	Yanisch-Peron *et al.*, (1985)
pBR325	Cloning vector	Ap^r, Cm^r, Tc^r	Ptentki *et al.*, (1981)
M13mp18	Cloning vector	Ap^r	Yanisch-Peron *et al.*, (1985)
M13mp19	Cloning vector	Ap^r	Yanisch-Peron *et al.*, (1985)
pRL271	Cargo plasmid	*sac*B, Cm^r, Em^r	(Black *et al.*, 1993)
pRL6	Shuttle vector	Cm^r, Nm^r	Elhai and Wolk, 1988
pRL528	Helper plasmid	Cm^r	Elhai and Wolk, 1988
pUC303	Shuttle vector	Sm^r, Cm^r	Kuhlemeier *et al.* (1983)
RP4	Conjugative plasmid	Ap^r, Km^r, Tc^r	Thomas and Smith (1987)
pHP45Ω	Cloning vector	Ap^r Sm^r/Spc^r	Prentki and Krisch (1984)

pAG75	pBR325: :7.5 kb *Hin*dIII cut *Anabaena* sp. PCC7120 chromosomal DNA	Cm^r, Ap^r	J. Scanlan (per. comm.)
pAG19	pBR325: :1.9 kb *Hin*dIII/*Cla*I cut *Anabaena* sp. PCC7120 chromosomal DNA	Cm^r, Ap^r	In this study
pAG24	pUC19: :2.4kb *Hin*dIII/*Hpa*I cut *Anabaena* sp. PCC7120 chromosomal DNA	Ap^r	In this study
pAG25	pUC19EX: :2.4kb *Hin*dIII/*Hpa*I cut *Anabaena* sp. PCC7120 chromosomal DNA	Ap^r	In this study
pAG39	pUC19: :3.9 kb *Hin*dIII/*Xba*I cut *Anabaena* sp. PCC7120 chromosomal DNA	Ap^r	In this study
pAG35	pUC19: :3.5 kb *Eco*RI cut *Anabaena* sp. PCC7120 chromosomal DNA	Ap^r	In this study
pAUG20	pAG75: : Ωinto *Hpa*I site (the orientation of Ω is same as *tal* and *zwf*).	Ap^r, Cm^r, Sm^r, Spc^r, Tc^r	In this study
pAUG401	*Nru*I digested pRL271: : 5.2 kb *Eco*RI digested interrupted *tal.* reverse orientation with Cm^r	*sacB*, Cm^r, Em^r, Sm^r, Spc^r	In this study
pDB	*Sal*I digested pUC19: :6 kb *Sal*I digested *Synechococcus* sp. PCC7942 chromosomal DNA	Ap^r	J. Scanlan (per. comm.)

pSDG2	*SalI*/*Pst*I digested pUC19: : *SalI*/*Pst*I digested *Synechococcus* sp. PCC7942 chromosomal DNA	Ap^r	In this study
pSDG4	pSDG2: :Ω into *BglI*	Ap^r Sm^r/Spc^r	In this study

2.3 Chemicals and biochemical buffers

Chemicals used in this study were all of analytical grade and were purchased from BDH, Sigma or Fisons. Commonly used buffers were as following (Maniatis *et al.*, 1982):

Tris Borate (TBE) 10x:

Tris base	108 g
Boric acid	55 g
0.5 M EDTA (pH 8.0)	40 ml

Final volume was brought to 1000 ml.

0.5 M EDTA pH 8.0:

186.1 g of disodium ethylenediaminetetraacetete was added to 800 ml water and dissolved before adjusting the pH to 8.0 with NaOH pellets. The final volume was brought to 1000 ml.

1 M Tris:

121.1 g Tris base was added to 800 ml water and dissolved. The pH was adjusted to the desired value (7.4, 7.6 and 8.0) by the addition of concentrated HCl.

2.4 Culture media and growth of bacterial strains

Anabaena sp. PCC7120 and *Synechococcus* sp. PCC7942 were grown at 30°C in BG11 medium (Rippka, 1988) in an illuminated orbital shaker under continuous light at an intensity of 60-70 $\mu E\ m^{-2}\ s^{-1}$. Cyanobacterial strains were maintained on BG11 medium solidified with 1.5% (w/v) agar. Contamination was checked by plating stock cultures onto BG11 medium supplemented with 10 mM glucose and 0.2% (w/v) yeast extract, and incubating at 30°C. The culture was taken to be axenic when no bacterial or fungal growth was observed on the plates after 3-4 days.

E. coli strains were grown in liquid media in Luria-Bertani (LB) (Maniatis *et al.*, 1982) or 2x YT (Bankier *et al.*, 1987) medium at 37°C overnight or by growing up to a desired optical density in a Gallenkamp shaking incubator at 200 rpm. The cultures were grown overnight on the appropriate agar dish and stored at 4°C until required in 3-4 weeks. 50% (v/v) glycerol was added to liquid cultures and aliquots were stored at -20°C for long term storage. Appropriate amounts of antibiotics were added to media where required. To make solid media, 1.5 % (w/v) Bactoagar (Difco) was added to the desired medium. The medium was autoclaved at 121°C and 15 psi for 15 mins.

BG11 medium (Rippka, 1988):

$NaNO_3$	1.500 g
$K_2HPO_4.3H_2O$	0.040 g
$MgSO_4.7H_2O$	0.075 g
$CaCl_2.2H_2O$	0.036 g
Citric Acid	0.006 g
$FeNH_4$ Citrate	0.006 g
$EDTA.Na_2Mg$	0.001 g

The volume was brought to 1000 ml after adding 1 ml of the trace elements and the pH was adjusted to 7.4 before the medium was autoclaved.

Trace elements:

H_3BO_3	2.860 g
$MnCl_2.4H_2O$	1.810 g
$ZnSO_4.7H_2O$	0.222 g
$NaMoO_4.H_2O$	0.391 g
$CuSO_4.5H_2O$	0.079 g
$Co(NO_3)_2.6H_2O$	0.049 g

The volume was brought up to 1000 ml.

Solid medium was prepared by making 2x BG11 and 2x 1.5% (w/v) Bactoagar in distilled water separately. After autoclaving at 121°C and 15 psi for 15 mins and cooling down to 50°C the stocks were mixed and poured in Petri dishes or in tissue-culture tubes. Before pouring appropriate amounts of antibiotics, where required, were added to the media.

Escherichia coli strains were grown either in Luria-Bertani (LB) medium (Maniatis *et al.*, 1982) or 2x YT medium. Strain TG1 was grown only in 2x YT

medium or in M9 minimal medium (Maniatis *et al.*, 1982) supplemented with thiamine to a final concentration of 2 μg/ml when competent cells were being prepared. Solid medium was prepared by adding Bactoagar to a final concentration of 1.5% (w/v).

2x YT medium (Bankier *et al.*, 1987):

Bactotryptone	10.0 g
Yeast extract	10.0 g
NaCl	5.0 g

The volume was brought up to 1000 ml by adding distilled water and sterilised by autoclaving.

Top Agar (Bankier *et al.*, 1987):

Bactotryptone	10 g
NaCl	8 g
Bactoagar	8 g

The volume was brought up to 1000 ml and autoclaved in 100 ml bottles.

M63 Minimal Medium (Miller, 1992):

For M63 minimal medium 10x stock solution of the salts was made as following:

10x M63 salts

KH_2PO_4	30g
K_2HPO_4	70g
$(NH_4)_2SO_4$	20g

After the salts were dissolved, 5 ml of 1 mg ml^{-1} $FeSO_4$ (final concentration 5 μg ml^{-1}) was added and the volume was brought to 1000 ml with distilled water.

100 ml 1x stocks were made by adding 10 ml of the 10x salts, supplemented with 0.1 ml 1M $MgSO_4$, 0.2 ml 50 mg ml^{-1} thiamine and 1 ml 20% (w/v) glucose. The volume was brought to 100 ml with distilled water. For M63 plates, 2x 1.5% (w/v) agar and 2x M63 salts were autoclaved separately. After autoclaving, the stocks were mixed and supplemented with $MgSO_4$, thiamine and glucose at the same final concentrations as above.

2.5 Preparation and storage of antibiotics

Antibiotics were used at the following concentrations (Maniatis *et al.*, 1982; Cai and Wolk, 1990) unless otherwise stated. The stocks were stored at -20°C.

Antibiotic	Stock concentration (mg ml^{-1})	Final concentration for *E. coli*(μg ml^{-1})
Ampicillin (Ap)	25	50
Chloramphenicol (Cm)	30	30
Streptomycin (Sm)	30	25
Spectinomycin (Spc)	5	50
Kanamycin (Km)	5	25
Neomycin (Nm)	5	50

When applicable, *Anabaena* sp. PCC7120 cells were selected against streptomycin/spectinomycin at a final concentration of 1.5 μg ml^{-1} in liquid medium and 2.5 μg ml^{-1} in solid medium while *Synechococcus* sp. PCC7942 cells

were selected against a final concentration of 10 μg ml^{-1} streptomycin and 1 μg ml^{-1} ampicillin.

2.6 Preparation of chromosomal DNA from cyanobacteria

For Southern blot hybridisation analyses, cyanobacterial chromosomal DNA was purified on a small scale.

2.6.1 Small scale chromosomal DNA purification from *Anabaena* sp. PCC7120

Total chromosomal DNA from *Anabaena* sp. PCC7120 was purified using the method described by Golden *et al.* (1985) as modified by Cai and Wolk (1990). Cells of *Anabaena* sp. PCC7120 in the mid-log phase of growth were harvested from a 20 ml liquid culture and suspended in a final volume of 400 μl in a microcentrifuge tube with 10 mM Tris-0.1 mM EDTA, pH 7.5. Then 150 μl of sterile glass beads (212-300 microns in size, from Sigma), 20 μl of 10% SDS, and 450 μl of phenol-chloroform were added. The mixture was subjected to a cycle of vigorous vortexing for 1 min followed by cooling on ice for 1 min for a total of four times. The resulting suspension was centrifuged at 12200 x g for 15 min, and the clear supernatant solution was transferred to a new microcentrifuge tube and phenol-chloroform extracted, and its DNA was ethanol precipitated. The DNA then resuspended in TE buffer pH 7.5 and stored at -20°C.

2.6.2 Small scale chromosomal DNA purification from *Synechococcus* sp. PCC7942

Small scale chromosomal DNA extraction from *Synechococcus* sp. PCC7942 was based on a method described by Lind *et al.* (1985).

20 ml of a *Synechococcus* sp. PCC7942 culture in late logarithmic phase was harvested and resuspended in 1 ml of solution 1 in an 1.5 ml Eppendorf tube then lysozyme was added to a final concentration of 10 mg ml^{-1}.

Solution 1:

Tris-HCl	0.05 M pH 8.0
Sucrose	20% (w/v)

After leaving the mixture at 37°C for one hour, 16 µl 30% Sarkosyl and 20 µl Proteinase K at a concentration of 5 mg ml^{-1} were added and kept at 65°C for two hours. Then equal volumes of phenol:chloroform and chloroform:isoamyl alcohol (24:1 v/v) were added respectively. The layers were separated by centrifuging for 5 mins in an Eppendorf centrifuge at 12000 x g. The chromosomal DNA was precipitated by adding 0.1 volume of 3 M sodium acetate pH 5.6 and 2 volumes of ethanol. The tube was left at -20°C overnight and centrifuged at 12000 x g in a microcentrifuge at room temperature for 15 mins. The pellet was vacuum dried and resuspended in a total volume of 100 µl TE buffer pH 8.0.

2.7 Preparation of plasmid DNA

2.7.1 Small scale plasmid DNA preparation from *Escherichia coli*

Small amounts of plasmid DNA for identification or screening of recombinant plasmids or restriction digestion and ligation were prepared by a modification of the alkaline lysis method (Maniatis *et al.*, 1982). The method used was as follows:

Solution I:

Glucose	50 mM
EDTA	10 mM
Tris-HCl	25 mM

Solution II:

NaOH	0.2 M
SDS	1.0 % (w/v)

Solution III:

Potassium acetate 5 M	60.0 ml
Glacial acetic acid	11.5 ml
Distilled water	28.5 ml

The pH of this solution was approximately 4.8.

A single colony inoculated in 10 ml LB or 2x YT medium was grown at 37°C overnight in a shaking incubator. Appropriate amounts of antibiotics were added at the selective concentrations. 1.5 ml of these cells were transferred in an 1.5 ml Eppendorf tube and harvested in an MSE micro centrifuge at 12000 x g for

5 mins and the medium was removed from the pellet by aspiration. The pellet was then resuspended in 150 µl of freshly made solution I containing 50 mM glucose, 10 mM EDTA, 25 mM Tris-HCl pH 8.0 and 5 mg ml^{-1} lysozyme and the tube was placed on ice for 5 mins. 200 µl of freshly made solution II containing 0.2 M NaOH and 1 % SDS was added and the contents were mixed by gentle inversion a few times before being placed on ice for 5 mins. Before gently vortexing 150 µl of ice-cold potassium acetate pH 4.8 (60 ml 5 M potassium acetate, 11.5 ml glacial acetic acid and 28.5 ml distilled water) was added to the tube and placed on ice for 5 mins. The contents were centrifuged at 12000 x g in a microcentrifuge for 5 mins and the supernatant was removed to a clean 1.5 ml Eppendorf tube. The supernatant was extracted using phenol:chloroform and then DNA was precipitated by the addition of ethanol. After incubating at -20°C for 15 mins the tube was centrifuged at 12000 x g for 20 mins. The DNA pellet was dried under vacuum for 10 mins and resuspended in 30 µl of TE buffer pH 8.0. 5-10 µl of the DNA solution was digested with appropriate restriction endonucleases. The remainder was stored at -20°C.

2.8 Quantitation of DNA

Double stranded DNA concentrations were measured using the spectrophotometric determination method described by Maniatis *et al.* (1982).

DNA which was contaminated with RNA was measured using ethidium bromide fluorescent quantitation. DNA to be measured was digested with a restriction enzyme, run on a 0.7 % agarose gel including ethidium bromide along

with a known amount of marker DNA. These two samples were compared with each other in respect to the fluorescence of ethidium bromide under UV illumination. Commercially supplied phage lambda DNA and pUC19 DNA were used as markers.

2.9 Phenol:chloroform extraction

To remove any protein contaminants from DNA solutions, phenol:chloroform (w/v) extraction was performed as follows (Maniatis *et al.*, 1982):

An equal volume of phenol:chloroform (1:1, w/v) was added to the DNA sample and mixed by either gently vortexing or shaking until an emulsion formed. After 5 mins the tube was centrifuged at 12000 x g for 4 mins. The top layer was removed carefully to a fresh tube avoiding any phenol contamination and then the DNA was ethanol precipitated.

2.10 Ethanol precipitation

Ethanol precipitation was used to concentrate the DNA from a diluted solution (Maniatis *et al.*, 1982). The DNA was precipitated by addition of 1/10 volume of 3 M sodium acetate pH 5.6 and 2 volumes of absolute ethanol. After mixing, the solution was stored at either-20°C for one hour or at -70°C for 15 mins. The DNA was then pelleted by centrifugation at 12000 x g for 20 mins in an Eppendorf centrifuge at 4°C. The pellet was washed with 70% ethanol to remove any of the remaining salt, dried in a vacuum desiccator and finally dissolved in the desired volume of TE buffer pH 8.0.

2.11 Restriction endonuclease digestion of DNA

Restriction endonucleases were purchased from Amersham International, Gibco-BRL and Pharmacia. Restriction digestion reactions were carried out according to the manufacturer's recommendations. The restriction buffers used were as supplied by manufacturers.

All restriction digestions were carried out at 37°C for 1 hour unless otherwise recommended by the manufacturer. *Hin*dIII or *Pst*I digested lambda DNA were used as molecular size markers to determine the size of digested DNA fragments.

2.12 Alkaline Phosphatase treatment of DNA

Alkaline phosphatase (AP) was purchased from Boehringer Mannheim. Prior to ligation vector DNA was treated with AP to remove the 5' phosphate group. This removal prevents self-ligation and increases the proportion of recombinant molecules obtained.

To a 20 µl volume of digestion mixture, 2.5 µl 10x AP buffer and 2 µl AP were added and incubated at 37°C for 30 mins. A further 2 µl AP was then added and incubation continued at 37°C for 30 min. The DNA was then purified by either phenol/chloroform extraction and ethanol precipitation or by using the Geneclean II Kit purchased from Bio 101 Inc. or Qiaex Gel Extraction Kit from

Qiagen according to the manufacturer's recommendations. The DNA was either used for ligation immediately or stored at -20°C until required.

2.13 Agarose gel electrophoresis

DNA fragments were separated in 0.7% (w/v) agarose (type II, Sigma) gels in TBE. For purification of a fragment after restriction endonuclease digestion, 1% agarose gels were routinely used.

The gel was prepared by boiling agarose in 1x TBE until all the agarose particles were melted. The solution was allowed to cool down to 50°C and ethidium bromide from a 10 mg ml^{-1} stock solution kept in dark was added to a final concentration of 0.5 μg ml^{-1}. Samples for analysis were prepared by adding 0.1 volume of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol and 15% Ficoll type 400 in water). Large gels were run at 70 mA overnight, and small gels at 80 mA for 1 to 2 hours.

2.14 Visualisation of DNA in agarose gels

DNA was visualised by fluorescence of ethidium bromide bound to the DNA upon irradiation with ultraviolet light (Maniatis *et al.*, 1982). Gels were examined over a longwave ultraviolet light box or photographed on a short wave ultraviolet light box using Polaroid P665 positive/negative film.

2.15 Purification of DNA from agarose gels

To recover a DNA sample from agarose gel, the region containing the sample was excised in as small a volume as possible. Than the DNA was recovered using either Geneclean II Kit from BIO 101 or Qiaex Gel Extraction Kit from Qiagen according to manufacturer's recommendations.

2.16 Ligation of DNA

T4 DNA ligase was purchased from Gibco-BRL and used according to the manufacturer's recommendations. 5x ligation buffer was supplied by the manufacturer. All ligations were carried out overnight at 15°C (Maniatis *et al.*, 1982). To optimise the ligation conditions, various concentrations of vector and insert DNA were used. Insert DNA concentrations were increased when blunt ended restriction sites were used for ligation.

2.17 Transformation of bacterial strains

2.17.1 Transformation of *Escherichia coli*

Transformation of *Escherichia coli* was carried out essentially using the calcium chloride method as described by Maniatis *et al.*, (1982) with some modifications as follows:

10 ml culture medium was inoculated with a single colony of the desired *E. coli* strain and incubated at 37°C in a shaking incubator at 250 rpm overnight. 50 µl of this overnight culture was then inoculated into 50 ml nutrient agar in a

250 ml conical flask and grown until the culture was in late exponential phase determined by measuring the optical density of the culture at 600 nm. Cultures typically had an OD_{600} of 0.4 to 0.7. The cells were harvested by centrifugation at 1200 x g at 4°C for 5 mins in Centra MP4R type (IEC) centrifuge after incubating the cells on ice for 20 mins. The cells were resuspended in 10 ml of 0.1 M $MgSO_4$, incubated on ice for 20 mins and then centrifuged as before. 10 ml of 0.1 M $CaCl_2$ was added to the cells, resuspended, incubated on ice for 20 mins and centrifuged as before. The cells were then resuspended in 1 ml of 0.1 M $CaCl_2$ and incubated on ice for one hour before mixing with the plasmid DNA. To a desired amount of DNA, 100 µl of the competent cells were added, briefly vortexed and incubated on ice for 45 mins. After incubating at 42°C for 3 mins the transformation mixture was plated out onto either LB or 2x YT agar plates containing selective antibiotics at required concentration and incubated at 37°C overnight.

2.17.2 Transformation of *Synechococcus* sp. PCC7942

Transformation of *Synechococcus* sp. PCC7942 was performed as described by Porter (1988). A 3-4 day grown culture was harvested and washed once and then resuspended in BG11 medium to give a final concentration of $1x10^9$ cells ml^{-1}. 10 µl DNA was added to 0.2 ml cells in clear polypropylene tubes and incubated in the light for 30 mins. The cells were then plated onto BG11 medium under non-selective conditions and incubated for 16-18 hours in the light. The antibiotic was then added underneath the agar at the desired concentration and

transformants were detected after 4-6 days. Transformant colonies were restreaked onto selective media.

2.17.3 Transformation of *Anabaena* sp. PCC7120 by electroporation

The method for transformation of *Anabaena* sp. PCC7120 by electroporation was based on the methods described by Thiel and Poo (1989) and Moser *et al.* (1993).

A 100 ml culture in BG11 medium was grown to a concentration of $2x10^7$ cell ml^{-1} in a 250 ml conical flask. An appropriate amount of culture was harvested to obtain a final concentration of $1x10^9$ cells ml^{-1}. Cells were harvested by centrifugation at 4000 x g for 15 mins at room temperature in a Centra MP4R type (IEC) centrifuge. The cells were washed twice with 10 ml of 1 mM HEPES buffer pH 7.2 and resuspended in a total volume of 1 ml. The transforming DNA was purified on a small scale and resuspended in 1 mM HEPES buffer. 40 µl of *Anabaena* sp. PCC7120 cells and the transforming DNA both in 1 mM HEPES were placed on the edge of an electroporation tube (Invitrogene Corporation) and mixed thoroughly by a tip. The tube was then given a single pulse (200Ω, 2.5 volts, 25 µF and time constant 4.5) in a BIO-RAD Gene Pulser. Immediately after the pulse, cells were washed with 5 ml BG11 medium and transferred into 20 ml of nonselective BG11 medium. After growing for 24 hours under low light intensity at 30°C with shaking at 175 rpm, cells were harvested, resuspended in 200 µl BG11 and plated on BG11 agar plates with selective concentrations of antibiotics. The plates were left at 30°C under a low light intensity until the

exconjugant colonies were grown. The exconjugant colonies were then restreaked on BG11 plates with selective concentrations of antibiotics to confirm stable antibiotic resistance.

2.18 Conjugation

Since no transformation systems exist for filamentous cyanobacterial strains to date, conjugation was performed for introducing recombinant plasmid DNA into *Anabaena* sp. PCC7120. The method used was based on the methods described by Wolk *et al.*(1984), Elhai and Wolk (1988) and McFarlane *et al.* (1987).

Escherichia coli S17.1 was used as the donor parent of biparental mating. The strain possesses a broad host range conjugative plasmid RP-4 integrated within and immobilised in the chromosome. Helper plasmid pDS4101, and another helper plasmid pRL528 with genes for *Ava*I and *Eco*47II methylation and the recombinant shuttle plasmid pAUG401 carrying an insertional mutation of the *tal* gene were transformed into this strain by using the $CaCl_2$ transformation method (see Section 2.11). The cells were grown overnight in 10 ml LB medium containing appropriate antibiotics. The cells were then diluted in 10 ml LB medium at a proportion of 0.25:10 (v/v) and allowed to grow until exponential growth phase was reached (OD_{600} = 0.5). 1.5 ml of the cells were harvested by centrifuging for 5 mins at 4000 x g in an EC Centra MP4R type bench centrifuge. The pellets were then washed with LB medium to remove any remaining antibiotics and finally resuspended in 60 µl LB.

Anabaena sp. PCC7120 cells were grown in BG11 medium for 4-5 days at 30°C. After vortexing a few times for partial fragmentation, 20 ml of the cells were harvested by centrifugation at 4000 x g in an EC Centra MP4R type bench centrifuge and washed with BG11. To obtain a 20 times concentration, the cells were finally resuspended in a total volume of 1 ml BG11. This stock was used for a series of dilutions of 1:1, 1:10 and 1:100.

Whatman cellulose acetate membrane filter (25 mm in diameter) with 0.45 μm pore size were autoclaved by wrapping in aluminium foil, transferred to a sterile Petri dish and dried under sterile conditions. The filter was placed on a metal disk with pores on the top of a filtration system. The *Anabaena* sp. PCC7120 and *E. coli* strain S17.1 cells prepared for mating as described above were mixed in an Eppendorf tube and poured onto the filter and then filtered using water pressure until the medium was completely filtered so that the cells contacted each other for performing conjugation. Different combinations of mating mixtures were designed to obtain the most efficient mating. This procedure is called biparental mating.

The triparental mating procedure for conjugation was also performed using two *E. coli* strains, which were HB101 and AEE101, and *Anabaena* sp. PCC7120 as parents. *E. coli* strain AEE101 contained a broad host range conjugative plasmid RP-4. The strain HB101 was transformed with helper plasmid pRL528 and recombinant shuttle plasmid pAUG401. All three parental strains were

prepared for mating as described above. Two of the *E. coli* strains were mixed in an Eppendorf tube 30 min before mating. 5 µl spots of each dilution of *Anabaena* sp. PCC7120 were applied to an empty sterile Petri dish. To each spot, a 5 µl of the *E. coli* mixture was applied and mixed thoroughly. 2 µl from each mixture was taken and applied to the filter as spots.

The filters with mating mixture were placed on the BG11 agar plates supplemented with 5% (v/v) LB. Then the plates were sealed with parafilm, covered with tissue and incubated at 20C for 4 hours to give time for conjugation. Eventually the filters were transferred onto BG11 agar plates and allowed to grow for 24 hours under light at 30°C. This step allows the potential conjugants to express antibiotic resistance genes. After growing on non-selective plates the filters were transferred onto BG11 agar plates supplemented with 2.5 µl spectinomycin and streptomycin and incubated under light at 30°C for two weeks. When green colonies were seen on the filters those colonies were patched onto BG11 agar plates with antibiotics, incubated until well grown colonies were observed. Some of the single colonies were then transferred onto BG11 agar plates supplemented with antibiotics and 5% sucrose to select double recombinants. The colonies grown on sucrose dish were taken as double recombinants and the others as single recombinants.

2.19 M13 cloning and sequencing techniques

The dideoxy chain termination DNA sequencing method of Sanger *et al.* (1977) was used for all sequencing studies. To sequence the *zwf* region of

Anabaena sp. PCC7120, the DNA fragments of interest were cloned into M13mp18 and M13mp19 to yield single-stranded templates. All subcloning into M13 and sequencing procedures were performed according to the methods described by Bankier *et al.* (1987), Messing and Bankier (1989) and Bankier and Barrell (1989). A kit for DNA Sequencing with Sequenase T7 DNA polymerase (USB) was used.

2.19.1 Generation of directed clones

Bacteriophage M13 was used to generate directed clones. To obtain sequences from the both ends the fragment was cloned into the multiple cloning site of M13. With the fragments digested with two different restriction enzymes M13mp18 and mp19, which are derivatives of M13 phage, were used. Fragments digested with a single restriction enzyme were cloned only into M13mp18. The methods for preparations of vector and fragment and ligation were as previously described (Sections 2.11, 2.12, 2.15, 2.16).

2.19.2 Generation of random clones

The DNA to be subcloned was self ligated at 15°C overnight before sonication to yield a random array of fragments. The self ligated DNA was sonicated twice for 60 secs at full power in 1.5 ml Eppendorf tube using a W-380 type cup-horn sonicator (sonic processor) from Heat Systems-Ultrasonics Inc with a water bath. The water was changed each time to prevent heating and the DNA solution in the tube was centrifuged briefly to collect the sample at the bottom of the tube. After running a small aliquot of sonicated fragment on a minigel to find

out whether the sonication produced suitable sizes, 20 µl of the fragments were end-filled using 2 µl of 2.5 mM deoxyribonucleotide triphosphates, 2.5 units of Klenow large fragments of DNA polymerase and 2.5 units of T4 DNA polymerase (purchased from GIBCO BRL). These end-filled fragments were run on a minigel. The gel slice containing the fragments from 200 base pairs (bp) to 1000 bp in size was excised and then the fragments were purified as described before (Sections 2.16 and 2.11). The end filled fragments were cloned into the *Sma*I site of M13mp18.

2.19.3 Transfection of M13 DNA into *Escherichia coli* strain TG1

E. coli strain TG1 cells were prepared for transfection using the method as described before (section 2.17.1). 150 µl of competent cells were used for each ligation reaction containing either directed or random clones of M13 DNA mixed gently and placed on ice for 30 mins. After 3 mins heat shock at 42°C, the cells were mixed with top agar and overlayed onto 2x YT medium and incubated at 37° C overnight.

2.19.4 Selection of recombinant M13 clones

Blue-white selection of colonies, where possible, was carried out by addition of 25 µl of 2 % X GAL and 2.5 % IPTG in top agar before mixing it with the transfected cells. If M13 DNA was digested by using more than one enzyme, and blue-white selection was not applicable, then all plaques were isolated and sequenced.

2.19.5 Preparation of template DNA

20 % PEG:

Polyethylene glycol (6000 M_r)	20 g
Sodium chloride	14.6 g

The volume was brought to 100 ml with water and the solution was sterilised by filtration.

Single stranded M13 DNA was purified from *E. coli* strain TG1 isolates according to the method of Bankier *et al.*, (1987). The plaques from an overnight grown *E. coli* TG1 lawn were transferred into 1.5 ml of a 1:100 diluted overnight culture of *E. coli* strain TG1 culture in 2x YT medium in a glass test tube by using a sterile toothpick. The tubes were incubated at 37°C for 5 hours in a 64 Environmental Incubator Shaker from New Brunswick Sci. shaking at 300 rpm. Then the individual cultures were transferred into 1.5 ml Eppendorf tubes and centrifuged for 5 min. After transferring the supernatant into new tubes, 200 µl of 20 % PEG was added and mixed by inverting the tube. The tubes were incubated at room temperature for 15 mins and then centrifuged at 12000 x g in a microfuge for 5 mins. The supernatant was removed, the tubes centrifuged again briefly and any remains of water and PEG were removed using an Eppendorf tip. 100 µl TE pH 8.0 was added to each tube and DNA dissolved before adding 100 µl phenol:chloroform. The tubes were mixed well, left at room temperature for 5 mins and then centrifuged as above for 2 mins after mixing again. 90 µl of top layer was removed into a clean tube before adding 0.1 volume of 3 M sodium acetate and 2 volumes of ethanol. The tubes were then mixed briefly and left at

-20°C overnight. Then the tubes were centrifuged at 12000 x g in a micro centrifuge and the supernatant was removed using a needle. After washing with 1 ml of 90 % ethanol and centrifuging for 5 mins, the pellet was dried in a vacuum desiccator for 10 mins. Finally, the pellet was dissolved in 30 µl sterile water , 4 to 7 µl of it was used for a sequencing reaction as template. The remaining template DNA was stored at -20°C.

2.19.6 Preparation of double stranded template DNA

The double stranded DNA to be sequenced was purified by using the small scale plasmid purification method. 25 µl of DNA was repurified by using Geneclean II Kit. 25 µl of glass beads were used and the DNA was recovered in 50 µl water. For denaturation of DNA the alkaline denaturation method of Haltiner *et al.* (1988) was used. 0.1 volume of 2M NaOH and 2 mM EDTA was added to the DNA and incubated at 37°C for 30 mins. The mixture was neutralised by adding 0.1 volume of 3 M sodium acetate pH 4.5 to 5.5 and the DNA was precipitated with 2 to 4 volumes of ethanol. The pellet was dissolved in 25 µl of water and 7 µl of it was used for a sequencing reaction. The other steps were same as those for single stranded templates.

2.19.7 Sequencing reactions

Sequencing reactions were performed using Sequenase Version 2.0 Sequencing Kit (USB and Amersham Life Science) according to the manufacturers instructions with some modifications. The priming reaction was prepared in a 0.5 ml Eppendorf tube by adding 1 µl of primer (at a final

concentration of 2 ng μl^{-1}), 2 µl of Sequenase buffer, 3 µl water and 4 µl of template DNA. The mixture was centrifuged for 5 secs and heated to 65°C and cooled down to 35°C slowly. When double stranded template or a lower quantity of single stranded DNA was used, no water was added and the template amount was increased to 7 µl. In another 0.5 ml Eppendorf tube, the labelling reaction was made by adding 0.4 µl labelling mix, 1 µl of 0.1 M DTT, 1.6 µl water, 1.78 µl TE pH 8.0, 0.5 µl ^{35}S dATP and 0.22 µl Sequenase version 2.0. The mixture was centrifuged for 5 seconds before transferring it to the priming reaction tube. Meanwhile 2.5 µl of dideoxynucleotides were dispensed onto one side of the wells of a Falcon Flexible Assay Plate (Becton Dickinson). Then 3.6 µl of the mixture of priming-labelling reactions was dispensed onto the other wall of each well. After centrifuging briefly in an MSE Mistral 2000 type centrifuge the reactions were incubated at 37°C for 10 mins and left at -20°C for 10 mins.

2.19.8 Gel electrophoresis

40% acrylamide:

Acrylamide (electrophoresis purity)	380 g
N,N'-methylene bisacrylamide	20 g

The volume was brought to 1000 ml with deionised water. After adding 20 g of mixed bed resin (Amberlite MB604 from BDH) the solution was stirred gently and filtered through Whatman filter paper. The stock was stored in the dark at 4°C.

0.5 x TBE 6% gel mix:

40% acrylamide	75 ml
10 x TBE	25 ml
Urea	230 g

The volume was brought to 500 ml with deionised water, stirred and filtered throught Whatman filter paper. The stock was stored at 4°C for 3-4 weeks.

5.0 x TBE 6% acrylamide mix (200 ml):

40% acrylamide	30 ml
10 x TBE	100 ml
Urea	92g
Bromophenol blue	5 mg

The volume was brought to 200 ml with deionised water, stirred and filtered through Whatman filter paper. The stock was stored for 3-4 weeks at -4°C.

25% AMPS:

0.125 g ammonium persulphate (Sigma) was dissolved in water and used freshly.

Two 55 cm x 20 cm glass plates were used for pouring the gradient gels. The inner sides of plates were cleaned and the inner side of one plate was silenised with 1.5 ml 2% dimethyldichlorosilene solution in 1,1,1-trichloroethane (BDH). The plates were assembled with a 0.5 cm wide 0.35 mm side spacer and 25 mm polyester tape. 45 ml 0.5x TBE gel mix was put in a beaker and 7 ml 5x TBE gel mix in another beaker. Then polymerisation was initiated by adding 80 µl of 25%

AMPS and 80 µl TEMED (N,N,N',N'-tetramethylethylenediamine from Sigma) in 45 ml 0.5 x TBE gel mix, 14 µl of each to 7 ml 5 x TBE gel mix. 8 ml of 0.5 x TBE gel mix was pipetted into a 25 ml pipette before pipetting 7 ml 5 x TBE gel mix with 2-3 air bubbles which allows a continuous gradient from 5% to 0.5%. After pouring this mixture in the assembled plates the gel plates were filled up with the remaining portion of the 0.5 x TBE gel mix. A 36 well comb was placed at the top of the gel and the gel allowed to polymerise for at least two hours.

The DNA in the sequencing reactions was denatured by adding 4 µl of stop solution to the plate followed by incubation at 80°C for 15 mins. 2 µl of the reactions were loaded onto the sequencing gel. TBE was the running buffer and the gels were run at 37 W for 2 to 3 hours using Bio-Rad model 3000 Xi type electrophoresis power supply.

2.19.9 Autoradiography

The sequencing gels were fixed in 10% acetic acid for 15 mins and transferred onto a sheet of 3 mm Whatman paper and dried in a Bio-Rad gel drier at 80°C for 30 mins. The dried gels were placed in X-ray cassettes in contact with Fuji RX-100 X-ray film overnight or for a longer exposure time as required. The X-ray films were developed using Kodak LX-24 developer and fixed using Kodak FX-40 fixer according to manufacturer's instructions.

2.19.10 Computer-assisted sequence analysis

The DNA sequence data was read from the autoradiographs by eye. The data was entered into the GCG (Program Manual for Wisconsin Package, Version 8, September 1994) sequence analysis programs (Devereux *et al.*, 1984). Analyses performed using these programs are described in Section 3.2.4.

2.20 Southern blotting procedure

The DNA was separated in a 0.8 % agarose gel, and then hydrolysed partially in 0.25 M solution of HCl at room temperature for 15 mins shaking on a Luckham R100/TW Rotatest type shaker. The DNA was then denatured in a solution of 0.5 M NaOH and 1.5 M NaCl for 2 x 15 mins before neutralising for 2 x 15 mins in a solution of 1 M Tris-HCl pH 7.4 and 3 M NaCl (Maniatis *et al.*, 1982).

20 x SSC:

NaCl	175.3 g
Sodium citrate	88.2 g

The salts were dissolved in 800 ml water, and the volume was adjusted to 1 l after adjusting the pH to 7.0 with a few drops of 10 N solution of NaOH. The solution was sterilised by autoclaving

20 x SSPE:

NaCl	174.0 g
NaH_2PO_4	27.6 g
EDTA	7.4 g

The salts were dissolved in 800 ml water. pH was adjusted to 7.0 with 10 N solution of NaOH. The volume was brought to 1 l and sterilised by autoclaving.

Southern transfer procedure was performed according to the method described by Maniatis *et al.* (1982). A glass plate which was longer and wider than the gel was placed in a bigger tray. The tray was filled to almost the top of the plate with 20 x SSC as transfer buffer. A pre-soaked 3 mm Whatman platform (wick) was placed on the glass plate so that the tips of the platform reached the bottom of the tray. All air bubbles in the platform were smoothed out with a glass rod. The gel was inverted and placed onto the platform. A Hybond-N nylon membrane (Amersham) of the same size as the gel was washed in 2 x SSC for 5 mins and placed onto the gel. Two pieces of 3 mm Whatman paper with same size as the membrane were placed on the membrane after washing with 2 x SSC and then surrounded with cling film to ensure that the buffer passed through the gel. Tissues were put at the top of the pile with a weight on the top. The blot was left overnight to ensure complete transfer of DNA to the membrane. Then the membrane was washed in 6 x SSC for 5 min at room temperature and dried between 3 mm Whatman paper.

The DNA was cross-linked to the membrane automatically in a UV Stratalinker 2400 type UV illuminator (Stratagene) for one cycle.

2.21 Preparation of radio-labelled DNA probes

The Nick translation method was used to radio-label DNA for use in hybridisation experiments. A commercially available Nick Translation Kit (Boehringer Mannheim) was used to complete the labelling according to recommendations of the manufacturer.

2.22 Hybridisation of Southern blots

Hybridisation of nick-translated DNA with homologous DNA immobilised on filters was carried out by the method described by Maniatis *et. al.* (1982).

2.23 Preparation of supernatant for enzyme assays

Exstracts of *Anabaena* sp. PCC7120 or *Synechococcus* sp. PCC7942 cells were prepared for enzyme assays according to the method described by Schaeffer and Stanier (1978). An extraction buffer was used for preparation of the supernatant. The extraction buffer was 0.05 M Tris-maleate supplemented with 0.1 % (v/v) β-mercaptoethanol and 0.01 M glucose-6-phosphate just before use. The pH was adjusted to 6.5 by adding 0.5 M maleic acid.

1 litre of *Anabaena* sp. strain PCC7120 or *Synechococcus* sp. PCC7942 cultures were harvested by centrifuging at 15000 x g for 10 mins in a Beckman centrifuge, washed once with the extraction buffer and resuspended in the same buffer. The cells were then disrupted by four passages through a French pressure cell operated at 14000 psi. The resulting crude extract was centrifuged at 12000 rpm in an EC Centra MP4R type bench centrifuge and the supernatant was pooled

in sterile Eppendorf tubes. The supernatant was centrifuged to remove membrane fragments at 86000 x g at 4°C for 30 mins. This supernatant was then stored at -20°C until use.

E. coli cells were prepared for assay of G6PDH activity according to the method described by Barnell *et al.*, (1990). To make potassium phosphate buffer 0.05 M K_2HP0_4 and KH_2PO_4 were each made and titrated against each other to obtain a pH of 6.8. 10 ml overnight grown cultures of *E. coli* strains TG1 and DF214 were harvested by centrifuging at 4000 x g at 4°C for 10 mins in a Centra MP4R type bench centrifuge (IEC) and washed with 1 ml of 0.05 M potassium phosphate buffer pH 6.8 in an Eppendorf tube. After sonicating in a Jacops type sonicator at the power level of 14 microns peak to peak for 30 seconds or until the contents of the tube became clear, the crude extract was centrifuged at 12000 x g in a EC Centra MP4R type bench centrifuge and the supernatant was pooled. The supernatant was used as G6PDH enzyme solution and stored at -20°C for approximately one month.

2.24 Enzyme assays

2.24.1 Glucose 6-phosphate dehydrogenase assay

The assay of *Anabaena* sp. PCC7120 and *Synechococcus* sp. PCC7942 G6PDH was performed according to the method described by Schaeffer and Stanier (1978). Tris-maleate buffer pH 7.4 supplemented with 0.1 % (v/v) β-mercaptoethanol and 10 mM $MgSO_4$ was used for the assay. A blank reaction

mixture was examined to check any endogenous NADP reduction along with a test mixture as follows:

	Test	Blank
Assay buffer	720 µl	750 µl
$NADP^+$ (50mM)	30 µl	-
Enzyme solution	100 µl	100 µl
G6P (100 mM)	150 µl	150 µl

G6PDH activity in nondenaturing polyacrylamide gels was detected as described by Scanlan *et al.* (1995). Supernatants (200 µg of protein per track) were loaded onto a gradient native polyacrylamide gel and run at 15 mA overnight (see Sections 2.25.2 and 2.25.3). The gel was then incubated in the dark at 30°C for 2 to 4 h in buffer containing 50 mM Tris-maleate (pH 6.5), 50 mM $MgCl_2$, 50 mM glucose-6-phosphate, 5 mM NADP, 0.5 mM nitroblue tetrazolium, and 0.1 mM phenylmethylsulfonyl fluoride.

G6PDH assays of *E. coli* TG1 and DF214 were performed according to the method described by Hylemon and Phibbs (1972). 0.05 M Tris-HCl buffer pH 8.0 supplemented with 0.1 % (v/v) β-mercaptoethanol and 10 mM $MgSO_4$ was used for assay reaction. The reaction mixtures were as follows:

	Test	Blank
Assay buffer	800 µl	820 µl
$NADP^+$ (50mM)	20 µl	-
Enzyme solution	100 µl	100 µl
G6P (100 mM)	80 µl	80 µl

Both in cyanobacterial and *E. coli* strains, rates of NADPH formation were measured at 340 nm for 3 mins in a PU8700 Series UV/Vis model scanning spectrophotometer (Philips) with a thermostatically controlled cuvette holder

maintained at 37°C for E. coli and 25°C for cyanobacterial strains. One unit of enzyme activity was defined as the formation of 1 μmol product in one min. Specific activity was expressed as units mg^{-1} protein. Protein was determined using the Bio-Rad Protein Assay kit according to manufacturer's instructions.

2.24.2 Transaldolase assay

Transaldolase assays of *Anabaena* sp. PCC7120 were performed according to the method described by Levering and Dijkhuizen (1990). Glyceraldehyde 3-phosphate produced in the reaction with fructose 6-phosphate and erythrose 4-phosphate as substrate is converted to dihydroxyacetone phosphate by triose phosphate isomerase. The formation of this compound is measured with NADH and α-glycerol-phosphate dehydrogenase. The reaction is followed spectrophotometrically by measuring the change in absorbance at 340 nm. The observed rate of NADH oxidation is, under appropriate conditions, proportional to the amount of transaldolase activity. Supernatants were prepared from *Anabaena* sp. PCC7120 as described for G6PDH above. The reagents used were as follows:

i) Tris-EDTA buffer, pH 8.0 (containing 200 mM Tris and 40 mM EDTA),

ii) NADH, 3 mM,

iii) D-Fructose 6-phosphate, 100 mM

iv) D-Erythrose 4-phosphate, 50 mM,

v) α-Glycerol-phosphate dehydrogenase (from rabbit muscle) 35 units ml^{-1} (Sigma) and

vi) Triose phosphate isomerase (from rabbit muscle), 200 units ml^{-1} (Sigma).

The assay mixture in a quartz cuvette (10 mm light path) contained Tris-EDTA buffer, 0.5 ml; NADH, 50 µl; fructose 6-phosphate, 50 µl; α-glycerol-phosphate dehydrogenase, 50 µl; triose phosphate isomerase, 50 µl; supernatant of *Anabaena* sp. PCC7120 as transaldolase solution, 10 µl; and water to adjust volume to 0.99 ml. This mixture was incubated at 30°C in a PU8700 Series UV/Vis model scanning spectrophotometer (Philips) for 5 min to record any endogenous NADH oxidation. Then transaldolase reaction was started by adding 10 µl of erythrose-4-phosphate in another cuvette including the reagents as above. The resulting rate of absorbance decrease at 340 nm, minus the erythrose 4-phosphate independent rate, was taken as the transaldolase activity. One unit of enzyme activity was defined as the formation of 1 µmol product in one min. Specific activity was expressed as units mg^{-1} protein. Protein was determined as described previously.

2.25. Electrophoresis of proteins

2.25.1. SDS discontinuous polyacrylamide gel electrophoresis (SDS disc PAGE)

Solution A: Acrylamide mix:

Acrylamide	60.0 g
Bis-acrylamide	1.6 g

The volume was brought to 200 ml with distilled water.

Solution B: Resolving gel buffer:

Tris base	36.3 g
1M HCl	48.0 ml

The volume was brought to 100 ml with distilled water after adjusting pH to 8.8.

Solution C: Stacking gel buffer:

Tris base	6.0 g
1M HCl	40.0 ml

The volume was brought to 100 ml with distilled water after adjusting pH to 6.8.

The acrylamide solution, resolving gel buffer and stacking gel buffer were filtered through Whatman No. 1 filter paper.

Solution D: Running buffer:

Glycine	8.64 g
Tris	1.8 g
SDS	0.6 g

The volume was brought to 600 ml with distilled water.

Solution E: Ammonium persulfate (AMPS): 75 mg of AMPS was dissolved in 5 ml of H_2O to make 1.5 % of AMPS and used freshly.

F Sample buffer:

Stacking gel buffer, solution B	2.5 ml
SDS (10%)	4.0 ml
β-mercaptoethanol	2.0 ml
Glycerol	2.0 ml

A small amount of tracking dye (bromophenol blue) was added to the sample buffer.

12.5% linear polyacrylamide gels were poured for electrophoresis of denatured proteins. For the resolving gel, 12.5 ml of solution A, 3.75 ml of solution B, 0.3 ml of 10% SDS and 113.3 ml distilled water were put in a conical flask and degassed using a water vacuum pump. Immediately before pouring, 150 µl of 15% AMPS and 15 µl of TEMED were added. After pouring 25 ml of the mixture, the gel was overlayered with water-saturated butan-2-ol and allowed to set. When the gel had set, the water-saturated butan-2-ol was washed off. For the stacking gel, 2.5 ml of solution A, 5 ml of solution C, 200 µl of 10% SDS and 12.2 ml distilled water was added in a conical flask. After degassing, 100 µl of 15% AMPS and 15 µl TEMED was added and then poured onto the top of the resolving gel before inserting an appropriate comb.

2.25.2. Native gradient PAGE

Solution A: Acrylamide mix (low Bis):

Acrylamide	60.0 g
Bis-acrylamide	0.3 g

The volume was brought to 100 ml with distilled water.

Solution B: Acrylamide Mix (high Bis):

Acrylamide	60.0 g
Bis-acrylamide	1.6 g

The volume was brought to 100 ml with distilled water.

Solution C: Resolving Gel Buffer was as described prevously (see Section 2.21.1 solution D).

Solution D: Stacking Gel Buffer was as described previously (see Section 2.21.1 solution E).

Solution E: Running buffer for native gradient gel

Tris base	12.0 g
Glycine	57.6 g

The volume was brought to 2 l with distilled water.

For the gradient gel, 30% and 4% polyacrylamide mixtures were made separately. For the 30% mix, 10 ml of solution A and 7.4 ml of 75% glycerol were put in a glass universal bottle and degassed. Before pouring 10 µl of TEMED and 100 µl of 10% AMPS were added. For the 4% mix, 3.32 ml solution B, 6.25 ml of solution C and 40.38 ml water were put in a conical flask and degassed. Before adding 20 µl TEMED and 200 µl AMPS, 5 ml of it were taken out to make "tooth former".

To pour the gradient native gel, silicon rubber tubing was passed through a peristaltic pump and attached to a 21Gx1.5 needle (Sabre, Sabre International Products Ltd.) which was used to pierce a size 41 Suba seal. A piece of plastic tubing with 1 mm internal diameter was passed through the Suba seal until it reaching to the bottom of the glass universal bottle. The other end of the tubing was taped to the assembled casting plates. The low concentration gel was pumped through to the 21G needle. The Suba seal was inserted into the universal bottle including the high concentration gel mix and a magnetic stirring bar. The peristaltic pump was then started to deliver the 25% to 50% of high concentration gel mix at full speed. The low concentration gel mix was then connected to the pump so that the two gel mixtures mix gradually in the universal bottle on a magnetic stirrer. The gel was then overlaid with water-saturated butan-2-ol and allowed to polymerise for at least 3 hours. After the gel had set, the butan-2-ol was washed off. To 5 ml of stacking gel (tooth former solution) from the 4 % polyacrylamide solution, 10 µl of TEMED and 100 µl of 10% were added and it was then poured on the top of the gel, inserting an appropriate comb.

2.25.3 Electrophoresis conditions

Sample buffer for SDS PAGE:

Stacking gel buffer, solution B	2.5 ml
SDS (10%)	4.0 ml
β-mercaptoethanol	2.0 ml
Glycerol	2.0 ml

A small amount of tracking dye (bromophenol blue) was added to the sample buffer.

Samples to be run in SDS PAGE were mixed with sample buffer (13.3 µl sample buffer for 40 µl), boiled in a water bath for 5 min and loaded on the gel. Samples to be run on a native gel were mixed with 30% glycerol (10 µl for 40 µl of sample) and loaded onto the gel. Electrophoresis was carried out using a Bio-Rad power/pack 300 model power supplier. SDS PAGE gels were run at 30 mA for 2 to 3 hours or at 7 mA overnight until the dye had reached to the bottom of the gel. Native gels were run at 15 mA overnight.

Electrophoresis calibration kits were supplied by BioRad or Sigma. A Sigma low molecular weight calibration kit was used as molecular weight marker for SDS PAGE. The markers (from Bio-Rad) were phosphorylase b (97.4 kDa) serum albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (31 kDa trypsin inhibitor (21.5 kDa) and lysosyme (14.4 kDa). A Pharmacia high molecular weight calibration kit containing the markers thyroglobulin (669 kDa), ferritin (440 kDa), catalase (232 kDa), lactate dehydrogenase (140 kDa) and albumin (67 kDa) were used as molecular weight markers for native PAGE.

Gels were stained in Coomassie blue solution for at least 10 mins. Coomassie blue solution contained 2 g Coomassie blue R-250, 45% methanol (v/v) and 10% glacial acetic acid (v/v) in 1000 ml. Gels were then destained in a

solution of 20% methanol (v/v) and 10% glacial acetic acid (v/v) (Hames and Rickwood, 1982).

2.25.4 Western bloting of proteins

Western boltting of proteins was carried out as described by Silman *et al.* (1995). Proteins were run in denaturing or non-denaturing polyacrylamide gels.

Transblot buffer:

9.08 g of Tris was dissolved in 400 ml distilled water and the volume was brought to 500 ml after adjusting to pH 8.3. Then 43.24 g of glycine and 1900 ml distilled water were added to the solution and dissolved. Finally 600 ml methanol was added to solution.

The gel from which proteins would be transferred to the membrane was washed in transblot buffer for 10 mins. A Hybond-C nitrocellulose membrane (Amersham) was cut to the same size as the gel along with 6 pieces of 3 mm Whatman paper. The membrane and papers were soaked in transblot buffer. Onto the one side of the plastic holder of the Bio-Rad blotting tank, a soaked sponge and three pieces of Whatman paper were placed. Than the gel, the membrane, three pieces of Whatman paper and a sponge were piled up respectively. Air bubbles were removed using a glass rod. The holder was then closed and clamped and transferred into the tank. The tank was filled with transblot buffer and connected to a 250/2.5 model Bio-Rad power supply. SDS gels were blotted at 300 mA for 2.5 hours while native gels were blotted at 150 mA overnight.

The membrane was transferred to a plastic box and stained with Ponceau S dye (0.5% Ponceau S dye [w/v] dissolved in 5% [v/v] trichloroacetic acid) for 10 mins. Once the protein bands were visualised the membrane was washed with distilled water. After the tracks and molecular weight markers were marked, the dye was removed by washing with phosphate-buffered saline (PBS) (140 mM sodium chloride, 20 mM sodium phosphate [pH 7.3]). The filter was then blocked with 100 ml of 2% Marvel (w/v) in PBS for 1 hour. Marvel solution was replaced with 20 ml of the fresh Marvel solution and primary antiserum was added to a final concentration of 1:1000 (v/v) for SDS gels and 1:100 (v/v) for native gels and left on a shaker overnight.

The filter was then washed three times with washing solution (0.1 % [v/v] Triton X-100 in PBS) and 10 ml of the washing solution and 30 µl of secondary antibody (horseradish peroxidase-conjugated goat anti-rabbit immunoglobulin G) (Amersham) was added and incubated with shaking for two hours. The membrane was washed twice with the washing solution and once with PBS for 10 mins for each wash before proceeding to the visualisation of immunoblotted proteins.

Developing solutions:

Solution A:

NaCl	1.5 g
1 M Tris pH 7.5	1.0 ml

The volume was brought to 50 ml with distilled water.

Solution B:

Chloronaphthol	30 mg
Methanol	10 ml

The volume was brought to 50 ml with distilled water.

After 50 µl of H_2O_2 was added to solution A both solution A and B were mixed and added to the blot. The blot was gently shaken until the bands were seen. Once the bands were visualised, the developing solution was removed and the blot was washed twice with distilled water. The blot was finally dried between Whatman paper.

2.26 Counting of cyanobacterial cells

Cyanobacterial cells were counted by using a cell counting chamber supplied by Weber England. The cells from five big squares were counted and cell number per ml were estimated by the formula as follows:

Cells ml^{-1} = Total cells counted from 5 big squares x 5 x 10^4.

Once the cell number per ml was estimated, the culture was either concentrated or diluted to give a desired concentration of cells. *Synechococcus* sp. PCC7942 cells were counted directly while *Anabaena* sp. PCC7120 filaments were sonicated briefly to obtain a countable filament size.

2.27 Dilution plating of *Synechococcus* sp. PCC7942 for dark viability analysis

Dark viability of wild-type of *Synechococcus* sp. PCC7942 and the mutant strains HK52 and HK55 was assayed by dilution plating on BG11 plates (Scanlan *et al.*, 1995). Flasks containing light-grown wild-type and mutant cells were wrapped with foil to exclude light and sampled daily, plated onto BG11 to give approximately 300 colonies per plate, and incubated at 30°C under constant illumination. For each strain, a light-grown culture was set up as positive control. Each experiment was performed in triplicate. The viable cell number of a light grown culture for a certain period of growth was taken as 100% viable, and the viable cell number of the dark grown culture of the same strain was compared to the wild-type.

Chapter 3

Sequencing of the *zwf* region of the genome of *Anabaena* sp. PCC7120

3.1. Introduction[1]

In this chapter, the sequencing studies on the *zwf* region of the genome of *Anabaena* sp. PCC7120 will be described. Firstly, studies on the determination of the nucleotide sequence of the *zwf* region will be summarised. Then, further analysis of nucleotide sequence of the region and amino acid sequence of the genes will be represented. When this study was started, G6PDH genes (*zwf*) had already been sequenced from eukaryotic sources such as human (Takizawa *et al.*, 1986) and *Saccharomyces cerevisiae* (Nogae and Johnston, 1990) and from prokaryotic sources such as *Zymomonas mobilis* (Barnell *et al.*, 1990), *E. coli* K-12 (Rowley and Wolf, 1991) and *Leuconostoc mesenteroides* (Lee *et al.*, 1991). The overall objective of the sequencing of the *zwf* region of *Anabaena* sp. PCC7120 was to facilitate analysis of the regulation of *zwf* in a filamentous, nitrogen fixing cyanobacterium and also to determine the gene organisation of the *zwf* region in a filamentous cyanobacterial strain. The first cyanobacterial *zwf* gene, from a unicellular strain *Synechococcus* sp. PCC7942, was cloned and sequenced in this laboratory (Scanlan *et al.*, 1992). The gene consisted of 1572 nucleotides encoding a polypeptide of 524 amino acids. Probing with the downstream part of this *Synechococcus* sp. PCC7942 *zwf* gene, a 7.5 kb *Hin*dIII fragment from *Anabaena* sp. PCC7120 was cloned into pBR325 to yield pAG75 (Figure 3.1) (J. Newman, personal communication). Further analysis of this clone showed that a 2.4 kb *Hpa*I/*Hin*dIII fragment contained the *zwf* gene.

[1] Most of the data represented in this chapter have been published in FEMS Microbilogy Letters 133 (1995) 187-193.

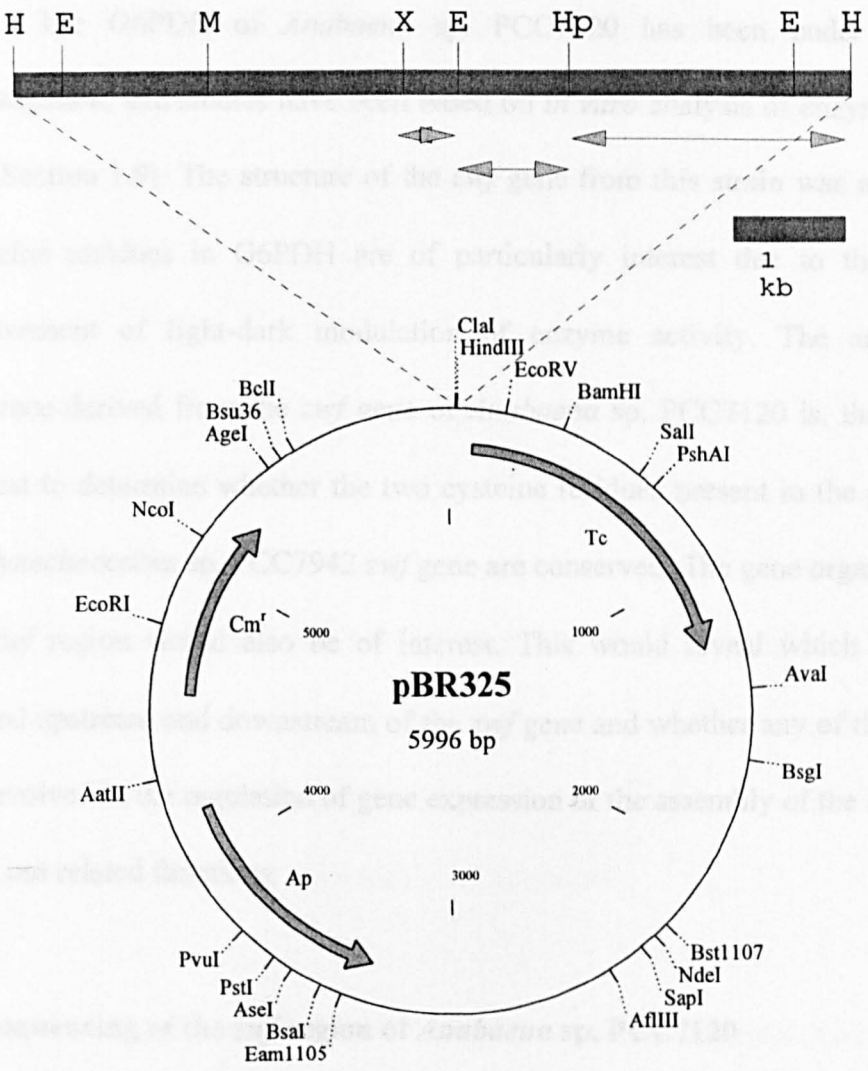


Figure 3.1. Physical map of the clone pAG75. Abbreviations: E, *Eco*RI; H, *Hin*dIII; Hp, *Hpa*I; M, *Mlu*I and X, *Xba*I. The bar indicates the size of the insert. The size of vector is indicated on the map. Arrows below the insert indicate the fragments which were sequenced as three separate sequencing projects.

The G6PDH of *Anabaena* sp. PCC7120 has been under extensive investigation, and studies have been based on *in vitro* analysis of enzyme activity (see Section 1.9). The structure of the *zwf* gene from this strain was not known. Cysteine residues in G6PDH are of particularly interest due to the possible involvement of light-dark modulation of enzyme activity. The amino acid sequence derived from the *zwf* gene of *Anabaena* sp. PCC7120 is, therefore, of interest to determine whether the two cysteine residues present in the product of the *Synechoccocus* sp. PCC7942 *zwf* gene are conserved. The gene organisation in the *zwf* region would also be of interest. This would reveal which genes are located upstream and downstream of the *zwf* gene and whether any of these genes are involved in the regulation of gene expression or the assembly of the enzyme or carry out related functions.

3.2 Sequencing of the *zwf* region of *Anabaena* sp. PCC7120

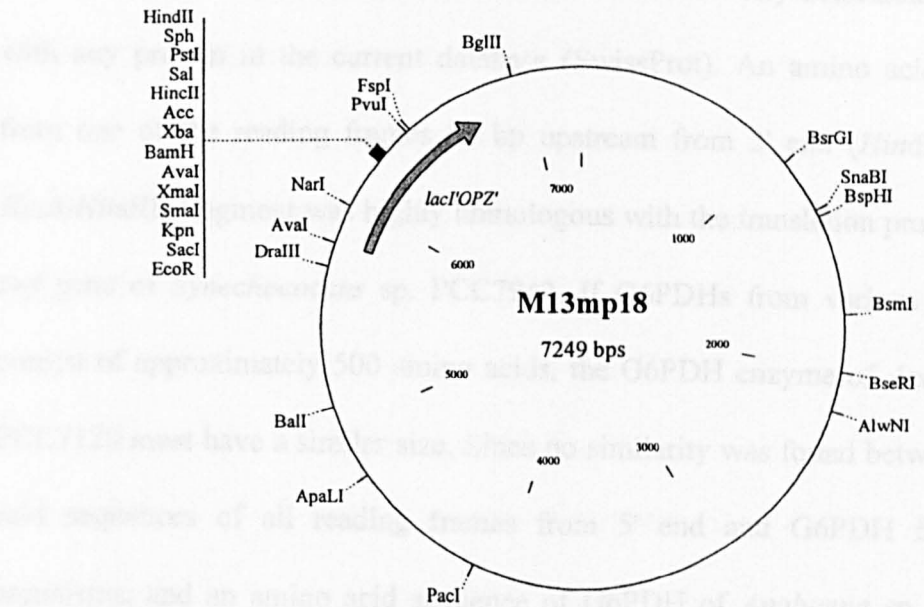
A 4,169 bp portion of the 7,500 bp insert in pAG75 representing the *zwf* region of the *Anabaena* sp. PCC7120 chromosome was sequenced. This sequencing involved three separate projects. Each project aimed to complete the sequence of a particular fragment on which a part or whole of a gene was located. The initial project aimed to sequence a 2.4 kb *Hpa*I/*Hin*dIII fragment which apparently contained the *zwf* gene. The second project aimed to sequence a *Eco*RI/*Hpa*I fragment from upstream of the 2.4 kb *Hpa*I/*Hin*dIII fragment. This strategy would reveal whether the region carried any genes. The third strategy involved completion of the nucleotide sequence of the *fbp* gene. To do this a 0.6

kb *Xba*I/*Eco*RI fragment was sequenced. To complete the sequence of a gene subsequently identified as coding for FBPase, some more sequence upstream from the *Xba*I site was determined. The fragments were linked using oligonucleotide primers.

3.2.1 Sequencing of the *zwf* gene of *Anabaena* sp. PCC7120

A 7.5 kb fragment of *Anabaena* sp. PCC7120 chromosomal DNA resulting from *Hin*dIII digestion had previously been cloned into pBR325 to yield pAG75 (Figure 3.1), and the 2.4 kb *Hin*dIII/*Hpa*I fragment of this clone was shown to include the *zwf* gene (J. Newman, personal communication). Plasmid DNA was then prepared from pAG75 on a small scale (see Section 2.7.1) and used for the purification of the required subfragments.

The initial strategy to obtain complete nucleotide sequence information on the *zwf* gene was to sequence the ends of this 2.4 kb *Hin*dIII/*Hpa*I fragment to determine whether the complete gene was carried on this fragment. The 2.4 kb *Hin*dIII/*Hpa*I fragment was purified from the clone pAG75 and cloned into the *Hin*dIII and *Hin*cII sites of both M13mp18 and mp19. (Restriction map of M13mp18 and multiple cloning sites of M13mp18 and mp19 is shown in Figure 3.2). Cloning of the same fragments into both M13mp18 and mp19 provided templates that contained the same fragments in opposite directions. Using these templates both ends of the 2.4 kb *Hin*dIII/*Hpa*I fragment would be sequenced. Approximately 150 bp from *Hpa*I site and 200 bp from *Hin*dIII site of the fragment were sequenced on M13mp18 and mp19 templates. Translation in all six



M13mp18 multiple cloning site

M13/pUC Forward Sequencing Primer

```
                                                                    3'-TGACCGGCAG CAAAATG-5'
5'-GAATTCGAGC TCGGTACCCG GGGATCCTCT AGAGTCGACC TGCAGGCATG CAAGCTTGGC ACTGGCCGTC GTTTTAC-3'
```

EcoRI SacI KpnI SmaI XmaI BamHI XbaI SalI AccI HincII PstI SphI HindIII

M13mp19 multiple cloning site

M13/pUC Forward Sequencing Primer

```
                                                                    3'-TGACCGGCAG CAAAATG-5'
5'-GCCAAGCTTG CATGCCTGCA GGTCGACTCT AGAGGATCCC CGGGTACCGA GCTCGAATTC ACTGGCCGTC GTTTTAC-3'
```

HindIII SphI PstI SalI AccI HincII XbaI BamHI SmaI XmaI KpnI SacI EcoRI

Figure 3.2: Restriction map of M13 mp18 RF and nucleotide sequence of the multiple cloning site of M13mp18 and mp19. The nucleotide sequence of M13/pUC forward sequencing primer is also shown.

reading frames from the 5' end (*Hpa*I site) did not show any detectable homology with any protein in the current database (SwissProt). An amino acid sequence from one of the reading frames 60 bp upstream from 3' end (*Hin*dIII) of this *Hpa*I/*Hin*dIII fragment was highly homologous with the translation product of the *zwf* gene of *Synechococcus* sp. PCC7942. If G6PDHs from various organisms consist of approximately 500 amino acids, the G6PDH enzyme of *Anabaena* sp. PCC7120 must have a similar size. Since no similarity was found between amino acid sequences of all reading frames from 5' end and G6PDH from other organisms, and an amino acid sequence of G6PDH of *Anabaena* sp. PCC7120 ended 60 bp upstream from the 3' end, the 2.4 kb *Hpa*I/*Hin*dIII fragment obviously carried the complete *zwf* gene.

To complete the nucleotide sequence of this 2.4 kb *Hpa*I/*Hin*dIII fragment, a random cloning and sequencing strategy was employed. The 2.4 kb *Hpa*I/*Hin*dIII fragment was self-ligated, sonicated, end filled and ligated into the *Sma*I site of M13mp18 (see Section 2.19.2). After transfecting into *E. coli* TG1 (section 2.19.3) and selecting white plaques (section 2.19.4), template DNAs were prepared as described in section 2.19.5. The sequences obtained from these templates gave a single contigous sequence of the 2.4 kb *Hpa*I/*Hin*dIII fragment. Almost all of the fragment was double-stranded except a 100 bp 3' part of this fragment. An oligonucleotide primer called Ola1 (see Table 3.1) was designed and used for completing double-stranded sequence of the region. Analysis of the sequence showed that the *zwf* gene of *Anabaena* sp. PCC7120 consisted of 1,527

nucleotides encoding a polypeptide of 509 amino acids. For further characterisation of the gene see Sections 3.2.4 and 3.5.

Table 3.1: Oligonucleotides used for joining and/or double-stranding the fragments.

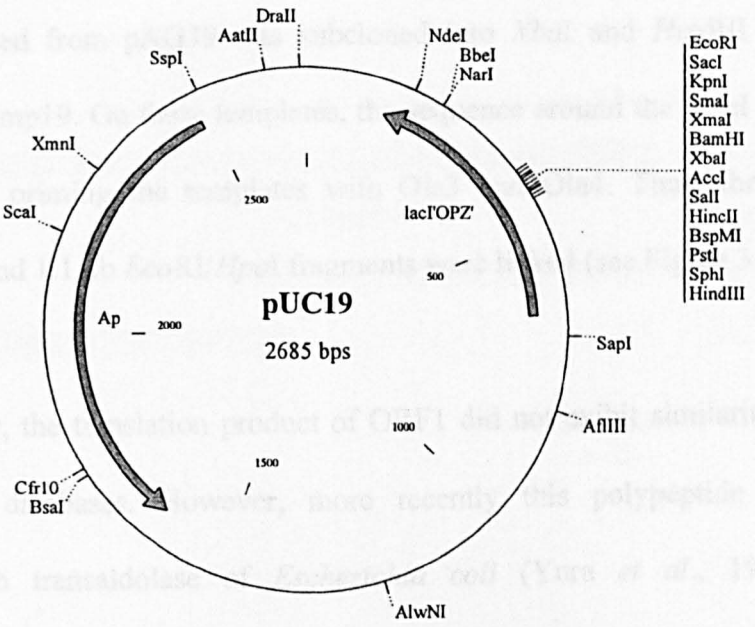
Name	DNA sequence(5' to 3')
Ola1	GCCTATTCTT AGATTGT
Ola2	GCACCACTGT AACGAGC
Ola3	TTAACGAAGC CCGCCGC
Ola4	TTACCCTCAG CTAGGCG
Ola5	GCGCGGCTAA GGCGACG
Ola6	GAGGAACCGT CTATTGG
Ola7	GGACGCAAGC AGATTGC
Ola8	TCAACGAGTC TACACAG

3.2.2 Sequencing of the *ORF1* upstream from the *zwf* gene

Translation of in all six frames of the nucleotide sequence of the 2.4 kb *Hpa*I/*Hin*dIII fragment indicated that one of the frames had a 200 amino acid sequence which was not interrupted by a stop codon. This frame started immediately from the 5' end of the 2.4 kb *Hpa*I/*Hin*dIII fragment and finished with a stop codon about 150 bp upstream from the *zwf* gene. The polypeptide encoded by this partial ORF exibited no detectable similarity in the contemporary

databases. It was decided to obtain the complete nucleotide sequence of this ORF, designated ORF1.

The strategy for sequencing of ORF1 was to subclone a 1.1 kb *Eco*RI/*Hpa*I fragment into pUC19 and to obtain the sequence by plasmid sequencing (Guo *et al.*, 1983; Murphy and Ward, 1989). The 1.1 kb *Eco*RI/*Hpa*I fragment of pAG75 was purified, end filled and subcloned into the *Sma*I site of pUC19 (for the structure of pUC19 see Figure 3.3) yielding pAG14 (see Figures 3.4 and 3.5). A 300 bp nucleotide sequence upstream from the 5' (*Hpa*I) end of the 2.4 kb *Hpa*I/*Hin*dIII fragment was obtained by using pAG14 as double-stranded DNA template. However, the start codon of the putative ORF1 was not reached by plasmid sequencing. A random cloning strategy was then applied. The *Eco*RI/*Bam*HI fragment of pAG14 was purified, sonicated and cloned into the *Sma*I site of M13mp18. The templates were prepared and used to complete the sequence of the 1.1 kb *Eco*RI/*Hpa*I fragment. Ola2 was used as primer to obtain double-stranded sequence of a region in the fragment (see Table 3.1). Thus the complete sequence information for ORF1 was obtained, the beginning of the gene being on the 1.1 kb *Eco*RI/*Hpa*I fragment and the last part on the 2.4 kb *Hpa*I/*Hin*dIII fragment. To link these two fragments, plasmid sequencing was performed using oligonucleotide primers (Table 3.1). Two oligonucleotides were designed for joining the two fragments, one of the oligonucleotides (Ola3) priming the coding strand upstream from the *Hpa*I site and the other (Ola4) priming the complementary strand downstream from the *Hpa*I site. A 4 kb *Xba*I/*Hin*dIII fragment of pAG75 was then subcloned into the *Xba*I/*Hin*dIII sites



M13/pUC19 Forward Sequencing Primer

```
5'-GT AAAACGACGG CCAGT-3'
5'-GT AAAACGACGG CCAGTGAATT CGAGCTCGGT ACCCGGGGAT CCTCTAGAGT CGACCTGCAG GCATGCAAGC TTGGCG-3'
```

EcoRI SacI KpnI SmaI XmaI BamHI XbaI SalI AccI HincII PstI SphI HindIII

Figure 3.3: Restriction map and multiple cloning site of pUC19. The nucleotide sequence of M13/pUC forward sequencing primer is also shown.

of pUC19 yielding the subclone pAG39 (Figure 3.4 and 3.5). This subclone was primed by Ola3 and Ola4 to determine the sequence at the *Hpa*I site. After failing to determine readable sequence by plasmid sequencing, the 4 kb *Xba*I/*Hin*dIII fragment purified from pAG39 was subcloned into *Xba*I and *Hin*dIII sites of M13mp18 and mp19. On these templates, the sequence around the *Hpa*I site was determined by priming the templates with Ola3 and Ola4. Thus, the 2.4 kb *Hpa*I/*Hin*dIII and 1.1 kb *Eco*RI/*Hpa*I fragments were linked (see Figure 3.6).

Initially, the translation product of ORF1 did not exibit similarity to any protein from databases. However, more recently this polypeptide exibited similarity with transaldolase of *Escherichia coli* (Yura *et al.*, 1992) and *Saccharomyces cerevisiae* (Schaaff *et al.*, 1990) in the database SWISS-PROT (Release 29, 1994). The polypeptide was 49% similar (28% identical) to *Saccharomyces cerevisiae* and 46% similar (22% identical) to *Escherichia coli* transaldolases. Consequently, the ORF1 was identified as encoding transaldolase on the basis of these similarity and designated *tal* (for further characterisation see section 3.2.4 and 3.4).

3.2.3 Sequencing of the *fbp* gene

The 5' (*Eco*RI) end of the 1.1 kb *Eco*RI/*Hpa*I fragment carried an incomplete ORF. The amino acid sequence encoded by this incomplete ORF was searched on the database SWISS-PROT using FASTA. It expressed similarity with the downstream part of FBPases from other prokaryotic and eukaryotic sources. The amino acid sequence was 47% identical to human FBPase (Solomon

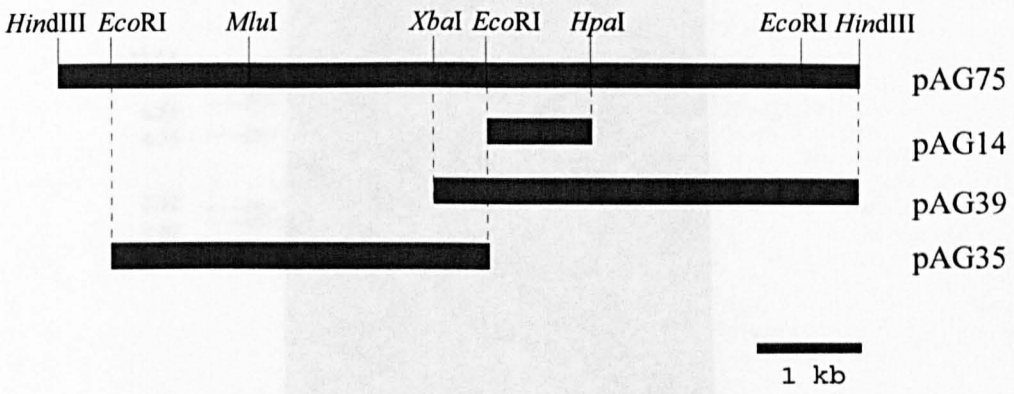


Figure 3.4: Subclones used for sequencing of the *zwf* region of the *Anabaena* sp. PCC7120 chromosome. For pAG75 see Figure 3.1. The others are pUC19-based subclones. pAG14 includes a 1.1 kb *Eco*RI/*Hpa*I fragment in the *Sma*I site. pAG39 includes a 4 kb *Xba*I/*Hin*dIII fragment in the *Xba*I/*Hin*dIII sites. pAG35 includes a 3.5 kb *Eco*RI fragment in the *Eco*RI site. (Also see Figure 3.5).

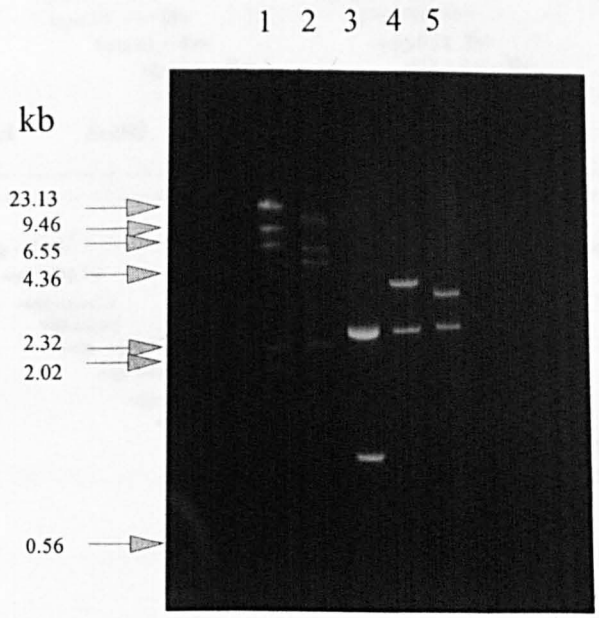


Figure 3.5: Restriction digestions of the clone pAG75 and subclones pAG14, pAG39 and pAG35. Lane 1, bacteriophage lambda DNA digested with *Hin*dIII used as molecular size marker. The sizes of the bands were shown on the left side of the picture. Lane 2, pAG75 digested with *Hpa*I and *Hin*dIII. The insert is seen as two bands at 2.4 and 5 kb region. The vector pBR325 is seen as the upper band. Lane 3, pAG14 digested with *Eco*RI and *Bam*HI. Lower band is the 1.1 kb insert and the upper the vector pUC19. Lane 4, pAG39 digested with *Xba*I and *Hin*dIII. The upper band is the 4 kb insert and the lower pUC19. Lane5, pAG35 digested with *Eco*RI. The upper band is the 3.5 kb insert and the lower pUC19. (Also see Figure 3.4).

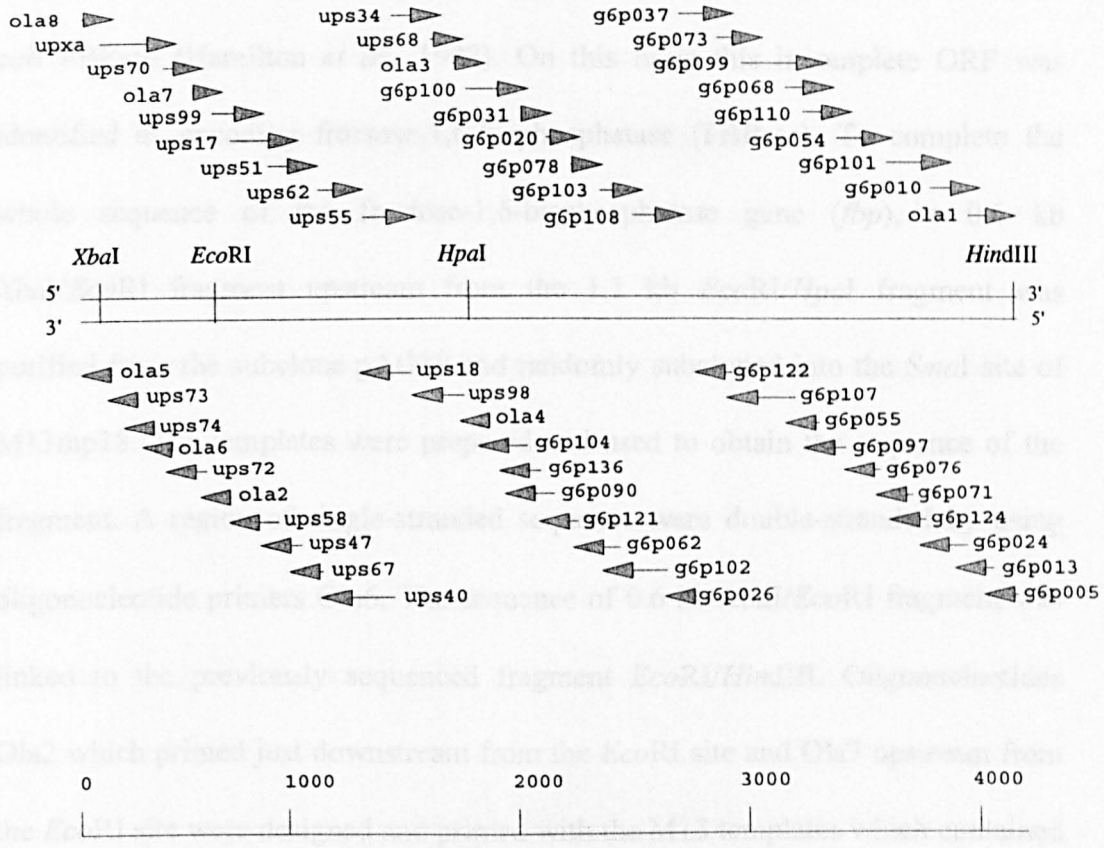


Figure 3.6: M13 templates used in the determination of the DNA sequence of the 4169 bp *zwf* region. The sequences from g6p005 to g6p136 were used to sequence the 2.4 kb *Hpa*I/*Hin*dIII fragment, from ups18 to ups 99 and upxa to sequence both the 1.1 kb *Eco*RI/*Hpa*I and 0.6 kb *Xba*I/*Eco*RI fragments. The sequences ola3 and ola4 were used for linking the 2.4 kb *Hpa*I/*Hin*dIII and 1.1 kb *Eco*RI/*Hpa*I fragments, ola2 and ola7 for linking the 1.1 kb *Eco*RI/*Hpa*I and 0.6 kb *Xba*I/*Eco*RI fragments. The sequences ola1 and ola64 were used for determining double-stranded sequences. The sequences ola5 and ola8 were used for obtaining the sequence of the region upstream of the *Xba*I site.

et al., 1988), 45% to *Spinacia oleracea* FBPase (Hur *et al.*, 1992) and 44% to *E. coli* FBPase (Hamilton *et al.*, 1988). On this basis this incomplete ORF was identified as encoding fructose-1,6-bisphosphatase (FBPase). To complete the whole sequence of the fructose-1,6-bisphosphatase gene (*fbp*), a 0.6 kb *Xba*I/*Eco*RI fragment upstream from the 1.1 kb *Eco*RI/*Hpa*I fragment was purified from the subclone pAG39 and randomly subcloned into the *Sma*I site of M13mp18. The templates were prepared and used to obtain the sequence of the fragment. A region of single-stranded sequence were double-stranded by using oligonucleotide primers Ola6. The sequence of 0.6 kb *Xba*I/*Eco*RI fragment was linked to the previously sequenced fragment *Eco*RI/*Hin*dIII. Oligonucleotides Ola2 which primed just downstream from the *Eco*RI site and Ola7 upstream from the *Eco*RI site were designed and primed with the M13 templates which contained the 4 kb *Xba*I/*Hin*dIII insert. The sequences obtained from these sequencing reactions were merged with the 0.6 kb *Xba*I/*Eco*RI and previously sequenced *Eco*RI/*Hin*dIII fragments (see Figure 3.6).

The nucleotide sequence information from the 0.6 kb *Xba*I/*Eco*RI fragment indicated that an approximately 50 bp part of the gene remained upstream from the *Xba*I site. Sequencing of this remaining part of the *fbp* gene upstream from the *Xba*I site was attempted by employing a double-stranded sequencing strategy. A 3.5 kb *Eco*RI fragment including the region with the beginning of the *fbp* gene was then subcloned into the *Eco*RI site of pUC19 to yield pAG35 (see Figures 3.4 and 3.5). A synthetic oligonucleotide Ola5 was used

to prime pAG35 just downstream of the *Xba*I site to obtain sequence of the region upstream from the *Xba*I site, but no readable sequence could obtain by sequencing from this double-stranded template pAG35. Then, the 3.5 kb *Eco*RI fragment was purified from the subclone pAG35 and subcloned into the *Eco*RI site of M13mp18 and templates were prepared. These templates were used to sequence the region upstream from the *Xba*I site priming by Ola5. This sequence, obtained from the templates which primed with Ola5, reached about 100 bp further upstream of the start codon of *fbp* gene. Some templates were not primed with Ola5 indicating that these templates carried inserts in the opposite direction. Another oligonucleotide Ola8 was designed and used to prime the templates in the opposite orientation and the sequence of double-stranded region was obtained. Thus the nucleotide sequence of the *fbp* gene was completed. For further analysis of the gene see Sections 3.2.4 and 3.3.

3.2.4 Merging the individual sequences and determination of the locations of the genes

For the analysis of the sequences determined, the GCG package version 7.2 (1992) and version 8 (1994) were used. Some databases are also available through the package: GenBank, GenPept, EMBL(Abridged), PIR-Protein, SWISS-PROT, PROSITE and Restriction Enzymes.

The Fragment Assembly programs of the package were used to enter and assemble overlapping nucleotide sequence fragments into one continuous

sequence. The GELENTER program was used to add sequences to the fragment assembly project and the overlapping sequences were merged using the GELOVERLAP program. The GELASSEMBLE program was used to put sequences together into a contiguous sequence. All individual nucleotide sequences determined from M13 and pUC19 templates were merged in one contiguous sequence which was 4,169 bp in size. This sequenced part of the *zwf* region of the *Anabaena* sp. 7120 chromosome was searched for possible open reading frames using the CODONPREFERENCE program. Three genes were recognised by this program as seen in Figure 3.7.

3.2.5 Identification, designation and organisation of the genes in the *zwf* region

Genes were identified on the basis of the similarity of their translation products with those of the genes from other organisms. The expected *zwf* gene of *Anabaena* sp. PCC7120 was translated and the gene product was compared to the *zwf* gene product of *Synechococcus* sp. PCC7942 which had already been sequenced and published (Scanlan *et al.*, 1992). This showed that these two gene products were 70% similar. The database search was then performed for gene product of the *zwf* gene of *Anabaena* sp. PCC7120 to confirm the similarity with the other G6PDHs from various organisms. This result confirmed that the *zwf* gene encodes the G6PDH enzyme of *Anabaena* sp. PCC7120. The product of ORF1 upstream of the *zwf* gene was searched on the database SWISS-PROT (Release 23, 1992). Since no significant similarity was found, the gene was thought to be a gene which might be involved in regulation of either the

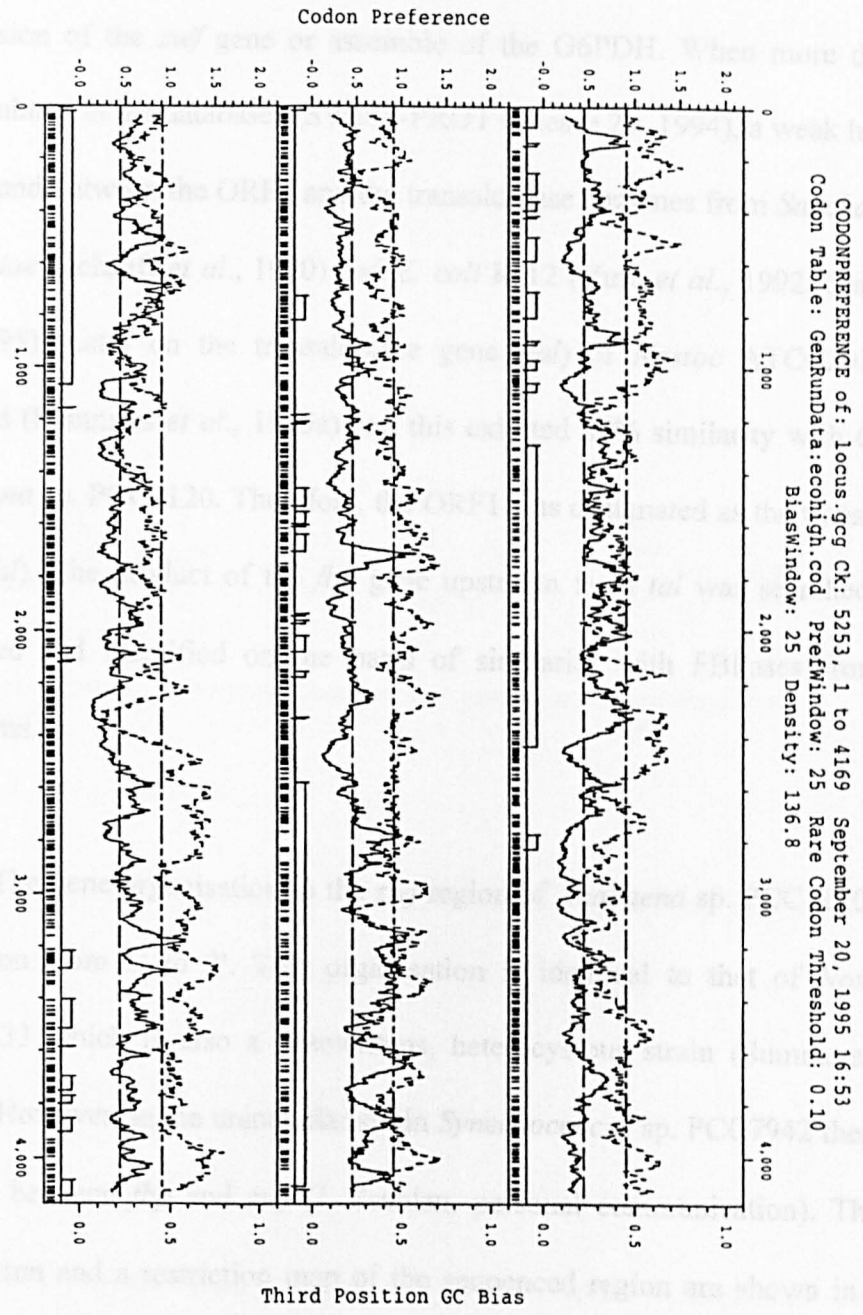


Figure 3.7: The genes located on the 4,169 bp *zwf* fragment of *Anabaena* sp. PCC7120. The sequence of the fragment was run on CODONPREFERENCE program. The *tal* gene is located on the first frame, *zwf* on the second and *fbp* on the third frame.

expression of the *zwf* gene or assemble of the G6PDH. When more data were accumulated in the databases (SWISS-PROT Release 29, 1994), a weak homology was found between the ORF1 and the transaldolase enzymes from *Saccharomyces cerevisiae* (Schaaff *et al.*, 1990) and *E. coli* K-12 (Yura *et al.*, 1992; Sprenger *et al.*, 1995). Later on the transaldolase gene (*tal*) of *Nostoc* ATCC29133 was reported (Summers *et al.*, 1995a) and this exibited 88% similarity with ORF1 of *Anabaena* sp. PCC7120. Therefore, the ORF1 was designated as the transaldolase gene (*tal*). The product of the *fbp* gene upstream from *tal* was searched on the databases and identified on the basis of similarity with FBPases from other organisms.

The gene organisation in the *zwf* region of *Anabaena* sp. PCC7120 is *fbp-tal-zwf* on from 5' to 3'. This organisation is identical to that of *Nostoc* sp. PCC29133 which is also a filamentous, heterocystous strain (Summers *et al.*, 1995a). However, in the unicellular strain *Synechococcus* sp. PCC7942 there is no *tal* gene between *fbp* and *zwf* (J. Scanlan, personal communication). The gene organisation and a restriction map of the sequenced region are shown in Figure 3.8.

The precise locations of the genes on the 4,169 bp Anabaena sp. PCC7120 fragment were determined (see Figure 3.9). The *fbp* gene is located between positions 21-1070 including the stop codon TAA. The *tal* gene starts at position 1291 and ends at 2436 including the stop codon TAG. The *zwf* gene is located

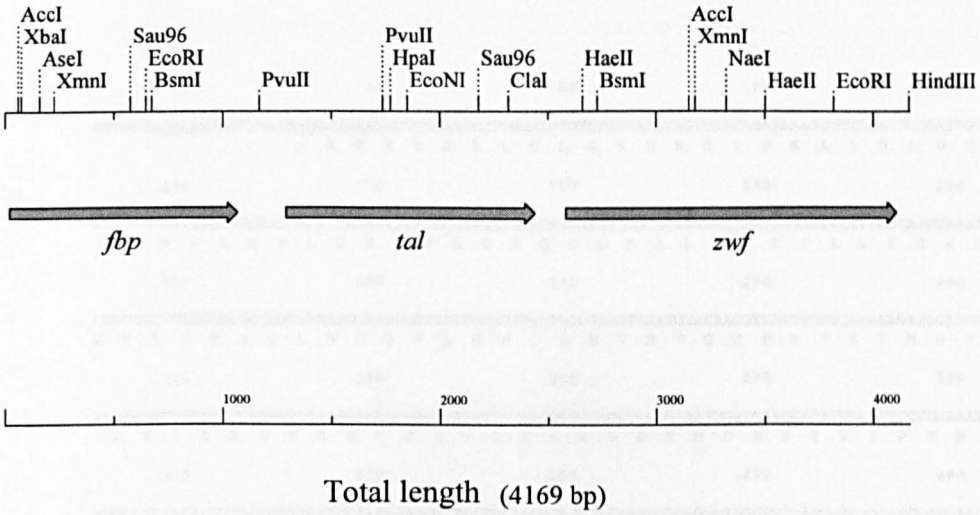


Figure 3.8: Restriction map and gene organisation in the *zwf* region of *Anabaena* sp. PCC7120 chromosome.

Figure 3.9: DNA sequence of the 4169 bp fragment and deduced amino acid sequences of the genes in the *zwf* region of *Anabaena* sp. PCC7120. The gene *fbp* is located between the positions 21 and 1070, *tal* between 1291 and 2436 and *zwf* between 2579 and 4108. Amino acid sequences are shown below the DNA sequence of the genes. A potential Shine-Dalgarno sequence 8 bp upstream from *fbp* and start codons are underlined.

```
        10                  30                  50                  70                  90
GTCGCTAGGAAGTACATAACATGGCTAAAGCCTCAGAATCCTTGGATTTGTCTGTCAACGAGTCTACAGATAAAGCTCTAGACCGCGATTGTACAACCTT
                  M  A  K  A  S  E  S  L  D  L  S  V  N  E  S  T  D  K  A  L  D  R  D  C  T  T  L

       110                 130                 150                 170                 190
ATCCCGTCACGTCCTCCAACAACTCCAAAGTTTTTCCGCAGATGCACAAGATTTGAGTGCATTAATGAATCGGATTGCTTTAGCAGGTAAATTAGTTGCC
 S  R  H  V  L  Q  Q  L  Q  S  F  S  A  D  A  Q  D  L  S  A  L  M  N  R  I  A  L  A  G  K  L  V  A

       210                 230                 250                 270                 290
CGTCGCCTTAGCCGCGCTGGTTTAATGGAAGGGGTTCTCGGATTTACTGGGGAGGTAAATGTACAAGGTGAATCCGTCAAAAAGATGGATGTCTATGCCA
R  R  L  S  R  A  G  L  M  E  G  V  L  G  F  T  G  E  V  N  V  Q  G  E  S  V  K  K  M  D  V  Y  A  N

       310                 330                 350                 370                 390
ATGACGTGTTTATCTCAGTGTTTAAGCAAAGCGGCTTAGTTTGTCGTCTTGCATCTGAGGAAATGGATGAACCATATTACATTCCCGAAAATTGCCCCAT
  D  V  F  I  S  V  F  K  Q  S  G  L  V  C  R  L  A  S  E  E  M  D  E  P  Y  Y  I  P  E  N  C  P  I

       410                 430                 450                 470                 490
AGGTCGTTACACCTTACTGTATGACCCAATAGACGGTTCCTCAAACACAGATACAAATCTCAGTTTGGGTTCCATTTTCTCTATCCGTCAACAGGAAGGT
 G  R  Y  T  L  L  Y  D  P  I  D  G  S  S  N  T  D  T  N  L  S  L  G  S  I  F  S  I  R  Q  Q  E  G

       510                 530                 550                 570                 590
GATGATAGTGATGGCCAAGCTAAAGACCTCCTCACCAACGGACGCAAGCAGATTGCGGCGGGGTTACATTTTATATGGGCCTAGCACGATGCTAGTTTATA
D  D  S  D  G  Q  A  K  D  L  L  T  N  G  R  K  Q  I  A  A  G  Y  I  L  Y  G  P  S  T  M  L  V  Y  T

       610                 630                 650                 670                 690
CAATGGGTACGGGAGTTCACTCATTTACCCTTGATCCCAGCTTAGGAGAATTCATCCTCAGCGAAGAAAATATCCGCATTCCCGACCACGGCGCAGTTTA
  M  G  T  G  V  H  S  F  T  L  D  P  S  L  G  E  F  I  L  S  E  E  N  I  R  I  P  D  H  G  A  V  Y

       710                 730                 750                 770                 790
CAGCGTTAATGAAGGTAACTTCTGGCAATGGGAAGAATCAATGCGGGAATATATCCGCTATGTTCACCGTACAGAAGGTTACACAGCTCGTTACAGTGGT
 S  V  N  E  G  N  F  W  Q  W  E  E  S  M  R  E  Y  I  R  Y  V  H  R  T  E  G  Y  T  A  R  Y  S  G

       810                 830                 850                 870                 890
GCAATGGTGAGCGATATCCATAGAATTTTGGTACAAGGTGGTGTGTTTCTCTACCCAGGGACGATTCAAAACCCTGAAGGTAAATTACGGTTGTTGTATG
A  M  V  S  D  I  H  R  I  L  V  Q  G  G  V  F  L  Y  P  G  T  I  Q  N  P  E  G  K  L  R  L  L  Y  E

       910                 930                 950                 970                 990
AATCTGCTCCTTTAGCCTTTTTAATTCAACAAGCTGGTGGTCGGGCAACTACGGGATTAGTAGATATCTTAGATGTTGTACCGAAAAAACTACACCAACG
  S  A  P  L  A  F  L  I  Q  Q  A  G  G  R  A  T  T  G  L  V  D  I  L  D  V  V  P  K  K  L  H  Q  R

      1010                1030                1050                1070                1090
TACACCTCTGATTATTGGTAGTAAAGAGGACGTAGCCAAGGTTGAATCATTTATTCAGAATGGTCATTAACAGGCAGGAGGCAGGGGGAGAATAAGAAAT
 T  P  L  I  I  G  S  K  E  D  V  A  K  V  E  S  F  I  Q  N  G  H

      1110                1130                1150                1170                1190
AATAATTCCTAACGCCCTACTCAGCACTTACTACTCAGCACTTTCTCAAGCTAAATGTGAGAGATTATCTCAGCTGGCCGATGTAATAAGTGATAGACAA

      1210                1230                1250                1270                1290
AGCAAAAATGCAAAAGTAACCATCAATACTTTGGTGTGTTTTGTCGAAGTCACTTAAAACATCTTTTACGGGGTGAAACAAACTGTAGCTATGGCCATCA
                                                                                  M  A  I  N
```

Figure 3.9 continued.

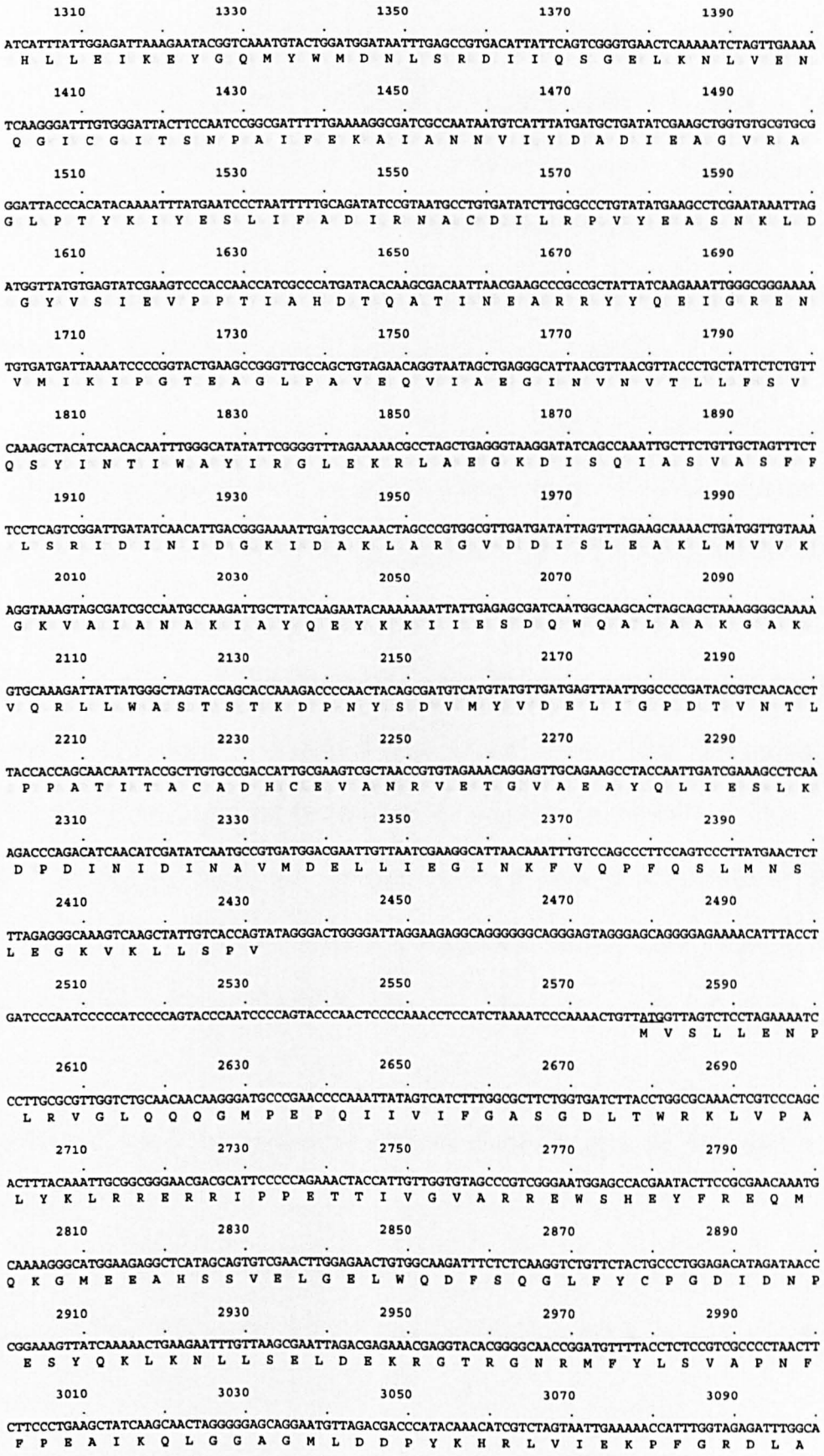
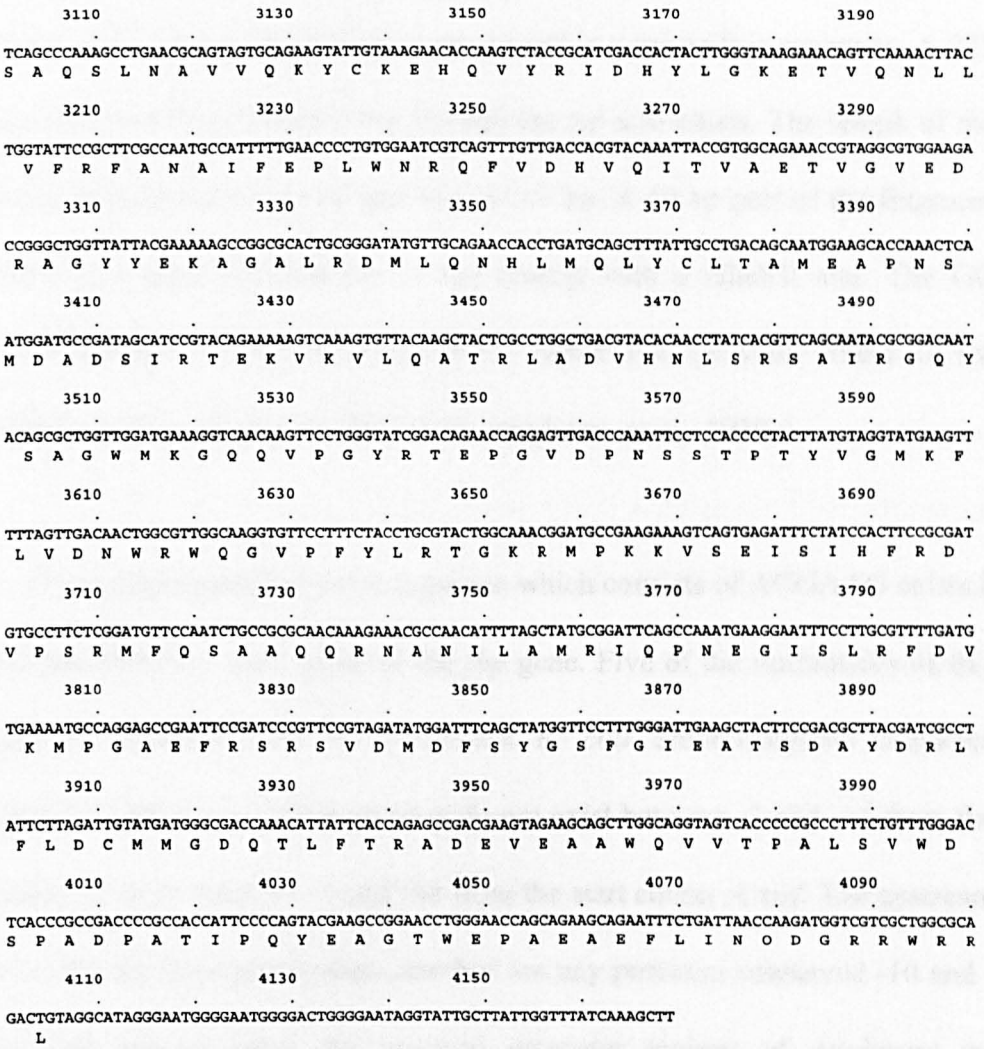


Figure 3.9 continued.



between positions 2579 and 4108 ending in a single stop codon TAG. A 20 bp sequence at the 5' end of the fragment sequenced is a noncoding sequence. A 223 bp noncoding sequence between the *fbp* and the *tal* also exists. The length of the noncoding sequence between *tal* and *zwf* is 143 bp. A 60 bp part of the fragment sequenced is located downstream of *zwf* ending with a *Hin*dIII site. The GC content of the region is 45.5% which is 3% higher than the value found for the whole genome of *Anabaena* sp. PCC7120 (Herdman, *et al.*, 1979a).

A potential Shine-Dalgarno sequence which consists of AGGAAG exists 8 bp upstream from the start codon of the *fbp* gene. Five of the nucleotides of this sequence are conserved in the consensus *E. coli* Shine-Dalgarno sequence (AGGAGG). Such conserved sequences do not exist between -1 and -15 from the start codon of *tal* or between -1 and -38 from the start codon of *zwf*. The upstream regions of the all three genes were searched for any potential conserved -10 and -35 promoter regions using the reported promoter regions of *Anabaena* sp. PCC7120 and other cyanobacterial strains (Curtis and Martin, 1994). There were no regions which were homologous to these promoter sequences.

3.3. DNA and amino acid sequence analysis of the *fbp* gene of *Anabaena* sp. PCC7120

The *fbp* gene of *Anabaena* sp. PCC7120 is located between position 21 and 1070 (see Figure 3.9). The gene consists of 1047 nucleotides which encode a

polypeptide of 349 amino acids. The gene starts with an ATG codon and stops with a stop codon of TAA.

The coding sequence of *fbp* encodes a polypeptide of 349 amino acids with an estimated molecular weight of 38,518 daltons. The polypeptide was compared with FBPases from various organisms using the PILEUP program (Figure 3.10). The GAP alignment program was used to find similarity between the FBPase of *Anabaena* sp. PCC7120 and FBPases from other organisms. GAP alignment analysis showed that the *fbp* gene of *Anabaena* sp. PCC7120 was 95% similar and 91% identical to the *fbp* gene of *Nostoc* sp. ATCC29133 (Summer *et al.*, 1995a). The gene is 81% similar and 67% identical to the *fbp* gene of *Synechococcus* sp. PCC7942 (Scanlan *et al.*, 1992), 67% similar and 45% identical to the *fbp* gene of *E. coli* (Hamilton *et al.*, 1988), and 62% similar and 41% identical to the *fbp* gene of the *Arabidopsis thaliana* chloroplast (Horsnell and Raines, 1991).

In mammalian FBPase, a lysine residue was shown to be involved in the catalytic mechanism (Ke *et al.*, 1989). This lysine residue is flanked by a highly conserved region and this region was designated as the active site and used as a signature pattern for FBPases. The amino acid sequence of the FBPase was searched on the MOTIFS program that revealed an active site which consisted of a consensus sequence as follows: GKLRLLYE. The active site residues are absolutely conserved in the FBPase of *Anabaena* sp. PCC7120 (Figure 3.10).

Figure 3.10: Alignement of FBPase from various organisms using the multiple sequence alignment program PILEUP. Abbreviations: An7120, *Anabaena* sp. PCC7120; N29133, *Nostoc* sp. ATCC29133; Eck-12, *E. coli* K-12; Stuber, *Solanum tuberosum* L. The '*' indicates a residue conserved in all three cyanobacterial FBPases and '#' indicates a residue conserved in all FBPases in the figure. Cysteine residues in the *Anabaena* sp. PCC7120 enzyme were highlighted and marked with an arrow over the line. Active site residues were highlighted and underlined.

```
        1                                                    50
An7120  .......... .......... .......... .......... ..........
N29133  .......... .......... .......... .......... ..........
Sy7942  .......... .......... .......... .......... ..........
Eck-12  .......... .......... .......... .......... ..........
Stuber  MAASAATTTS SHLLLSSSRH VASSSQPSIL SPRSLFSNNG KRAPTGVRNH

                                          ↓
        51                                                  100
An7120  .MAKASESLD LSVNESTDKA LDR.....DC TTLSRHVLQQ LQSFSADAQD
N29133  .MAKTPESLE SSINEITDRA LDR.....DC TTLSRHVLQQ LQSFSPDAQD
Sy7942  .MAQSTTS.. ....ETHTRD LDR.....DC TTLSRHVLEQ LQSFSPEAQD
Eck-12  .......... .......... .........M KTLGEFIVEK QHEFSHATGE
Stuber  QYASGVRCMA VAADASETKT AARKKSGYEL QTLTGWLLRQ EMKGEIDA.E
         **    *              ***     ** ** ***** * *****  ***
        101                                                 150
An7120  LSALMNRIAL AGKLVARRLS RAGLMEGVLG FTGEVNVQGE SVKKMDVYAN
N29133  LSALMNRIAL AGKLVARRMS RAGLMEGVLG FTGEVNVQGE SVKKMDVYAN
Sy7942  LAALMQRIGL AAKLIARRLS HAGLVDDALG FTGEINVQGE AVKRMDVYAN
Eck-12  LTALLSAIKL GAKIIHRDIN KAGLVD.ILG ASGAENVQGE VQQKLDLFAN
Stuber  LTIVMSSISL ACKQIASLVQ RAG.ISNLTG VQGAINIQGE DQKKLDVISN
        # *** *# # * #  *** *  ##*    *# **#* #*### ** *####*
        151          ↓                  ↓                  200
An7120  DVFISVFKQS GLVCRLASEE MDEPYYIPEN CPIGRYTLLY DPIDGSSNTD
N29133  DVFISVFKQS GLVCRLASEE MENPYYIPEN CPIGRYTLLY DPIDGSSNTD
Sy7942  QVFISVFRQS GLVCRLASEE MEKPYYIPEN CPIGRYTLLY DPLDGSANVD
Eck-12  EKLKAALKAR DIVAGIASEE .EDEIVVFEG CEHAKYVVLM DPLDGSSNID
Stuber  EVFSNCLRSS GRTGIIASEE EDVPVAVEES YS.GNYVVVF DPLDGSSNID
         ****** ** ******#### *  *****#* *****#**** ######## #
```

Figure 3.10 continued.

```
          201                                                        250
An7120    TNLSLGSIFS I.RQQEGDDS DGQAK..... ........DL LTNGRKQIAA
N29133    NNLSLGSIFA I.RQQEGTDS DGKAT..... ........DL LANGRKQLAA
Sy7942    VDLNVGSIFA V.RRQEFYDE SHEAK..... ........DL LQPGDRQIAA
Eck-12    VNVSVGTIFS IYRRVTPVGT PVTEE..... ........DF LQPGNKQVAA
Stuber    AAVSTGSIFG IYSPNDECIV DDSDDISALG SEEQRCIVNV CQPGNNLLAA
            *  #*##    * **  *     *                ** *  #  * ##
          251                                                        300
An7120    GYILYGPSTM LVYTMGTGVH SFTLDPSLGE FILSEENIRI PDHGAVYSVN
N29133    GYILYGPCTM LVYTIGKGVH SFVLDPSLGE FILTEENIRI PNHGSVYSVN
Sy7942    GYVLYGASTL LVYSMGQGVH VFVLDPSLGE FVLAQSDIQL PNSGQIYSVN
Eck-12    GYVVYGSSTM LVYTTGCGVH AFTYDPSLGV FCLCQERMRF PEKGKTYSIN
Stuber    GYCMYSSSVI FVLTLGKGVF SFTLDPMYGE FVLTQENIEI PKAGRIYSFN
          ## *#*  *  *#*  # ##*  # *###*#* # #      * #  #  ##*#
          301                                                        350
An7120    EGNFWQWEES MREYIRYVHR TE.....GYT ARYSGAMVSD IHRILVQGGV
N29133    EGNFWQWEES IREYIRYVHR TE.....GYS ARYSGAMVSD IHRILVQGGV
Sy7942    EGNFWQWPEG YRQYIREMHR RE.....AYS GRYSGALVAD FHRILMQGGV
Eck-12    EGNYIKFPNG VKKYIKFCQE EDKSTNRPYT SRYIGSLVAD FHRNLLKGGI
Stuber    EGNYQMWDDK LKKYIDDLKD PG.PTGKPYS ARYIGSLVGD FHRTLLYGGI
          ####### *   * ##*  **  *      #   ##*#* # #  ##*# *###*
          351                                                        400
An7120    FLYPGTIQNP EGKLRLLYES APLAFLIQQA GGRATTGLVD ILDVVPKKLH
N29133    FLYPGTIQNP EGKLRLLYET APLAFLIEQA GGRATTGLVN ILDVVPKKLH
Sy7942    FLYPETVKNP TGKLRLLYEA APMAFLAEQA GGKASDGQKP ILLRQPQALH
Eck-12    YLYPSTASHP DGKLRLLYEC NPMAFLAEQA GGKASDGKER ILDIIPETLH
Stuber    YGYPRDAKSK NGKLRLLYEC APMSFIVEQA GGKGSDGHSR VLDIQPTEIH
          **## *  **  ########  *# *#*  ## ## *  #    *#   #  *#
          401                  426
An7120    QRTPLIIGSK EDVAKVESFI QNGH..
N29133    QRTPLIIGSK EDVAKVESFI QNGH..
Sy7942    ERCPLIIGSA ADVDFVEACL AESVP.
Eck-12    QRRSFFVGND HMVEDVERFI REFPDA
Stuber    QRVPLYIGST EEVEKLEKYL A.....
           # *****#   *#  *#
```

A unique region of 12-16 amino acids was reported to be involved in light regulation of chloroplast FBPase activity (Marcus *et al.*, 1988). Two cysteine residues are present in this region. These cysteine residues were reported to be involved in the formation of a disulphide bridge in the dark resulting in inactivation of the enzyme. The disulphide bridge is reduced in the light and the enzyme is reactivated. The alignment of the FBPase of *Anabaena* sp. PCC7120 with FBPases from both *Solanum tuberosum* L. and *Arabidopsis thaliana* chloroplasts revealed that this unique region was not conserved in the FBPase of *Anabaena* sp. PCC7120, raising questions about light regulation of the enzyme via the reduction of a disulphide bridge formed between cysteine residues in the enzyme (Figure 3.11). Therefore, the FBPase enzyme in *Anabaena* sp. PCC7120 is probably not subject to light modulation via the formation/reduction of disulphide bridges in a similar manner to the chloroplast enzyme. However, three cysteine residues Cys-24, Cys-108 and Cys-125 are present in the deduced amino acid sequence of the *fbp* gene of *Anabaena* sp. PCC7120. To understand whether or not these cysteine residues in the *Anabaena* sp. PCC7120 FBPase enzyme are involved in regulation of the enzyme, further analysis of these residues is needed.

3.4 DNA and amino acid sequence analysis of the *tal* gene of *Anabaena* sp. PCC7120

The *tal* gene of *Anabaena* sp. PCC7120 is located between nucleotide positions 1291 and 2436 of the sequenced region (see Figure 3.9). The gene

```
An7120    SIRQQEGDDSDGQA.KDLLTNGRK.Q.............IAAGYIL
N29133    AIRQQEGTDSDGKA.TDLLANGRK.Q.............LAAGYIL
Sy7942    AVRRQEFYDESHEA.KDLLQPGDR.Q.............IAAGYVL
Escoli    SIYRRVTPVGTPVTEEDFLQPGNK.Q.............VAAGYVV
Athchl    GIYSPNDECIVDDS.DDISALGSEEQRCIVNVCQPGNNLLAAGYIL
Stuber    GIYNPNDECLADH.GDDST.LDNIEQKCIVNVCQPGTNLLAAGYCM
Spichl    GIYSPNDECIVDSDHDDESQLSAEEQRCVVNVCQPGDNLLAAGYCM
Consen                                -C-VNVCQPG-NL
```

Figure 3.11. Comparison of the region involved in the light-regulation of activity of chloroplast FBPase enzymes with *Anabaena* sp. PCC7120 (An7120, this study), *Nostoc* sp. ATCC29133 (N29133, Summers *et al.*, 1995a). *Synechococcus* sp. PCC7942 (Sy7942, J. Scanlan, personal communication), *E. coli* (Escoli, Hamilton *et al.*, 1988), *Arabidopsis thaliana* chloroplast (Athchl, Horsnell and Raines, 1991), *Solanum tuberosum* chloroplast (Stuchl, Koßmann *et al.*, 1992), and spinach chloroplast (Spichl, Marcus *et al.*, 1988) FBPases. Consen is the consensus sequence of the insertion region of the chloroplast FBPases.

comprises 1143 nucleotides which encode a polypeptide of 381 amino acids. The gene starts with an ATG codon and stops with a TAG codon.

The *tal* gene encodes a protein of 381 amino acids with an estimated molecular weight of 42,116 daltons. The polypeptide was compared to transaldolases from various other organisms using the PILEUP and GAP alignment programs (Figure 3.12). The results showed that the transaldolase gene of *Anabaena* sp. PCC7120 was closely matched to that of *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995a), these two enzymes being 92% similar and 83% identical to each other. However, the transaldolase enzyme did not exibit similarity with the transaldolases from other organisms as well as it did to that of *Nostoc* sp. ATCC29133. The enzymes of *Anabaena* sp. PCC7120 and *E. coli* (Yura *et al.*, 1988) matched with ten gaps, and also the homology was lower (46% similar and 22% identical). A similar pattern was found when the gene was compared with the enzyme of *Saccharomyces cerevisiae* (Schaaff *et al.*, 1990). The enzymes matched with twelve gaps, and they were overall 49% similar and 28% identical to each other. Consequently the transaldolase enzyme of *Anabaena* sp. PCC7120 is very highly homologous to *Nostoc* sp. ATCC29133 but the homology is rather lower with the transaldolases of *E. coli* and *Saccharomyces cerevisiae*. Database searches showed that there has not yet been an active site identified for transaldolase enzymes.

Figure 3.12: Alignment of transaldolases from various organisms using the multiple sequence alignment program of PILEUP. Abbreviations: An7120, *Anabaena* sp. PCC7120; N29133, *Nostoc* sp. ATCC29133; Eck-12, *E. coli* K-12; Scerev, *Saccharomyces cerevisiae*. The '*' indicates a residue conserved in two cyanobacterial transaldolases and '#' indicates a residue conserved in all transaldolases in the figure.

```
1                                                          50
An7120  ..........  MAINHLLEIK EYGQMYWMDN LSRDIIQSGE LKNLVENQGI
N29133  ..........  MSTNHLLEIK QYGQSIWMDN LTRDIIQSGE LKDMVENQGI
Eck-12  ..........  ..TDKLTSLR QYT.......  .T.VVADTGD IAAMKLYQPQ
Scerev  MSEPAQKKQK VANNSLEQLK ASG.......  .TVVVADTGD FGSIAKFQPQ
                    *  **#****  ***  **** * ******#* **  ***#**
        51                                                 100
An7120  CGITSNPAIF EKAIANNVIY DADIEAGVRA GLPTYKIYES LIFADIRNAC
N29133  SGITSNPAIF EKAIAGNAIY DADIEAGVRA GLPIYKIYES LVFADIRNAC
Eck-12  DATTNPSLIL ..NAAQIPEY RKLIDDAVAW A....KQQSN DRAQQIVDAT
Scerev  DSTTNPSLIL ..AAAKQPTY AKLIDVAVEY G....KKHGK TTEEQ.ENAV
         **#****#* ****# * *# ***#***#** *** *#**** * ******#*
        101                                                150
An7120  DILRPVY..E ASNKLDGYVS IEVPPTIAHD TQATINEARR YY.....QEI
N29133  DILRPVY..E ASNRLDGYVS IEVPPTIAHD TEATISEARR YF.....QEI
Eck-12  DKLAVNIGLE ILKLVPGRIS TEVDARLSYD TEASIAKAKR LIKLYNDAGI
Scerev  DRLLVEFGKE ILKIVPGRVS TEVDARLSFD TQATIEKARH IIKLFEQEGV
        #*#****  #  *** **#**# *##******# # #*# *#** *      ***
        151                                                200
An7120  GRENVMIKIP GTEAGLPAVE QVIA.EGINV NVTLLFSVQS YINTIWAYIR
N29133  GRENLMIKIP GTEPGLPAVE QVIA.EGMNV NITLLFSVES YVNTAWAYVR
Eck-12  SNDRILIKLA STWQGIRAAE QL.EKEGINC NLTLLFSF.. ..........
Scerev  SKERVLIKIA STWEGIQAAK ELEEKDGIHC NLTLLFSF.. ..........
        **** *###** *#* #**#** **** *# ** # ######* * * ** *** *
        201                                                250
An7120  GLEKRLAEGK DISQIASVAS FFLSRIDINI DGKIDAKLAR GVDDISLEAK
N29133  GLEKRLAEGK DISRIASVAS FFLSRIDSNI DAKIDAKLKK GVDDIAVEAK
Eck-12  .AQARACAEA GVFLISPFVG RILDWYKANT DKK.EYAPAE DPGVVSVSEI
Scerev  .VQAVACAEA QVTLISPFVG RILDWYKSST G.K.DYKGEA DPGVISVKKM
        ********** *** #***** **#**** ** * #*****   *****  ***
        251                                                300
An7120  LMVVKGKVAI ANAKIAYQEY KKIIESDQWQ ALAAKGAKVQ RLLWASTSTK
N29133  LKAVKGKVAI ANAKLAYQEY KKIIRSERWQ ELVAKGATVQ RLLWASTSTK
Eck-12  YQYYKE.... .......HGY ETVVMGASFR NIGE...... ..........
Scerev  YNYYKK.... .......YGY KTIVMGASFR STDE...... ..........
        *  *#***** **** ****# **** *  **  * **** ** **********
        301                                                350
An7120  DPNYSDVMYV DELIGPDTVN TLPPATITAC ADHCEVANRV ETGVAEAYQL
N29133  DPNYNDVMYI EELVGPDTVN TVPPATIKAC ADHCNVSDRL ETDVNNAYTL
Eck-12  .........I LELAGCDRL. TIAPALLKEL AESEGAIERK LSYTGEVKAR
Scerev  .........I KNLAGVDYL. TISPALLDKL MNSTEPFPRV LDPVSAKKEA
        **** ****   *# #*#*** # *##** ** **** *  #  ** *  ** *
        351                                                400
An7120  IESLKDPDIN IDI......N AVMDELLIEG INKFVQPFQS LMNSLEGKVK
N29133  IENLKDPDIN IDL......D TVMDELLVEG IDKFIOPFQS LMNSLEDKIK
Eck-12  PARIT....E SEFLWQHNQD PMAVDKLAEG IRKFAIDQEK LEKMIGDLL.
Scerev  GDKISYISDE SKFRFDLNED AMATEKLSEG IRKFSADIVT LDLIEKKVT.
        ** ******* **          ****** ## # *# ***** #****** * *
        401
An7120  LLSPV
N29133  VLSPV
Eck-12  .....
Scerev  A....
         ****
```

3.5 DNA and amino acid sequence analysis of the *zwf* gene of *Anabaena* sp. PCC7120

The downstream part of the 4169 bp fragment contained the *zwf* gene which encodes the G6PDH enzyme. The gene starts with an ATG codon at position 2579 and ends at position 4108 with a TAG stop codon. The gene consists of 1527 nucleotides.

The gene encodes a polypeptide of 509 amino acids with an estimated molecular weight of 58,289 daltons which roughly corresponds to the monomer of G6PDH (see section 1.9.2). The deduced amino acid sequence of the G6PDH of *Anabaena* sp. PCC7120 was compared with the G6PDHs from various other organisms using the PILEUP program (Figure 3.13). GAP alignment analysis of the G6PDH of *Anabaena* sp. PCC7120 was also performed using the G6PDH sequences from individual organisms. The homology among G6PDHs of cyanobacteria including *Anabaena* sp. PCC7120 was very high. The deduced amino acid sequence of G6PDH of *Anabaena* sp. PCC7120 was 96% similar (94% identical) to that of *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995a), and 82% similar (69% identical) to that of *Synechococcus* sp. PCC7942 (Scanlan *et al.*, 1992). However, the homology was low between that of *Anabaena* sp. PCC7120 and both *E. coli* and *Solanum tuberosum* L. cytosolic G6PDHs. The *Anabaena* sp. PCC7120 enzyme was 60% similar (41% identical) to the *E. coli* enzyme (Rowley and Wolf, 1991) and 62% similar (39% identical) to the *Solanum tuberosum* L. enzyme (Graeve *et al.*, 1994). The deduced amino acid

Figure 3.13: Alignment of G6PDHs from various organisms using the multiple sequence alignment program PILEUP. Abbreviations: An7120, *Anabaena* sp. PCC7120; N29133, *Nostoc* sp. ATCC29133; Eck-12, *E. coli* K-12; Athpla, *Arabidopsis thaliana* chloroplast. The '*' indicates a residue conserved conserved in all three cyanobacterial G6PDHs and '#' indicates a residue conserved in all G6PDHs in the figure. Cysteine residues in the *Anabaena* sp. PCC7120 enzyme were highlighted and marked with an arrow over the line. Active site residues were highlighted and underlined.

```
        1                                                    50
An7120  .........M VSLLENPLRV GLQQQGMPEP QI..IVIFGA SGDLTWRKLV
N29133  .........M VSLLENPLRV GLQQQGMPEP QI..IVIFGA SGDLTWRKLV
Sy7942  ........MT PKLLENPLRI GLRQDKVPEP QI..LVIFGA TGDLTQRKLV
Eck-12  .......... .......... .MAVTQTAQA CD..LVIFGA KGDLARRKLL
Athpla  MAASWCIEKR GSIRNDSFR. ..DNDNIPET GCLSIIVLGA SGDLAKKKTF
                     *******  ** *   *** **   ***##  ###* *#**
        51                                                   100
An7120  PALYKLRRER RIPP.ETTIV GVARREWSHE YFREQMQKGM EEAHSSVELG
N29133  PALYKLRRER RIPP.ETTIV GVARREWSHE YFREQMQKGM EEAHPDVDLG
Sy7942  PAIYEMHLER RLPP.ELTIV GVARRDWSDD YFREHLRQGV EQFGGGIQAE
Eck-12  PSLYQLEKAG QLNP.DTRII GVGRADWDKA AYTKVVREAL ETFMKETIDE
Athpla  PALFNLYRQG FLQSNEVHIF GYARTKISDD DLRSRIRGYL SQ...GKENE
        #* *    ** * ** * *#* #**#* **   ****    *  *
        101          ↓                                       150
An7120  ELWQDFSQGL FYCPGDIDNP ESYQKLKNLL SELD...EKR GTRGNRMFYL
N29133  ELWQDFSQGL FYSPGDIDNP ESYQKLKTLL SELD...EKR GTRGNRMFYL
Sy7942  EVWNTFAQGL FFAPGNIDDP QFYQTLRDRL ANLD...ELR GTRGNRTFYL
Eck-12  GLWDTLSARL DFCNLDVNDT AAFSRLGAML DQKN...RI. .....TINYF
Athpla  GEVSEFLQLI KYVSGSYDSA EGFTSLDKAI SEHEFSKNST EGSSRRLFYF
        * *  * *** *  ** ** *   ** #   *   **   * * ****** *#*
        151                                                  200
An7120  SVAPNFFPEA IKQLGGAGML DD..PYKHRL VIEKPFGRDL ASAQSLNAVV
N29133  SVAPSFFPEA IKQLGSGGML ED..PYKHRL VIEKPFGRDL ASAQSLNQVV
Sy7942  SVAPRFFGEA AKQLGAAGML AD..PAKTRL VVEKPFGRDL SSAQVLNAIL
Eck-12  AMPPSTFGAI CKGLGEAKLN AK..PA..RV VMEKPLGTSL ATSQEINDQV
Athpla  ALPPSVYPSV CRMIKSYCMN KSDLGGWTRT VVEKPFGKDL ASSEQLSSQI
        ***# ** **  ****  ***  *  *   #* # ###*#**#  *** **
```

Figure 3.13 continued.

```
         201↓                                                      250
An7120   QKYCKEHQVY RIDHYLGKET VQNLLVFRFA NAIFEPLWNR QFVDHVQITV
N29133   QKYCKEHQVY RIDHYLGKET VQNLLVFRFA NAIFEPLWNR QFVDHVQITV
Sy7942   QNVCRESQIY RIDHYLGKET VQNLLVFRFA NAIFEPLWNR QYIDHVQITV
Eck-12   GEYFEECQVY RIDHYLGKET VLNLLALRFA NSLFVNNWDN RTIDHVEITV
Athpla   GELFDEPQIY RIDHYLGKEL VQNLLVLRFA NRFFLPLWNR DNIDNIQIVF
         *  * # # # #########* #*######### #**#***#** *  #***#**
         251                                 ↓                300
An7120   AETVGVEDRA GYYEKAGALR DMLQNHLMQL YCLTAMEAPN SMDADSIRTE
N29133   AETVGVEDRA GYYESAGALR DMLQNHLMQL YCLTAMEAPN AMDADSIRTE
Sy7942   AETVGLEGRA GYYETAGALR DMVQNHLMQL FSLTAMEPPN SLGADGIRNE
Eck-12   AEEVGIEGRW GYFDKAGQMR DMIQNHLLQI LCMIAMSPPS DLSADSIRDE
Athpla   REDFGTEGRG GYFDEYGIIR DIIQNHLLQV LCLVAMEKPV SQKPEHIRDE
         *###* #### #### *###* ## ######*#   **##* ##    ** ## #
         301                                                350
An7120   KVKVLQ.ATR LADVHNLSRS AIRGQYSAGW MKGQQVPGYR TEPGVDPNSS
N29133   KVKVLQ.ATR LADVHNLSRS AVRGQYSAGW MKGQAVPGYR TEPGVDPNST
Sy7942   KVKVVQ.ATR LADIDDLSLS AVRGQYKAGW MNGRSVPAYR DEEGADPQSF
Eck-12   KVKVPEVSSP HRPLQRTRKN RTRAIYCVP. .QGKKVPGYL EEEGANKSSN
Athpla   KVKVLQSMLP IEDEE..... VVLGQY.... ......EGYK DDPTVPNNSN
         #### * **# ***   ** * * ***# *** * *  ** #*  * * ** #
         351                                                400
An7120   TPTYVGMKFL VDNWRWQGVP FYLRTGKRMP KKVSEISIHF RDVPSRMFQS
N29133   TPTYVAMKFL VDNWRWKGVP FYLRTGKRMP KKVSEIAIHF REVPSRMFQS
Sy7942   TPTYVAMKLL VDNWRWQGVP FYLRTGKRMP KKVTEIAIQF KTVPHLMFQS
Eck-12   TETFVAIRVD IDNWRWAGVP FYLRTGKRLP TKCSEVVVYF KTPELNLFKE
Athpla   TPTFATMVLR IHNERWEGVP FIMKAGKALN SRKAEIRVQF KDVPGDIFRC
         #*#** ** * **#*## ### #*********** *** #* * #   **  *#**
         401                                                450
An7120   AAQQRN.ANI LAMRIQPNEG ISLRFDVKMP GAEFRS..RS VDMDFSY.GS
N29133   AAQQTN.ANI LTMRIQPNEG ISLRFDVKMP GAEFRT..RS VDMDFSY.GS
Sy7942   ATQKVNSPNV LVLRIQPNEG VSLRFEVKTP GSSQRT..RS VDMDFRY.DT
Eck-12   SWQDL.PQNK LTIRLQPDEG VDIQVLNKVP GLDHKHNLQI TKLDLSYSET
Athpla   QKQGR...NE FVIRLQPSEA MYMKLTVKKP GLEMSTVQS. .ELDLSYGQR
         * #  *  #  *  #*###*#*  **** *# # #   *   ** ***#* #
         451              ↓                                 500
An7120   FGIEATSDAY DRLFLDCMMG DQTLFTRADE VEAAWQVVTP ALSVWDSPAD
N29133   FGIQATSDAY DRLFLDCMMG DQTLFTRADE VEAAWQVVTP ALSVWDAPAD
Sy7942   AFGSPTQEAY SRLLVDCMLG DQTLFTRADE VEASWRVVTP LLESWDDPRQ
Eck-12   FNQTHLADAY ERLLLETMRG IQALFVRRDE VEEAWKWVDS ITEAWAMDND
Athpla   YQGVVIPEAY ERLILDTIRG DQQHFVRRDE LKAAWEIFTP LLH.RIDNGE
             *  ##  ##  *** # *#*****#*## *** # ****  *  ** *
         501                                              546
An7120   PATIPQYEAG TWEPAEAEFL INQDG..RR. .WRRL .... ......
N29133   PTTIPQYEAG TWEPEQAELL INQDG..RR. .WRRL..... ......
Sy7942   AAGISFYEAG TWEPAEAEQL INRDGAVGV. .VSRIPATOL NSSGDV
Eck-12   AP..KPYQAG TWGPVASVAM ITRDG..RS. .WNEFE.... ......
Athpla   VKPIP.YKPG SRGPAEADEL LQNAGYVQTH GYIWIPPTL. ......
               # *# ***#  ** * ** *#         *
```

sequence from *Anabaena* sp. PCC7120 was therefore consistently homologous to the enzymes from the other organisms and the gene was designated *zwf*.

A lysine residue was identified as a reactive nucleophile associated with the activity of the G6PDH enzyme (Jeffery *et al.*, 1989). The sequence surrounding this lysine residue is composed of the amino acids DHYLGKE and is totally conserved among various organisms. This pattern can be used as a signature pattern for G6PDH. This active site pattern in the *Anabaena* sp. PCC7120 enzyme is located between amino acid positions 196 and 202, and is absolutely conserved (Figure 3.13).

The oxidation/reduction of a disulphide bridge between two cysteine residues in the enzyme has been proposed to be involved in light modulation of cyanobacterial G6PDHs (Cossar *et al.*, 1984) (also see section 1.9.3.4). Four cysteine residues are present in the *Anabaena* sp. PCC7120 G6PDH at positions 101, 187, 265 and 445. Cysteines 187, 265 and 445 are also conserved in the *Nostoc* sp. ATCC29133 G6PDH. Only two cysteines are found in *Synechococcus* sp. PCC7942 G6PDH, Cys-187 and Cys-445. These two cysteines Cys-187 and Cys-445 are, therefore, conserved in all cyanobacterial G6PDHs which have been sequenced to date, consistent with the possibility that they are involved in the light modulation of the enzyme.

3.6 Conclusions

A 4,169 bp region around *zwf* gene of *Anabaena* sp. PCC7120 was sequenced on clone pAG75 which included a 7.5 kb *zwf* fragment in pBR325. The whole sequence was completed by sequencing three subfragments and joining these fragments by using oligonucleotide primers. All the sequence data were obtained on single-stranded or double-stranded templates by using the dideoxy chain termination method. Most of the templates were from single-stranded M13mp18 or mp19 subclones, double-stranded templates were occasionally used. Subfragments were cloned into M13 either directly into two restriction sites in the multiple cloning site of the vector or randomly into a single site (which was usually the *Sma*I site). Three subfragments were subcloned into pUC19 to handle a particular fragment more easily. These subclones were pAG14, pAG39 and pAG35 containing 1.1 kb *Eco*RI/*Hpa*I, 4 kb *Xba*I/*Hin*dIII and 3.5 kb *Eco*RI fragments, respectively.

The nucleotide sequence of the *zwf* gene was obtained by sequencing a 2.4 kb *Hpa*I/*Hin*dIII subfragment which was cloned into M13mp18 and mp19. The *zwf* gene was located in the downstream part of the fragment. The upstream part of the fragment also contained part of an ORF for which no homology was found initially with any gene in the current databases. When more data were accumulated in the databases, this ORF was identified as the transaldolase gene and designated *tal*. The upstream part of the 1.1 kb *Eco*RI/*Hpa*I fragment also included the downstream part of the *fbp* gene. The whole sequence of the gene

was completed by sequencing a 0.6 kb *Xba*I/*Eco*RI fragment cloned randomly into M13mp18. A small part of the gene upstream from *Xba*I site was sequenced using oligonucleotide primers on M13mp18 templates. Thus, the sequence of a 4,169 bp region of the *Anabaena* sp. PCC7120 *zwf* chromosome was completed. The organisation of the three genes in this region was *fbp-tal-zwf* from 5' to 3'. The GC content of the region is 45.5% which is 3% higher than the value was found for the whole genome of *Anabaena* sp. PCC7120.

The *fbp* gene of *Anabaena* sp. PCC7120 consists of 1047 nucleotides which encode a polypeptide of 349 amino acids. The deduced amino acid sequence of the gene has a high homology to the FBPases from various organisms. An active site region with a reactive lysine residue is absolutely conserved. Three cysteine residues exist in the enzyme, which are Cys-24, Cys-108 and Cys-125. These cysteine residues may not be involved in the light-modulation of the enzyme since a 13 amino acid region with two cysteines which is involved in the light modulation of chloroplast FBPases is not conserved in the *Anabaena* sp. PCC7120 FBPase.

The transaldolase gene, *tal*, of *Anabaena* sp. PCC7120 is composed of 1143 nucleotides which encode a polypeptide of 381 amino acids. The deduced amino acid sequence of the gene has a very high homology with that of *Nostoc* sp. ATCC29133 but the homology to transaldolase enzymes from *E. coli* and *Saccharomyces cerevisiae* is rather poor.

The *zwf* gene of *Anabaena* sp. PCC7120 consists of 1527 nucleotides encoding a polypeptide of 509 amino acids. The deduced amino acid sequence of the gene is highly homologous to the G6PDH from other organisms. The cyanobacterial G6PDHs are rather more homologous to each other. An active site pattern with a reactive lysine residue is absolutely conserved in the enzyme of this organism. Four cysteine residues exist in the *Anabaena* sp. PCC7120 G6PDH, which are Cys-101, Cys-187, Cys-265 and Cys-445. Two of these cysteines, Cys-187 and Cys-445, are conserved in all three of the cyanobacterial G6PDHs which have been sequenced to date (*Anabaena* sp. PCC7120, *Nostoc* sp. ATCC29133 and *Synechococcus* sp. PCC7942). This is consistent with the possibility that cysteine residues may be involved in the light modulation of G6PDH.

The arrangement of the *fbp*, *tal* and *zwf* genes is the same as that reported for *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995a). However, in the unicellular strain *Synechococcus* sp. PCC7942 there is no *tal* gene between *fbp* and *zwf* (J. Scanlan, personal communication). Thus, the gene organisation at *zwf* regions of the filamentous, heterocystous strains and the unicellular strain is not identical.

Chapter 4

Complementation and mutagenesis of the *zwf* gene of *Anabaena* sp. PCC7120

4.1. Introduction

The *zwf* gene of *Anabaena* sp. PCC7120 has already been characterised by nucleotide sequencing and comparison of the nucleotide and amino acid sequences with various other organisms (see Chapter 3). In this chapter, the attempts to achieve genetic complementation of a *zwf*-deficient *E. coli* strain DF214 by the *zwf* gene of *Anabaena* sp. PCC7120 will be described. In addition, the attempt at mutagenesis of the *zwf* gene of *Anabaena* sp. PCC7120 will also be discussed.

Since there were no previous reports describing the use of a cyanobacterial *zwf* gene to complement a *zwf*-deficient *E. coli* strain, it was of interest to complement such a mutant using the *zwf* fragment of *Anabaena* sp. PCC7120. An *E. coli* strain deficient for both the phosphoglucose isomerase gene (*pgi*) and for *zwf* (*E. coli* DF214) was reported by Vinopal *et al.* (1975). Mutagenesis of these genes had been performed by insertion of the phage Mu into the *pgi* gene and deletion of the *zwf* gene. The strain *E. coli* DF214 had apparently lost both phosphoglucose isomerase and G6PDH activity and could not metabolise glucose unless one of these genes was complemented by a foreign gene. The objective was to determine whether a *zwf* fragment of *Anabaena* sp. PCC7120 could complement the mutant enabling this strain to metabolise glucose and grow on a glucose-containing minimal medium.

It was also of interest to produce a *zwf* null mutant of *Anabaena* sp. PCC7120. The phenotypic analysis of such a mutant would allow some insights

into the role and regulation of the enzyme *in vivo*. It would also reveal the role of G6PDH in heterocysts by following any change in nitrogen fixation ability and the cells ability to survive under nitrogen-starved conditions. Site-directed mutagenesis of the conserved cysteine residues was also proposed, to determine whether a disulphide bridge was involved in the light-modulation of the enzyme. To replace the genomic copy of the *zwf* gene with a cloned copy of the gene in which cysteine codons would be modified, a deletion mutation of the *zwf* gene was first needed. Deletion-insertion mutagenesis of the *zwf* was, therefore, the primary target for the mutagenesis strategy. Sequencing of the *zwf* region of *Anabaena* sp. PCC7120 has already given more information about the restriction sites which could be used for deletion of most of the gene. Consequently a strategy to delete most of the *zwf* gene of *Anabaena* sp. PCC7120 was undertaken.

4.2. Complementation of *Escherichia coli* DF214 with *zwf* fragments of *Anabaena* sp. PCC7120

E. coli strain DF214 is a double mutant of phosphoglucose isomerase and glucose-6-phosphate dehydrogenase (G6PDH) (Vinopal *et al.*, 1975). This double mutant can not metabolise glucose either via fermentation because of its lack of a functional phosphoglucose isomerase or via the oxidative pentose phosphate pathway because of its lack of a functional G6PDH. When the cells of *E. coli* DF214 grow on a minimal medium supplemented with glucose, the cells should be able to grow only if one of the *pgi* or *zwf* genes are being expressed *in trans*. The clone pAG75 which included a 7.5 kb *zwf* fragment of *Anabaena* sp. PCC7120 and three subclones which included smaller *zwf* fragments of *Anabaena*

sp. PCC7120 were used to attempt to complement the *zwf* deficient mutant *E. coli* DF214.

4.2.1. Complementation of *E. coli* DF214 with pAG75

The plasmid pAG75 is a pBR325 based clone including a 7.5 kb *zwf* fragment of *Anabaena* sp. PCC7120 (see Section 3.1). The *zwf* gene is located in the downstream part of the insert (see Section 3.2.5). The cells of *E. coli* DF214 were grown in a minimal medium (see Section 2.4) supplemented with casein hydrolysate and L-tryptophan and uracil (as described in Section 2.4). Competent cells were transformed with pAG75 DNA as described in Section 2.17.1. Transformants were plated on M63 minimal plates containing ampicillin and plates were incubated for about 24 hours. Positive (*E. coli* TG1) and negative controls (*E. coli* DF214) were inoculated simultaneously. The colonies resulting from transformation with pAG75 were still smaller than wild-type strain *E. coli* TG1. Both of the control cells had grown on M63 medium without antibiotic selection.

E. coli DF214 cells when transformed with pAG75 were grown on M63 minimal medium containing ampicillin and gave approximately 100 colonies per plate. Surprisingly, negative control *E. coli* DF214 cells when grown on the minimal medium without ampicillin, grew as well as the transformants did. This result raised questions about positive selection of complemented cells on glucose containing minimal medium. Obviously, having ampicillin resistance, pAG75 transformed *E. coli* DF214 cells presumably also contained the *zwf* fragment in

the cell. The double mutant *E. coli* DF214 could only grow on glucose containing M63 medium if one of the mutations in these genes was leaky or if supplements of the minimal medium such as 1% casein could be used as an energy source or both. To understand why the mutant cells were growing on minimal medium, M63 minimal plates were prepared and supplemented with a concentration range of casein hydrochloride, from 0% to 1%, and the mutant cells were inoculated onto these plates. The mutant cells grew on all the plates in approximately the same number, but the colony sizes were dramatically reduced when the casein concentrations were lower than 0.05%. This suggested that casein was being used as both an energy and carbon source, but the cells grown on 0% casein containing minimal medium also suggested that one or both of the mutations were leaky.

A single colony from each of the control strains and five colonies from pAG75 transformed *E. coli* DF214 were further analysed by assaying G6PDH activity. Each colony was grown in two different 10 ml cultures one of which was M63 minimal medium and the other was 2x YT medium. This would reveal whether G6PDH activity changed when the cells grew in nutrient agar medium. The cells were prepared for G6PDH assays as described in Section 2.23 and G6PDH activity was assayed as described in Section 2.24.1. The results showed that there was no difference between the G6PDH activity from M63 grown and from nutrient agar grown cells. Therefore, all the experiments were performed by using 2x YT medium, transformants were selected against antibiotic resistance, and G6PDH assays were performed for determination of possible complementation of *E. coli* DF214. Wild-type G6PDH activity was 121.7 units

mg^{-1} protein. The activity in the *zwf* mutant *E. coli* DF214 was 2.6 units mg^{-1} protein, corresponding to 2% of wild-type activity (see Figure 4.1). This result explained how the mutant cells grew on minimal medium without casein. This low level G6PDH activity may correspond to a limited function of the *zwf* gene in the cell. It may also correspond to a low level of phosphoglucose isomerase activity allowing glucose to enter the metabolism through which NADH was produced. In fact, a 2.8% phosphoglucose isomerase activity was reported in the mutant *E. coli* DF214 cells (Vinopal *et al.*, 1975). As seen in Figure 4.1, G6PDH activity in pAG75 transformed *E. coli* DF214 cells was similar to that of the mutant except one colony which has 15.4% of wild-type activity, implying that the mutant might not complemented by the *Anabaena* sp. PCC7120 *zwf* fragment.

To understand whether the product of the *zwf* gene of *Anabaena* sp. PCC7120 was translated in *E. coli* DF214, Western blotting analysis was employed against the enzymes from the wild-type *E. coli* DF214, the mutant and pAG75 transformed *E. coli* cells. An antibody was produced against the protein encoded by part of the *Anabaena* sp. PCC7120 *zwf* gene in this laboratory (Sundaram *et al.*, in press). This antibody was used for the detection of protein encoded by the *zwf* gene of *Anabaena* sp. PCC7120 when transformed into *E. coli* DF214. Electrophoresis of proteins was as described in Sections 2.25.1 and 2.25.3. The Western blotting procedure was as described in Section 2.25.4. The result of the Western blot showed that the antibody reacted with a protein of approximately 70 kDa from the supernatant of *Anabaena* sp. PCC7120, but no protein of similar molecular weight to this protein was observed in the pAG75

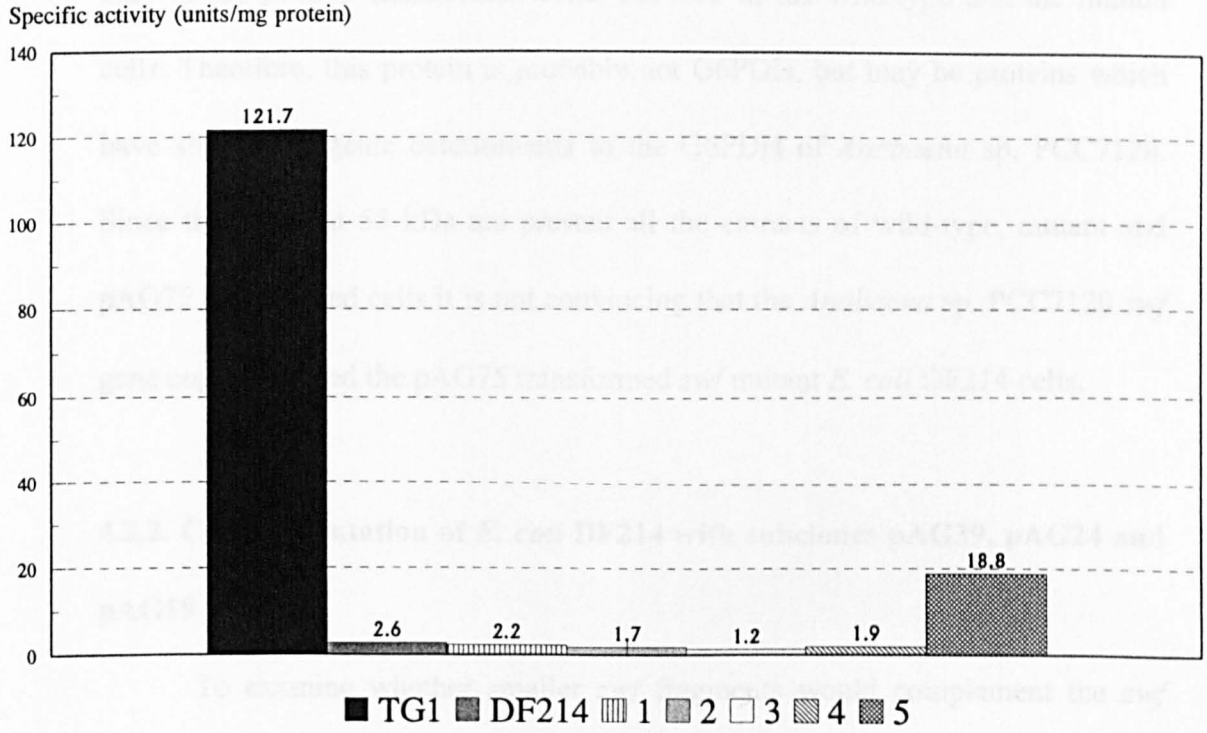


Figure 4.1: G6PDH activities in extracts from *E. coli* TG1 as wild type, *E. coli* DF214 as *zwf* mutant, and from pAG75 transformed *E. coli* DF214 (1-5). The numbers over the bars are the G6PDH activities (units/mg protein).

transformed *E. coli* DF214 (see Figure 4.2). Two proteins, one of 55 kDa and the other of 31 kDa reacted with the antibody. The 55 kDa band corresponds to the the size of *Anabaena* sp. PCC7120 G6PDH. The 31 kDa band was observed not only in the pAG75 transformed cells, but also in the wild-type and the mutant cells. Therefore, this protein is probably not G6PDH, but may be proteins which have similar antigenic determinants to the G6PDH of *Anabaena* sp. PCC7120. Since the bands at 55 kDa are present all the extracts of wild-type, mutant and pAG75 transformed cells it is not convincing that the *Anabaena* sp. PCC7120 *zwf* gene copmlemented the pAG75 transformed *zwf* mutant *E. coli* DF214 cells.

4.2.2. Complementation of *E. coli* DF214 with subclones pAG39, pAG24 and pAG19

To examine whether smaller *zwf* fragments would complement the *zwf* mutant *E. coli* DF214, three *zwf* subfragments from the clone pAG75 were subcloned into pUC19 and pBR325. The subclone pAG39 was a pUC19-based subclone including a 4 kb *Xba*I/*Hin*dIII *zwf* fragment and has been characterised in Section 3.2.2 (see also Figure 4.3 and 4.4). A 2.5 kb *Hpa*I/*Hin*dIII *zwf* fragment was also subcloned into the *Hin*cII and *Hin*dIII sites of pUC19 to yield pAG24 (see Figure 4.3 and 4.4). A deletion strategy was employed for producing a subclone with a 1.9 kb *Cla*I/*Hin*dIII *zwf* fragment as insert. A *Cla*I site was present near the end of the *tal* gene which is unique to the 7.5 kb *zwf* fragment which was cloned into the *Hin*dIII site of pBR325. Another *Cla*I site unique for pBR325 also exists at 3 bp upstream from the *Hin*dIII site into which the 7.5 kb fragment was cloned (for the structure of pBR325, see Figure 3.1). Digestion of

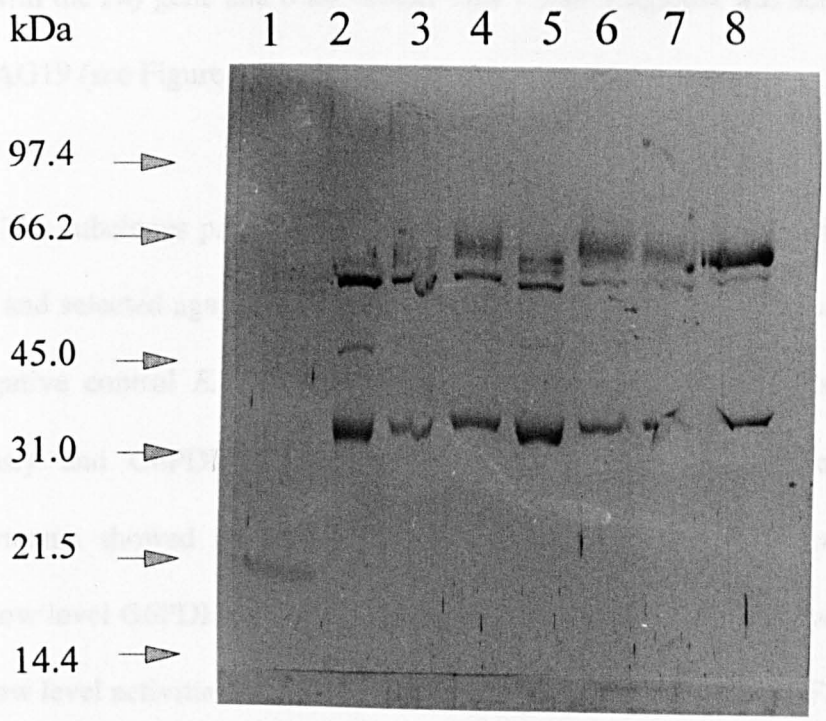


Figure 4.2: Western blot of pAG75 transformed *E. coli* DF214. 50 µl protein was loaded to each line of a SDS-polyacrylamide gel. Lane 1 is the supernatant from *Anabaena* sp. PCC7120 used as a control for cyanobacterial G6PDH. The band at 55 kDa has been characterised as G6PDH (Sundaram *et al.*, in press). Lane 2 is supernatant from *E. coli* TG1 used as wild type in respect to *zwf*. Lane 3 is the supernatant from the *zwf* mutant *E. coli* DF214. Lanes 4-8 are the supernatants from the cells transformed with pAG75. The supernatants was same as those used for G6PDH assay (see Figure 4.1).

pAG75 with *Cla*I produced two fragments; one 5.6 kb insert part of the clone without the *zwf* gene and the other of 7.9 kb which included a 1.9 kb part of the insert with the *zwf* gene and 6 kb vector. This 7.9 kb fragment was self-ligated to yield pAG19 (see Figure 4.3 and 4.4).

The subclones pAG19, pAG24 and pAG39 were transformed into *E. coli* DF214 and selected against ampicillin. The positive control was *E. coli* TG1 and the negative control *E. coli* DF214. Supernatants were prepared as described previously and G6PDH activities were assayed (Figure 4.5). None of the transformants showed higher activities than negative control, *E. coli* DF214. These low level G6PDH activities from pAG19, pAG24 and pAG39 were similar to the low level activities obtained from pAG75 transformed *E. coli* DF214 cells.

The results suggested that the *zwf* gene of *Anabaena* sp. PCC7120 did not complement the *zwf* mutant *E. coli* DF214 cells. Many cyanobacterial genes have been reported to be subject to regulation by activator proteins (Curtis and Martin, 1994). Absence of these possible activator protein(s) in *E. coli* cells may affect the initiation of transcription from an *Anabaena* sp. PCC7120 *zwf* fragment.

4.3 Mutagenesis of the *zwf* gene of *Anabaena* sp. PCC7120

As mentioned previously (see Section 4.1), a deletion-insertion strategy was employed for the mutagenesis of the *zwf* gene of *Anabaena* sp. PCC7120. To delete the major part of the *zwf* gene a smaller subfragment, the 2.4 kb *Hpa*I/*Hin*dIII fragment, would be subcloned into a suitable vector. This 2.4 kb

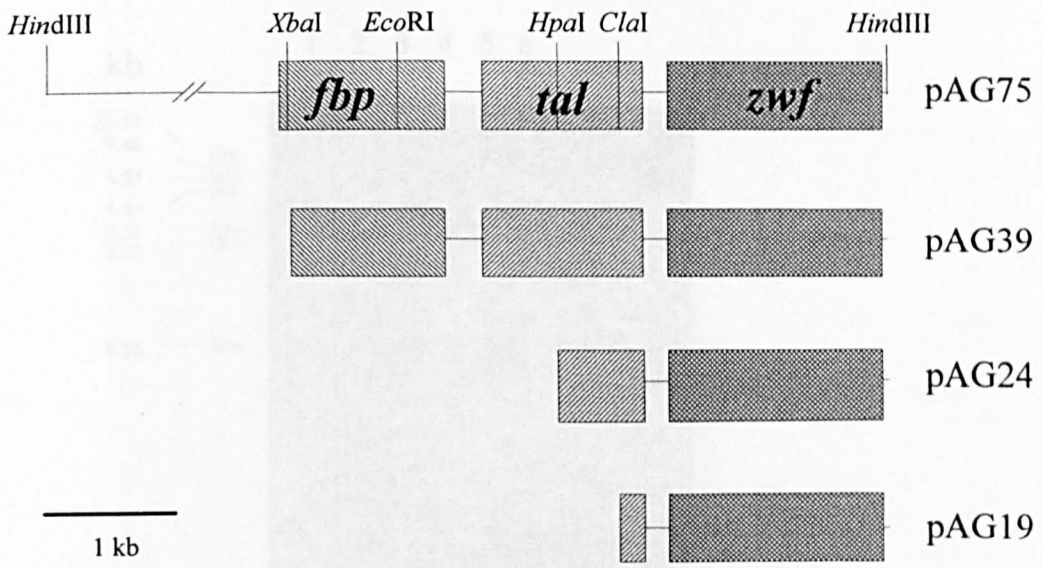


Figure 4.3: Inserts of pAG75, pAG39, pAG24, pAG19. The clone pAG75 contains a 7.5 kb *Hin*dIII fragment. The subclone pAG39 a 4 kb *Xba*I/*Hin*dIII fragment, pAG24 a 2.4 kb *Hpa*I/*Hin*dIII fragment, and 1.9 kb *Cla*I/*Hin*dIII fragment. pAG75 and pAG19 are pBR325 based clones; pAG39 and pAG24 are pUC19 based clones.

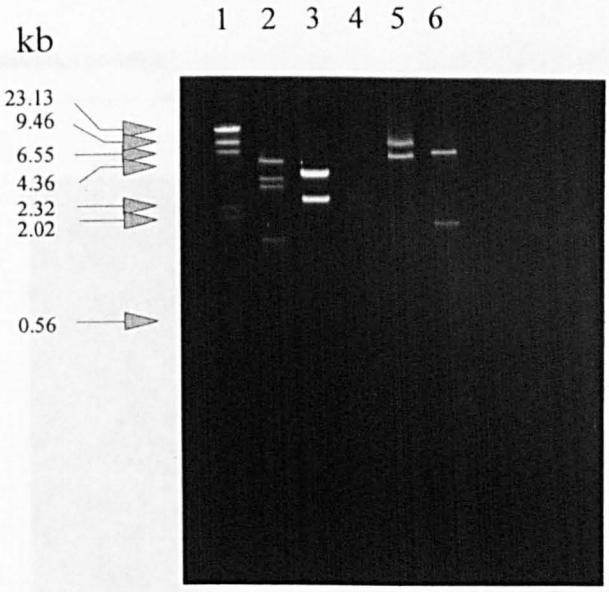


Figure 4.4: Restriction digestion of pAG75, pAG39, pAG24 and pAG19. The first lane is *Hin*dIII digested phage λ DNA used as molecular size marker. Lane 2 is *Eco*RI digested pAG75. The bands on the top and at the bottom are the parts of the vector pBR325 with small parts of the insert. The bands at the middle are parts of the insert. Lane 3 is *Eco*RI digested pAG39. The upper band is the insert and the lower band the vector pUC19 with a part of the insert. Lane 4 is *Eco*RI digested pAG24. The lower band is the insert and the upper the vector, pUC19 with a part of the insert. Lane 5 is *Cla*I digested pAG75. The upper band is the vector pBR325 and a 1.9 kb part of the insert containing *zwf*, and the lower band the 5.6 kb part of insert. Lane 6 is *Cla*I and *Hin*dIII digested pAG19. The lower band is 1.9 kb insert and the upper band is the vector pBR325.

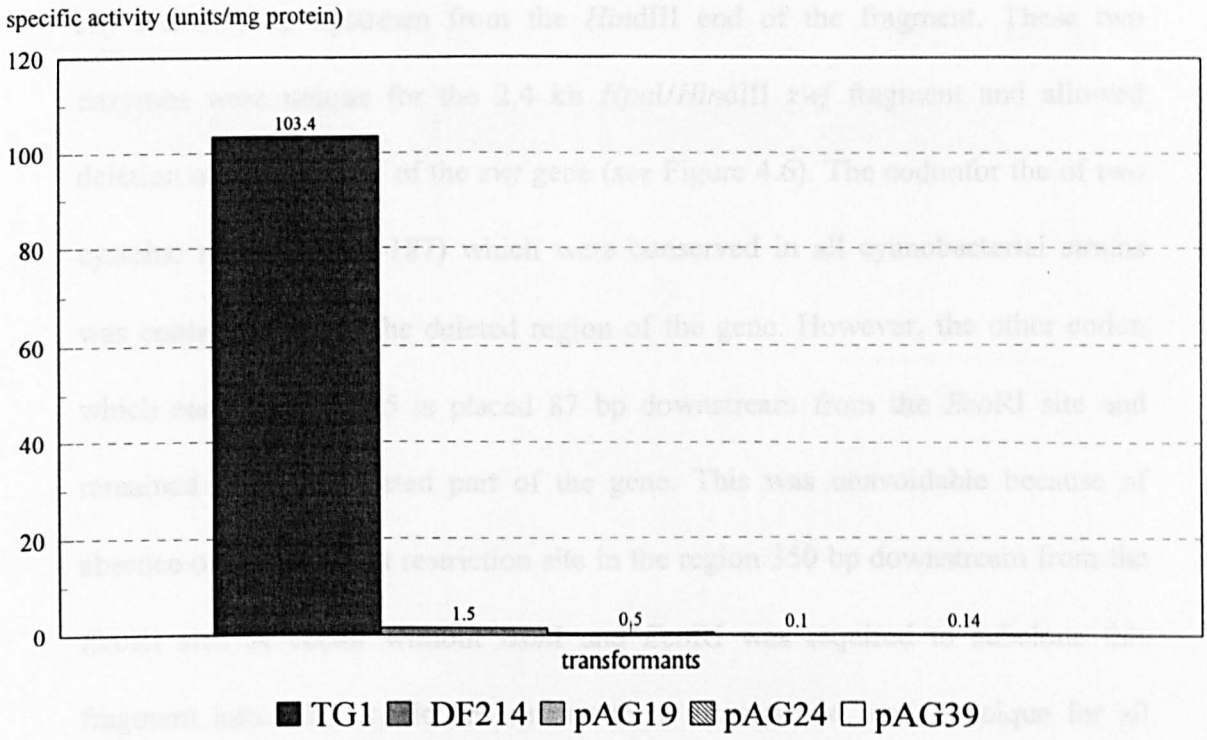


Figure 4.5: G6PDH activities from *E. coli* TG1 as wild type, *E. coli* DF214 as *zwf* mutant, and from *E. coli* DF214 transformed with pAG19, pAG24 and pAG39. The numbers over the bar are the G6PDH activities (units/mg protein).

*Hpa*I/*Hin*dIII fragment contained the *zwf* gene and also carried two convenient restriction sites which were *Bsm*I and *Eco*RI. The *Bsm*I site was located 147 bp downstream from the start codon of *zwf* and 950 bp downstream from the *Hpa*I site of the fragment; the *Eco*RI site was located 287 bp upstream from the end of *zwf* and 350 bp upstream from the *Hin*dIII end of the fragment. These two enzymes were unique for the 2.4 kb *Hpa*I/*Hin*dIII *zwf* fragment and allowed deletion of the majority of the *zwf* gene (see Figure 4.6). The codonfor the of two cysteine residues (Cys-187) which were conserved in all cyanobacterial strains was contained within the deleted region of the gene. However, the other codon which encodes Cys-445 is placed 87 bp downstream from the *Eco*RI site and remained in the undeleted part of the gene. This was unavoidable because of absence of a convenient restriction site in the region 350 bp downstream from the *Eco*RI site. A vector without *Bsm*I and *Eco*RI was required to subclone this fragment into. This would ensure that these sites would remain unique for all subclones. There is no *Bsm*I site and only one *Eco*RI site in pUC19, and therefore, pUC19 was deemed to be the most convenient vector to subclone the 2.4 kb *Hpa*I/*Hin*dIII fragment into after deleting the *Eco*RI site.

4.3.1 Construction of the subclone pAG25 for deletion of *zwf*

The *Eco*RI site is the first site in the multiple cloning site of pUC19. Deletion of this site would leave *Eco*RI a unique restriction site in a subclone containing the 2.4 kb *Hpa*I/*Hin*dIII fragment. To delete this site, pUC19 was digested with the restriction enzymes *Eco*RI and *Bam*HI which cut pUC19 at the multiple cloning site (for the structure of the multiple cloning site of pUC19 see

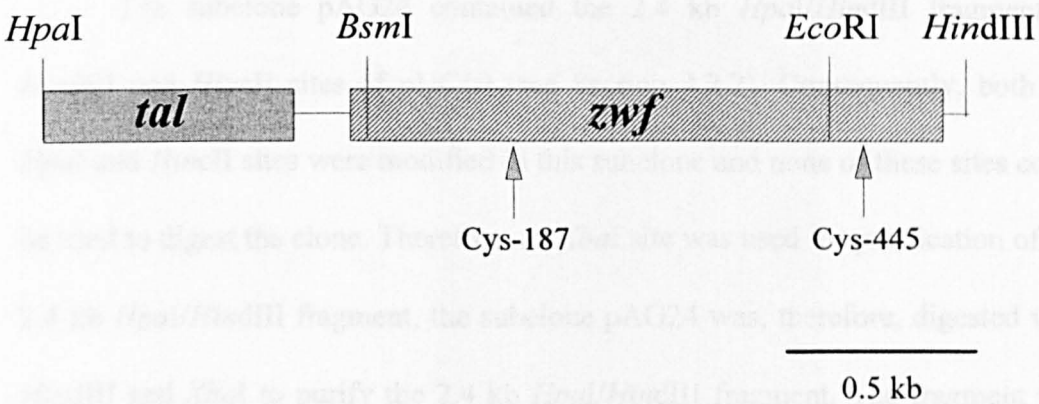


Figure 4.6 Physical map of the 2.4 kb *Hpa*I/*Hin*dIII fragment of *Anabaena* sp. PCC7120. The downstream part of *tal* gene is seen upstream from *zwf*. Positions of the cysteines which are conserved in cyanobacterial strains are shown on the gene. *Bsm*I and *Eco*RI sites are unique for this fragment.

Figure 3.3). The fragment was then end-filled and self-ligated to yield pUC19x. The restriction sites for *Eco*RI, *Bam*HI and the others between these two sites in the multiple cloning site were not represented in pUC19x (see Figure 4.7).

The subclone pAG24 contained the 2.4 kb *Hpa*I/*Hin*dIII fragment in *Hin*dIII and *Hin*cII sites of pUC19 (see Section 4.2.2). Consequently, both the *Hpa*I and *Hin*cII sites were modified in this subclone and none of these sites could be used to digest the clone. Therefore, an *Xba*I site was used for purification of the 2.4 kb *Hpa*I/*Hin*dIII fragment, the subclone pAG24 was, therefore, digested with *Hin*dIII and *Xba*I to purify the 2.4 kb *Hpa*I/*Hin*dIII fragment. The fragment was then subcloned into the *Hin*dIII and *Xba*I sites of pUC19x to yield pAG25 (see Figure 4.7 and Figure 4.8). *Eco*RI and *Bsm*I sites were unique for this subclone (see Figure 4.7).

4.3.2 Deletion-insertion mutation of the *zwf* gene of *Anabaena* sp. PCC7120

Double digestion with *Bsm*I and *Eco*RI of pAG25 produced two fragments, one was a 1.1 kb internal *zwf* fragment and the other a 3.9 kb fragment consisting of vector and the remaining part of the insert. The *Bam*HI and *Eco*RI ends of this 3.6 kb fragment were filled to obtain blunt ends. This was necessary to ligate a *Sma*I digested interposon Ω (Prentki and Krisch, 1984) into these filled sites (for characterisation of the interposon, Ω, see Section 5.3). Several attempts to insert the interposon Ω into the *Bsm*I/*Eco*RI deleted *zwf* gene in pAG25 were

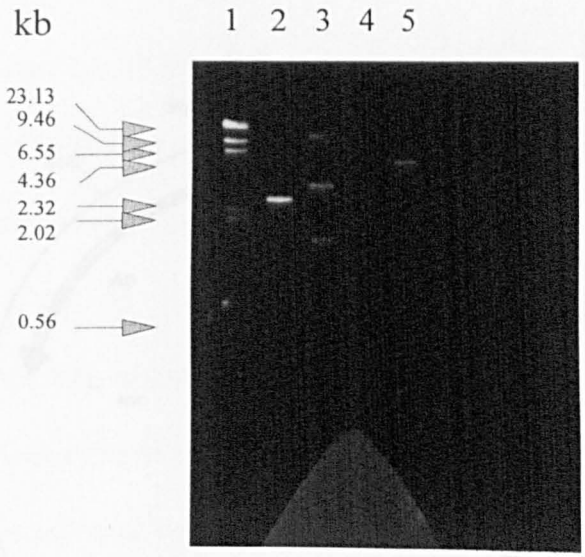


Figure 4.7: Restriction digestion of pUC19, pUC19x, pAG24 and pAG25. The first lane is *Hin*dIII digested λ DNA used as molecular size marker. Lane 2 is *Eco*RI digested pUC19. Lane 3 is *Eco*RI digested pUC19x in which the *Eco*RI site was destroyed. Undigested DNA confirms the deletion of the *Eco*RI site. Lane 4 is *Eco*RI digested pAG24 which contains the 2.4 kb *Hpa*I/*Hin*dIII fragment as does pAG25. Two bands indicate the existence of two *Eco*RI sites, one in the vector pUC19 and the other in the insert. Lane 5 is *Eco*RI digested pAG25. The single band confirms a single *Eco*RI site in the insert.

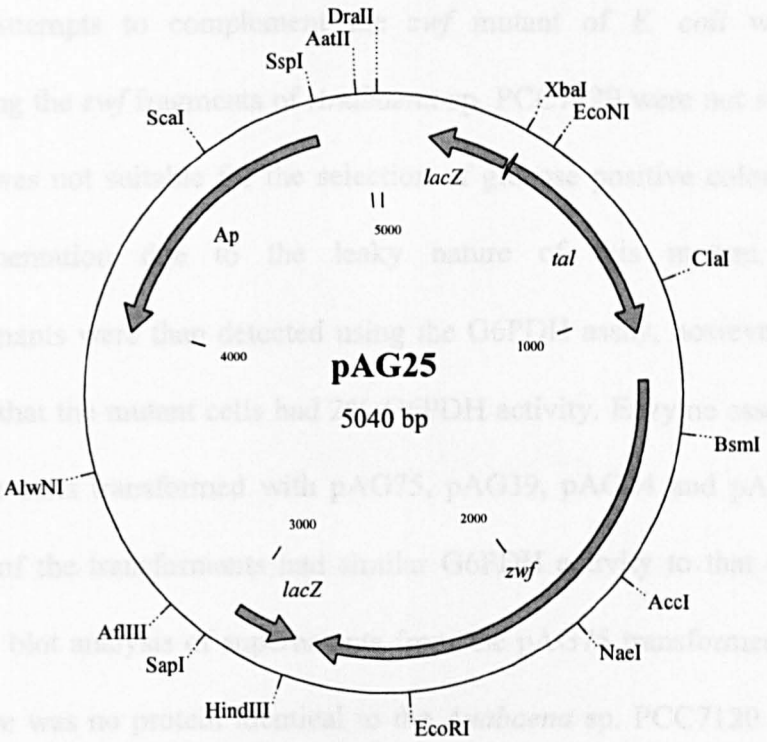


Figure 4.8: Physical map of pAG25. A *Xba*I/ *Hin*dIII fragment was sucloned in pUC19x. *Bsm*I and *Eco*RI are unique for the subclone and result in deletion of the majority of the *zwf* gene. A part of the *tal* gene is seen upstream from *zwf*.

not successful. The reasons behind this failure may be due to an incomplete end filling of the cohesive ends produced by *Bsm*I and *Eco*RI.

4.4 Conclusions

Attempts to complement the *zwf* mutant of *E. coli* with subclones containing the *zwf* fragments of *Anabaena* sp. PCC7120 were not successful. The mutant was not suitable for the selection of glucose positive colonies indicating complementation due to the leaky nature of this mutant. Presumptive transformants were than detected using the G6PDH assay, however, such assays showed that the mutant cells had 2% G6PDH activity. Enzyme assays of extracts from the cells transformed with pAG75, pAG39, pAG24 and pAG19 indicated that all of the transformants had similar G6PDH activity to that of the mutant. Western blot analysis of supernatants from the pAG75 transformed cells showed that there was no protein identical to the *Anabaena* sp. PCC7120 G6PDH being produced. All of these results suggested that the *zwf* gene of *Anabaena* sp. PCC7120 was not translated in *E. coli* DF214.

A strategy to produce a deletion-insertion mutation of the *zwf* gene of *Anabaena* sp. PCC7120 was employed. A subclone, pAG25, was constructed to produce unique *Bsm*I and *Eco*RI sites in the subclone such that use of these sites allowed deletion of most of the *zwf* gene. However, all the attempts to ligate a *Sma*I digested interposon, Ω, were not successful probably because of incomplete end-filled *Bsm*I and *Eco*RI sites.

Chapter 5

Mutagenesis of the transaldolase gene (*tal*) of *Anabaena* sp. PCC7120

5.1 Introduction

In this chapter, studies on the mutagenesis of the transaldolase gene (*tal*) of *Anabaena* sp. PCC7120 will be described. To investigate the role of transaldolase in growth, survival and nitrogen fixation, the *tal* gene of *Anabaena* sp. PCC7120 was mutated. Mutagenesis of *tal* involved insertion of the interposon Ω into the gene and re-introduction of this gene into *Anabaena* sp. PCC7120 cells. The phenotypic characterisation of the mutant cells will also be included in this chapter.

Transaldolase is one of the components of the OPP cycle which is the major route of catabolism of endogenous glycogen or exogenous carbohydrates. The pathway is also a major reducing power supplier to nitrogenase in heterocysts (see Sections 1.6.2.1 and 1.7). The transaldolase enzyme is not well characterised in cyanobacteria. Two and a half decades ago, Latzko and Gibbs (1969) assayed transaldolases from the cyanobacterium *Tolypothrix tenuis* (*Calothrix* sp. PCC7101) and from spinach. Cyanobacterial cells had four times higher activity than spinach cells (1.55 and 0.4 μmol mg/protein/hour, respectively). Later, it was suggested that the enzyme was a component involved in carbon dissimilation through the Oxidative Pentose Phosphate (OPP) cycle in cyanobacteria (Smith, 1982). Recently, two transaldolase genes (*tal*) have been sequenced from *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995a) and *Anabaena* sp. PCC7120 (this study, see Section 3.4). These genes were identified on the basis of similarity of their

translation products to the transaldolases from *E. coli* (Yura *et al.*, 1988; Sprenger *et al.*, 1995) and from *Saccharomyces cerevisiae* (Schaaff *et al.*, 1990).

Therefore, the overall objective of the mutagenesis of the *Anabaena* sp. PCC7120 *tal* gene was the further characterisation of the transaldolase enzyme in a filamentous, heterocystous cyanobacterium and to understand whether the enzyme was important for growth and survival of the cells in the presence or absence of combined nitrogen.

5.2 Strategy for mutagenesis of *tal* of *Anabaena* sp. PCC7120

The strategy for mutation of the *tal* gene of *Anabaena* sp. PCC7120 was interruption of the gene by inserting an interposon (Ω) into the gene. A *HpaI* site was available in the middle of the gene and unique for clone pAG75. This site was used for insertion of the interposon into the gene. A 5.2 kb *Eco*RI fragment containing the mutated *tal* gene was isolated and transferred to a cyanobacterial cargo plasmid which provided some convenient features for selection of mutant cells. This subclone was transferred into the wild-type *Anabaena* sp. PCC7120 cells, and single and double recombinants were selected. Finally, the phenotypes of single and double recombinants were determined.

5.3 Interruption of the *tal* gene of *Anabaena* sp. PCC7120

The Ω fragment is a 2.0 kb segment consisting of an antibiotic resistance gene (*aadA* conferring Sm^r/Spc^r phenotype) of the R100.1 plasmid. The *aadA*

gene was flanked by short inverted repeats carrying transcription and translation termination signals and synthetic polylinkers (Prentki and Krisch, 1984) (see Figure 5.1). Insertion of this fragment into a gene terminates transcription and hence protein synthesis prematurely and also provides the positive selection of the mutants by conferring Sm/Spc resistance.

The clone pAG75 was digested with *Hpa*I, the unique restriction site being in the middle of the *tal* gene. The plasmid pHP45Ω (Prentki and Krisch, 1984) was digested with *Sma*I to purify the 2.0 kb Ω fragment. The linearized pAG75 and the Ω fragments were blunt ended (created by digestion with *Hpa*I and *Sma*I, respectively). The Ω fragment was then ligated into the *Hpa*I site of pAG75 yielding pAUG20 (see Figures 5.2 and 5.3). pAUG20 was transformed into *E. coli* TG1 and selection made for Sm/Spc resistance which was conferred by the Ω fragment inserted in the *tal* gene. pAUG20 was then digested with *Bam*HI to confirm that the Ω fragment was in the insert. A 2.0 kb fragment was produced by this digestion confirming the existence of the fragment in the construct (see Figure 5.3 lane 4). In plasmid pAG75 a 3.2 kb *Eco*RI fragment carried the downstream part of *fbp*, the entire *tal* gene and the majority of the *zwf* gene. In plasmid pAUG20, the Ω fragment was additionally ligated into the *tal* gene, making the size of this *Eco*RI fragment 5.2 kb (see Figure 5.3 lane 5). An *Eco*RI fragment of 5.2 kb containing the interrupted *tal* gene was available for purification and further subcloning studies.

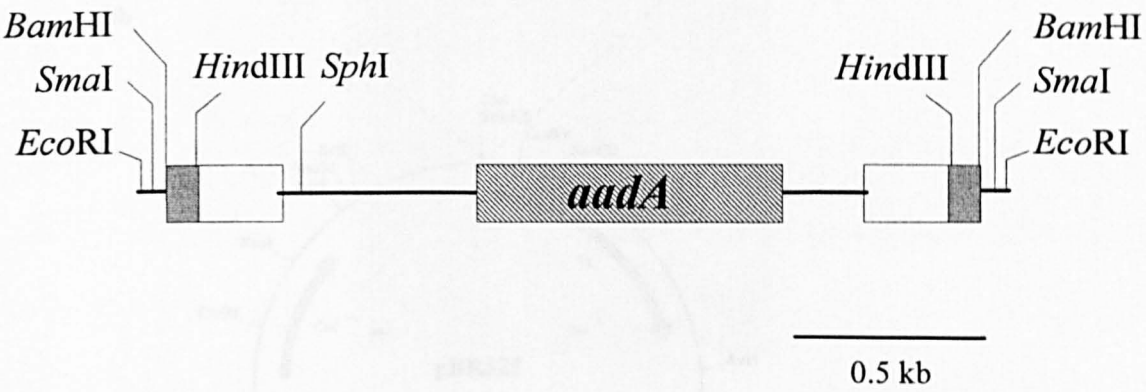


Figure 5.1: Physical map of the Ω fragment flanked by short inverted repeates carrying the T4 transcriptional termination signals (empty rectangles), the translational stop codons (blackened rectangles) and the polylinker. The *aadA* gene confers streptomycin/spectinomycin resistance (after Prentki *et al.*, 1991).

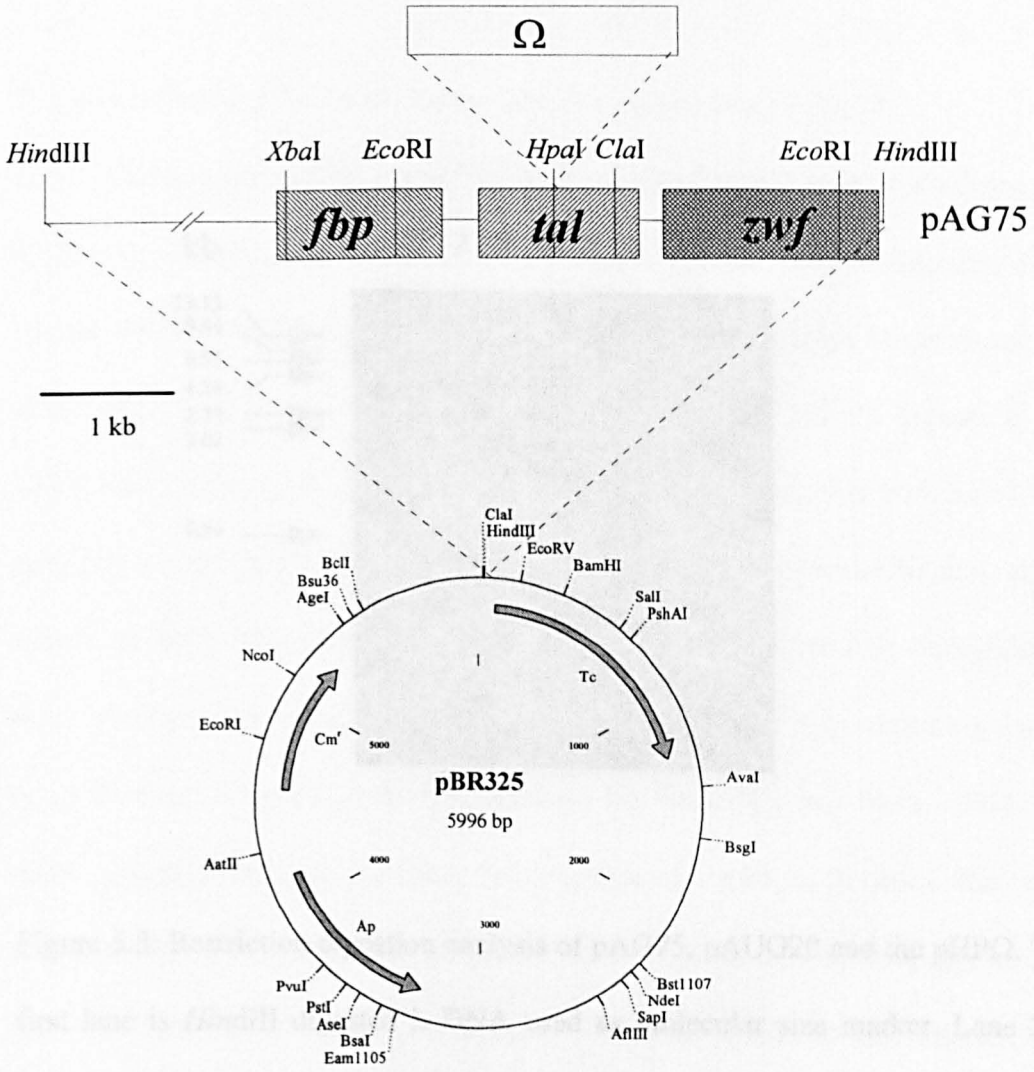


Figure 5.2: Physical map of the construct pAUG20. *Sma*I digested Ω fragment inserted into the *Hpa*I site in the middle of the tal gene. The *Hpa*I site is unique to the clone pAG75. The bar indicates the size of the insert and the Ω fragment. The size of the vector is indicated inside the circle.

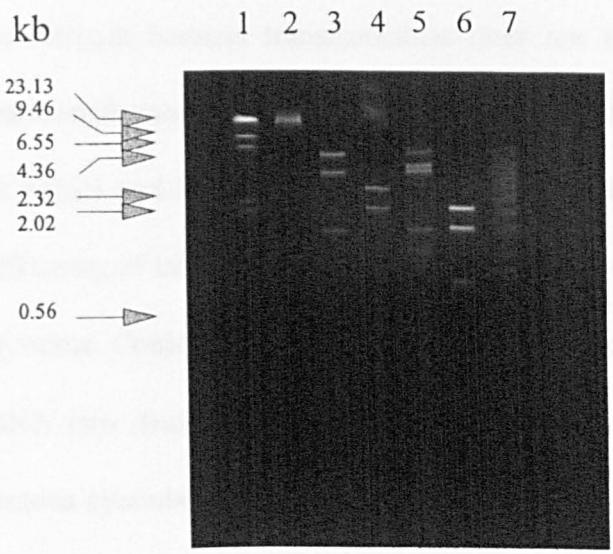


Figure 5.3: Restriction digestion analysis of pAG75, pAUG20 and the pHPΩ. The first lane is *Hin*dIII digested λ DNA used as molecular size marker. Lane 2 is *Hpa*I digested pAG75 producing 13.5 kb fragment with blunt ends. Lane 3 is *Hpa*I and *Hin*dIII digested pAG75. Lane 4 is *Bam*HI digested pAUG20. The presence of the 2 kb Ω fragment is confirmed in the construct. Lane 5 is *Eco*RI and *Pst*I digested pAUG20. The upper band is an *Eco*RI fragment that contains the *tal* gene interrupted by the Ω fragment. Lane 6 is *Sma*I and *Pst*I digested pHPΩ. The upper band is Ω fragment with blunt ends produced by *Sma*I. Lane 7 is 1 kb molecular size markers.

5.4. Subcloning of the *tal* mutation into the cargo vector pRL271

Introduction of DNA into filamentous cyanobacteria such as *Anabaena* sp. PCC7120 is difficult because transformation does not occur efficiently. Even though electrotransformation has been reported to be applicable to *Anabaena* sp. strain M131 (Thiel and Poo, 1989) and to *Nostoc* sp. PCC7121 (Moser *et al.*, 1993) the efficiency of introduction was reportedly poor, possibly because of host restriction systems. Conjugation is, therefore, probably the most efficient way to introduce DNA into *Anabaena* sp. PCC7120. Isolation of double recombinants from filamentous cyanobacteria is also difficult because single crossover events occur far more frequently than double crossover events (Cai and Wolk, 1990). For these reasons, the *tal* mutation was transferred into a plasmid that was subsequently transferable into *Anabaena* sp. PCC7120 cells by conjugation and that also provided a system for positive selection of double recombinants. A conditionally lethal gene (*sacB*) has been used to isolate double recombinants in Gram-negative bacteria including the cyanobacterium *Anabaena* sp. PCC7120 (Cai and Wolk, 1990). *sacB* is a gene of *Bacillus subtilis* coding for the exoenzyme levansucrase, production of which is induced by sucrose (Gay *et al.*, 1983). Expression of the gene is lethal to Gram-negative bacteria including cyanobacteria due to the production of toxic compounds from sucrose.

A cargo plasmid, pRL271, was designed for gene replacement in *Anabaena* sp. PCC7120 (Cai, 1991; Black *et al.*, 1993). The plasmid has a *bom*

(*oriT*) region for initiation of conjugal transfer and lacks a cyanobacterial replicon. Thus, the plasmid can survive only if it integrates into the chromosome after being transferred into *Anabaena* sp. PCC7120. This vector also provides the *sacB* gene for positive selection of double recombinants, as well as genes encoding antibiotic resistance to chloramphenicol and erythromycine. The 5.2 kb *Eco*RI fragment from pAUG20 containing the *tal* gene (interrupted by the Ω fragment) was purified by digesting with *Eco*RI and *Pst*I (see Figure 5.3 lane 5). The ends of this fragment were filled and the fragment subcloned into the *Nru*I site of pRL271 to yield pAUG401 (see Figure 5.4). This construct provided the *bom* (*oriT*) region for conjugal transfer and also *sacB* for positive selection of double recombinants.

Restriction digestion analysis was carried out to confirm the presence of the 5.2 kb *Eco*RI fragment in the plasmid pAUG401. *Eco*RI restricts pRL271 at three *Eco*RI sites in the vector (see Figure 5.5 lane 8). Since the ends of the 5.2 kb *Eco*RI insert were deleted by end-filling, *Eco*RI digestion of pAUG401 should produce three fragments, one of 9.55 kb which includes the insert and parts of the plasmid upstream and downstream from the *Nru*I site, one of 1.52 kb and one of 0.43 kb which were internal fragments of the plasmid (Figure 5.4). However, *Eco*RI digestion of pAUG401 produced four bands two of which were 1.52 and 0.43 kb (the expected size of fragments of the vector pRL271). The sizes of the other two fragments were 8.5 kb and 2.4 kb (see Figure 5.5 lane 4). This result indicated that an extra *Eco*RI site had been created in pAUG401. Further

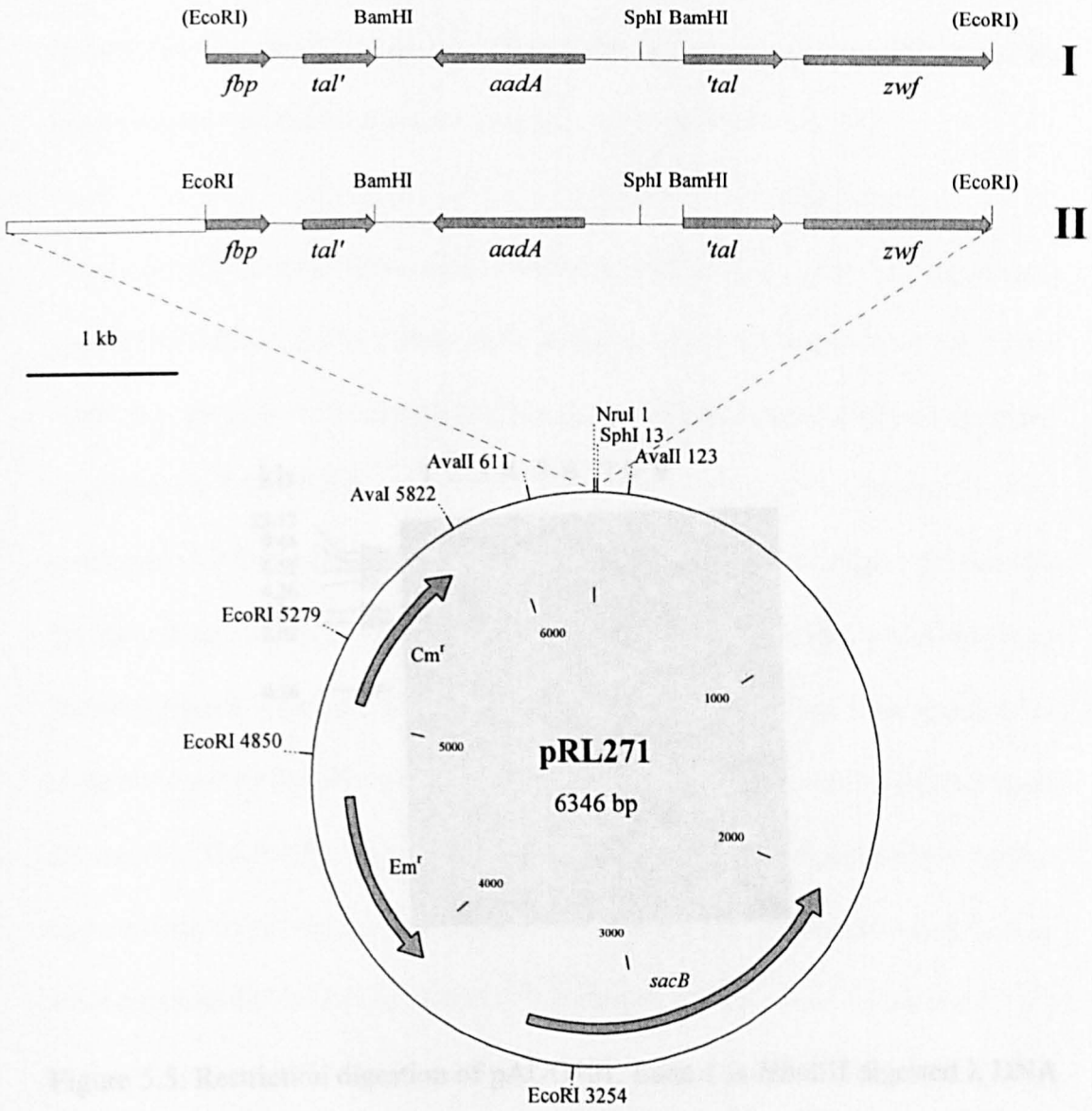


Figure 5.4: Physical map of cargo plasmid pAUG401. (I) is the 5.2 kb *Eco*RI fragment carrying *tal* mutation. (II) is the insert which was subcloned into the *Nru*I site of pRL271. On the left hand side, an artificial addition to the insert is seen (for details see text). *Eco*RI sites in brakets indicate the end-filled sites. The bar corresponds the molecular size of the inserts. The size of the vector is indicated inside.

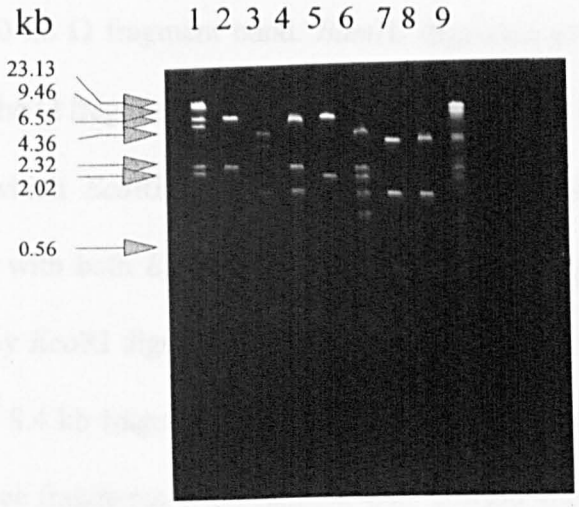


Figure 5.5: Restriction digestion of pAUG401. Lane 1 is *Hin*dIII digested λ DNA used as molecular size marker. Lane 2 is *Sph*I digested pAUG401. Lane 3 is *Eco*RI digested pAG75. Lanes 4 and 5 are pAUG401digested with *Eco*RI and *Bam*HI, respectively. Lane 6 is pAUG401 double-digested with *Bam*HI and *Eco*RI. Lane 7 is *Eco*RI digested pRL271. Lane 8 is also pRL271 double-digested with *Eco*RI and *Bam*HI. Lane 9 is 1 kb molecular size markers.

characterisation of pAUG401 was carried out to determine if the 5.2 kb *Eco*RI fragment carrying the *tal* mutation was present in the construct.

*Bam*HI does not have a resriction site in pRL271 (Figure 5.5, see lanes 8 and 9) and restricts the Ω at both ends producing a 2.0 kb fragment. If the 5.2 kb insert is present in the construct pAUG401, *Bam*HI digestion would be expected to produce a 2.0 kb Ω fragment band. *Bam*HI digestion of the construct indeed confirmed that the Ω fragment was present in the construct (see Figure 5.5 lane 5). To determine which *Eco*RI fragment carried the Ω fragment, pAUG401 was double-digested with both *Eco*RI and *Bam*HI. Three fragments were identical to those obtained by *Eco*RI digestion, being 2.4 bp, 1.52 bp and 0.43 bp (see Figure 5.5 lane 6). The 8.4 kb fragment obtained following *Eco*RI digestion was further digested into three fragments using *Bam*HI. One of these fragments was 2.0 kb in size corresponding to the Ω fragment. The other fragments were 5.3 kb and 1.1 kb in size. The 5.3 kb fragment corresponded to the part of the insert downstream from the *Hpa*I site and the vector part between the *Nru*I and the *Eco*RI sites. The 1.1 kb fragment corresponded to the 1.1 kb part of the insert upstream from the *Hpa*I site. Apparently, the *Eco*RI site in the *fbp* end of the insert was not repaired during end filling and this site had been ligated with another fragment. This fragment might be a broken part of the 5.2 kb *Eco*RI fragment caused during purification of the fragment from an agarose gel. The *Eco*RI sites from the complete fragment and the broken fragment perhaps were incompletely end filled and ligated to each other during the subcloning of the 5.2 kb fragment into the

pRL271. The size of this unexpected fragment, 2.4 kb, suggested that the size of this broken part of the insert was approximately 1.3 kb from either end of the insert. Southern blot analysis of single recombinants showed that this 2.5 kb fragment hybridised with the probe which was the 3.2 kb *Eco*RI fragment of *Anabaena* sp. PCC7120 indicating that the extra piece of DNA in pAUG401 was indeed the broken part of the insert.

Restriction analysis of pAUG401 showed that the 5.2 kb insert carrying the *tal* mutation had been successfully subcloned into the replacement vector pRL271 despite the existence of an artificial fragment in construct pAUG401. This was determined by single digestion with *Eco*RI and double digestion with *Eco*RI and *Bam*HI. This plasmid, pAUG401 was now ready to be introduced into *Anabaena* sp. PCC7120 resulting in recombination and thus mutagenesis of the wild-type *tal* gene in the chromosome.

5.5 Introduction of the plasmid pAUG401 into *Anabaena* sp. PCC7120 cells

In this section, studies on the introduction of pAUG401 into *Anabaena* sp. PCC7120 cells will be summarised. Electrotransformation was the first attempted method for DNA introduction into *Anabaena* sp. PCC7120 cells. Subsequently, conjugation was employed for the introduction by both biparental and triparental mating.

5.5.1 Introduction of the plasmid pAUG401 into *Anabaena* sp. PCC7120 by electroporation

Electroporation was used to introduce DNA into the filamentous cyanobacterial strains *Anabaena* sp. strain M131 (Thiel and Poo, 1989) and *Nostoc* sp. PCC7121 (Moser *et al.*, 1993). The *Anabaena* sp. PCC7120 cells were transformed with the plasmid pAUG401 by electroporation as described in Section 2.17.3. The plasmid pRL6 possesses a cyanobacterial replicon and can replicate in *Anabaena* sp. PCC7120 (Wolk *et al.*, 1984). This plasmid was used as a positive control for the efficient DNA transfer into the cells. Several attempts to transfer both pAUG401 and pRL6 into the *Anabaena* sp. PCC7120 failed, possibly for two reasons. Firstly, it may be the failure to transfer DNA into the cells by electroporation. The second reason may be due to restriction of transferred DNA by host restriction systems. This is not surprising since *Anabaena* sp. PCC7120 has a well-known restriction system, and this host restriction reduces the efficiency of survival of DNA dramatically in *Anabaena* sp. PCC7120 cells (Elhai and Wolk, 1988). In fact, there are six *Ava*II and three *Ava*I sites in the Ω fragment and an *Ava*II site also exists in pRL6 (Wolk *et al.*, 1984). Attempts to introduce the cargo plasmid pAUG401 and positive control pRL6 into *Anabaena* sp. PCC7120 were not successful, again probably because of inefficient transfer or restriction of transferred DNA by *Anabaena* sp. PCC7120 restriction system.

5.5.2 Introduction of the cargo plasmid pAUG401 into *Anabaena* sp. PCC7120 cells by conjugation

Conjugation has been widely used for gene transfer into filamentous cyanobacteria (Wolk *et al.*, 1984; Elhai and Wolk, 1988) (see Section 1.5.4.3.1). A triparental mating procedure was described by Elhai and Wolk in 1988. One of the three parents is *E. coli* bearing the conjugal plasmid. The second parent is the other *E. coli* bearing the cargo plasmid and helper plasmid. The third parent is the target cyanobacterium, *Anabaena* sp. PCC7120 in this study. A biparental conjugation method was described by McFarlane *et al.* (1987) for *Anabaena* sp. ATCC27893. In this case one of the parents was the *E. coli* strain S17-1 which contains a derivative of RP4 (RP4-2::Mu-Km::Tn7) integrated into the chromosome, the cargo plasmid and helper plasmid are transformed into this strain. The other parent is the cyanobacterial strain, *Anabaena* sp. PCC7120.

Initially, a biparental conjugation system was employed for transfer of pAUG401 into the *Anabaena* sp. PCC7120 cells. *E. coli* strain S17-1 had already been transformed with the helper plasmid pDS4101. This helper plasmid carried the *mob* gene encoding a DNA-nicking protein and the ampicillin resistance gene. The DNA-nicking protein specifically recognises the *bom* site (for basis of mobility) of the cargo plasmid. As mentioned in Section 5.5.1, both *Ava*I and *Ava*II enzymes present in *Anabaena* sp. PCC7120 may restrict these sites in pAUG401 and the DNA transferred into the cell may not, therefore, be able to survive. The plasmid pRL528 carries genes for both *Ava*I and *Ava*II methylases as

well as the *bom* gene and chloramphenicol resistance gene (Elhai and Wolk, 1988). This plasmid was also transformed into *E. coli* S17-1 and selecting against chloramphenicol and ampicillin. The cargo plasmid pAUG401 was then transformed into this *E. coli* strain and selected against streptomycin and spectinomycin. After these transformations, *E. coli* S17-1 had the following phenotypes i) RP4 derivative in the chromosome, ii) helper function from pDS4101 and pRL528, iii) *Ava*I and *Ava*II methylases from pRL528. A positive control was obtained by transforming pRL6 into the *E. coli* S17-1 bearing the helper plasmids. This plasmid carries a cyanobacterial replicon and the genes for chloramphenicol and neomycin resistance. The biparental conjugation procedure was as described in Section 2.18. Exconjugant *Anabaena* sp. PCC7120 cells were inoculated onto BG11 medium containing streptomycin and spectinomycin and positive control cells on BG11 medium containing neomycin. No colonies grew on the test medium while a few yellowish colonies grew on the control medium. This result suggested that at least the cargo plasmid pAUG401 was not efficiently transferred into the recipient *Anabaena* sp. PCC7120.

After unsuccessful attempts to transfer pAUG401 into the *Anabaena* sp. PCC7120 cells by electroporation and biparental conjugation, a triparental conjugation procedure was employed as described in Section 2.18. *E. coli* strain AEE101 was one of the two *E. coli* parent strains and carried the conjugative plasmid RP-4 (D. H. Hodgson, personal communication). The other *E. coli* strain HB101 was transformed with the helper plasmid pRL528 carrying the *mob* gene

and the methylase genes for *Ava*I and *Ava*II. These cells were then also transformed with pAUG401. The recipient parent *Anabaena* sp. PCC7120 was prepared and conjugation was performed as described in Section 2.18. Similarly, pRL6 was also conjugated into the *Anabaena* sp. PCC7120 as a positive control. The pAUG401 exconjugant *Anabaena* sp. PCC7120 cells were selected against streptomycin and spectinomycin and pRL6 exconjugants against neomycin. Approximately 20 colonies on streptomycin and spectinomycin medium and a large number of colonies on neomycin medium grew. This result showed that gene transfer into the *Anabaena* sp. PCC7120 cells by triparental conjugation had been successfully completed. The pAUG401 exconjugants were then transferred onto fresh streptomycin and spectinomycin medium successively to eliminate possible pseudo-exconjugants.

5.6 Selection and segregation of single and double recombinants

Streptomycin and spectinomycin resistant exconjugant colonies may be either single or double recombinants. If these colonies were inoculated onto the sucrose containing medium only double recombinant colonies would survive. This is because *sacB* gene expression is lethal for *Anabaena* sp. PCC7120 in the presence of sucrose. The single recombinants carry pAUG401 containing *sacB* in the chromosome. To select double recombinants, twenty exconjugant colonies were grown on both sucrose-containing and sucrose-free media. Only two colonies grew on sucrose containing medium indicating that these colonies did not carry any single recombinant chromosome in the cell. However, both single and

double recombinants may well carry wild-type copies of the chromosome in the same cell. These wild-type copies were segregated from double and single recombinant cells by breaking the fragments into single cells by sonication and successively growing against streptomycin and spectinomycin. The single recombinant cells carrying pAUG401 in the chromosome was named *Anabaena* sp. HK28 and double recombinant cells carrying the interrupted copy of *tal* *Anabaena* sp. HK29.

5.7 Characterisation of single recombinant strain HK28 and double recombinant strain HK29

In this section, studies on the characterisation of HK28 and HK29 will be summarised. Initially, Southern blot analysis was employed to confirm interruption of the *tal* gene in the chromosome by a double crossover event and the complete segregation of wild-type copies of the chromosome. Then, phenotypic characterisations were carried out by assaying transaldolase activity. G6PDH activity from these mutant strains was also assayed to determine if there was a promoter region between *tal* and *zwf* in *Anabaena* sp. PCC7120.

5.7.1 Southern blot analysis of the strains HK28 and HK29

Chromosomal DNA was purified from wild-type and mutant strains HK28 and HK29 as described in Section 2.6.1. Southern blotting was performed as described in Section 2.20. The 3.2 kb *Eco*RI fragment from pAG75 carrying the *tal* gene was used as homologous hybridisation probe to detect both mutated and

wild-type copies of this region in wild-type and mutant cells. Chromosomal DNAs from wild-type and mutant strains HK28 and HK29 cells were digested with *Eco*RI and then hybridised with the 3.2 kb *Eco*RI fragment. The probe would be expected to hybridise with a 3.2 kb fragment from the wild-type chromosome and a 5.2 kb fragment from the chromosome of the double recombinant strain HK29, (corresponding to the insertion of the 2.0 kb Ω fragment into the *tal* gene) (see Figure 5.6). Two sets of fragments may hybridise with the probe depending on the location of the single crossover. If the crossover occurred upstream from the Ω fragment, three bands, of approximately 8.5 kb, 3.2 kb and 2.4 kb in size, would be expected to hybridise with the probe. If the crossover occurred downstream from the Ω fragment, then hybridisation would be expected to occur with fragments of 6.5 kb, 5.2 kb and 2.4 kb in size.

The result of Southern blot analysis showed that a 3.2 kb fragment from the wild-type hybridised with the probe as predicted (Figure 5.7). Three fragments were detected of 6.5, 5.2 and 2.4 kb in size indicating that the crossover occurred downstream from the Ω fragment. A 5.2 kb fragment was detected from all four colonies of the double recombinant strain HK29 that were analysed. This indicated that the Ω fragment had been successfully inserted into the *tal* gene in the chromosome of *Anabaena* sp. PCC7120. This Southern blot analysis also showed that wild-type copies of the chromosome were completely segregated from the single recombinant strain HK28 while two of the four double recombinant

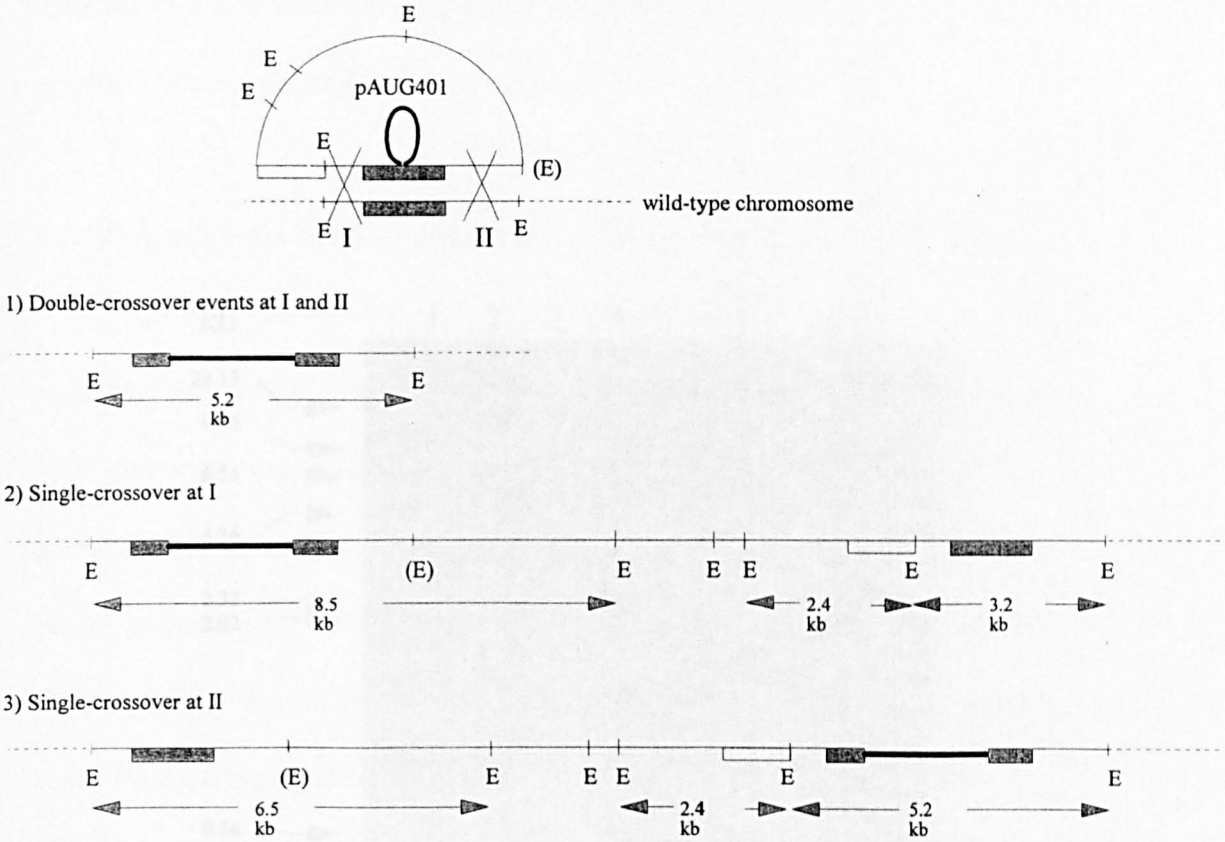


Figure 5.6: Integration of pAUG401 in the *Anabaena* sp. PCC7120 chromosome by single and double crossover events. All crossover possibilities are schematically represented. The sizes are arbitrary (for correct sizes, see Figure 5.4). The sizes of the *Eco*RI fragments of the recombinants which was expected to hybridise with the probe (3.2 kb *Eco*RI fragment of *Anabaena* sp. PCC7120 carrying *tal* gene) are shown below the chromosome. Dark rectangle represents the tal gene. The empty rectangle represent the artificial fragment in pAUG401 (see Section 5.4). The Ω fragment is shown as thick bar.

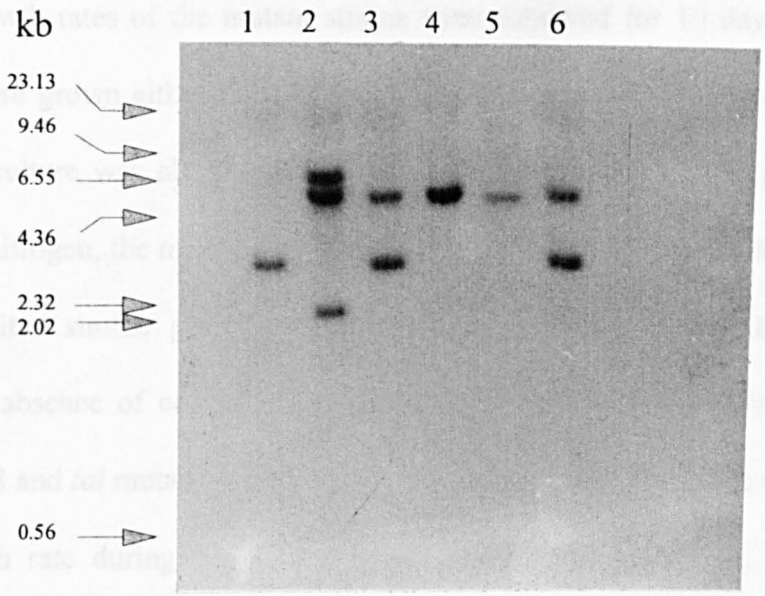


Figure 5.7: Southern blot analysis of wild-type, single and double recombinants. Chromosomal DNAs were digested with *Eco*RI. Lane 1 is wild-type, lane 2 single recombinant strain HK28. Lanes 2,4,5 and 6 are double recombinants, *tal* mutant strain HK29. All the double recombinants have mutated copies of *tal* (5.2 kb), but two of them also have wild-type copy of the gene (lanes 3 and 6).

colonies still had wild type copies of the chromosome. This was shown by the presence of a 3.2 kb fragment in the chromosome (see Figure 5.7).

5.7.2 Growth rates of wild-type and the mutant strains HK28 and HK29

Growth rates of the mutant strains were followed for 10 days when the mutants were grown either in the presence or absence of combined nitrogen. A wild type culture was also inoculated as a positive control. In the presence of combined nitrogen, the *tal* mutant strains HK29 and the single recombinant strain HK28 exibited similar growth rates compared to the wild-type. (Figure 5.8). Growth in absence of combined nitrogen revealed that the single recombinant strain HK28 and *tal* mutant strain HK29 grew at a similar growth rate to the wild-type growth rate during exponental phase (Figure 5.9). However, the mutant strains HK28 HK29 reached to stationary phase relatively faster than the wild-type. (Figure 5.9). This implies that reduction of growth rates of these strains was not because of the transaldolase mutation. If it was the case, the single recombinant strain HK28 would have grown at the same rates as the wild-type strain because this single recombinant strain has a complete copy of the *tal* gene upstream from the insertion site of the Ω fragment (see Section 5.7.1 and Figure 5.6). This result implies that transaldolase does not have significant effect on the growth of the cells of *Anabaena* sp. PCC7120. Faster transition of cells of the mutant strains HK28 and HK29 in absence of combined nitrogen may be because of reduced G6PDH activity in these strains (see Sections 5.7.3 and 5.7.4)

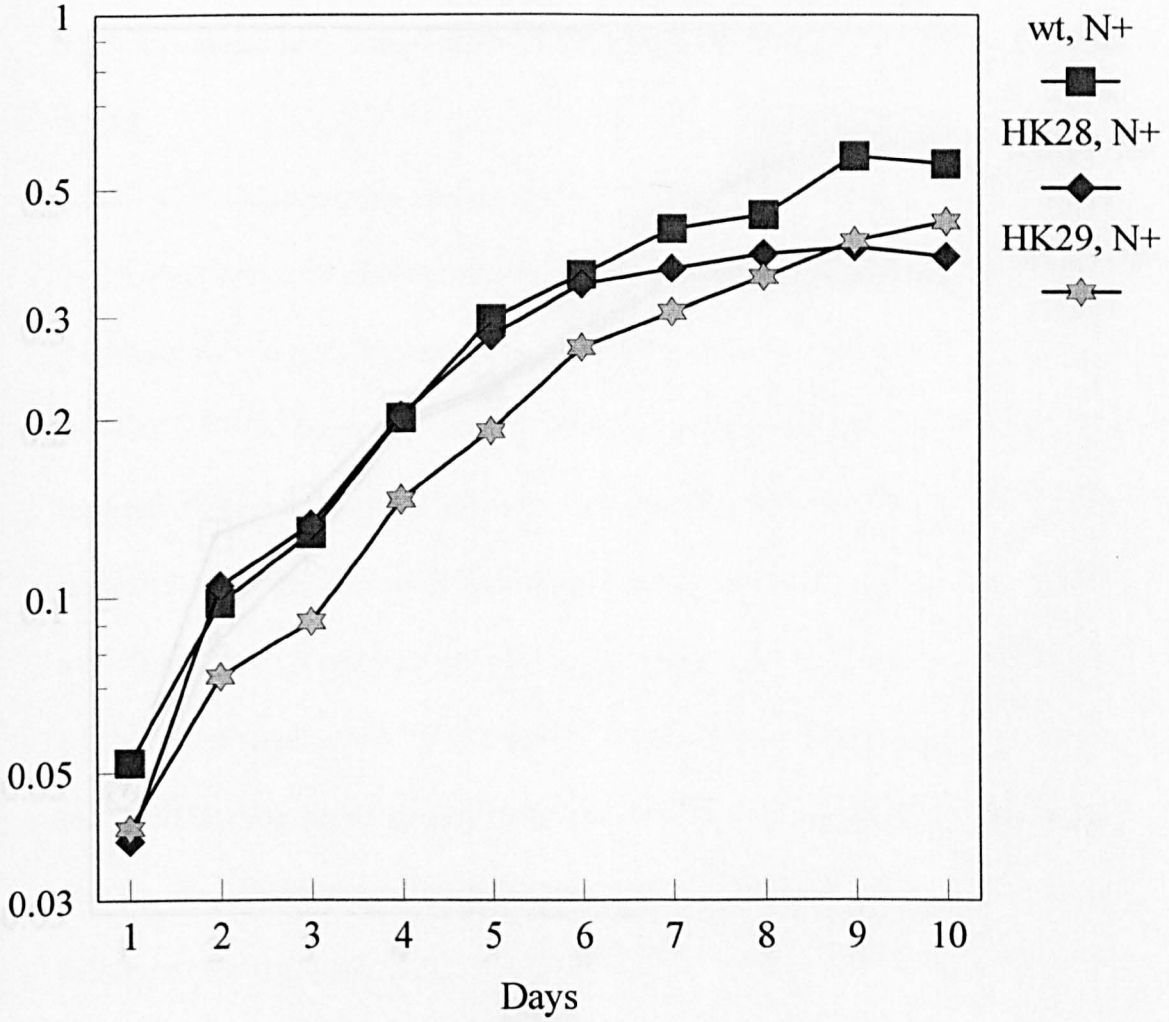


Figure 5.8: Growth of wild-type, single recombinant strain HK28 and *tal* mutant strain HK29. Cells were grown in the presence of combined nitrogen (N+) under continuous illumination.

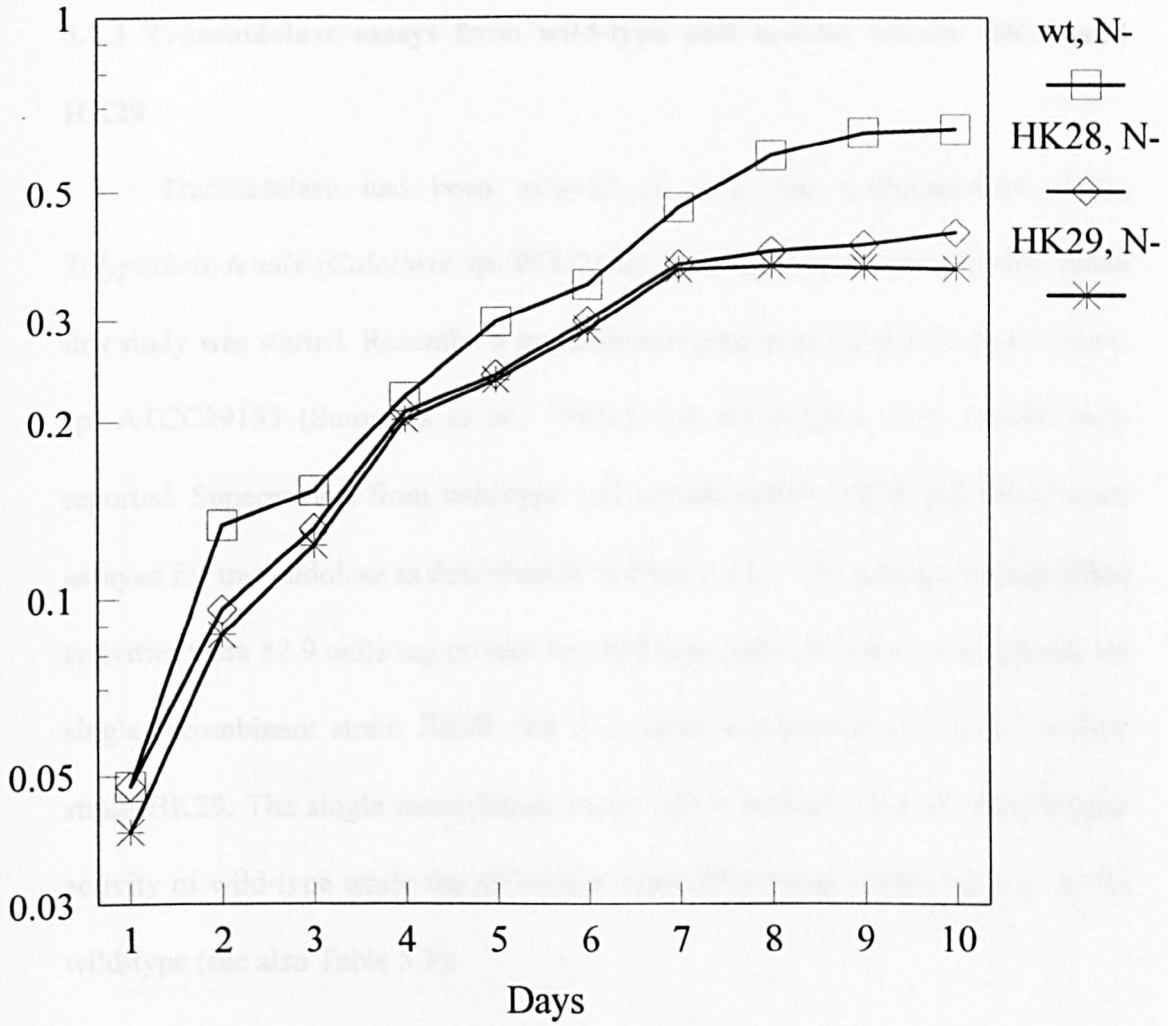


Figure 5.9: Growth of wild-type, single recombinant strain HK28 and *tal* mutant strain HK29. Cells were grown in the absence of combined nitrogen (N-) under continuous illumination.

5.7.3 Transaldolase assays from wild-type and mutant strains HK28 and HK29

Transaldolase had been assayed in only one cyanobacterial strain, *Tolypothrix tenuis* (*Calothrix* sp. PCC7101) by Latzko and Gibbs (1969) when this study was started. Recently, a transaldolase gene was sequenced from *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995a), but no enzyme assay results were reported. Supernatants from wild-type and mutant strains HK28 and HK29 were assayed for transaldolase as described in Section 2.24.2. The average transaldolase activities were 82.9 units/mg protein for wild-type cells, 60.1 units /mg protein for single recombinant strain HK28 and 91.1 units /mg protein for the *tal* mutant strain HK29. The single recombinant strain HK28 retained 72% of transaldolase activity of wild-type while the *tal* mutant strain HK29 had similar activity to the wild-type (see also Table 5.1).

The results of the transaldolase assays and measurements of growth rate for wild type and mutant cells suggested that *Anabaena* sp. PCC7120 may have more than one gene which encodes transaldolase. The multigenic nature of cyanobacterial genomes is a well known phenomenon. Martin *et al.* (1993) reported three copies of the glyceraldehyde-3-phosphate dehydrogenase gene in *Anabaena variabilis* (ATCC29413). Similarly, two copies of the allophycocyanin genes from *Calothrix* sp. PCC 7601 and a duplicated copy of the gene of D2 protein of the photosystem II reaction centre (*psbD*) (Bustos and Golden, 1992)

were reported. These reports support the possibility for the existence of more than one gene encoding transaldolase in *Anabaena* sp. PCC7120.

Table 5.1: Transaldolase activities in wild-type, single recombinant strain HK28 and *tal* mutant strain HK29. Activities are expressed as specific activity (units/mg protein).

Strain	Measurements	Average (% WT activity)
wild-type	84.3 68.3 96.3	82.9
HK28	54.3 50.7 75.4	60.1 (72.4%)
HK29	80.4 76.4 116.5	91.1 (109.9%)

5.7.4 G6PDH activity in wild-type and the mutant strains HK28 and HK29

In *Nostoc* sp. ATCC29133, *zwf* has been reported to be in same operon as *tal* (Summers *et al.*, 1995b). The operon consists of five genes encoding proteins of the OPP cycle including *fbp*, *tal*, *zwf* and *opcA*. The genes could be co-transcribed in groups. The *zwf* gene was transcribed by more than one promoter which are located between *fbp* and the start codon of *zwf* itself. Being a filamentous, heterocystous strain (like *Nostoc* sp. ATCC29133), *Anabaena* sp. PCC7120 may have a similar operon in this region. Indeed, the gene organisation

is similar in this region in both of these two organisms. G6PDH might be affected in three possible ways in the single recombinant strain HK28 and *tal* mutant HK29. Firstly, G6PDH may be completely inactivated if all the genes in the region are co-transcribed by a single promoter due to insertion of transcription and translational stop codons into the *tal* gene upstream from *zwf*. Secondly, G6PDH may be as active as in the wild-type if *zwf* is transcribed by an independent operon located downstream of the insertion site of the Ω fragment. Thirdly, G6PDH may have an activity less than wild-type if *zwf* is transcribed by more than one promoter located upstream and downstream of the insertion site. Some of this signals may start upstream from the insertion site of the Ω fragment and some other signals may start downstream from the insertion site of the Ω fragment. The *zwf* gene of *Nostoc* sp. ATCC29133 is transcribed by more than one signals (Summers *et al.*, 1995b).

G6PDH activity was assayed in the supernatants of wild-type, single recombinant strain HK28 and *tal* mutant strain HK29 as described in Section 2.24.1. The results supported the third possibility. The activities were 136.6 units/mg protein for the wild-type cells, 33.7 units/mg protein for the single recombinant strain HK28 and 46.8 units/mg protein for the *zwf* mutant strain HK29. The *tal* mutant strain HK29 retained 34.1 % of the wild-type activity while the single recombinant strain HK28 retained 24.6% of wild-type G6PDH activity (see also Table 5.2). These results are consistent with the results obtained from *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995b). The *zwf* gene of *Anabaena* sp.

PCC7120 was, therefore, transcribed by more than one promoter which are located upstream and downstream of the insertion site of the Ω fragment, that is the *Hpa*I site in the middle of the *tal* gene. The reduction of G6PDH activity in mutant strains HK28 and HK29 may be involved in the rapid transition into the stationary phase in absence of combined nitrogen (see Section 5.7.2).

Table 5.2: G6PDH activities in wild-type, single recombinant strain HK28 and *tal* mutant strain HK29. Activities are expressed as specific activity (units/mg protein).

Strains	Measurements	Average (% WT activity)
wild-type	174.8 122.5 112.5	136.6
HK28	30.9 35.8 34.6	33.7 (24.6%)
HK29	40.1 58.2 42.1	46.8 (34.1%)

5.8 Conclusions

A strategy for mutagenesis of the transaldolase gene (*tal*) of *Anabaena* sp. PCC7120 was employed. The interposon Ω was inserted into the middle of the *tal* gene yielding pAUG20. A 5.2 kb fragment carrying the *tal* mutation was transferred into *Anabaena* sp. PCC7120 cargo plasmid pRL271 to yield

pAUG401. This cargo plasmid, pAUG401 was conjugated into *Anabaena* sp. PCC7120 to allow both single and double recombination with the wild-type chromosome. A single recombinant (*Anabaena* sp. strain HK28) and a double recombinant (*Anabaena* sp. strain HK29) were isolated and phenotypic characteristics were analysed. The *tal* mutant strain HK29 had similar transaldolase activity to the wild-type (109.9 % of wild type) while the single recombinant strain HK28 retained 72.4% of transaldolase activity of wild-type. No significant difference was observed between the growth rates of the wild-type and the mutant strains HK28 and HK29 in the presence of combined nitrogen. However, the strains HK28 and HK29 reached to the stationary phase earlier than wild-type in absence of combined nitrogen. G6PDH activities from the strains HK28 and HK29 were 24.6% and 33.7% of wild-type activity, respectively, implying that *zwf* is transcribed by more than one promoter which are located both upstream and downstream of the insertion site of the Ω fragment. This low level of G6PDH activity may be the reason why the mutant strains HK28 and HK29 reach to stationary phase in absence of combined nitrogen faster than wild-type.

Chapter 6

Mutagenesis of the *opcA* gene of *Synechococcus* sp. PCC7942

6.1 Introduction

In this chapter studies on the mutagenesis of the *opcA* gene of *Synechococcus* sp. PCC7942 will be described. Also, the phenotypic characterisation of the *opcA* mutant cells will be included. An insertion/inactivation approach was carried out for mutagenesis of this gene. The Ω fragment was used for inactivation of the gene.

Nucleotide sequence analysis had shown that the *opcA* gene was located downstream from the *zwf* gene of both *Synechococcus* sp. PCC7942 (J. Scanlan, personal communication) and *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995a). The gene was thought to be involved in the regulation of G6PDH activity, but no further characterisation of the gene had been undertaken when this study was started. There was no identifiable DNA-binding motif indicating that the gene was unlikely to be involved in control of expression of the *zwf* gene. Therefore, it was hypothesised that *opcA* may be required for assembly or activation of functional G6PDH from *Synechococcus* sp. PCC7942. To understand whether it was required for production of the full functional form of G6PDH, the *opcA* gene of *Synechococcus* sp. PCC7942 was mutated and the phenotypes of the mutant cells were characterised. There were two reasons for choosing the *Synechococcus* sp. PCC7942 *opcA* gene for mutagenesis. Firstly, the gene had already been sequenced from *Synechococcus* sp. PCC7942, and an *opcA* gene from *Anabaena* sp. PCC7120 had not been cloned. Secondly, DNA transfer into *Synechococcus* sp. PCC7942 cells can be done by transformation, and segregation of wild-type

chromosomes from single and double recombinants is much easier and faster than in a filamentous strain.

6.2 Strategy for mutagenesis of *opcA* gene of *Synechococcus* sp. PCC7942

An insertion/inactivation strategy was carried out for mutagenesis of the *opcA* gene. A clone, pDB, contained a 6.0 kb fragment of *Synechococcus* sp. PCC7942 including the *opcA* gene (J. Scanlan, personal communication) (Figure 6.1). Being of large size a smaller fragment was subcloned into pUC19 and the gene was interrupted using this subclone. The construct carrying the *opcA* mutation was then transformed back into the *Synechococcus* sp. PCC7942 cells and single and double recombinants were selected.

6.3 Interruption of the *opcA* gene of *Synechococcus* sp. PCC7942

The clone pDB contained half of the *zwf* gene, *opcA*, and two genes which encode cytochrome b_6 and subunit IV of the cytochrome b_6f complex. There were two *Bgl*II sites in the clone, one in the middle of the *opcA* gene and the other further downstream. If a smaller fragment was subcloned, the *Bgl*II site in the *opcA* gene would be unique for this subclone, and this site could, therefore, be used for the insertion of the Ω fragment into the gene. There were two *Pst*I sites in the clone pDB, one was 300 bp downstream from the stop codon of the *opcA* gene and the other in the multiple cloning site of pUC19. When pDB was digested with *Pst*I two fragments were produced. One of these fragments was the 3.8 kb insert which carried the region of the insert downstream from the *opc*A gene. The other was a 4.8 kb fragment which carried both the 2.2 kb insert (containing the *opcA*

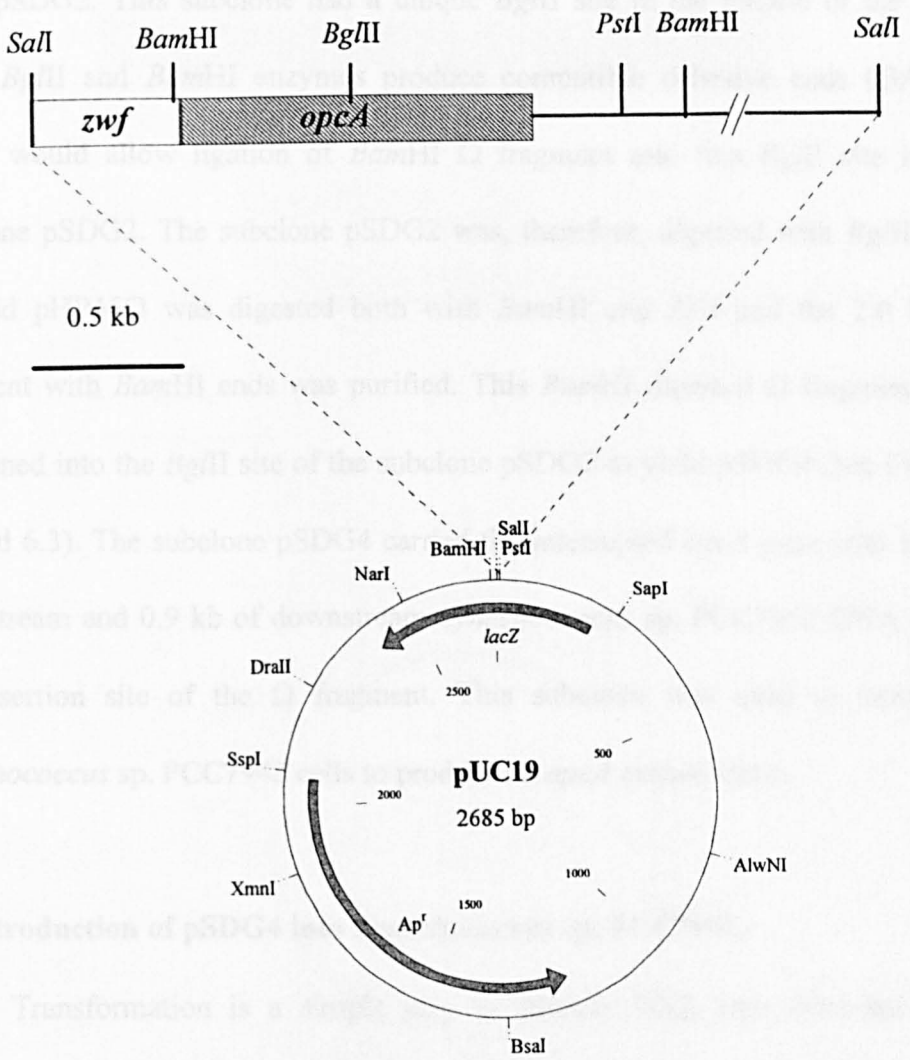


Figure 6.1: Physical map of the clone pDB. The total length of insert is 6.0 kb. The bar indicates the size of insert. The size of the vector is shown inside of the map.

gene) and 2.68 kb of the vector, pUC19. This 4.8 kb fragment was self-ligated to yield pSDG2. This subclone had a unique *Bgl*II site in the middle of the *opcA* gene. *Bgl*II and *Bam*HI enzymes produce compatible cohesive ends (GATC). These would allow ligation of *Bam*HI Ω fragment into this *Bgl*II site in the subclone pSDG2. The subclone pSDG2 was, therefore, digested with *Bgl*II. The plasmid pHP45Ω was digested both with *Bam*HI and *Pst*I and the 2.0 kb Ω fragment with *Bam*HI ends was purified. This *Bam*HI digested Ω fragment was subcloned into the *Bgl*II site of the subclone pSDG2 to yield pSDG4 (see Figures 6.2 and 6.3). The subclone pSDG4 carried the interrupted *opcA* gene with 1.3 kb of upstream and 0.9 kb of downstream *Synechococcus* sp. PCC7942 DNA, from the insertion site of the Ω fragment. This subclone was used to transform *Synechococcus* sp. PCC7942 cells to produce an *opcA* mutant strain.

6.4 Introduction of pSDG4 into *Synechococcus* sp. PCC7942

Transformation is a simple way to transfer DNA into *Synechococcus* strains including the strain PCC7942 (see Section 1.5.4.3.2). *Synechococcus* sp. PCC7942 cells were transformed with pSDG4 as described in Section 2.17.2. The shuttle plasmid pUC303 (Kuhlemeier *et al.*, 1983) was also transformed into *Synechococcus* sp. PCC7942 cells as a control. This plasmid pUC303 contains a cyanobacterial replicon and an antibiotic resistance gene encoding streptomycin resistance. Transformed cells were inoculated onto streptomycin-containing media, *Synechococcus* sp. PCC7942 cells which had not been transformed were also inoculated onto streptomycin-containing media as a negative control. Over

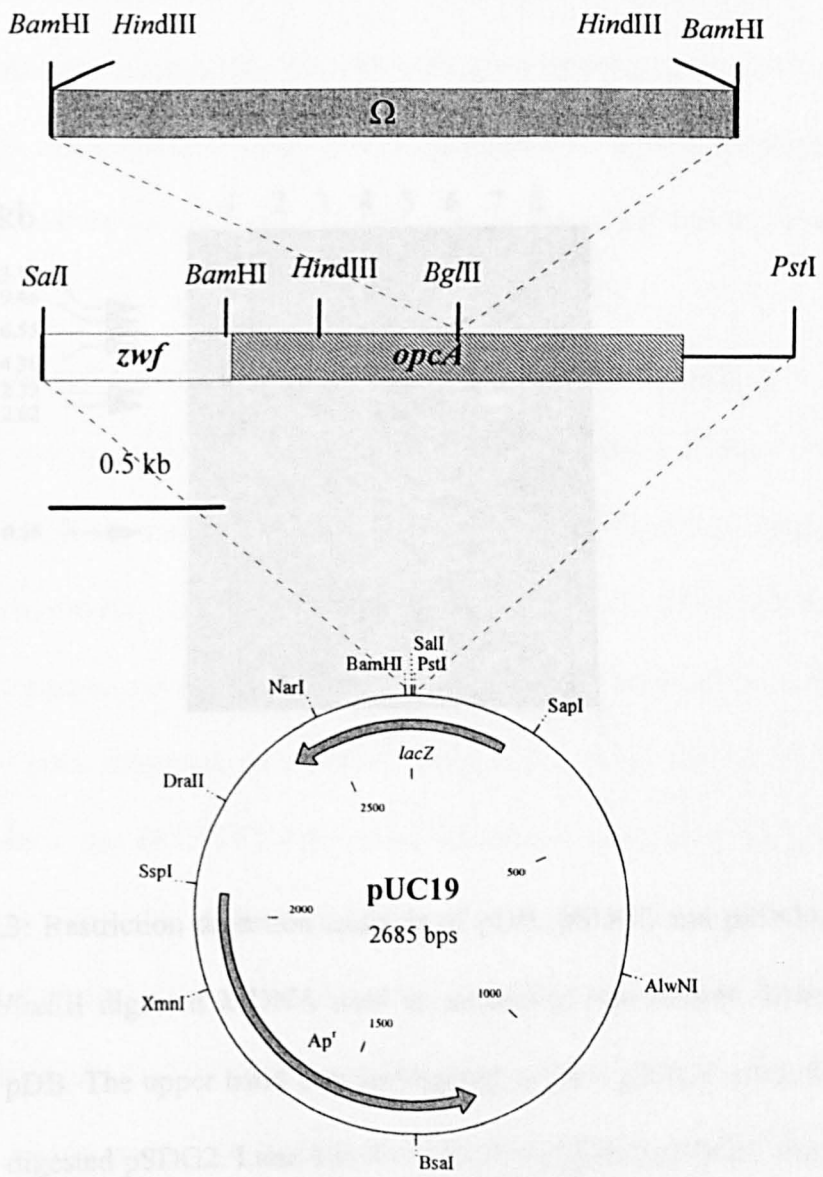


Figure 6.2: Physical map of the construct pSDG4. *Bam*HI digested Ω fragment was inserted into the *Bgl*II site of the clone pSDG2. The bar indicates the sizes of the insert and the Ω fragment. The size of the vector is shown inside the circle.

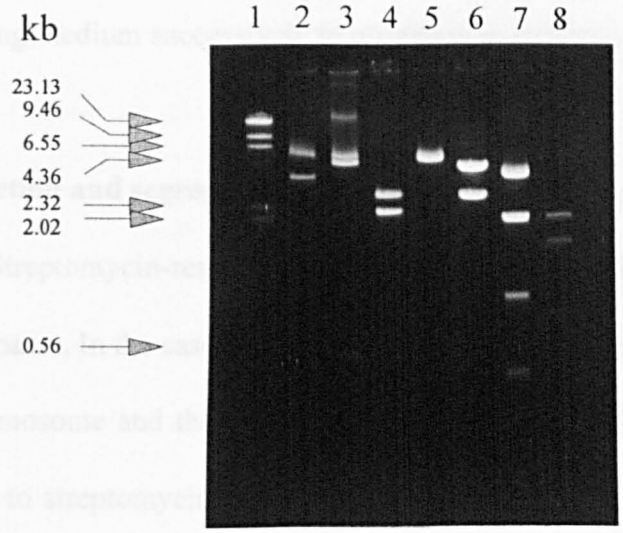


Figure 6.3: Restriction digestion analysis of pDB, pSDG2 and pSDG4. The first lane is *Hind*III digested λ DNA used as molecular size marker. Lane 2 is *Pst*I digested pDB. The upper band was self-ligated to yield pSDG2. Lane 3 is *Bam*HI and *Sal*I digested pSDG2. Lane 4 is *Sal*I and *Pst*I digested pSDG2. The bands are 2.7 kb vector and 2.2 kb insert. Lane 5 is *Bgl*II digested pSDG2 which confirms the unique *Bgl*II site in the subclone. Lane 6 is *Sal*I and *Pst*I digested pSDG4. The lower band is the vector. The upper band is the insert carrying the Ω fragment (see also lane 4). Lane 7 is *Bam*HI digested pSDG4. The 2.0 kb band confirms the presence of the Ω fragment in the construct. Lane 8 is *Bam*HI and *Pst*I digested pHP45Ω. The Ω fragment is the upper band.

200 colonies per plate from both pSDG4 and pUC303 transformed cells grew on streptomycin-containing media while no cells grew on the negative control plates. The pSDG4 transformant colonies were then transferred onto fresh streptomycin-containing medium successively to eliminate possible pseudo-transformants.

6.5 Selection and segregation of single and double recombinants

Streptomycin-resistant colonies may be either single or double recombinants. In the case of a single crossover, pSDG4 would have integrated into the chromosome and these cells would have been ampicillin resistant as well as resistant to streptomycin. If double recombination had occurred, the cells would have only been streptomycin-resistant. Fifty colonies of pSDG4 transformed *Synechococcus* sp. PCC7942 cells were inoculated onto both ampicillin and streptomycin containing medium and also only streptomycin containing medium. Only two of the fifty colonies grew on the ampicillin and streptomycin-containing medium indicating that only these two colonies were single recombinants and the rest were double recombinants. These two single recombinant colonies and four double recombinant colonies were successively transferred onto fresh antibiotic media for segregation of wild-type copies of the chromosome from mutant cells. A single recombinant clone carrying pSDG4 in the chromosome was named *Synechococcus* sp. HK52 and a double recombinant clone carrying the interrupted copy of *opcA*, *Synechococcus* sp. HK55.

6.6 Characterisation of mutant strains HK52 and HK55

In this section, studies on the characterisation of both single recombinant and double recombinant strains HK52 and HK55 will be summarised . Firstly, Southern blot analysis of the strains was carried out to confirm the interruption of the *opcA* gene in the chromosome by a double crossover event and also segregation of all wild-type copies of the chromosome. Then, G6PDH activities were assayed in cell-free extracts of these strains. G6PDH activity was also detected by activity staining in non-denaturing polyacrylamide gels. Western blot analysis was carried out on blots from both denaturing and non-denaturing gels using an antiserum raised against a truncated G6PDH from *Anabaena* sp. PCC7120 to determine whether any change in G6PDH synthesis or molecular assembly had occurred in the *opcA* mutant cells.

6.6.1 Southern blot analysis of the strains HK52 and HK55

Chromosomal DNA was purified from wild-type and mutant strains HK52 and HK55 as described in Section 2.6.3; Southern blots were performed as described in Section 2.20. The subclone pSDG2 was used as a probe. Chromosomal DNA from wild-type and mutant strains was double-digested with *Sal*I and *Bam*HI and then hybridised with radio-labelled pSDG2. The probe would be expected to hybridise with 0.5 and 1.9 kb fragments of wild-type chromosomal DNA (see Figure 6.4). Two fragments (0.5 kb and 3.9 kb in size) would be expected to hybridise from the chromosomal DNA of the double recombinant strain HK55. Two sets of fragments may hybridise with the probe depending on the location of the single crossover. If the crossover occurred upstream from the Ω

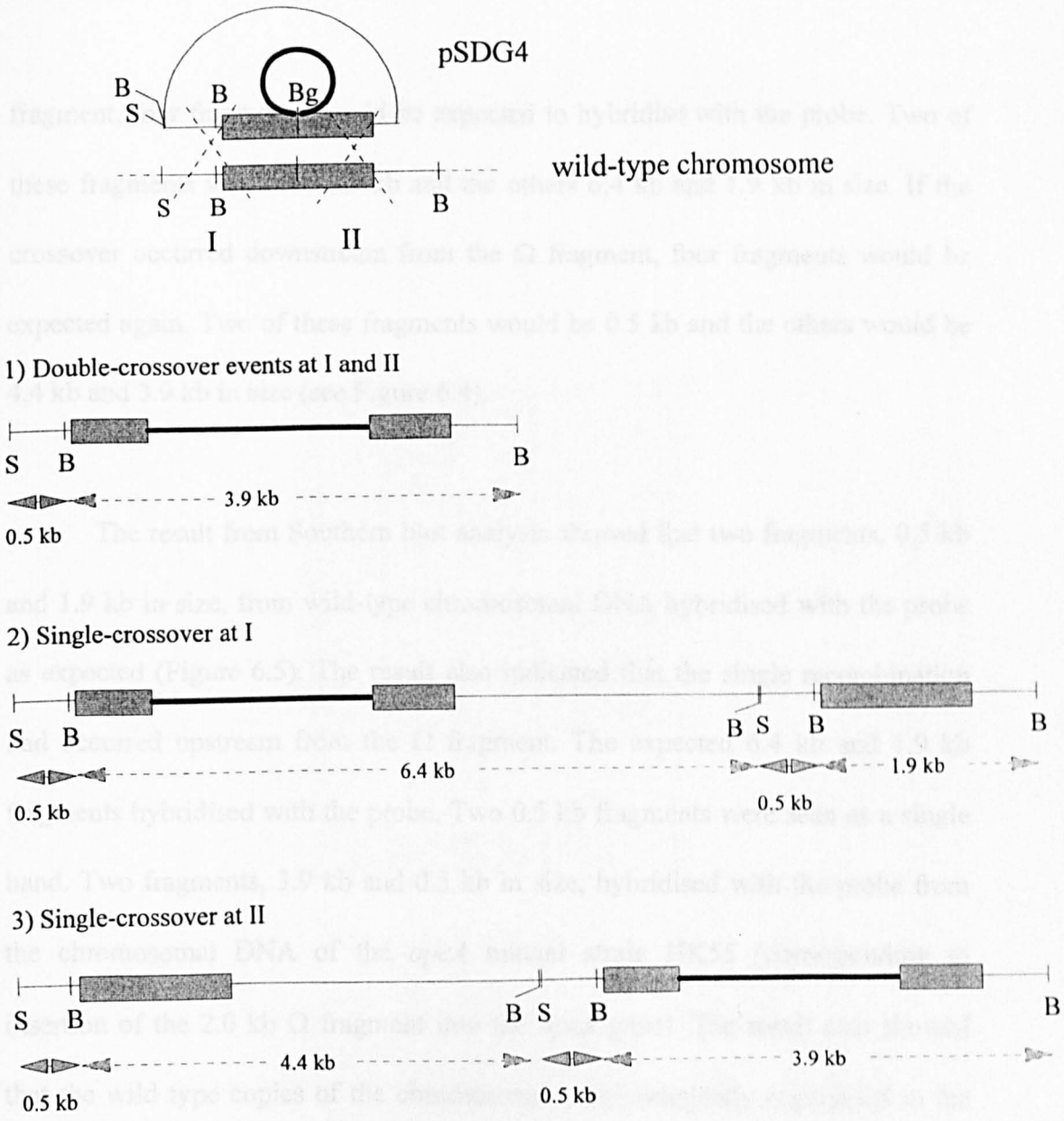


Figure 6.4: Integration of pSDG4 in the *Synechococcus* sp. PCC7942 chromosome by single and double crossover events. All crossover possibilities are schematically represented. The sizes are arbitrary (for correct sizes, see Figure 6.2). The sizes of the *Bam*HI/*Sal*I fragments of the recombinants are shown below the chromosome. Dark rectangle represents the *opcA* gene. The Ω fragment is shown as a thick bar. Abbreviations: B, *Bam*HI; Bg, *Bgl*II; S, *Sal*I.

fragment, four fragments would be expected to hybridise with the probe. Two of these fragments would be 0.5 kb and the others 6.4 kb and 1.9 kb in size. If the crossover occurred downstream from the Ω fragment, four fragments would be expected again. Two of these fragments would be 0.5 kb and the others would be 4.4 kb and 3.9 kb in size (see Figure 6.4).

The result from Southern blot analysis showed that two fragments, 0.5 kb and 1.9 kb in size, from wild-type chromosomal DNA hybridised with the probe as expected (Figure 6.5). The result also indicated that the single recombination had occurred upstream from the Ω fragment. The expected 6.4 kb and 1.9 kb fragments hybridised with the probe. Two 0.5 kb fragments were seen as a single band. Two fragments, 3.9 kb and 0.5 kb in size, hybridised with the probe from the chromosomal DNA of the *opcA* mutant strain HK55 (corresponding to insertion of the 2.0 kb Ω fragment into the *opcA* gene). The result also showed that the wild type copies of the chromosome were completely segregated in the mutant strain HK55.

6.6.2 Growth rates of wild-type and the mutant strains HK52 and HK55

Growth rates of the wild-type, the single recombinant strain HK52 and the *opcA* mutant strain HK55 were followed for 14 days under continuous illumination. Strain HK52 was grown in media containing both streptomycin and ampicillin, whilst HK55 was grown in media containing only streptomycin. Wild-

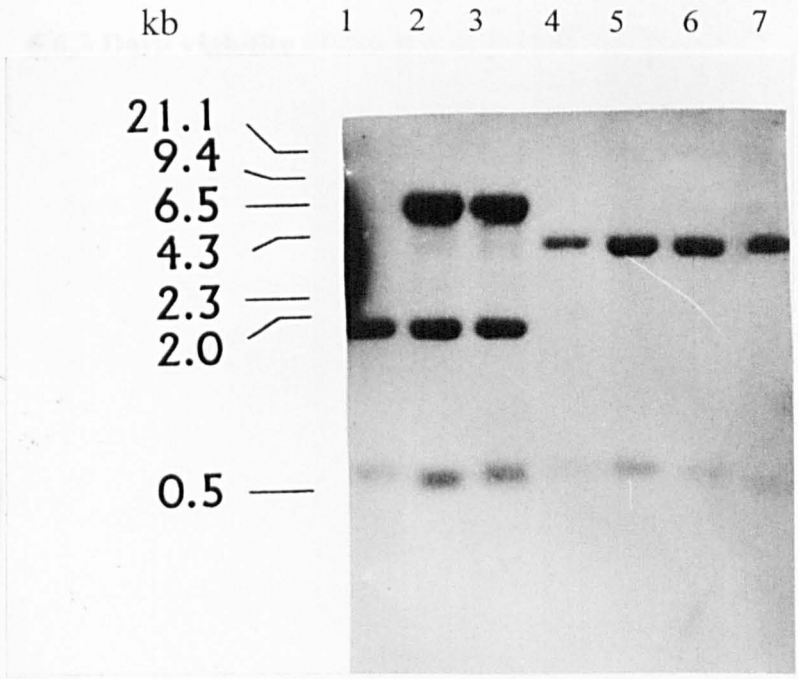


Figure 6.5: Southern blot analysis of wild-type, single and double recombinants. Chromosomal DNA was digested with both *Sal*I and *Bam*HI. Lane 1 is wild-type, lanes 2 and 3 single recombinant strain HK52. Lanes 4, 5, 6 and 7 are double recombinants, *opcA* mutant strain HK55. All the double recombinants have mutated copies of *opcA* (3.9 kb). None of the double recombinants have a wild-type copy of the *opcA* gene.

type and the mutant strains HK52 and HK55 exhibited a similar growth rate (see Figure 6.6).

6.6.3 Dark viability of the strains HK52 and HK55

The dark viability of wild-type, single recombinant and *opcA* mutant cells was assayed by dilution plating on BG11 medium as described in Sections 2.26 and 2.27. The maintenance of viability after relatively prolonged periods in the dark was investigated (Table 6.1). The *opcA* mutant strain HK55 showed zero viability after 72 h dark incubation while 47% of wild type cells were still viable. The single recombinant strain HK52 cells showed an unstable viability rate throughout the incubation period in the dark. Scanlan *et al.* (1995) suggested that the slow loss of viability of the *zwf* mutant *Synechococcus* sp. PCC7942 cells in the dark may be because of the increasing starvation for an essential metabolite, probably a C_5 intermediate. Similarly, the loss of the viability of the *opcA* mutant cells, in which G6PDH activity decreased 98.6% (see Section 6.6.4), may be because of the starvation for such a metabolite in.

6.6.4 Glucose-6-phosphate dehydrogenase assays in extracts of the mutant strains HK52 and HK55

As mentioned earlier, the *opcA* gene was considered to be involved in the assembly of the functional form of the G6PDH. If this is the case, no G6PDH activity or a low level of activity would be expected in the *opcA* mutant strain HK55 due to the interruption of the gene in this strain. Similarly, the single recombinant strain HK52 would be expected to have similar G6PDH

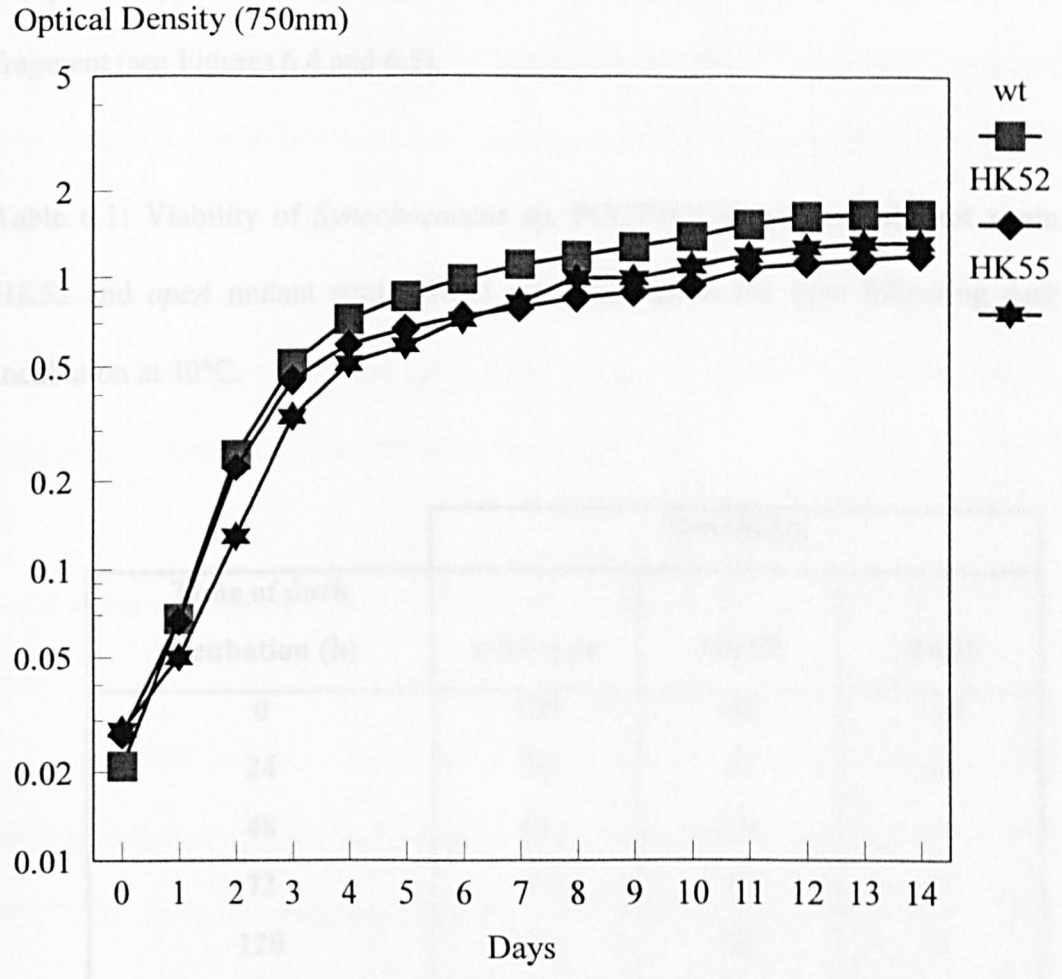


Figure 6.6: Growth of wild-type, single recombinant strain HK52 and *opcA* mutant strain HK55. Cells were grown in BG11 medium and under continuous illumination.

activityn with the *opcA* mutant strain HK55, because the strain does not have a complete copy of the *opcA* gene upstream from the insertion site of the Ω fragment (see Figures 6.4 and 6.5).

Table 6.1: Viability of *Synechococcus* sp. PCC7942, single recombinant strain HK52 and *opcA* mutant strain HK55 after transfer to the light following dark incubation at 30°C.

	%viability		
Time of dark incubation (h)	**wild-type**	**HK52**	**HK55**
0	100	100	100
24	78	32	26
48	81	18	5
72	47	29	0
120	24	62	0
216	5	15	0

G6PDH activities in wild-type, HK52 and HK55 were 110.9, 27.0 and 1.5 units/mg protein, respectively. Only 1.4% of the wild-type activity was retained in the *opcA* mutant strain HK55. This was consistent with the prediction that the gene affected the functional form of G6PDH. The single recombinant strain HK52 retained 24% of the wild-type activity. Since this strain does not have the complete gene upstream from the interruption point, it is unclear why the strain

has 24% G6PDH activity. Summers *et al.* (1995b) reported a low level transcription of the *opcA* gene in *zwf* mutant strain of *Nostoc* sp. ATCC29133 while no independent *opcA* trancript existed in wild-type cells. They explained this with the possibility that this trancript may be produced by a fortuitous promoter caused by the *zwf* inactivation. This 25% activity in the single recombinant strain HK52 implies the presence of a weak promoter, rather than a fortuitous one, upstraem of the *opcA* gene, which may operate when transcription signals are not received from the other promoters.

6.6.5 Determination of glucose-6-phosphate dehydrogenase activity in native gels

G6PDH activities were detected, using non-denaturing polyacrylamide gels, in cell-free extracts from a range of cyanobacteria including *Anabaena* sp. PCC7120, *Synechococcus* sp. PCC7942, *Plectonema boryanum* sp. PCC73110, *Synechocystis* sp. PCC6803, *Nostoc* MAC PCC8009 and marine strain *Synechococcus* sp WH7803 (Sundaram *et al.*, in press). To quantify the activities *in situ* and also to obtain the sizes of the active enzyme, G6PDH activities were detected in native gels as described in Section 2.24.1. The results showed that *in situ* activities in the wild-type and the mutant strains HK52 and HK55 were consistent with spectrophotometrically determined activities (see Figure 6.7). However, a 150 kDa activity band was detected in the single recombinant strain HK52. It is currently unclear what the function of this band is in this strain.

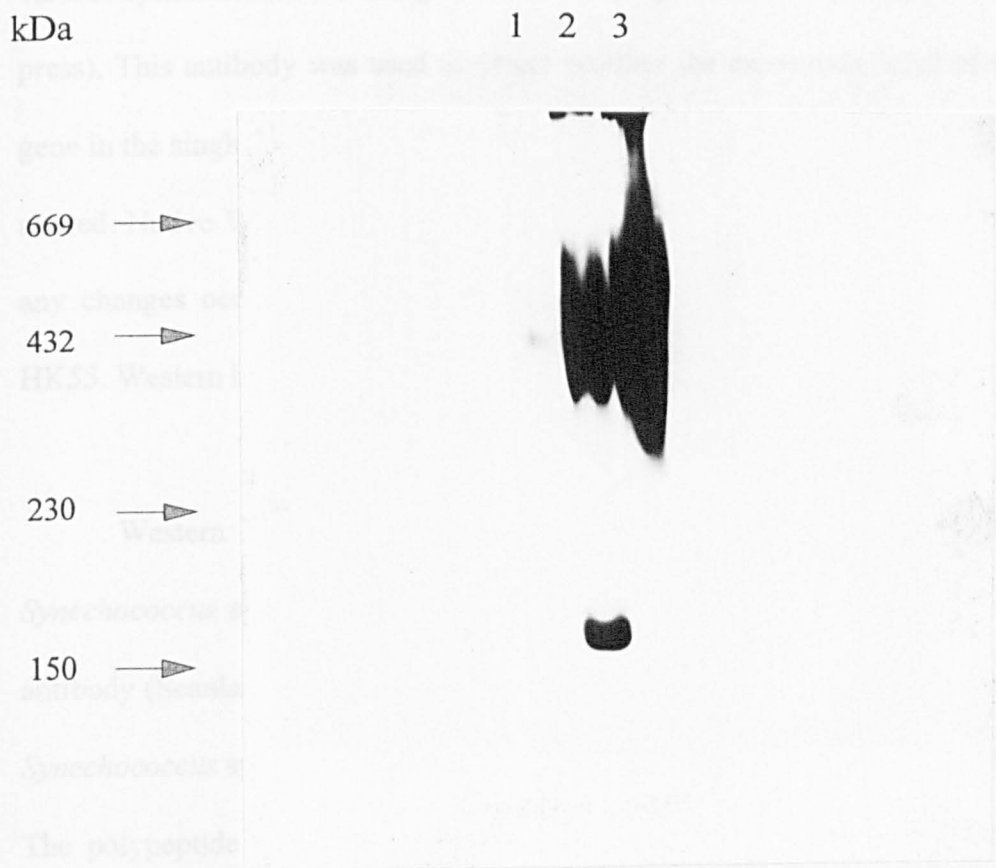


Figure 6.7: G6PDH activity staining of a native polyacrylamide gel of supernatants of wild-type, single recombinant strain HK52 and the *opcA* mutant strain HK55. 200 µl protein was loaded each lane. Lanes: 1, wild-type; 2, HK52; 3, HK55.

6.6.6 Western blotting analysis of the strains HK52 and HK55

An antibody produced against part of the G6PDH of *Anabaena* sp. PCC7120 and used to detect both denatured and native forms of G6PDHs from various cyanobacteria including *Synechococcus* sp. PCC7942 (Sundaram *et al.*, in press). This antibody was used to detect whether the expression level of the *zwf* gene in the single recombinant strain HK52 and the *opcA* mutant strain HK55 was altered. Native Western blot analysis was also carried out to investigate whether any changes occurred in the functional form of G6PDH in strains HK52 and HK55. Western blotting was performed as described in Section 2.25.4.

Western blot analysis of G6PDH from wild-type and *zwf* mutant of *Synechococcus* sp. PCC7942 had shown that a 55 kDa protein reacted with this antibody (Scanlan *et al.*, 1995). A 55 kDa polypeptide is represented in wild-type *Synechococcus* sp. PCC7942 and the mutant strains HK52 and HK55 (Figure 6.8). The polypeptide is present slightly higher in wild-type than strains HK52 and HK55. The other bands were also reacted with preimmune serum and are, therefore, disregarded. This result suggested that the *zwf* gene is expressed in the *opcA* mutant cellsand that the *opcA* gene is not essential for expression of the *zwf* gene. However, it may affect the quantity of expression. In nondenaturing gels, the antibody reacted with two protein bands from the wild-type extract which were of about 430 kDa in size (Figure 6.9). The upper band in strain HK52 was much weaker than that of wild type. In the *opcA* mutant strain HK55, the upper band completely disappeared and lower band was significantly reduced. The reduction or absence of the protein or proteins in the upper band resulted in a

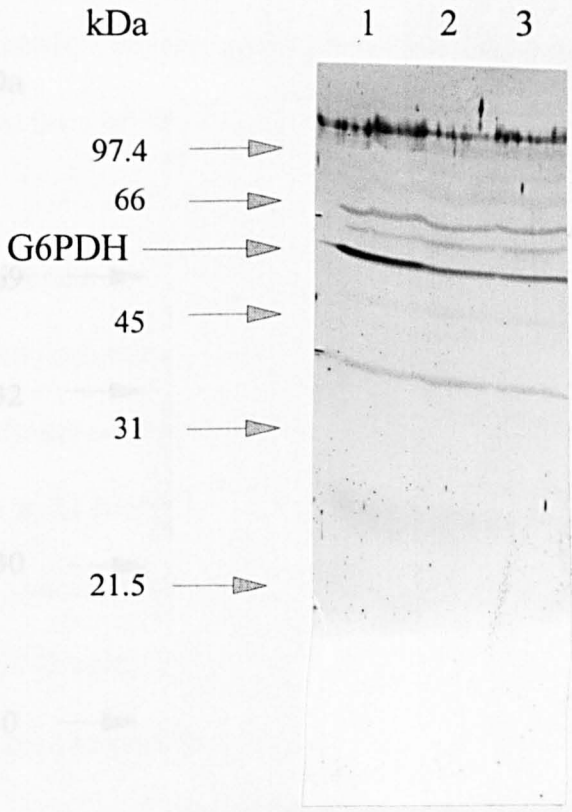


Figure 6.8: Western blot of the supernatants from wild-type, single recombinant strain HK52 and the *opcA* mutant HK52 in SDS-polyacrylamide gel. 50 µl of protein was loaded each lane. The blot was probed with antibody raised against a truncated *Anabaena* sp. PCC7120 G6PDH. Lanes: 1, wild-type; 2, HK52; 3, HK55.

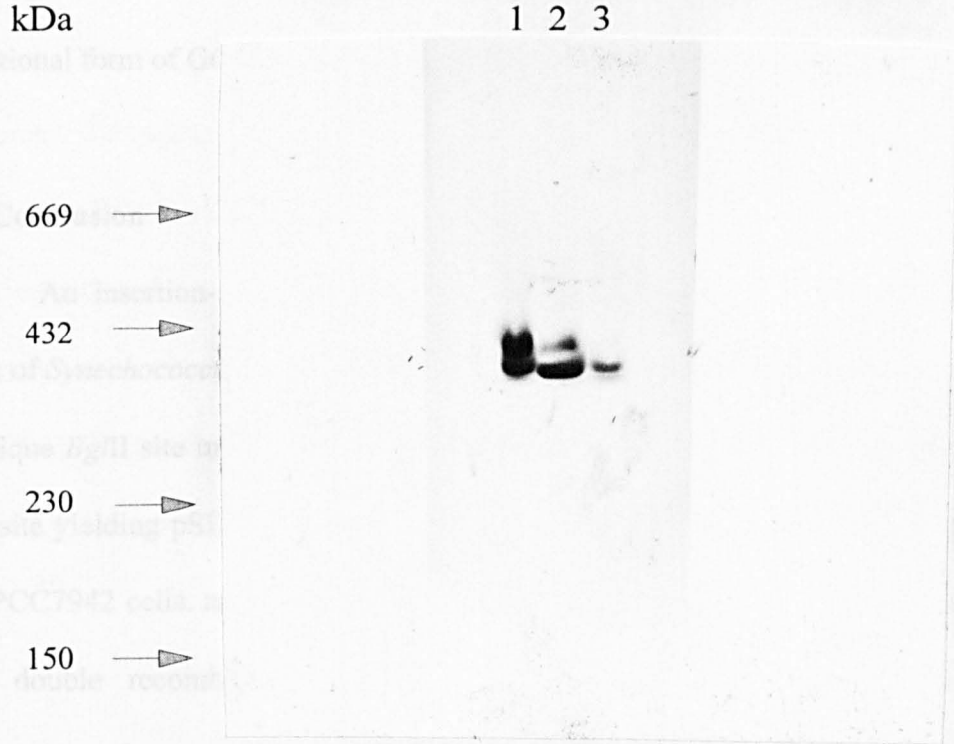


Figure 6.9: Western blot of supernatants from wild-type, single recombinant strain HK52 and *opcA* mutant strain HK55 in a native polyacrylamide gel. 200 µl protein was loaded each lane. The blot was probed with antibody raised against a truncated *Anabaena* sp. PCC7120 G6PDH. Lanes: 1, wild-type; 2, HK52; 3, HK55.

significant reduction in G6PDH activity. In *Nostoc* sp. ATCC29133, RNA blot analysis showed that the *opcA* gene did not affect transcription of *zwf* (Summers *et al.*, 1995b). This result is also supported by the lack of an identifiable DNA-binding motif. The results support the involment of the *opcA* gene in synthesis of functional form of G6PDH rather than in control of expression of *zwf*.

6.7 Conclusion

An insertion-inactivation approach was carried out to mutate the *opcA* gene of *Synechococcus* sp. PCC7942. A subclone, pSDG2, was produced to create a unique *Bgl*II site in the middle of the gene. The interposon Ω was inserted into this site yielding pSDG4. This subclone was transformed into the *Synechococcus* sp. PCC7942 cells, and then single recombinant cells (*Synechococcus* sp. HK52) and double recombinant cells (*Synechococcus* sp. HK55) were selected. Phenotypes of these strains were further characterised and compared with those of the wild-type strain. Growth rates of the wild-type and the strains HK52 and HK55 were similar throughout the 14 day growth period. The *opcA* mutant cells showed zero viability after a 72 h dark incubation period, whereas 47% of the wild-type cells were still viable. The single recombinant cells showed an unstable viability pattern throughout the dark incubation period. G6PDH activity of the *opcA* mutant strain HK 55 was 98.6% lower than that of wild-type strain while the single recombinant strain HK52 retained 24% of the wild-type activity. This result was confirmed by G6PDH activity staining of a nondenaturing gel. A lower activity band at 150 kDa which was not represented in either in the wild-type or in

double recombinant cells was also detected in single recombinant cells. Western blot analyses in denaturing gels suggested that G6PDH was expressed in in wild-type and mutant cells, with slightly higher quantity in wild-type, implying that the *opcA* mutation may not affect the expression of *zwf*. Native Western blot analysis indicated that two protein bands in the 430 kDa region strongly reacted with the antibody. The upper band was dramatically reduced in the single recombinant strain HK52. In the double recombinant strain HK55, the upper band completely disappeared and the lower band was dramatically reduced. The overall results suggested that the *opcA* gene affected G6PDH activity in *Synechococcus* sp. PCC7942. However, the gene is not essential for expression of the *zwf* gene. The product of the *opcA* gene perhaps affects assembly of the G6PDH.

Chapter 7

Conclusions

The OPP cycle is the main route for carbohydrate dissimilation in vegetative cells of cyanobacteria. In heterocysts, the cycle supplies reducing power to nitrogenase, thus, playing an important role in nitrogen fixation. G6PDH is the first enzyme of the cycle and one of the most important regulatory targets of cyanobacterial carbon metabolism (Stanier and Cohen-Bazire, 1977). Some questions on the function and regulation of the OPP cycle and G6PDH were still not completely answered when this study started. How do the OPP cycle and G6PDH function in the vegetative cells and heterocysts? Does a thiol/disulphide exchange mechanisms control the light-dark regulation of G6PDH activity or not? How do the different molecular forms of G6PDH assemble? To answer such questions, biochemical studies were undertaken, based largely on *in vitro* analysis of enzyme activity (Schaeffer and Stanier, 1978; Udvardy *et al.*, 1983; Cossar *et al.*, 1984). However, more recently molecular genetic studies on the G6PDH and the other OPP cycle enzymes have been established (Scanlan *et al.*, 1992; Summers *et al.*, 1995b) and further genetic studies would help to find answers to these questions about both the function and regulation of the OPP cycle and G6PDH in vegetative cells and heterocysts of cyanobacteria. This study provides some answers to these questions and generates a new dimension with regard to the assembly of the functional form of G6PDH.

The analysis in Chapter 3 showed that three genes are located in the *zwf* region: *fbp*, *tal* and *zwf* encoding FBPase, transaldolase and G6PDH, respectively. These three enzymes are all the components of the OPP cycle. The presence of *fbp*

and *tal* upstream from *zwf* strongly suggests that the OPP cycle in this organism may function to completely oxidise glucose-6-phosphate to CO_2 with the concomitant production of maximal amounts of NADPH rather than oxidation to ribulose-5-phosphate. The absence of a *tal* gene between *fbp* and *zwf* in *Synechococcus* sp. PCC7942 genome implies that the OPP cycle may play a special role in filamentous, heterocystous cyanobacteria. This result supports the idea of the OPP cycle being the supplier of reducing power to nitrogenase in heterocysts.

Sequence analysis of the *fbp* gene showed that the *Anabaena* sp. PCC7120 enzyme is more similar to the cytosolic form of plant FBPase rather than the chloroplast form. A region of 12-16 amino acids which was reported to be involved in light regulation of the chloroplast FBPase activity (Marcus *et al.*, 1988) is not present in the *Anabaena* sp. PCC7120 FBPase (Chapter 3). Consequently, it seems likely that if the cyanobacterial FBPases are subject to light dark regulation it is not via a thiol/disulphide exchange mechanism. It may be light-regulated by regulatory metabolites such as AMP rather than a thiol/disulfide exchange mechanism.

The results of the sequence analyses of the *zwf* gene of *Anabaena* sp. PCC7120 support the idea of the light-dark control of G6PDH activity by a thiol/disulphide exchange mechanism (Chapter 3). Four cysteine residues are present in the *Anabaena* sp. PCC7120 G6PDH, Cys-101, Cys-187, Cys-265 and

Cys-445. Two of these cysteines, Cys-187 and Cys-445, are absolutely conserved in three cyanobacterial *zwf* genes from *Anabaena* sp. PCC7120 (Chapter 3), from *Nostoc* sp. ATCC29133 and from *Synechococcus* sp. PCC7942. This reinforces the likelihood of their role in the regulation of enzyme activity. Site-directed mutagenesis of these two conserved cysteine residues would reveal if such a mechanism is involved in light regulation of G6PDH. However, this work could not be carried out because of time limitation.

Mutagenesis analysis in Chapter 5 suggested that there may be more than one transaldolase gene in *Anabaena* sp. PCC7120. The mutation of the *tal* gene should negatively affect the growth and survival of the cells in the absence of combined nitrogen, because *tal* mutant cells would not have a complete OPP cycle operation. However, the *tal* mutant cells showed a similar level of transaldolase activity to that of wild-type, and growth rate of the mutant strain HK29 in the presence and absence of combined nitrogen are not significantly different from wild type. All these results support the presence of more than one transaldolase gene in *Anabaena* sp. PCC7120. The results of mutagenesis of the *tal* gene also implied that the *tal* gene and the *zwf* gene are in same operon in the *Anabaena* sp. PCC7120 genome. The *zwf* gene is located downstream from the *tal* gene in the *Anabaena* sp. PCC7120 chromosome. If these genes are in the same operon as in *Nostoc* sp. ATCC29133, the mutagenesis of the *tal* gene would be expected to affect the *zwf* gene as well. This effect would be either complete inactivation or partial decrease in G6PDH activity. The *tal* mutant cells retained only 34.1% of

the wild type G6PDH activity. This implies that these two genes are located in a single operon and transcribed by more than one promoter. This result is consistent with that reported for the genes in the *zwf* region of *Nostoc* sp. ATCC29133 (Summers *et al.*, 1995b).

The results in Chapter 6 provided a new dimension for research into G6PDH in cyanobacteria. Deficiency of a gene, *opcA*, downstream from the *zwf* gene *Synechococcus* sp. PCC7942 resulted in 98.6% reduction in G6PDH activity. Western blot analysis implies that the product of the gene affects the assembly of the functional form of G6PDH rather than expression of the *zwf* gene. G6PDH was reported to be an oligomeric enzyme (Schaeffer and Stanier, 1978; Udvardy *et al.*, 1983; Cossar *et al.*, 1984; Sundaram *et al.*, in press). Oligomers with higher molecular weight are more active than the lower molecular weight forms. The results of mutagenesis of the *opcA* gene established a new dimension in the functional assembly of G6PDH. Even though the precise mechanism of the involvement of the *opcA* gene is not known yet, the gene appears to be needed for assembly of the functional form of G6PDH in *Synechococcus* sp. PCC7942.

Some important work related to this study could not be completed because of time limitation and is yet to be done. One of these topics is the site-directed mutagenesis of conserved cysteine residues of cyanobacterial G6PDH. Obviously, the question whether the enzyme is light regulated via a thiol/disulphide exchange mechanism would only be answered after modification of cysteine codons of a

cyanobacterial *zwf* gene, possibly those of *Synechococcus* sp. PCC7942 because this strain has only two cysteine residues in the G6PDH. Reduction of G6PDH activity in the *tal* mutant strain *Anabaena* sp. HK29 implies that the genes in the *zwf* region of *Anabaena* sp. PCC7120 are located in a single operon and regulated in a similar way to that in *Nostoc* sp. ATCC29133. However, no sequence information is yet available downstream from the *zwf* gene and it is not known if an *opcA* gene is present in the *Anabaena* sp. PCC7120 genome. It is of interest to find the nucleotide sequence downstream from the *zwf* gene of *Anabaena* sp. PCC7120. Further investigations on the precise role of the *opcA* gene in assembly of the functional form of G6PDH or in regulation at gene expression level is another work to be completed. Amino acid sequence analysis of the proteins from the active form of G6PDH can be determined to understand if the product of *opcA* is a part of the functional form and thus, if the functional form is a heterologous aggregate. For this purpose, an antibody can also be produced against the translation product of the *opcA* gene and then Western blot analysis can be performed using this antibody as well as that raised against G6PDH of *Anabaena* sp. PCC7120.

The research described in this thesis has, I hope, advanced our understanding of the nature of dissimilatory carbon metabolism in cyanobacteria. However, to my mind, it is more important that it has asked new questions regarding protein assembly that may be of much more general biological significance.

REFERENCES

Adams, D. G. and Carr, N. G. (1981). The developmental biology of heterocyst and akinete formation in cyanobacteria. CRC Crit. Rev. Microbiol. **9**: 45-100.

Alam, C., Curtis, S., Gleason, F. K., Gerami-Nejad, M. and Fuchs, J. A. (1989). Isolation, Sequence, and expression in *Escherichia coli* of an unusual thioredoxin gene from the cyanobacterium *Anabaena* sp strain PCC7120. J. Bacteriol. **171**: 162-171.

Allen, M. M. (1984). Cyanobacterial cell inclusions. Ann. Rev. Microbiol. **38**: 1-25.

Anderson, L. E., Nehrlich, S. C. and Champigny, M. (1978). Light modulation of enzyme activity. Activation of the light effect mediators by reduction and modulation of enzyme activity by thiol-disulphide exchange? Plant Physiol. **61**: 601-605.

Austin, P. A., Ross, I. S. and Mills, J. D. (1992). Light/dark regulation of photosynthetic enzymes within intact cells of the cyanobacterium *Nostoc* sp. *Mac.* Biochem. Biophys. Acta 1099:226-232.

Balogh, A., Borbély, G., Cséke, C., Udvardy, J. and Farkas, G. L. (1979). Virus infection affects the molecular properties and activity of glucose-6-P dehydrogenase in Anacystis nidulans, a cyanobacterium. FEBS Letters **105**: 158-162.

Bankier, A. T., Weston, M. K. and Barrell, B. G. (1987). Random cloning and sequencing by the M13/dideoxynucleotide chain termination method. In R Wu (ed.) Methods in Enzymology **155**: 51-93.

Bankier, A. T. and Barrell, B. G. (1989). Sequencing single-stranded DNA using the chain-termination method. In: CJ Howe and ES Ward (ed.) Nucleic acids sequencing. A practical approach, pp. 37-78. IRL Press, Oxford.

Barnell, W. O., Yi, K. C. and Conway, T. (1990). Sequence and genetic organisation of *Zymomonas mobilis* gene cluster that encodes several enzymes of glucose metabolism. J. Bacteriol. **172**: 7227-7240.

Barry, B. A., Boerner, R. J. and de Paula, J. C.(1994). The use of cyanobacteria in the study of the structure and function of photocystem II. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 217-257. Kluwer Academic Publishers, Dordrecht.

Bishop, R. H. (1979). Regulatory characteristics of a fructose bisphosphatase from the blue-green bacterium *Anacystis nidulans*. Arch. Biochem. and Biophys. **196**: 295-300.

Black, T. A., Cai, Y. and Wolk, C. P. (1993). Spatial expression and autoregulation of *hetR*, a gene involved in the control of heterocyst development in *Anabaena*. Mol. Microbiol. **9**: 77-84.

Böhme, H. (1987). Regulation of electron flow to nitrogenase in a cell-free system from heterocysts of *Anabaena variabilis*. Biochim. Biophys. Acta 891:121-128.

Brock, T. D. (1973). Evolutionary and ecological aspects of the cyanophytes. In: NG Carr and BA Whitton (ed.) The Biology of Blue-green algae, pp. 486-500. Blackwell Scientific Publications, Oxford.

Broedel Jr., S. E. and Wolf Jr., R. E. (1991). Groth-phase-dependent induction of 6-phosphogluconate dehydrogenase and glucose-6-phosphate dehydrogenase in the cyanobacterium *Synechococcus* sp. PCC7942. Gene **109**: 71-79.

Buchanan, B. B. (1980). The role of light in the regulation of chloroplast enzymes. Ann. Rev. Plant Physiol. **31**: 341-374.

Buikema, W. J. and Haselkorn, R. (1993). Molecular genetics ofcyanobacterial development. Ann. Rev. Plant Physiol. Plant Mol. Biol. **44**: 33-52.

Bustos, S. A. and Golden, S. S. (1992). Light-regulated expression of the *psbD* gene family in *Synechococcus* sp. strain PCC7942: evidence for the role of duplicated *psbD* genes in cyanobacteria. Mol. Gen. Genet. **232**: 221-230.

Cai, Y. and Wolk, C. P. (1990). Use of conditionally lethal gene in *Anabaena* sp. strain PCC7120 to select for double recombinants and entrap insertion sequences. J. Bacteriol. **172**: 3138-3145.

Cai, Y. (1991). Molecular genetic approaches towards the understanding of heterocyst differentiation and pattern formation in the cyanobacterium *Anabaena* sp. A desertation for the degree of doctor of phylosophy. Michigan State University.

Carr, N. G. (1973). Metabolic control and autotrophic physiology. In: NG Carr and BA Whitton (ed.) The Biology of Blue-green algae, pp.39-65. Blackwell Scientific Publications, Oxford.

Carrasco, J. L., Chueca, A., Prado, F.E., Hermoso, R., Làzaro, J.J., Ramos, J.L., Sahrawy, M. and Gorgé, J. L. (1994). Cloning, structure and expression of a pea cDNA clone coding for a phytosynthetic fructose-1,6-bisphosphatase with

some features different from those of the leaf chloroplast enzyme. Planta **193**: 494-501.

Carr, N. G. and Mann, N. H. (1994). The oceanic cyanobacterial picoplankton. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 27-48. Kluwer Academic Publishers, Dordrecht.

Castenholz, R. W. and Waterburry, J.B. (1989a). Oxigenic Photosynthetic Bacteria. Group 1 Cyanobacteria. Preface. In: JT Staley, MP Bryant, Pfenning N and JG Holt (ed.) Bergey's Manual of Systematic Bacteriology vol. 3, pp. 1710-1727. William&Wilkins Publications, Baltimore.

Castenholz, R. W. and Waterburry, J.B. (1989b). Taxa of the cyanobacteria. In: JT Staley, MP Bryant, Pfenning N and JG Holt (ed.), Bergey's Manual of Systematic Bacteriology vol. 3, pp. 1727-1728. William&Wilkins Publications, Baltimore.

Castenholz, R. W. (1989a). Subsection III. Order Oscillatoriales In: Staley, JT, Bryant, MP, Pfenning, N and Holt, JG (ed.), Bergey's Manual of Systematic Bacteriology vol. 3, pp. 1771-1780. William&Wilkins Publications, Baltimore.

Castenholz, R. W. (1989b).Subsection IV. Order Nostocales. In: JT Staley, MP Bryant, Pfenning N and JG Holt (ed.), Bergey's Manual of Systematic Bacteriology vol. 3, pp. 1780-1793. William&Wilkins Publications, Baltimore.

Castenholz, R. W. (1989c). Subsection V. Order Stigonematales. In: JT Staley, MP Bryant, Pfenning N and JG Holt (ed.) Bergey's Manual of Systematic Bacteriology vol. 3, pp. 1794-1799. William&Wilkins Publications, Baltimore.

Clancey, C. J. and Gilbert, H. F. (1987). Thiol/disulphide exchange in the thioredoxine-catalysed reductive activation of spinach chloroplast fructose-1,6-bisphosphatase. Kinetics and thermodynamics. J. Biol. Chem. **262**:13545-13549.

Codd, G. A. and Poon, G. K. (1988). Cyanobacterial toxins. In: LJ Roger and JR Gallon (ed.) Biochemistry of the Algae and Cyanobacteria pp. 284-296. Clarendon Press, Oxford.

Copeland, L. and Turner, J. F. (1987). The regulation of glycolysis and the pentose phosphate pathway. In DD Davies (ed.) The Biochemistry of Plants vol. 11, Biochemistry of Metabolism, pp. 107-128. Academic Press, San Diego.

Cossar, J. D., Rowell, P. and Stewart, W. P. (1984). Thioredoxin as a modulator of glucose-6-phosphate dehydrogenase in a N_2-fixing cyanobacterium. Journal of Genaral Microbiology 130:991-998.

Cséke, C., Balogh, A. and Farkas, G. L. (1981). Redox modulation of glucose-6-phosphate dehydrogenase in *Anacyctis nidulans* and its uncoupling by phage infection. FEBS Letters **126**: 85-88.

Curtis, S. E. and Martin, J. A. (1994). The transcription apparatus and the regulation of transcription initiation. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 613-639. Kluwer Academic Publishers, Dordrecht.

Devereux, J., Haeberli, P. and Smithies, O. (1984). A comprehensive set of sequence analysis programs for VAX. Nuclic Acids Research **12**: 387-395.

Doolittle, W. F. (1982). Molecular evolution. In: NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 307-331. Blackwell Scientific Publications, Oxford.

Douglas, S. E. (1994). Chloroplast origins and evolution. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 91-118. Kluwer Academic Publishers, Dordrecht.

Elhai, J. Wolk, C. P. (1988). Conjugal transfer of DNA to cyanobacteria. In: Methods in Enzymology 167, Cyanobacteria pp. 747-754. In: L Parker and AN Glazer (ed.) Academic Press Inc., San Diego.

Falkner, G., Horner, F., Werdan, K. and Heldt, H. W. (1976). pH changes in the cytoplasm of blue-green alga *Anacystis nidulans* caused by light-dependent proton flux into the thylakoid space. Plant Physiol. **58**: 717-718.

Fay, P. (1973). The heterocyst. In: Carr NG and Whitton BA(ed.) The Biology of Blue-green algae, pp. 238-259. Blackwell Scientific Publications, Oxford.

Fickenscher, K. and Scheibe, R. (1986). Purification and properties of the cytoplasmic glucose-6-phosphate dehydrogenase from pea leaves. Arch. Biochem. and Biophys. **247**: 393-402.

Fogg, G. G. (1987). Marine planktonic cyanobacteria. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 513-534. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Freedman, R. B., Hirst, T. R. and Tuite, M. F. (1994). Protein disulphide isomerase: building bridges in protein folding. Trends in Biochemical Sci. **19**: 331-336.

Frieden, C. (1979). Slow transitions and hysteretic behavior in enzymes. Ann. Rev. Biochem. **48**: 471-489.

Gantt, E. (1994). Supramolecular membrane organisation. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 119-138. Kluwer Academic Publishers, Dordrecht.

Gay, P. le Coq, D. Steinmetz, M. Ferrari, E. and Hoch, J. A. (1983). Cloning structural gene *sacB*, which codes for exoenzyme levansucrase of *Bacillus subtilis*: expression of the gene in *Escherichia coli*. J. Bacteriol. **153**: 1424-1431.

GCG (1992). Program manual for the Wisconsin package, Version 7.2, Genetic Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711.

GCG (1994). Program manual for the Wisconsin package, Version 8, Genetic Computer Group, 575 Science Drive, Madison, Wisconsin, USA 53711.

Gendel, S., Straus, N., Pulleyblank, D. and Williams, J. (1983). Shuttle cloning vectors for the cyanobacterium *Anacystis nidulans*. J. Bacteriol. **156**: 148-154.

Gibson, C. E. and Smith, R. V. (1982). Freshwater plancton. In: NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 463-489. Blackwell Scientific Publications, Oxford.

Giovannoni, S. J., Turner, S., Olsen, G. J., Barns, S., Lane,D. J. and Pace, N. R. (1988). Evolutionary relationship among cyanobacteria and green chloroplasts. J. Bacteriol. **170**: 3584-3592.

Glazer, A. N. (1987). Phycobilisomes: assembly and attachment. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 69-94. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Gleason, F. K. (1994). Thioredoxins in cyanobacteria: Structure and redox regulation of enzyme activity. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 714-729. Kluwer Academic Publishers, Dordrecht.

Golbeck, J. H. (1992). Structure and function of photosystem I. Annu. Rev. Plant Pysiol. Plant Mol. Biol. **43**: 293-324.

Golbeck, J. H. (1994). Photosystem I in cyanobacteria. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 319-360. Kluwer Academic Publishers, Dordrecht.

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

Golden, S. S. (1988). Mutagenesis of cyanobacteria by classical and gene-transfer-based methods. In: L Parker and AN Glazer (ed.) Methods of Enzymology 167, Cyanobacteria pp. 714-727 Academic Press Inc., San Diego.

Graeve, K., von Schaewen, A. and Scheibe, R. (1994). Purification, characterisation, and cDNA sequence of glucose-6-phosphate dehydrogenase from potato (*Solanum tuberosum* L.). Plant Journal **5**: 353-361.

Grigorieva, G. and Shestakov, S. (1982). Transformation in the cyanobacterium *Synechocystis* sp. 6803. FEMS Microbiology Letters **13**: 367-370.

Grossman, A. and McGowan, R. E. (1975). Regulation of glucose-6-phosphate dehydrogenase in blue-green algae. Plant Physiol. **55**: 658-662.

Grossman, A. R., Schaefer, M. R., Chiang, G. G. and Collier, J. L. (1993). The phycobilisome, a light-harvesting complex responsive to environmental conditions. Microbiol. Rev. **57**: 725-749.

Guo, L., Yang, R. C. A. and Wu, R. (1983). An improved strategy for rapid direct sequencing of both strands of long DNA molecules cloned in a plasmid. Nucleic Acids Res.. **11**: 5521-5540.

Haltiner, M., Kempe, T. and Tjian, R. (1985). A novel strategy for constructing clustered point mutations. Nucleic Acids Res. **13**: 1015-1025.

Hames, B. D. and Rickwood, D. (1982). Gel electrophoresis of proteins: a practical approach. IRL Press, Oxford.

Herdman, M., Janvier, M., Waterbury, J. B., Rippka, R.and Stanier, R. Y. (1979a). Deoxyribonucleic acid base composition of cyanobacteria. J. Gen. Microbiol. **111**: 63-71.

Herdman, M., Janvier, M., Rippka, R., Stanier, R. Y. and Mandel, M. (1979b). Genome size of cyanobacteria. Journal of General Microbiology **111**: 72-85.

Herdman, H., (1982). Evolution and genetic properties of cyanobacterial genomes. In: NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 263-305. Blackwell Scientific Publications, Oxford.

Herdman, M. (1987). Akinetes: structure and function. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 227-250. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Heyer, H., Stal, L. and Krumbein, W. E. (1989). Simultaneous heterolactic and acetate fermentation in the marine cyanobacterium *Oscillatoria limosa* incubated anaerobically in the dark. Arch. Microbiol.**151**: 558-564.

Ho, K. K. and Krogmann, D. W. (1982). Photosynthesis. In: NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 191-214. Blackwell Scientific Publications, Oxford.

Horsnell, P. R. and Raines, C. A. (1991). Nucleotide sequence of a cDNA clone ending chloroplast fructose-1,6-bisphosphatase from *Arabidopsis thaliana.* Plant Mol. Biol. **17**: 185-1991.

Houmard, J. Capuano, V., Coursin, T. and Tandeau de Marsac, N. (1988). Genes encoding core components of the phycobilisome in the cyanobacterium *Calothrix* sp. PCC7601: Occurrence of a multigene family. J. Bacteriol. **170**: 5512-5521.

Houmard, J. and Tandeau de Marsac, N. (1988). Cyanobacterial genetic tools: current status.In: L Parker and AN Glazer (ed.) Methods in Enzymology 167, Cyanobacteria pp. 808-847, Academic Press Inc., San Diego.

Hughes, S. G., Bruce, T. and Murray, K. (1980). The isolation and characterisation of sequence-specific endonuclease from *Anabaena subcylindrica.* Biochem. J. 185:59-63.

Hughes, S. G. and Murray, K. (1980). The nucleotide sequences recognized by endonucleases *AvaI* and *AvaII* from *Anabaena variabilis.* Biochem. J. **185**: 65-75.

Hunter, C. N. and Mann, N. H. (1992). Genetic manipulation of photosynthetic prokaryotes. In: P Fay and C Van Baalen (ed.) Photosynthetic Prokaryotes Biothecnology Handbooks vol. 6, pp. 153-179. Plenium Press, New York.

Hur, Y., Unger, E. A. and Vasconcelos,A. C. (1992). Isolation and characterisation of a cDNA encoding cytosolic fructose-1,6-bisphosphatase from spinach. Plant Mol. Biol. **18**: 799-802.

Hylemon, P. B.and Phibbs Jr., P.V. (1972). Independant regulation of hexose catabolizing enzymes and glucose transport activity in *Pseudomonas aeruginosa.* Biochem. Biophys. Res. Comm. **48**: 1041-1048.

Ihlenfeldt, M. J. A., Gibson, J. (1975). CO2 fixation and its regulation in *Anacystis nidulans* (*Synechococcus*). Arch. Microbiol. **102**: 13-21.

Jeffery, J., Wood, I., Macleod, A., Jeffery, R.and Jornwall, H. (1989). Glucose-6-phosphate dehydrogenase. Characterisation of a reactive lysine residue in the *Pichia jadinii* enzyme reveals a limited structural variation in a functionally significant segment. Biochem. Biophys. Res. Comm. **160**: 1290-1295.

Kallas, T. (1994). The cytochrome b_6f complex. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 259-317. Kluwer Academic Publishers, Dordrecht.

Ke, H., Thorpe, C. M., Seaton, B. A., Limpscomb, W. N. and Marcus, F. (1989). Structure refinement of fructose-1,6-bisphosphatase and its fructose 2,6-bisphosphate complex at 2.8 Å resolution. J. Mol. Biol. **212**: 513-539.

Kerby, N. W. and Stewart, D. P. (1988). The biotechnology of microalgae and cyanobacteria. In: LJ Roger and JR Gallon (ed.) Biochemistry of the Algae and Cyanobacteria pp. 319-334 edited by Clarendon Press, Oxford.

Khayat, E., Harn, C. and Daie, J. (1993). Purification and light-dependent molecular modulation of the cytosolic fructose-1,6-bisphosphatase in sugerbeet leaves. Plant Physiol. **101**: 57-64.

Koßmann, J., Müller-RÖber, B., Dyer, T. A., Raines, C. A., Sonnewald, U. and Willmitzer, L. (1992). Cloning and expression analysis of the plastidic fructose-1, 6-bisphosphatase coding sequence from potato: circumstantial evidence for the import of hexose into chloroplasts. Planta **188**: 7-12.

Krogmann, D. W. (1973). Photosynthetic reactions and components of thylakoids. In: NG Carr and BA Whitton (ed.) The Biology of Blue-green algae, pp. 80-98. Blackwell Scientific Publications, Oxford.

Kuhlemeier, C. J., Borrias, W. E., van den Hondel, C. A. M. J. J. and van Arkel, G. A. (1981). Vectors for cloning in cyanobacteria: Construction and characterisation of two recombinant plasmids capable of transformation to *Escherichia coli* K12 and *Anacystis nidulans* R2. Mol. Gen. Genet. **184**: 249-254.

Kuhlemeier, C. J., Thomas, A. A. M., van der Ende, A., van Leen, R. W., Borrias, W. E., van den Hondel, C. A. M. J. J. and van Arkel, G. A. (1983). A host-vector system for gene cloning in the cyanobacterium *Anacystis nidulans* R2. Plasmid **10**: 156-163.

Labarre, J., Chauvat, F. and Thuriaux, P. (1989). Insertional mutagenesis by random cloning of antibiotic resistance genes into the genome of the cyanobacterium *Synechocystis* strain PCC 6803. J. Bacteriol. **171**: 3449-3457.

Lambert, G. R. and Carr, N. G. (1984). Resistance of DNA from filamentous and unicellular cyanobacteria to restriction endonuclease cleavage. Biochim. Biophys. Acta **781**: 45-55.

Latzko, E. and Gibbs, M. (1969). Enzyme activities of the carbon reduction cycle in some photosynthetic organisms. Plant Physiol. **44**: 295-300.

Leach, C. K. and Herdman, M. (1973). Structure and function of nucleic acids. In: NG Carr and BA Whitton (ed.) The Biology of Blue-green algae, pp. 186-200. Blackwell Scientific Publications, Oxford.

Lee, W. T., Flynn, T. G., Lyons, C. and Levy, H. R. (1991). Cloning of the gene and amino acid sequence for glucose 6-phosphate dehydrogenase from *Leuconostoc mesenteroides*. The Journal of Biological Chemistry **266**: 13028-13034.

Levering, P. R. and Dijkhuizen, L. (1990). Transaldolase isoenzymes from *Arthrobacter* P1. In: MF Lidstrom (ed.) Methods in Enzymology vol. 188, Hydrocarbons and Methylotrophy, pp. 405-411. Academic Press, San Diego.

Levy, H. R. (1979). Glucose-6-phosphate dehydrogenases. Advances in Enzymology **48**: 97-192.

Lex, M. and Carr, N. G. (1974). The metabolism of glucose by heterocysts and vegetative cells of *Anabaena cylindrica*. Arch. Microbiol. **101**: 161-167.

Lind, L. K., Kalla, S. R., Lönneborg, A., Öquist, G. and Gustafsson, P. (1985). Cloning of the β-phycocyanin gene from *Anacystis nidulans*. FEBS Letters **188**: 27-32.

Lubberding, H. J. and Bot, P. V. M. (1984). The influence of temperature on glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase and the regulation of these enzymes in a mesophilic and a thermophilic cyanobacterium. Arch. Microbiol. **137**: 115-120.

Maniatis, T., Fritsch, E. F. and Sambrook, J. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Marcus, F., Moberly, L. and Latshaw, S. P. (1988). Comparative amino acid sequence of fructose-1,6-bisphosphatases: Identification of a region unique to the light-regulated chloroplast enzyme. Proc. Natl. Acad. Sci. USA **85**: 5379-5383.

Margheri, M. C. and Allotta, G. (1993). Homoacetic fermentation in the cyanobacterium *Nostoc* sp. strain Cc from *Cycas circinalis*. FEMS Microbiology Letters **111**: 213-218.

Martin, W., Brinkmann, H., Savonna, C. and Cerf, R. (1993). Evidence for a chimeric nature of nuclear genomes: Eubacterial origin of eukaryotic glyceraldehyde-3-phosphate dehydrogenase genes. Proc. Natl. Acad. Sci. USA **90**: 8692-8696.

Matthijs, H. C. P. and Lubberding, H. J. (988). Dark respiration in cyanobacteria. In: LJ Roger and JR Gallon (ed.) Biochemistry of the Algae and Cyanobacteria pp. 132-145, Clarendon Press, Oxford.

McFarlane, G. J. B., Machray, G. C. and Stewart, W. D. P. (1987). A simplified method for conjugal gene transfer into the filamentous cyanobacterium *Anabaena* sp. ATCC 27893. Journal of Microbiological Methods **6**: 301-305.

Messing, J. and Bankier, A. T. (1989). The use of single-stranded DNA phage in DNA sequencing. In: CJ Howe and ES Ward (ed.) Nucleic acids sequencing. A practical approach, pp. 1-36. IRL Press, Oxford.

Miller, J. H. (1992). A short course in bacterial genetics. A laboratory manual and handbook for *E. coli* and related bacteria. Cold Spring Harbor Laboratory Press.

Moezelaar, R. and Stal, L. J. (1994). Fermentation in the unicellular cyanobacterium *Microcystis* PCC7806. Arch. Microbiol. **162**: 63-69.

Morand, L. Z., Cheng, R. H., Krogmann, D. W. and Ho, K. K. (1994). Soluble electron transfer catalysts of cyanobacteria. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 381-407. Kluwer Academic Publishers, Dordrecht.

Moser, D. P., Zarko, D. and Kallas, T. (1993). Characterisation of a restriction barrier and electrotransformation of the cyanobacterium *Nostoc* PCC 7121. Arch. Microbiol. **160**: 229-237.

Murphy, G. and Ward, E. S. (1989). Sequencing of double-stranded DNA. In: CJ Howe and ES Ward (ed.) Nucleic acids sequencing. A practical approach, pp. 99-115. IRL Press, Oxford.

Nicholas, J. M. and Adams, D. G. (1982). Akinetes. In: NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 387-412. Blackwell Scientific Publications, Oxford.

Nogae, I. and Johnston, M. (1990). Isolation and characterisation of the *ZWF1* gene of Saccharomyces cerevisiae, encoding glucose-6-phosphate dehydrogenase. Gene **96**: 161-169.

Ormerod, J. G. (1992). Physiology of the photosynthetic procaryotes In: Mann NH and Carr, NG (ed.) Photosynthetic Prokaryotes Biothecnology Handbooks vol. 6, pp. 93-121. Plenium Press, New York.

Padan, E. and Cohen, Y. (1982). Anoxygenic photosynthesis. In: NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 215-235. Blackwell Scientific Publications, Oxford.

Peschek, G. A. (1987). Respiratory electron transport. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 119-161. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam, New York, Oxford.

Porter, R. D. (1986). Transformation in cyanobacteria. CRC Crit. Rev. in Microbiology **13**: 111-132.

Porter, R. D. (1988). DNA transformation. In: L Parker and AN Glazer (ed.) Methods of Enzymology 167, Cyanobacteria pp. 703-712, Academic Press Inc., San Diego.

Prentki, P., Karch, F., Iida, S. and Meyer, J. (1981). The plasmid cloning vector pBR325 contains a 482 base-pair-long inverted duplication. Gene **14**: 289-299.

Prentki, P. and Krisch, H. M. (1984). *In vitro* insertional mutagenesis with a selectible DNA fragment. Gene **29**: 303-313.

Prentki, P., Binda, A. and Epstein, A. (1991). Plasmid vectors for selecting IS1-promoted deletions in cloned DNA: sequence analysis of omega interposon. Gene **103**: 17-23.

Richardson, L. L. and Castenholz, R. W. (1987). Enhanced survival of the cyanobacterium Oscillatoria terebriformis in darkness under anaerobic conditions. App. Env. Microbiol. **53**: 2151-2158.

Richmond, A. (1992). Mass culture of cyanobacteria. In: P Fay and C Van Baalen (ed.) Photosynthetic Prokaryotes, Biothecnology Handbooks vol. 6 pp. 181-210. Plenium Press, New York.

Rippka, R. (1988). Isolation and purification of cyanobacteria. In: L Parker and AN Glazer (ed.) Methods of Enzymology 167, Cyanobacteria pp. 3-27. Academic Press Inc., San Diego.

Rippka, R., Deruells, J., Waterburry, J. B., Herdman, M. and Stanier, R.Y. (1979). Generic assignments, strain histories and priperties of pure cultures of cyanobacteria. J. Gen. Microbiol. **111**: 1-61.

Rowell, P., Darling, A. J., Amla, D. V. and Stewart, W. D. P. (1988). Thioredoxin and enzyme regulation. In: LJ Roger and JR Gallon (ed.) Biochemistry of the Algae and Cyanobacteria pp. 201-216 Clarendon Press, Oxford.

Rowell, P. and Kerby, N. W. (1992). Potential and commercial applications for photosynthetic prokaryots. In: P Fay and C Van Baalen (ed.) Photosynthetic Prokaryotes Biothecnology Handbooks vol. 6, pp. 233-266. Plenium Press, New York.

Rowley, D., and Wolf Jr, R. E. (1991). Molecular characterisation of the *Escherichia coli* K-12 *zwf* gene encoding glucose 6-phosphate dehydrogenase. J. Bacteriol. **173**: 968-977.

Rutherford, A. W. (1989). Photosystem II, the water-splitting enzyme. TIBS **14** January 1989: 227-232.

Sanger, F., Nicklen, S. and Coulsen, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**: 5463-5467.

Saunders, V. A. (1992). Genetics of the photosynthetic prokaryotes. In: P Fay and C Van Baalen (ed.) Photosynthetic Prokaryotes, Biothecnology Handbooks vol. 6, pp. 121152. Plenium Press, New York.

Scanlan, D. J., Newman, J., Sebaihia, M., Mann, N. H. and Carr, N.G. (1992). Cloning and sequencing analysis of the glucose-6-phosphate dehydrogenase gene from the cyanobacterium *Synechococcus* PCC7942. Plant Mol. Biol. **19**: 877-880.

Scanlan, D. J., Sundaram, S., Newman, J., Mann, N. H. and Carr, N. G. (1995). Characterisation of a *zwf* mutant of *Synechococcus* sp. strain PCC7942. J. Bacteriol. **177**: 2550-2553

Schaaff, I., Hohmann, S. and Zimmermann, F. K. (1990). Molecular analysis of the structural gene for yeast transaldolase. Eur. J. Biochem. **188**: 597-603.

Schaeffer, F. and Stanier, R. Y. (1978). Glucose-6-phosphate dehydrogenase of *Anabaena* sp. Kinetic and molecular properties. Arch. Microbiol. **116**: 9-19.

Schnarrenberger, C., Flechner, A. and Martin, W. (1995). Enzymatic evidence for a complete oxidative pentose phosphate pathway in chloroplasts and an incomplete pathway in cytosol of spinach leaves. Plant Physiol. **108**: 609-614.

Schopf, J. W. and Walter, M. R. (1982). Origin and early evolution of cyanobacteria: the geological evidence. In: NG Carr and BA Whitton (ed.) The

Biology of Cyanobacteria, pp. 543-564. Blackwell Scientific Publications, Oxford.

Schmetterer, G. (1994). Cyanobacterial respiration. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 409-435. Kluwer Academic Publishers, Dordrecht.

Schnarrenberger, C., Flechner, A. and Martin, W. (1995). Enzymatic evidence for a complete oxidative pentose phosphate pathway in chloroplasts and incomplete pathway in the cytosol of spinach leaves. Plant Physiol. **108**: 609-614.

Schrautemeier, B. Böhme, H. and Böger, P. (1994). In vitro studies on pathways and regulation of electron transport to nitrogenase with a cell-free extract from heterocysts of *Anabaena variabilis*. Arch. Microbiol. **137**: 14-20.

Sherman, L., Bricker, T. Guikema, J and Pakrasi, H. (1987). The protein composition of the photosynthetic complexes from the cyanobacterial thylakoid membrane. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 1-33. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Sidler, W. A. (1994). Phycobilisome and phycobiliprotein structures. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 139-216. Kluwer Academic Publishers, Dordrecht.

Silman, N. J., Carr, N. G. and Mann, N. H. (1995). ADP-ribosylation of glutamine synthetase in the cyanobacterium *Synechocystis* sp. strain PCC 6803. J. Bacteriol. **177**: 3527-3533.

Simon, R. D. (1987). Inclusion bodies in the cyanobacteria: cyanophycin, polyphosphate, polyhedral bodies. In: P Fay and C Van Baalen (ed.) The

Cyanobacteria, pp. 199-225. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Simon, R., Priefer, U. and Puhler, A. (1983). A broad host range mobilisation system for *in vivo* genetic engineering: transposon mutagenesis in Gram negative bacteria. Biotechnology November 1993, 784-791.

Simpson, G. G. (1990). Studies on the regulation of glucose-6-phosphate dehydrogenase from the N2-fixing cyanobacterium *Anabaena variabilis*. PhD thesis, University of Dundee.

Smith, A. J. (1973). Synthesis of metabolic intermediates. In: NG Carr and BA Whitton (ed.) The Biology of Blue-green algae, pp. 1-38. Blackwell Scientific Publications, Oxford.

Smith, A. J. (1982). Modes of cyanobacterial carbon metabolism. In: NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 47-85. Blackwell Scientific Publications, Oxford.

Solomon, D. H., Raynal, M., Tejwani, G. A. and Cayre, Y. E. (1988). Activation of the fructose-1,6-bisphosphatase gene by 1,25-dihydroxyvitamin D_3 during monocystic differentiation. Proc. Natl. Acad. Sci. USA **85**: 6904-6908.

Sprenger, G. A., Schorken, U., Sprenger, G. and Sham, H. (1995). Transaldolase B of *Escherichia coli* K-12: Cloning of its gene, *talB*, and characterisation of the enzyme from recombinant strains. J. Bacteriol. **177**: 5930-5936.

Srivastava, D. K. and Anderson, L. E. (1983). Isolation and characterisation of light- and dithiothreitol-modulable glucose-6-phosphate dehydrogenase from pea chloroplasts. Biochim. Biophys. Acta **724**: 359-369.

Stanier, R. Y. and Cohen-Bazire, G. (1977). Phototrophic prokaryotes: the cyanobacteria. Ann. Rev. Microbiol. **31**: 225-274.

Stanier, R. Y., Sistrom, W. R., Hansen, T. A., Whitton, B. A., Castenholz, R. W., Pfennig, N., Gorlenko, V. N., Kondratieva, E. N., Eimhjellen, K., E., Whittenbury, R., Gherna, R. L. and Truper, H. G. (1978). Proposal to place the nomenculature of cyanobacteria (blue-green algae) under the rules of the International Code of Bacteria. International Journal of Systematic Bacteriology **28**: 335-336.

Stewart, W. D. P. (1980). Some aspects of structure and function in N_2-fixing cyanobacteria. Ann. Rev. Microbiol. **34**: 497-536.

Stewart, A. C. (1988). Molecular biology of photosynthetic reaction centres. In LJ Roger and JR Gallon (ed.) Biochemistry of the Algae and Cyanobacteria pp. 105-120, Clarendon Press, Oxford.

Summers, M. L., Meeks, J. C., Chu, S. and Wolf Jr, R. E. (1995a). Nucleotide sequence of an operon in *Nostoc* sp. strain ATCC29133 encoding four genes of the oxidative pentose phosphate cycle. Plant Physiol. **107** : 267-268.

Summers, M. L., Wallis, J. G., Campbell, E. L. and Meeks, J. C. (1995b). Genetic evidence of a major role for glucose-6-phosphate dehydrogenase in nitrogen fixation and dark growth of the cyanobacterium *Nostoc* sp. strain ATCC29133. J. Bacteriol. **177**: 6184-6194.

Sundaram, S., Scanlan, D. J., Carr, N. G. and Mann, N. H. (in press). Multiple molecular forms of glucose-6-phosphate dehydrogenase in cyanobacteria.

Sutcliffe, J. and Church, G. M. (1978). The clevage site of the restriction endonuclease *Ava*II. Nucleic Acids Res. **5**: 2313-2319.

Tabita, F. R. (1987). Carbon dioxide fixation and its regulation in cyanobacteria. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 95-117. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Tabita, F. R. (1994). The biochemistry and molecular regulation of carbon dioxide metabolism in cyanobacteria. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 437-467. Kluwer Academic Publishers, Dordrecht.

Takizawa, T., Huang, I., Ikuta, T. and Yoshida, A. (1986). Human glucose-6-phosphate dehydrogenase: Primary structure and cDNA cloning. Proc. Natl. Acad. Sci. USA **83**: 4157-4161.

Tandeu de Marsac, N., Houmard, J. (1987). Advances in cyanobacterial molecular genetics. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 251-302. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Tandeu de Marsac, N., Houmard, J. (1993). Adaptation of cyanobacteria to environmental stimuli: New steps towards molecular mechanisms. FEMS Microbiology Reviews 104: 119-190.

Thiel, T. and Poo, H. (1989). Transformation of a filamentous cyanobacterium by electroporation. J. Bacteriol. **171**: 5743-5746.

Thiel, T. (1994). Genetic analysis of cyanobacteria. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 581-611. Kluwer Academic Publishers, Dordrecht.

Thomas, C. M. and Smith, C. A. (1987). In compatibility group P plasmids: Genetics, evolution, and use in genetic manupulation. Ann. Rev. Microbiol. 41:77-101

Udvardy, J., Gode, M. M. and Farkas, G. L. (1982). Regulatory properties of a fructose-1,6-bisphosphatase from the cyanobacterium *Anacystis nidulans*. J. Bacteriol. **151**: 203-208.

Udvardy, J., Borbely, G., Juhasz, A. and Farkas,G. L. (1984). Thioredoxins and the redox modulation of glucose-6-phosphate dehydrogenase in *Anabaena* sp. strain PCC7120 vegetative cells and heterocysts. J. Bacteriol. **157**: 681-683.

Udvardy, J., Juhàsz, A. and Farkas, G. L. (1983). Interactions between hysteretic regulation and redox modulation of glucose-6-phosphate dehydrogenase from *Anacystis nidulans*. FEBS Letters 152:97-100.

Vinopal, R. T., Hillman, J. D., Schulman, H. Reznikoff, W. S. and Fraenkel, D. G. (1975). New phosphoglucose isomerase mutants of *Escherichia coli*. J. Bacteriol. **122**: 1172-1174.

Walsby, A. E. (1987). Mechanisms of buoyancy regulation by planktonic cyanobacteria with gas vesicles. In: P Fay and C Van Baalen (ed.) The Cyanobacteria, pp. 377-392. Elsevier Science Publisher B. V. (Biomedical Division), Amsterdam.

Ward, E. S. and Howe, C. J. (1989). Troubleshooting in chain-termination DNA sequencing. In: CJ Howe and ES Ward (ed.) Nucleic acids sequencing. A practical approach, pp. 79-97. IRL Press, Oxford.

Waterbury, J. B. and Rippka, R. (1989). Subsection I. Order Chroococales. In: Staley JT Bryant MP, Pfenning N and Holt JG (ed.) Bergey's Manual of Systematic Bacteriology, vol. 3 pp. 1728-1746. William & Wilkins publications, Baltimore.

Waterbury, J. B. (1989). Subsection II.Order Pleurocapsales In: Staley JT Bryant MP, Pfenning N and Holt JG (ed.) Bergey's Manual of Systematic Bacteriology, vol. 3 pp. 1747-1770. William & Wilkins publications, Baltimore.

Wertman, K., Wyman, A. R. and Botstein, D. (1986). Host/vector interactions which affect the ability of recombinant phage lambda clones. Gene **49**: 253-262.

Whitton, B. A. (1992). Diversity, ecology, and taxonomy of the cyanobacteria. In: P Fay and C Van Baalen (ed.) Photosynthetic Prokaryotes, Biothecnology Handbooks vol. 6, pp. 1-52. Plenium Press, New York.

Willetts, N. and Skurray, R. (1980). The conjugation system of F-like plasmids. Ann. Rev. Genet. 14:41-76.

Willetts, N. and Wilkins, B. (1984). Processing of plasmid DNA during bacterial conjugation. Microbiol. Rev. **48**: 24-41.

Williams, J. G. K. (1988). Construction of specific mutations in photosystem II photosynthetic reaction center by gentic engineering methods in *Synechocystis* 6803. In L Parker and AN Glazer (ed.) Methods in Enzymology 167, Cyanobacteria pp. 766-778. Academic Press Inc., San Diego.

Wilmotte, A. (1995). Molecular evolution and taxonomy of Cyanobacteria. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 1-25. Kluwer Academic Publishers, Dordrecht, Boston, London.

Wittrup, K. D. (1995). Disulphide bond formation and eukaryotic secretory productivity. Current Opinions in Biotechnology **6**: 203-208.

Woese, C. R. (1987). Bacterial evolution. Microbiol. Rev. **51**: 221-171.

Wolk, C. P. (1980). Heterocysts, ^{13}N, and N_2-fixing plants. In: Newton WE and Orme-Johnson WH (ed.) Nitrogen Fixation vol. II, pp.279-292 University Park Press, Baltimore.

Wolk, C. P. (1982). Heterocysts. In NG Carr and BA Whitton (ed.) The Biology of Cyanobacteria, pp. 359-386. Blackwell Scientific Publications, Oxford.

Wolk, C. P., Vonshak, A., Kehoe, P. and Elhai, J. (1984). Constraction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria. Prot. Natl. Acad. Sci. USA **81**: 1561-1565.

Wolk, C. P., Ernst, A. and Elhai, J. (1994). Heterocyst metabolism and development. In: DA Bryant (ed.) The Molecular Biology of Cyanobacteria, pp. 769-823. Kluwer Academic Publishers, Dordrecht.

Yanisch-Perron, C., Vieira, J. and Messing, J.(1985). Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and Puc19 vectors. Gene **33**: 103-119.

Yee, B. C., de la Torre, A., Crawford, N. A., Lara, C. , Charlson, D. E. and Buchanan, B. B. (1981). The ferredoxin/thioredoxin systhem of enzyme regulation on a cyanobacterium. Arch. Microbiol. **130**: 14-18.

Yura, T., Mori, H., Nagai, H., Nagata, T., Ishihama, A., Fujita, N., Isono, K., Mizobuchi, K. and Nakata, A. (1992). Systematic sequencing of the *Escherichia coli* geneome: analysis of the 0-2.4 region. Nucleic Acids Res. **20**: 3305-3308.