INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

CHARACTERIZATION OF A PLASMID FROM THE CYANOBACTERIUM *MICROCYSTIS AERUGINOSA*

by

MARGARET M. WALLACE

A dissertation submitted to the Graduate Faculty in Biochemistry
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy, The City University of New York

1997

This manuscript has been read and accepted for the Graduate Faculty in Biochemistry in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

4/17/97
Date

Co-Chair of Examining Committee
Professor Thomas Schmidt-Glenewinkel, Hunter College

4/17/97
Date

Co-Chair of Examining Committee
Professor Shirley Raps, Hunter College

April 24, 1997
Date

Executive Officer
Professor Horst Schulz

Professor Joseph Krakow, Hunter College

Professor Rivka Rudner, Hunter College

Professor Peter Lipke, Hunter College

Professor Eleanor Wurtzel, Lehman College

Professor Barbara Zilinskas, Rutgers University

Supervising Committee

The City University of New York

Abstract

CHARACTERIZATION OF A PLASMID FROM

THE CYANOBACTERIUM *MICROCYSTIS AERUGINOSA*

by

Margaret M. Wallace

Advisers: Professor Shirley Raps
Professor Thomas Schmidt-Glenewinkel

The complete nucleotide sequence of the plasmid pMa025 from the unicellular, toxin producing cyanobacterium *Microcystis aeruginosa* UV025 was determined. The plasmid is 8018 bp in length and has a 62.3 % G+C content.

Nineteen potential open reading frames (ORFs) were identified. The deduced amino acid sequence from a 1700 bp segment (5000-6700) of pMa025 was significantly similar to three putative cyanobacterial plasmid replication proteins. Within this region ORF M encodes a putative plasmid replication protein, RepA, that is significantly similar to a hypothetical plasmid replication protein from *Synechococcus* sp. strain PCC7942 plasmid pUH24.

Several regions of pMa025 showed significant similarity to putative genes and hypothetical proteins involved in bacterial polyketide and fatty acid synthesis. pMa025 may be involved in the biosynthesis of secondary metabolites produced by *Microcystis*.

A gene-transfer system, based on the shuttle vector pMaLIND7, was developed.

Acknowledgments

No man is an island. Nowhere is this more evident than in the world of multi-disciplinary scientific research. Fortunately, assistance has always be available. This work could not have been completed without the help and support of many individuals.

First, I wish to thank all the members of my committee: Dr. Joseph Krakow, Dr. Rivka Rudner, Dr. Peter Lipke, Dr. Thomas Schmidt-Glenewinkel, Dr. Eleanor Wurtzel and Dr. Barbara Zilinskas for their guidance. I would also like to express my gratitude to Dean Erwin Fleissner.

Next I would like to thank my electronic pen pals, Drs. Lynn Miller and Eric Cabot, at GCG. Their input was invaluable during the data analysis.

This manuscript could not have been prepared without the assistance of Ms. Violeta Zekaj.

The "Big Three" of cyberspace at Hunter: Dr. Lloyd Williams, Mr. Vadim Lekmanov and Mr. Raphael Rios deserve particular mention. A special thanks to Raphael who has been both a teacher and a friend.

Most importantly, I wish to thank my mentor, Dr. Shirley Raps, who took a chance on a returning student and turned her into a scientist. Dr. Raps never let me give up and long before the advertisement logo became popular she urged me to "just do it."

Lastly, I wish to dedicate this thesis to my mother, Mrs. George Wallace Jr.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

This study deals with the characterization of a plasmid from the cyanobacterium *Microcystis aeruginosa*. The cyanobacteria are Gram-negative prokaryotes that carry out oxygenic photosynthesis in a manner analogous to that of higher plants. While in green plants the components of the photosynthetic apparatus are encoded by both nuclear and chloroplast genes, those of the cyanobacteria are encoded by a single genome. Thus, in cyanobacteria, molecular studies of photosynthesis are not complicated by the nuclear-organelle interaction present in higher plants. Several genera of cyanobacteria have been shown to be transformable (Porter,1986; Tandeau de Marsac et al.,1987; Thiel,1994), a characteristic that has been exploited in the development of gene-transfer systems. These properties make the cyanobacteria attractive candidates for the study of photosynthesis at the molecular level.

Several cyanobacteria produce peptides responsible for toxic "blooms." The genus *Microcystis* produces microcystins which are potent hepatotoxins. The most prevalent, microcystin-LR (Namikoshi et al.,1992), is synthesized by the planktonic *M. aeruginosa*. Microcystin-LR is a cyclic heptapeptide which causes massive structural disorganization of the liver, intrahepatic bleeding and shock (Eriksson et al.,1989; Hooser et al.,1989,1991). The toxin is a potent inhibitor of type 1 and 2A protein phosphatases (Honkanen, et al., 1990; Mac Kintosh et al.,1990). The hepatic structural disorganization is thought to result from an imbalance in the level of phosphorylation and dephosphorylation of cytoskeletal proteins (Carmichael, 1994). Concomitant with its role as a phosphatase inhibitor, microcystin-LR is a strong tumor promoter in rat liver

(Nishiwaki-Matsushima et al.,1992). Although the incidence of water based toxicosis has been largely restricted to animals, several recent reports have implicated microcystins in human illness and death (Turner et al.,1990; Carmichael,1994; Toxic Cyanobacteria Web Site, 1996). Thus, knowledge of the synthesis and regulation of microcystins has ecological and epidemiological relevance.

Understanding and exploiting the molecular genetics of cyanobacteria would also contribute significantly to the field of public health. Mosquitos are vectors in several diseases. Construction of transgenic cyanobacteria expressing larvicidal genes from *Bacillus sphaericus* (Tandeau de Marsac et al.,1987) and *Bacillus thuringiensis* (Murphy et al.,1992; Soltes-Rak et al.,1995) shows promise for the control of insect larvae in aquatic habitats.

Manipulating the molecular genetics of cyanobacteria has potential use in agriculture, ecology, toxicology and public health.

Cyanobacterial Molecular Genetics

A. Plasmids

Studies of cyanobacterial molecular genetics have been hampered by the lack of suitable gene-transfer systems. Minimally, this requires a means of transferring DNA carried on a vector into a host, and selection for successful transfer by a change in phenotype. The choice of transfer system (natural, chemical or electro-transformation. conjugation), vector (cloning, shuttle, cargo) and marker (antibiotic resistance) depends upon the capabilities of the host and the purpose of the introduced DNA (Table I, Thiel,1994). Cloning vectors that do not replicate in cyanobacteria potentially allow

integration of introduced gene(s) into the bacterial chromosome. Vectors that replicate autonomously in cyanobacteria permit the establishment of stable merodiploids. Biphasic vectors that contain replicons for a cyanobacterium and *E. coli* are the most versatile. In these shuttle vectors the enterobacterium serves as the primary host for cloning and DNA manipulation (taking advantage of widely developed molecular biological techniques) followed by transfer of the vector into the cyanobacterium for molecular studies.

Pursuant to this, cyanobacteria have been screened for the presence of plasmids. Approximately 60% of those tested contain one or more (up to eight) plasmids (Tandeau de Marsac et al.,1987; Houmard et al.,1988). Sizes range from 1.3 kb (Potts,1984) to >100 kb (Rebière et al.,1986). Covalently closed circular DNAs have been isolated from unicellular (Asato et al.,1973; Restaino et al.,1975; Roberts et al.,1976; Friedberg et al.,1979; Lau et al.,1979,1980; van den Hondel et al.,1979,1989; Lambert et al.,1982. Hauman et al.,1982; Vakeria et al.,1985; Schwabe et al.,1988; Bose et al.,1990; Raps,1990; Tominaga et al.,1993,1995; Yang et al.,1993,1994) and filamentous (Simon,1978; Friedberg et al.,1979; Lambert et al.,1984; Felkner et al.,1988; Bose et al.,1990; Perkins et al.,1992; Vachhani et al.,1992) cyanobacteria.

Plasmid-encoded traits are believed to confer adaptations to local environments and to be rapidly disseminated, via horizontal transmission, in a bacterial population (Eberhard,1989). Indeed, in addition to intraspecific transfer, there is evidence for interspecific (Lau et al.,1979,1980; van den Hondel et al.,1979; Felkner et al.,1988) and intergeneric (van den Hondel et al.,1979, Lau et al.,1980; Felkner et al.,1988) transmission in cyanobacteria.

Although many characteristics such as resistance to antibiotics, high salt concentrations or heavy metals; toxin and gas vacuole production; and restriction modification enzyme synthesis have been assumed to be plasmid borne, there is as yet no definitive proof for such functions (Bose et al.,1990; Tandeau de Marsac et al.,1987). Other than genes for a sugar-nonspecific nuclease, *nucA* (Muro-Pastor et al.,1994), and cysteine biosynthesis (Nicolson et al.,1995), cyanobacterial plasmids are all phenotypically cryptic.

B. Transformation

Transformation is the transfer of naked DNA into cells. Competence refers to the ability of cells to incorporate exogenous DNA, a characteristic that is either present naturally or which can be induced chemically or electrically.

Only unicellular strains of cyanobacteria from the genera *Synechococcus* (Shestakov et al.,1970; Orkwiszewski et al.,1974; Stevens et al.,1980; Chauvet et al.,1983; Golden et al.,1984; Buzby et al.,1985; Essich et al.,1990) and *Synechocystis* (Devilly et al.,1977; Grigorieva et al.,1982) have been shown to be transformable (for a review see Porter,1986). All but one are naturally competent. The exception, *Synechocystis* sp. strain PCC6308, cannot be transformed without prior $CaCl_2$ treatment (Devilly et al.,1977).

Transfer of donor DNA from one organism into recipient cells of a different species or genus is termed heterospecific or heterogeneric transformation, respectively. Reciprocal heterospecific transformation between strains of *Synechocystis* has been demonstrated (Grigorieva et al.,1982). Nonreciprocal heterospecific transformation

between two *Synechococcus* strains has also been shown (Matsunaga et al.,1990). Reciprocal heterogeneric transformation between *Synechococcus* and *Synechocystis* strains has been reported (Devilly et al.,1977; Stevens et al.,1986). Such transfers occur between genetically related organisms, i.e. similar mole % G+C, and may be used to resolve taxonomic ambiguities.

C. Shuttle Vectors

Ideally, a shuttle vector would fulfill several requirements: 1) be small in size; 2) have replicons for both *E. coli* and a particular cyanobacterial strain; 3) contain one or more selectable markers (usually antibiotic resistance); 4) have a cloning site(s) for the incorporation of exogenous DNA; 5) be stably maintained in the transformants.

E. coli plasmids cannot replicate autonomously in cyanobacteria (Simon,1978; van den Hondel et al.,1980a,b; Kuhlemeier et al.,1981; Gendel et al.,1983a). Cyanobacterial plasmids cannot be maintained in *E. coli* (van den Hondel et al.,1980; Kuhlemeier et al.,1983). However, there is evidence that cloning vectors based on the broad-host-range IncQ/P4 plasmids RSF1010 and pKT230 (Bagdasarian et al.,1981) replicate in cyanobacteria (Kreps et al.,1990; Sode et al.,1992; Mermet-Bouvier et al.,1993). These gene-transfer systems involve conjugal transfer between *E. coli* and the cyanobacterium and, depending on the IncQ derivative, the addition of an IncP helper plasmid (Kreps et al.,1990) which codes for mobilization and transfer functions. Daniell et al. (1986) reported the successful transformation of a unicellular cyanobacterium, *Anacystis nidulans* 6301, with the *E.coli* plasmid, pBR322. However. transformation was more efficient in permeaplasts than intact cells, and required

prolonged (>18hr) contact between donor DNA and recipient. Although in these systems the donor plasmids shuttled between the hosts, transfer techniques were complicated and relatively inefficient.

Conjugal based gene-transfers systems have proven invaluable for molecular genetic studies in the non-transformable filamentous cyanobacteria (Wolk et al.,1984; Cobley et al.,1993; Vachhani et al.,1993). However, for the transformable unicellular cyanobacteria, biphasic vectors with replicons for both *E.coli* and a particular cyanobacterium would be useful. In these simpler systems, a single plasmid would shuttle efficiently between the respective hosts.

Such vectors have been developed for *Synechococcus* sp. PCC7942 (*Anacystis nidulans* R-2). This species harbors two cryptic plasmids of approximately 8 and 49 kb designated pUH24 and pUH25, respectively (van den Hondel et al.,1980) or pANS and pANL, respectively (Gendel et al.,1983a). In *Synechococcus* sp. strain PCC7942 the first shuttle vectors were constructed by adding antibiotic resistance genes to pUH24 either by *in vivo* transposon mutagenesis (van den Hondel et al.,1980) followed by *in vitro* ligation to an *E.coli* plasmid (Kuhlemeier et al.,1981) or by *in vitro* ligation of a transposon to an *E. coli* vector followed by transformation and cointegration with pUH24 (Sherman et al.,1982). In some instances, these chimeric plasmids were limited by their inability to replicate efficiently, in both *E. coli* and *Synechococcus* sp. PCC7942 and by the possession of few known unique restriction sites for cloning. Subsequently, several shuttle vectors were developed using either pANS(L) or deletion derivatives and *E. coli* plasmids (Gendel et al.,1983a,b,1987; Golden et al.,1983; Laudenbach et al.,1983,1985;

Lau et al.,1985). Many of these shuttle vectors contain several cloning sites (some within the *lacZ* gene, thus permitting rapid blue/white screening in the presence of X-gal) and multiple antibiotic resistance markers.

Shuttle systems for several other cyanobacteria have been developed: *Synechococcus* sp. strain PCC7002 (*Agmenellum quadruplicatum* PR-6; Buzby et al.,1983); *Anacystis nidulans* 6311 (Friedberg et al.,1983); *Synechococcus* sp. NKBG042902 (Matsunaga et al.,1990); *Synechocystis* sp. strain PCC6803 (Chauvat et al.,1986); *Calothrix* sp. strain PCC7601 (Cobley et al.,1993); *Anabaena* and *Nostoc* sp. strain PCC7524 (Wolk et al.,1984; Schmetterer et al.,1988; Murray et al.,1991) and *Plectonema* and *Anabaena* PCC7120 (Walton et al.,1993).

Since transformation efficiency is inversely proportional to plasmid size (Hananhan,1983), several shuttle vectors have been streamlined by using only that portion of an endogenous cyanobacterial plasmid that supports autonomous replication (Lau et al.,1985; Gendel et al.,1987,1991; Laudenbach et al.,1985; Schmetterer et al.,1988; Murray et al.,1991).

Shuttle vector stability has been improved by judicious placement of cyanobacterial sequences in the hybrid to prevent recombination of the shuttle vector with either the resident plasmid (Kuhlemeier et al.,1981; Golden et al.,1983) or the chromosome (Williams et al.,1983; van der Plas et al.,1990). The use of plasmid-cured strains (Kuhlemeier et al.,1983,1985; Gendel et al.,1991) combined with selection for markers on both the vector and cloned gene (Kuhlemeier et al.,1985) also enhances shuttle vector stability.

Characterization of Cyanobacterial Plasmids

A. Sequences

Sequences from several cyanobacterial plasmids have been reported (Wickrema, 1989; Perkins et al.,1992; van der Plas, et al.,1992; Walton et al.,1992,1993; Schaefer et al.,1993; Tominaga et al.,1993a,b,1995; Yang, et al.,1993,1994; Kurokawa et al.,1994). Wickrema (1989) sequenced a 1.5 kb plasmid, pGL3, from the filamentous cyanobacterium *Plectonema boryanum* PCC6306. Sequence analysis indicated 26 potential open reading frames (ORFs). Two of these, ORFs 1 and 2, were preceded by promoter and Shine-Dalgarno consensus sequences similar to those for *E.coli* and were postulated to code for proteins of 40 and 122 amino acids, respectively. The sequence of the 8 kb plasmid, pUH24, from *Synechococcus* sp. PCC7942 has been reported (van der Plas et al.,1992). There are 8 putative ORFs, two of which, F and E, overlap and are presumed to code for proteins, Rep A and B, involved in plasmid replication. While definitive proof for their actual coding function requires data from expression and protein sequencing studies, these two putative genes occupy part of a 3.6 kb fragment shared by all functional *Synechococcus* sp. PCC7942 cloning vectors (Kuhlemeier et al.,1983,1985; Gendel et al.,1983b,1987,1991; Lau et al.,1985). In addition, this 3.6 kb region contains features such as inverted and direct repeats and A-T rich stretches commonly found in prokaryote plasmid origins of replication (Scott,1984).

Interestingly, sequences from several small plasmids from diverse genera (Perkins et al.,1992; Walton et al.,1993, Tominaga et al.,1993a,b,1995; Yang et al.,1993,1994; Kurokawa et al.,1994) show homologies to proteins involved in replication and structural

motifs found in plasmids from Gram-positive bacteria that replicate by a rolling circle mechanism. A 4.2 kb plasmid, pRF1, from *Plectonema* sp. strain PCC6402 was sequenced by Perkins and Barnum (1992). Analysis revealed 7 putative ORFs. The predicted amino acid sequence of one, ORF C, showed homology to several plasmid replication (Rep) proteins from Gram-positive bacteria and to ORF 1 from plasmid pCA2.4 from *Synechocystis* sp. strain PCC6803 (Yang et al.,1993). In addition, regions upstream of ORF 1 contained promoter and Shine-Dalgarno consensus sequences similar to those in *E.coli*. Although the sequence of a second *Synechocystis* PCC6803 plasmid, pCB2.4, lacked overall homology to pCA2.4 (Yang et al.,1994), a 37 bp region of similarity was found. This region contained a putative "nicking" site used by replication proteins encoded by plasmids that replicate by a rolling circle mechanism.

A similar finding was obtained for two plasmids pMA1 (2.3kb) and pMA2 (5kb) from *Microcystis aeruginosa* f. *aeruginosa* Kűtzing (Tominaga et al.,1993a,1995). In both plasmids, a putative origin of replication (*ori*) region was identified which was A+T rich (63%) and contained a 19 bp consensus sequence characteristic of plasmids from Gram-positive organisms that replicate via a rolling circle mechanism. In pMA1, promoter and ribosome binding sites were identified upstream of an ORF which codes for a presumptive replication protein homologous to those found in Gram-positive bacteria. Although Shine-Dalgarno consensus sequences were identified upstream of two putative ORFs in pMA2, neither showed homology to any protein in the GenBank, SWISS-PROT or PIR databases.

Two 1.5 kb plasmids from the LPP group of cyanobacteria, pPF1 from

Phormidium foveolarum, and pGL3 from *Plectonema boryanum* sp. strain PCC6306. show overall sequence homology including a potential "nicking" cassette (Walton et al.,1993; Tominaga et al.,1993b; Kurokawa et al.,1994).

Lastly, the origins of replication from plasmids pFdA and pDU1, from *Fremyella diplosiphon* UTEX 481 and *Nostoc* sp. strain PCC7524, respectively, were characterized (Schaefer et al.,1993). The regions showed significant structural similarities. Each contained one large ORF encoding a presumptive Rep protein of similar size and predicted amino acid sequence, and an upstream region of dyad symmetry.

The sequence homologies observed among diverse cyanobacterial groups underscore the possibility of interspecific or intergeneric plasmid transfer. Because of their horizontal transmission, plasmid encoded genes could confer rapid adaptation to changing local environments (Eberhard,1989).

B. Coding functions

Few studies have demonstrated coding functions for plasmid borne genes. Expression in *E.coli* of the putative gene, *repA*, from *Synechocystis* sp. strain PCC6803 plasmid, pCA2.4 produced a protein with the expected N terminal sequence and predicted molecular mass (Yang et al.,1993). An 1800 base transcript from the 2.8 kb plasmid, pMA1, was detected in *Microcystis aeruginosa* strain HUB5-2-4 (Schwabe et al.,1990). Expression of this transcript in *E.coli* yielded two polypeptides of 43 and 50 kD. These studies indicate that, at least for some genes, cyanobacterial transcription and translation signals are recognized in *E.coli*.

Only two studies have definitively demonstrated coding functions for plasmid

borne genes. The 410 kb-α megaplasmid from *Anabaena* sp. strain PCC7120 carries the *nucA* gene which encodes a sugar non-specific nuclease (Muro-Pastor et al.,1994). This marker could be conjugally transferred and expressed in a *nucA*$^-$ strain. A 3.8 kb fragment from the large (49kb) endogenous plasmid, pANL, from *Synechococcus* sp. strain PCC7942 encodes two genes, *srpG & H,* for cysteine biosynthesis (Nicholson et al.,1995). Transcripts are regulated by sulfur availability and the positive transcription factor, Cys R. The *srpG* and *srpH* genes were able to complement *E.coli cysKcysM* and *cysE* mutants, respectively.

Current Experiments

The toxin producing cyanobacterium *Microcystis aeruginosa* UV027, which has been used in our laboratory to study phycobilisome organization, does not harbor a plasmid. However, a related strain UV025 possesses an 8 kb plasmid, pMa025 (Raps,1990). A hybrid plasmid, pMaL, has been constructed by cloning pMa025 into the Sal I site of pBluescript II and is being maintained in *E. coli* XL1-Blue. Southern hybridization failed to detect homology between UV027 chromosomal DNA and ^{32}P-labeled pMa025 (Wallace et al.,1995). Preliminary evidence indicates transformation to ampicillin resistance in both UV025 and UV027.

The large size of the recombinant plasmid (11 kb) limits its use as a shuttle vector. In order to streamline the hybrid plasmid, sequences extraneous to the cyanobacterial replicon and plasmid maintenance will be eliminated. This goal has been approached by: 1) determining the complete nucleotide sequence of pMa025, followed by a search of GenBank-EMBL for similarity to regions known to be involved in plasmid replication:

2) translation of the sequence followed by a search of the SWISS-PROT, PIR(R) and GenPept databases for similarity to proteins involved in plasmid replication and stability; 3) transformation of UV027 to carbenicillin resistance with the presumptive shuttle vector.

Such a vector will enable the introduction of altered genes into *Microcystis* and thereby expedite the study of the molecular genetics of photosynthesis and of toxin production. Wild-type copies of chromosomal genes can be "knocked out" using integration platforms (discussed in Thiel,1994). Altered genes can then be introduced, via a shuttle vector, and their functions assessed. In this system, the primary cloning and DNA manipulation would be done in *E.coli* (using widely developed molecular biological techniques) followed by transfer to *Microcystis* for study.

CHAPTER I. SEQUENCING CYANOBACTERIAL PLASMID, pMA025

METHODS

I. Constructs

A. pMaUC

Initially, the recombinant plasmid, pMaUC was used. In order to sequence both strands of pMA025 using the -40 Universal primer, the plasmid was cloned into the Sal I site of pUC19 (NEB) in two different orientations. The orientation of pMa025 was verified by double-digesting the hybrid plasmids with Hind III (cuts pUC19) and Sac II (cuts pMa025). Results of the digests were monitored electrophoretically. Plasmids showing bands of approximately 9 & 1 kb and 6 & 4 kb were designated pMaUC 1 & II, respectively.

pMaUC I & II were each amplified in *E.coli* XL1-Blue (Stratagene). Plasmids were isolated by alkaline lysis. Supercoiled plasmids were purified by equilibrium centrifugation in CsCl-EtBr density gradients (Sambrook et al.,1989) and stored in TE at -70°C.

pMaUC was severely limited in the number of restriction sites suitable for the production of nested deletions. Digestion with one endonuclease, PspA I, required high concentrations of enzyme (>10U/μg DNA), prolonged incubation times (16hr), stabilization with BSA (Gonzalez et al.,1977) and was extremely inefficient. The only manufacturer of PspA I, Stratagene, was unable to supply the enzyme for long periods of time (>6 mo.) on several occasions. In addition, very few deletions were obtained using the ExoIII/Mung Bean Nuclease kit (Stratagene). As a result, a minimal amount of

sequence was obtained. These constructs were, therefore, abandoned and will not be discussed further.

B. pMaL

Recombinant plasmids, pMaLI & pMaLII, were made by ligating linearized (SalI) pBluescript II SK- (Stratagene) to Sal I linearized pMa025. The orientation of pMa025 in these constructs was determined by sizing restriction fragments produced by digestion with Sac II. Constructs producing bands of approximately 9.4 & 1.6kb and 6.6 & 4.4kb were named pMaL I & II, respectively (Figure 1).

pMaLI and pMaLII were each amplified in *E. coli* XL1-Blue. Supercoiled plasmids were isolated by alkaline lysis, purified by CsCl-EtBr equilibrium centrifugation (Sambrook et al., 1989) and stored in TE (pH 8.0) at -70° C.

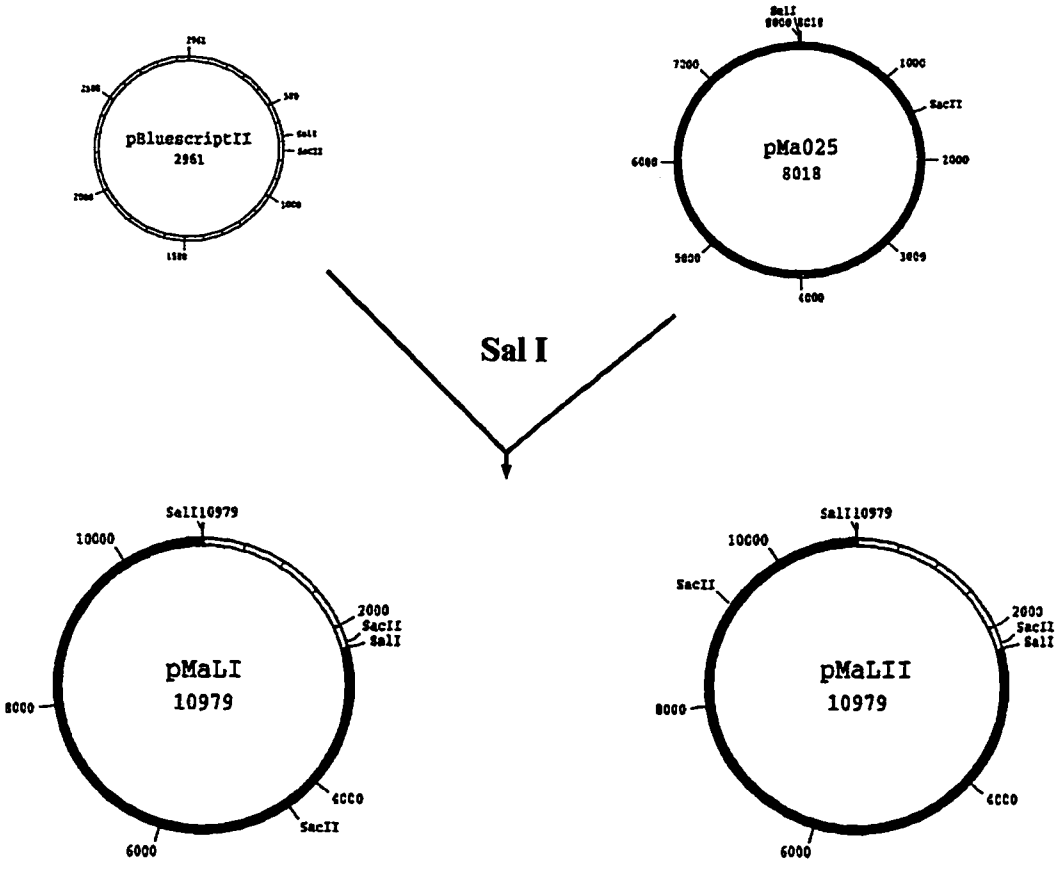


Figure 1a. Construction of pMaLI and pMaLII.

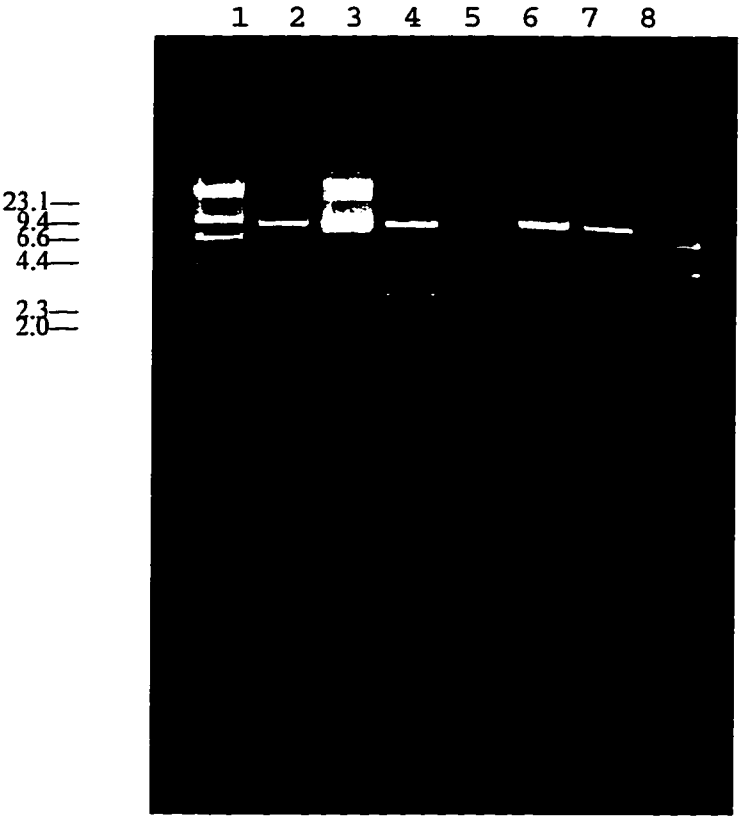


Figure 1b. Agarose gel electrophoresis of recombinant plasmids pMaLI and pMaLII. 0.8% agarose was used. The gel was run at 40V, 1.5 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 U.V. transilluminator.
Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lane 2, pMaUC; Lane 3, pMaLI; Lane 4, Sal I digest of pMaLI; Lane 5, Sac II digest of pMaLI; Lane 6, pMaL II; Lane 7, Sal I digest of pMaLII; Lane 8, Sac II digest of pMaLII.

II. Sequencing Using Nested Deletions

Because of limitations in resolving sequences longer than 200-300 bases, the large size of the insert (8kb) required the use of a method(s) which would result in fragments amenable to current techniques. One such method, the Exo III/ Nuclease technique, produces a series of unidirectional nested deletions of predictable size.

Exo III hydrolyzes ds DNA processively and uniformly in the 3'→5' direction. Products are 5' mononucleotides and residual ss DNA. ExoIII digests ds DNA with blunt ends or 5' overhangs, but will not digest ds DNA from 3' overhangs or 5' overhangs that have been filled in with α thiodeoxynucleotide derivatives (α thio-dNTPs). Since the enzyme can also initiate digestion at nicks, a high concentration of supercoiled DNA is required.

Briefly, supercoiled DNA is double digested to completion with two restriction enzymes that have sites located between the target DNA (site A) and the primer binding site (site B). Digestion at site A will result in a blunt end or a 5' overhang which is the substrate for Exo III; digestion at site B will produce either a 3' overhang or a 5' overhang which can be filled in with α thio-dNTPs. The double digested DNA is then treated with Exo III for various periods of time, aliquots are removed, and the reaction is terminated. Treatment with S1 Nuclease creates blunt ends which are ligated. The recircularized deletions are used to transform *E. coli*. The resulting deletion library is then screened for subclones of the requisite length (Figure 2).

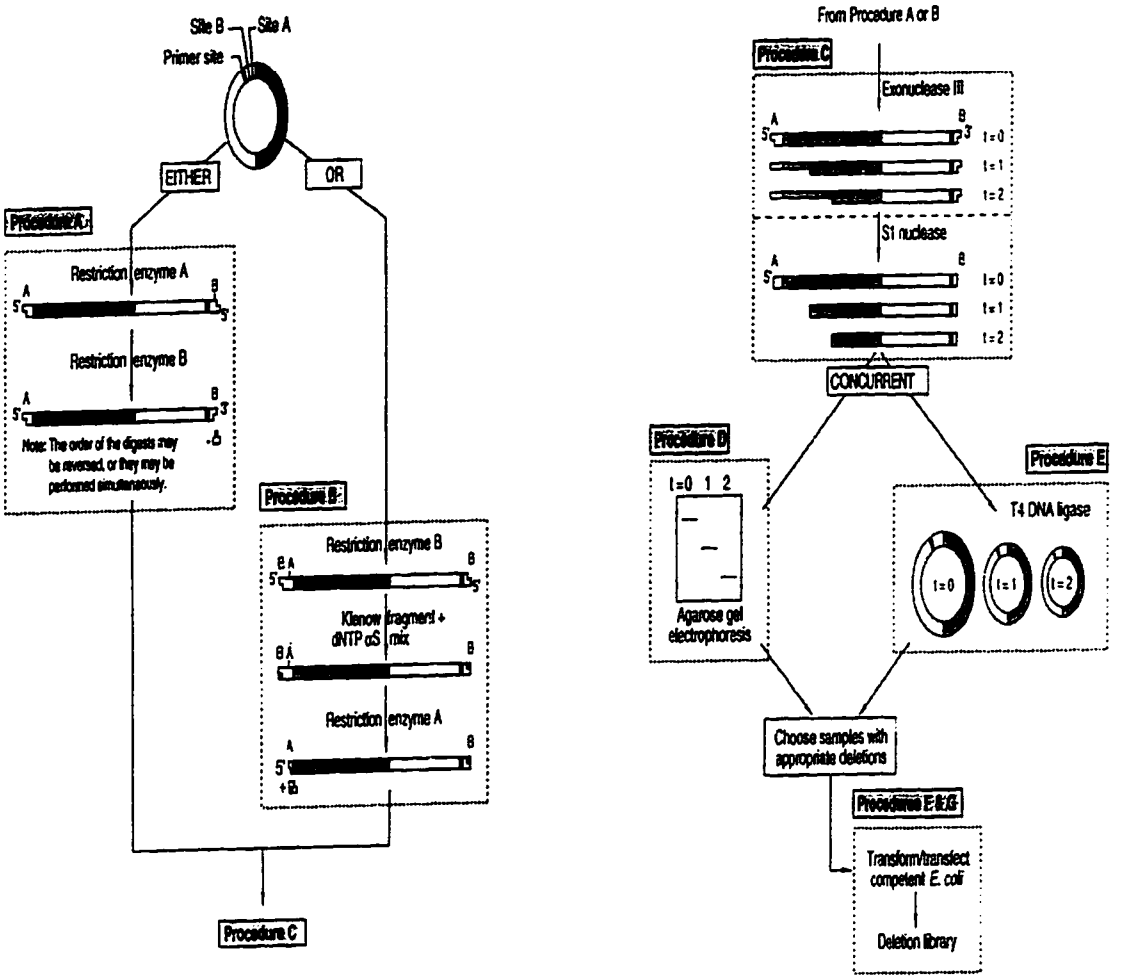


Figure 2. Nested deletion protocol.

A. Digestion & End Protection of Recombinant Plasmids

pMaLI was digested with Xba I (NEB) and end-protected with α thio-dNTPs. Digestion with Sma I produced a blunt end, the substrate for ExoIII.

Ten μg pMaLI was digested with Xba I (15U/μg DNA). Reaction conditions were: 1X One-Phor-All buffer PLUS (Pharmacia); total volume 60 μl; 37°C; 3 hrs. The enzyme was heat inactivated at 65°C for 20 minutes. Completeness of digestion was verified electrophoretically (0.8% agarose, 40V, 30 min, 1X TAE).

The DNA was end-protected by incubating the Xba I digest with 0.15U Klenow fragment of Pol I (Pharmacia), 3μl dNTPαS mix (400μM each of dATPαS, dCTPαS, dGTPαS & dTTPαS) in 1X Klenow buffer at 37°C for 15 min. The reaction was terminated by heat denaturation of the enzyme (65°C, 20 min.). The Klenow fragment must be removed to prevent residual polymerase/exonuclease activity from interfering with subsequent steps. The DNA was precipitated in 7.5 volumes (vol) of ethanol at -70°C for 10 min. Two vol NaCl (250mM)/glycogen (250ng/μl) was added as a carrier. Precipitated DNA was collected by centrifugation at 4°C for 20 min, washed in 3.5 vol 70% ethanol, reprecipitated and vacuum dried. The pellet was dissolved in 55 μl 1X One-Phor-All Buffer PLUS, resulting in a buffer that was compatible with subsequent enzymatic activities.

Linearized, end-protected DNA was digested with Sma I (6U/μg DNA, NEB) at 30°C for 3 hrs. This was followed by heat inactivation at 65°C for 20 min.

B. Preparation of Nested Deletions

Several pilot studies were conducted to determine the optimal conditions for the production of the nested deletions. The ds Nested Deletion kit (Pharmacia) was used with the following modifications: 1) the procedure was scaled up 3-fold; 2) the salt concentration was adjusted to 25mM NaCl during the ExoIII digestion; 3) the reaction was run at 37°C; 4) blunt-end ligation was carried out at 21-23°C for 4 hrs using T4 DNA ligase from NEB.

An equal vol of 2X ExoIII buffer was added to the double-digested, end-protected DNA. The final buffer was 1X ExoIII (66.6mM Tris-HCl (pH8.0), 0.66 mM $MgCl_2$), 25mM NaCl. The NaCl was provided by carry over from the One-Phor-All Buffer PLUS used in the Sma I digest. The mixture was equilibrated at 37°C, 3-5 min. A 2μl aliquot (time (t)=0) was removed and added to a chilled microfuge tube containing 3μl S1 nuclease /buffer mix. Three μl Exo III (300U) was added to the mixture. Thirty 2μl aliquots were removed at one minute intervals and added to the chilled S1 nuclease/buffer mix. The combination of high salt (50mM KAc (pH4.6), 0.42M NaCl, 1.66mM $ZnSO_4$) and low pH and temperature of the S1 nuclease/buffer mix terminated the Exo III reaction. Two sizing standards, linearized pMaLI and pBluescript II, were also incubated with S1 nuclease/buffer.

Samples from the 31 time points (t=0 to t=30), plus the sizing standards, were blunt-ended by incubating the mixtures at room temperature (22°C) for 30 min. This activated the S1 nuclease (1.5U/reaction) already present in the mix. The reaction was terminated by the addition of 1μl S1 Stop Solution (303mM Tris base, 50mM EDTA) to

each tube, followed by incubation at 65°C for 10 min.

Half (3µl) of the resultant blunt-ended DNA was analyzed electrophoretically (0.8% agarose, 24V, 6 hr., 1X TAE). The remainder was kept on reserve pending verification of the results.

C. Ligation

The time points were ligated using T4 DNA ligase (NEB). Each reaction contained: 3µl Exo III/S1 nuclease digest; 8U T4 DNA ligase; 1X ligase buffer; total vol 20 µl. Reactions were incubated at RT (21-23°C) for 4 hrs. Xba I and Sma I digested pMaLI were included as sticky and blunt-ended controls, respectively.

D. Deletion Library

A deletion library was constructed by using ligated plasmids from 30 deletion time points (t=1-30) to transform competent (Hanahan,1983) *E.coli* XL1-Blue cells to carbenicillin (CAR) resistance. Since the yield of the ligated plasmids was expected to be rather low, the entire 20 µl reaction was used to transform 200 µl *E.coli* (OD_{550} 0.61 ≈ $8x10^7$ cells/ml) cells. One-fifth of the total reaction, i.e. 200 µl, was plated. Carbenicillin (100µg/ml) was used in lieu of ampicillin since the ß-lactam ring of the former is more resistant to cleavage, and thus, the number of false positive (satellite) colonies was reduced.

Six colonies from transformants from each of the 30 time points were streaked onto LB/CAR/X-GAL/IPTG plates. These 180 streaks, corresponding to 6 possible deletions per time point, constituted the deletion library. Plasmid DNA from 4 streaks

from each plate was prepared (FMC SpinBind Miniprep), linearized (XhoI) and sized. The deletions were "ordered" and sequenced. Glycerol stocks, in duplicate, were stored at -70°C.

III. Sequencing

The chain-termination method (Sanger et al.,1977) of DNA sequencing involves the *in vitro* synthesis of a DNA strand by a DNA polymerase using ssDNA as a template. Synthesis is initiated at the site of an oligonucleotide primer and terminated by the incorporation of a dideoxy- nucleoside triphosphate (ddNTP) which does not support chain elongation. Four separate reactions, each with a specific dNTP/ddNTP ratio, provide complete sequence information. Incorporation of a radio-labeled dNTP (α ^{32} P-dATP or α ^{35} S-dATP) permits autoradiographic visualization of the sequence. Several pilot studies were done in order to optimize each step in the sequencing protocol.

A. Denaturing ds DNA and Annealing Template & Primer

Pilot studies indicated that, of the protocols available, sequence results were optimized using plasmid DNA prepared by the Quick Denature (USB) technique. This procedure has the advantage of being rapid and efficient. The elimination of an ethanol precipitation step reduces DNA loss, and thus, smaller quantities of template are required for preparation.

1N NaOH (2µl) was added to a mixture of ds DNA and primer (total reaction vol 11µl). Primer and plasmid DNA were used in a molar ratio of 4:1, i.e., 2.0:0.5 pmoles. This stoichiometry results in the best effective sequencing range. The alkali lowered the

Tm such that the DNA was denatured after a brief (10 min) incubation at 37°C. Addition of an equimolar amount of HCl neutralized the alkali. Plasmid Reaction Buffer was added (to fix the pH at 7.5) and the mixture was incubated at 37°C for 10 min to allow primer-template annealing.

B. Primers

The first strand of the 8kb insert was sequenced from the deletion library using the T3 primer (Stratagene). Gaps were filled in using custom-made primers (RCMI, Midland, Great American Gene Co.) designed using the GCG program Prime. Primers were selected according to the following criteria: 1) start point at least 50 bases upstream of the 3'end of the previously sequenced region; 2) 18-20 nucleotides in length; 3) approximately 50% GC content; 4) Tm between 55 & 65°C; 5) no long A/T runs; 6) G/ C at 3 out of 5, particularly the last, 3' terminal positions; 6) absence of secondary structure. In some instances, Prime failed to identify any potential primers. In these cases either one or more of the program's default criteria were modified or the primer was selected by eye. In the latter instance, primer binding uniqueness and lack of secondary structure were checked using the GCG programs, FindPatterns and StemLoop, respectively.

The second strand was sequenced from the opposite direction using the -40 Universal primer (USB) and custom-synthesized primers. The template DNA was the hybrid construct, pMaLI.

In every set of sequencing reactions, pBluescript II was included as a positive control, sequenced from either the T3 or -40 Universal primers.

C. Labeling

The labeling reaction both extends the primer and adds a radiolabeled dATP analogue. [α-^{35}S]dATP was chosen because its low energy ß particles produce a sharper image than [α-^{32}P]dATP on photographic film. The 5X labeling mix (7.5 μM each dGTP, dCTP and dTTP) was diluted 1:5 with double-distilled sterile water. Sequenase 2.0 (genetically modified T7 DNA polymerase, USB) was diluted 1:8 with enzyme dilution buffer. Pyrophosphatase (0.0075U) was included in the latter to avoid any sequence-dependent reversal of the DNA polymerase reaction, i.e., pyrophosphorolysis.

The following reagents were added to the annealed primer-template (on ice): 1 μl 0.1 M DTT, 2 μl diluted labeling mix, 0.5 μl [α-^{35}S]dATP (≈12 μCi/μl), 2 μl (3.25U) diluted Sequenase 2.0. The mixtures were incubated at 19-20°C for 5 min.

In those cases where template secondary structure produced gel artifacts, Sequenase and pyrophosphatase were diluted in glycerol buffer. The stabilization of enzymatic activities allowed the labeling reactions to be run at 37°C.

D. Termination Reactions

2.5 μl of each termination mixture was transferred to individual microfuge tubes. The concentration of each termination mixture was 80 μM for each of the four dNTPs and 8 μM for a particular ddNTP, i.e., a 10/1 dNTP/ddNTP ratio. The mixtures were equilibrated at 37°C for 2-3 min. 4.5 μl of the labeled reaction was transferred to each termination tube and incubated at 37° C for 8 min. When Sequenase and pyrophosphatase were stabilized by glycerol, the termination reactions were run at 50°C for 5 min.

If template secondary structure was refractory to sequencing at elevated temperatures, the labeling and termination reactions were run at the usual temperatures (19-20° & 37°C), followed by a terminal deoxynucleotidyl transferase (TdT) chase. In regions of high secondary structure Sequenase pauses and dissociates from the template. This results in incorrect termination i.e., with a deoxy instead of a dideoxy nucleotide. This artifact is evident as bands across four lanes in the sequencing gel. TdT is a DNA polymerase that nonspecifically adds nucleotides to a free 3′ hydroxyl group. The enzyme adds hundreds of nucleotides, in a relatively short period of time, to the nonspecifically terminated fragments. These long products remain at the top of the sequencing gel and no longer interfere with reading of the sequence. A 1mM solution of all four dNTPs (Pharmacia) was made in Sequenase Enzyme Dilution Buffer. TdT was added to the dNTP solution to a final concentration of 2U/μl. After the termination step, 1μl of the TdT/dNTP mix was added to each tube. The reactions were incubated 30 min at 37°C.

All reactions were stopped by the addition of 4 μl of Stop Solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol FF).

IV. Denaturing Gel Electrophoresis

5% polyacrylamide-7M urea denaturing gels were prepared as follows: 10 ml 50% acrylamide stock solution (Long Ranger, FMC), 42g urea, 20 ml 5X TBE, double-distilled deionized H_2O to give a total volume 99.5ml. 500 μl 10% ammonium persulfate and 50 μl TEMED were added to initiate polymerization. The gel was cast and allowed to set overnight. When sequencing reactions were run with glycerol-stabilized

Sequenase, gels were prepared with a glycerol-tolerant (tris-taurine-EDTA) buffer.

The sequencing samples were heat denatured at 70-80° C for 5 min and 1.5 µl was loaded onto the gel. To extract the maximum amount of information, gels were double-loaded, running 2hr 15min and 1hr 45min for the first and second loadings, respectively.

The Long Ranger gel solution contains a modified acrylamide monomer and a co-monomer with a novel cross-linker which does not precipitate urea and thus eliminated the need to fix the gel in acetic acid-methanol.

Gels were dried on a slab dryer for 1.5-2.0 hr at 70°C. Drying is necessary to prevent diffusion of the sequence bands, since exposure to photographic film is done at room temperature for ^{35}S labeled products.

To maximize visualization of the sequences, the gel was exposed to Kodak XAR film for 48 hrs and, if necessary, reexposed for 72-96 hrs.

The autoradiograph was developed using Kodak's X-OMAT M20 processor.

V. Analysis of Sequence Data

A. Sequence Assembly, Composition & Restriction Pattern

Sequences were read into a tape recorder and transcribed at least twice for each fragment. All analyses were performed using the programs of the GCG Sequence Analysis Software Package.

Raw sequence data were first entered as text files using the SeqEd program. Fragments were analyzed for overlaps and assembled into contigs using the Fragment Assembly Programs (based on Staden, R.,1980): GelStart, GelEnter, GelMerge and

GelAssemble. The resulting contigs were written to fragment files using the SeqOut command.

The final sequence was analyzed for G+C and trinucleotide content using the program Composition.

A circular restriction map was generated using MapSort and PlasmidMap.

B. Database Searches

Since the sequence (8018kb) was too large to be handled by some programs, nine overlapping subsets of 4kb each were created using the program Assemble. These files were then used as input for further analyses: similarity to sequences in nucleic acid and amino acid databases using the FastA (Pearson et al.,1988) and BLAST (Altschul et al.,1990; Gish et al.,1993) algorithms, respectively. FastA searches the GenEMBL nucleotide sequence databases. BLAST(X) translates both strands of the query sequence in all 6 reading frames and searches the non- redundant GenBank+EMBL+PDB+ SwissProt+SPupdate+PIR databases. The blast searches were performed at NCBI using the BLAST network. By handling the data in this manner, one could "walk" around the entire plasmid.

More refined BLAST searches were done on those regions of pMa025 that showed significant similarity ($P<10^{-3}$) to sequences in the database. BLAST searches were also done on putative ORFs (discussed below).

C. ORFs

The entire sequence and the "bridging sequence" from 7kb-3kb were analyzed for open reading frames coding for at least 75 amino acid residues using the Map program. Since there is evidence in cyanobacteria that GTG is used as an initiation codon (Lomax et al.,1987; Conley et al.,1988; Golden et al.,1988; Reddy et al.,1988; Yasui et al.,1988; Meng et al.,1989; van der Plas et al.,1992; Tominaga et al.,1993a,1995), the translation table used by Map was modified to accept both (A/G)TG as start codons. These potential ORFs were screened and putative ORFs were identified using additional criteria. For a particular strand, if an ORF was completely contained within (Kaneko et al.,1995) or overlapped another by more than 10% (see Golden et al.,1988; van der Plas et al.,1992), the longer was chosen unless the shorter was preceded by potential translation and/or transcription initiation signals (see below) or showed significant similarity to sequence(s) in the database.

There is evidence that the nonsense triplet, TGA, is a "leaky" termination signal. Of the three nonsense codons, TGA exhibits the weakest pairing with the 3' end of 16S rRNA (Tinoco et al.,1971; Uhlenbeck et al.,1971; discussed in Shine et al.,1974). Read-through in suppressor$^-$ strains has been demonstrated both *in vivo* (Sambrook et al.,1967; Roth,1970; Ferretti,1971; Moore et al.,1971; Weiner et al.,1971) and *in vitro* (Model et al.,1969; Hirsch et al.,1971). Although several aminoacyl-tRNA species have been reported to bind (U/T)GA (Söll et al.,1965; Caskey et al.,1968) the evidence for t-RNA^{Trp} (Hirsch et al.,1971) appears to be the strongest.

The possibility of read-through at the opal termination signal was investigated by re-screening the pMa025 sequence for ORFs by adding a second modification to the translation table used by Map. The nonsense triplet TGA was programed to be read as a sense codon specifying the amino acid tryptophan. Potential ORFs were screened as described above and putative ORFs were used to search the amino acid database using BLAST.

The molecular weights (MW) of the hypothetical proteins encoded by the putative ORFs was determined using the program PeptideSort.

D. Transcription & Translation Initiation Signals

Most cyanobacterial promoters (>70%) have -10 elements that conform in both position and sequence to the consensus hexamer (TANNNT, N=(A/T/G/C)) of *E.coli* σ^{70} promoters (see Curtis, S.E.,1994). The -35 consensus element (TTGNNN) is less conserved, occurring in less than 50% of cyanobacterial promoters (see Curtis, 1994).

The program FindPatterns was used to identify motifs in pMa025 conforming to either the -10 *E.coli* consensus element alone or in conjunction with the -35 consensus element. Specifically, the pMa025 sequence was searched with two patterns: TANNNT and TTGNNN (N) {16,24}TANNNT. The spacer between the -35 and -10 elements was allowed to vary from 16 to 24 nucleotides; no mismatches were allowed. Once a -10 element was identified, the downstream region was examined for a possible transcription initiation site- i.e, a purine 7$\pm$ 1 bases downstream of the consensus hexamer.

Regions upstream of the translation start codon were examined manually for potential ribosome binding sites (rbs) using either the canonical *E.coli* Shine-Dalgarno

(1974) sequence (A/G)GG(A/G)GGT or the sequence (T (A/G)(A/G)GG(A/G)GGT) complementary to the 3' end of 16S rRNA in *Synechococcus* sp. PCC6301 (Tomioka et al.,1983). G-U/T pairing was permitted; a minimum match of 4 bases was required; the spacer between the initiation codon and the rbs was allowed to vary from 2 to 16 nucleotides.

E. Terminator Sequences, Direct & Inverted Repeats

The pMa025 sequence was searched for prokaryotic rho-independent RNA polymerase termination sequences using the program Terminator (Brendel,1984).

The putative intergenic regions of the pMa025 sequence were screened for direct and inverted repeats using the programs Repeat and StemLoop, respectively. The parameters used to find direct repeats were : minimum repeat window=8; minimum stringency=8; range 50 and 100. The program thus reported direct repeats of 8 or more bases, that were identical, and that occurred within either 50 or 100 bases of each other.

Strictly speaking, the program StemLoop is used to find inverted repeats, i.e., regions of secondary structure in RNA. However, it was used , with suitable modification of the program default parameters, to indicate regions of likely secondary structure in noncoding regions of the plasmid sequence. The program was directed to find inverted repeats meeting the following minimum criteria: minimum stem length=15; minimum bonds/stem=30; minimum loop size=6; maximum loop size=30. Thus, inverted repeats of 15 or more bases separated by 6 to 30 residues were reported. Although the program does not calculate ΔG values *per se*, a rough estimate of the stability of the secondary structure can be obtained from the minimum bonds/stem value.

StemLoop scores G-U/T, A-T and G-C as 1,2 and 3 bonds, respectively. By setting the minimum bonds/stem value at 30, it was assumed that imperfect inverted repeats with the potential to form weak-moderate secondary structures in the DNA would be found.

F. Plasmid Replication

The sequences of both *Microcystis* plasmids have been reported to contain a putative nicking cassette, CTTGATA (Tominaga et al.,1993a,1995), indicative of a rolling circle mechanism of replication. FindPatterns was used to search the pMa025 sequence was for this motif.

A preliminary study was done in an attempt to localize the origin of replication (*ori*) in pMa025. Restriction digested pMa025 was probed, under low-stringency, with a ^{35}S-labeled Hind III fragment (≈1.6 kb) from the shuttle vector pUC104 (gift of Dr. W. Borrias). This fragment contained the putative *ori* from *Synechococcus* PCC7942 plasmid pUH24 (van der Plas et.al.,1992).

G . Sequence Comparisons & Shuffle Test

Regions of pMa025 exhibiting significant similarity to sequences in the database at either the amino acid (BLAST, $P<10^{-3}$) or the nucleotide (FastA, >50% identity over >1kb) level were aligned. Pairwise comparisons between pMa025 and database sequences were done using the alignment programs Bestfit (Smith et al.,1981) or Gap (Needleman et al.,1970). PileUp (Feng et al.,1987) was used to create multiple sequence alignment(s). The choice of program was dictated by the strength and/or coextensiveness of the alignment (seen in the FastA or BLAST results). Bestfit is a local alignment tool

and was used to find the best segment of similarity between two sequences. Gap uses a global alignment algorithm and was used to determine the overall similarity between two sequences. PileUp is an extension of Gap that was used for multiple sequence alignments.

The significance of pairwise comparisons can be evaluated using a statistical method, the shuffle test (Doolittle,1986). This was implemented by using the command line parameter -ran=100 in both Gap and Bestfit alignment programs. The test consists of seven steps: 1) the best alignment and quality score from a pairwise comparison is generated; 2) the second sequence is "shuffled" such that the number and composition of the residues remains the same, but the order is randomized; 3) the best alignment and quality score between the first sequence and the "shuffled" sequence is determined; 4) the "shuffling" and comparison steps are repeated 100 times; 5) the mean quality score and standard deviation from the "shuffled" comparisons is computed; 6) the z score (z=quality score- mean quality score/standard deviation) is manually calculated; 7) the significance of the original comparison is evaluated. According to Doolittle (1986), for sequences over 100 residues in length, a z score below 3 is probably not significant, a score between 3 and 5 is marginal. Scores between 6 and 10 are probably significant. Scores above 10 are highly significant.

RESULTS

I. Nested Deletions

The electrophoretic analysis of ExoIII/S1 Nuclease treated DNA from the thirty-one deletion time points (t=0-30) is shown in Figure 3. Since these results confirmed the creation of a series of nested deletions, the DNA was ligated and screened.

A series of 23 nested deletions, encompassing the entire 8 kb insert, is shown for the XhoI linearized plasmids (Figure 4). Roughly, this corresponded to an average difference of 350 bp between successive deletions. However, since ExoIII activity declines over time, this prediction (based on electrophoretic mobilities) probably under and over estimated the extent of the deletions at the beginning and end of the series, respectively. Indeed, it required 5, 3, 4, 4, 5 and 4 additional primers to close the gap in sequence between the undeleted plasmid and deletion 1, between deletions 1 and 2, 4 and 5, 6 and 7, 8 and 9, and 15 and 16, respectively. In contrast, only 1, 1, 2 and 1 additional primers were required to span the gaps between deletions 12 and 13, 17 and 18, 19 and 20, and 20 and 21, respectively. Deletions 21 through 23 did not require additional primers: the fragment sequences overlapped considerably. While gap distance between successive deletions certainly contributed to the need for additional primers, the presence of template secondary structure was another factor. The fifty bp regions surrounding the junctions between deletions 6-7 and 8-9 have G+C contents of 68.6% and 66.7%, respectively.

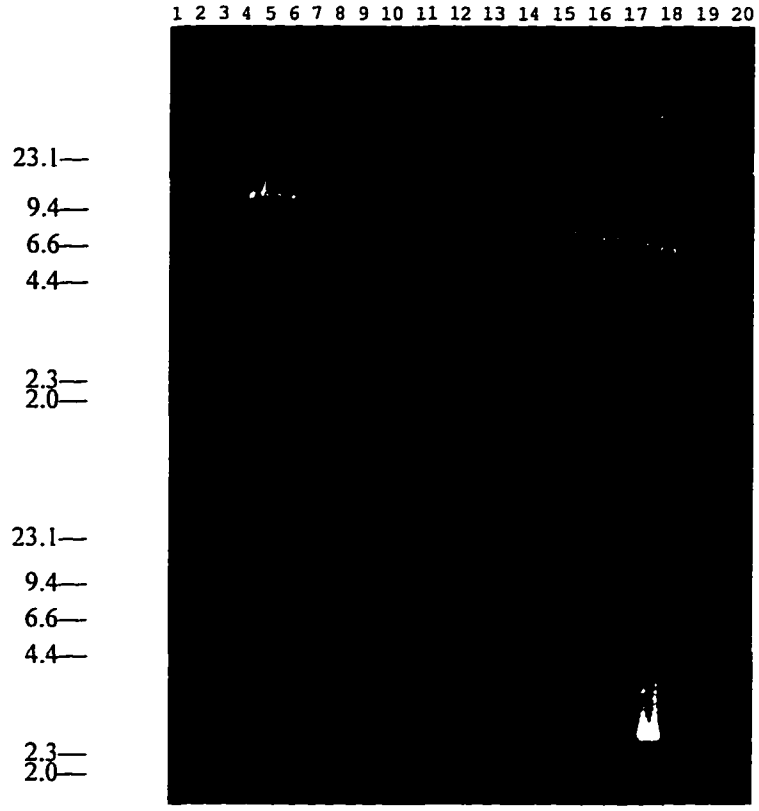


Figure 3. Agarose gel electrophoresis of ExoIII/S1 nuclease digested pMaLI. 0.8% agarose was used. The gel was run at 25V, 6.0 hr. The DNA was stained with ethidium bromide and the fluorescence photographed on a 312 nmU.V. transilluminator.

Top: Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lane 2, 1 kb DNA ladder; Lane 3, linear pMaLI; Lanes 4-20, deletions 0-16.

Bottom: Lane 1, Hind III digest of lambda DNA; Lane 2, 1 kb DNA ladder: Lanes 3-16, deletions 17-30; Lane 17, linear pBluescript II; Lanes 18-20, empty.

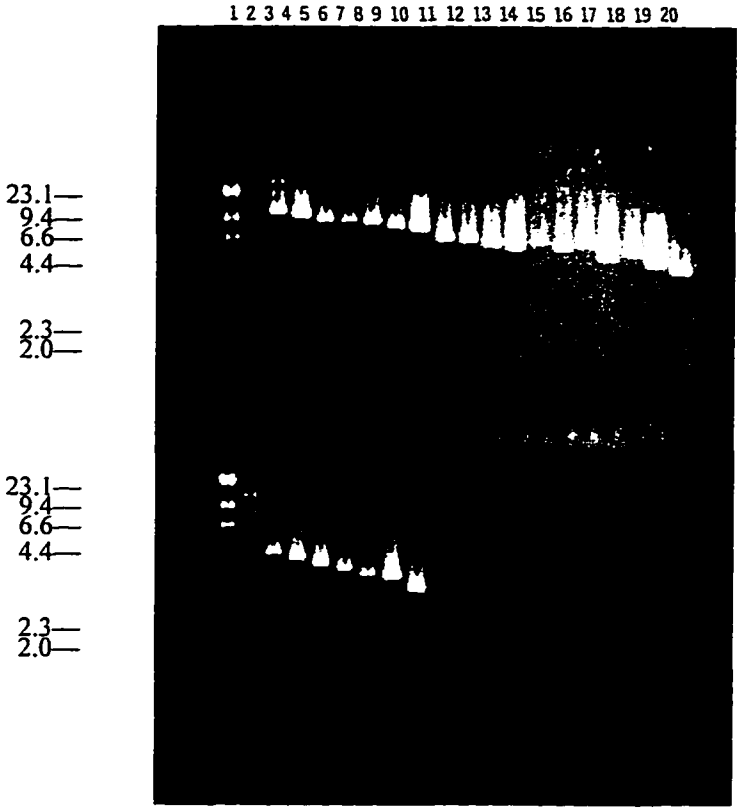


Figure 4. Agarose gel electrophoresis of the nested deletions of pMaLI. 0.8% agarose was used. The gel was run at 25V, 6.0 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.

Top: Lane 1, Hind III digest of lambda DNA; Lane 2, 1 kb DNA ladder. fragment sizes (kb) are indicated at the left;
Lane 3, Xho I digest of pMaLI; Lanes 4-20, Xho I digest of deletions 1-17.

Bottom: Lane 1, Hind III digest of lambda DNA; Lane 2, 1 kb DNA ladder: Lanes 3-8, Xho I digest of deletions 18-23; Lane 9, Xho I digest of pBluescript II; Lanes 10-20, empty.

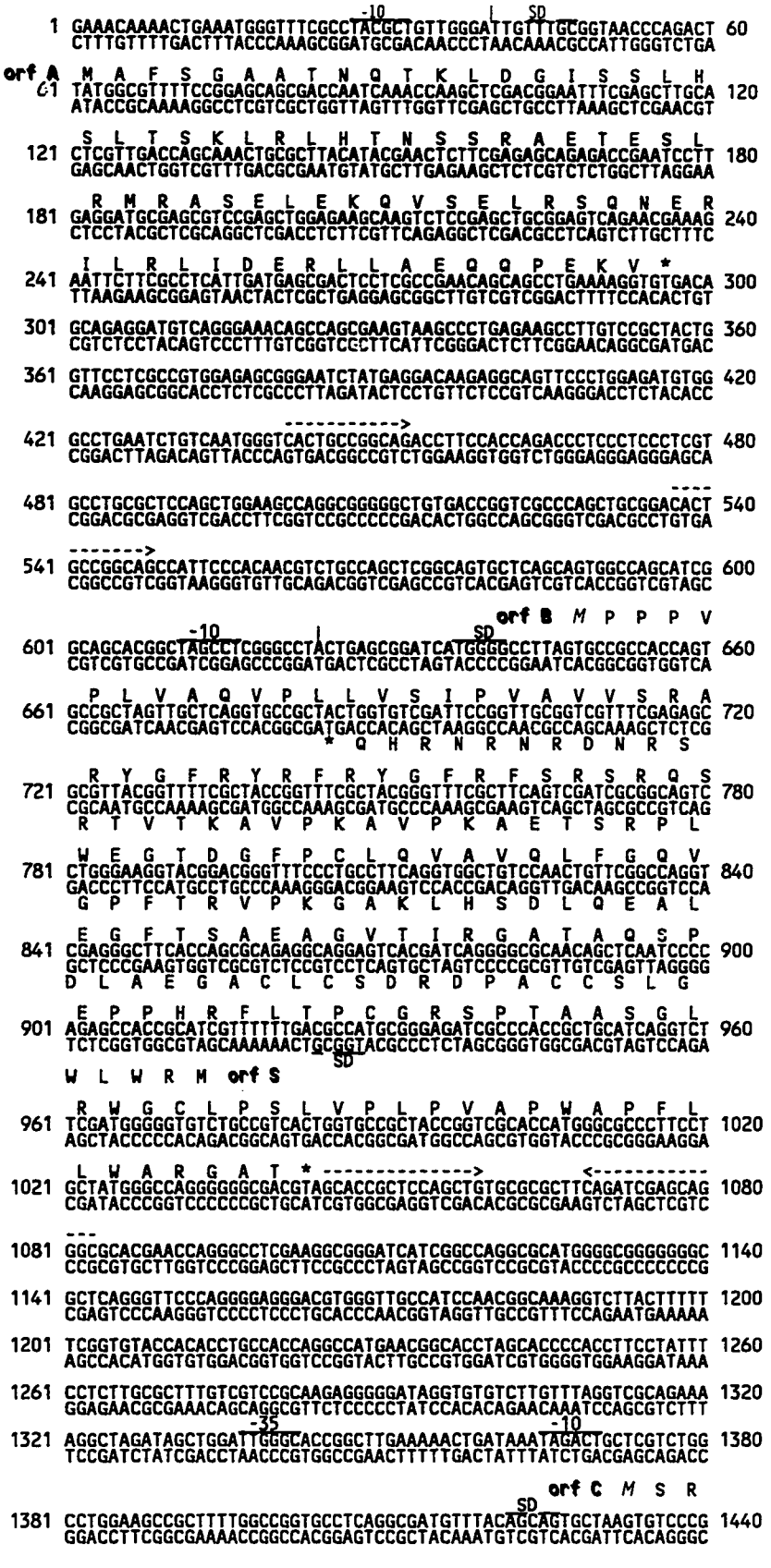
II. pMa025

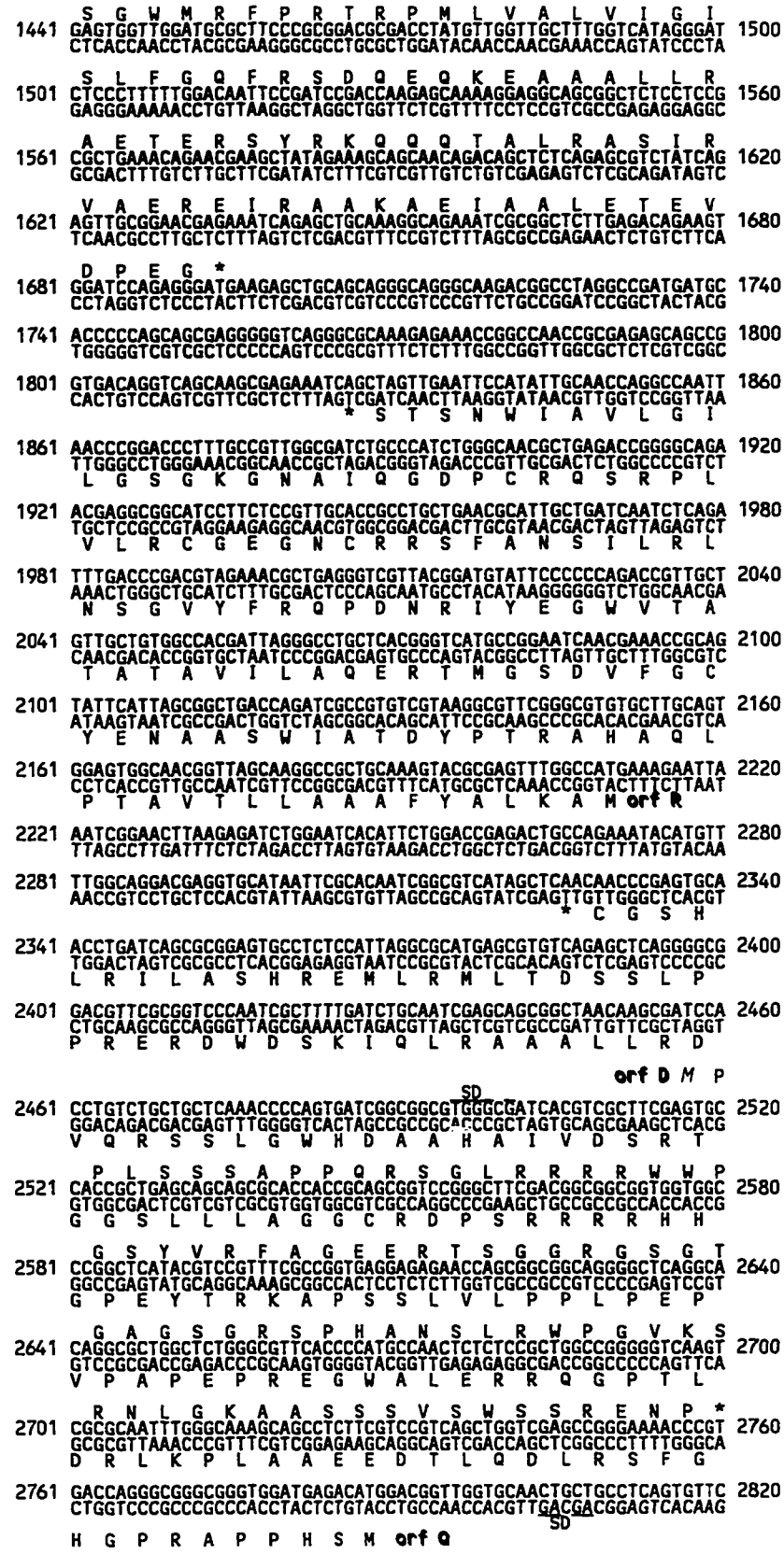
A. Sequence, Composition & Restriction Pattern

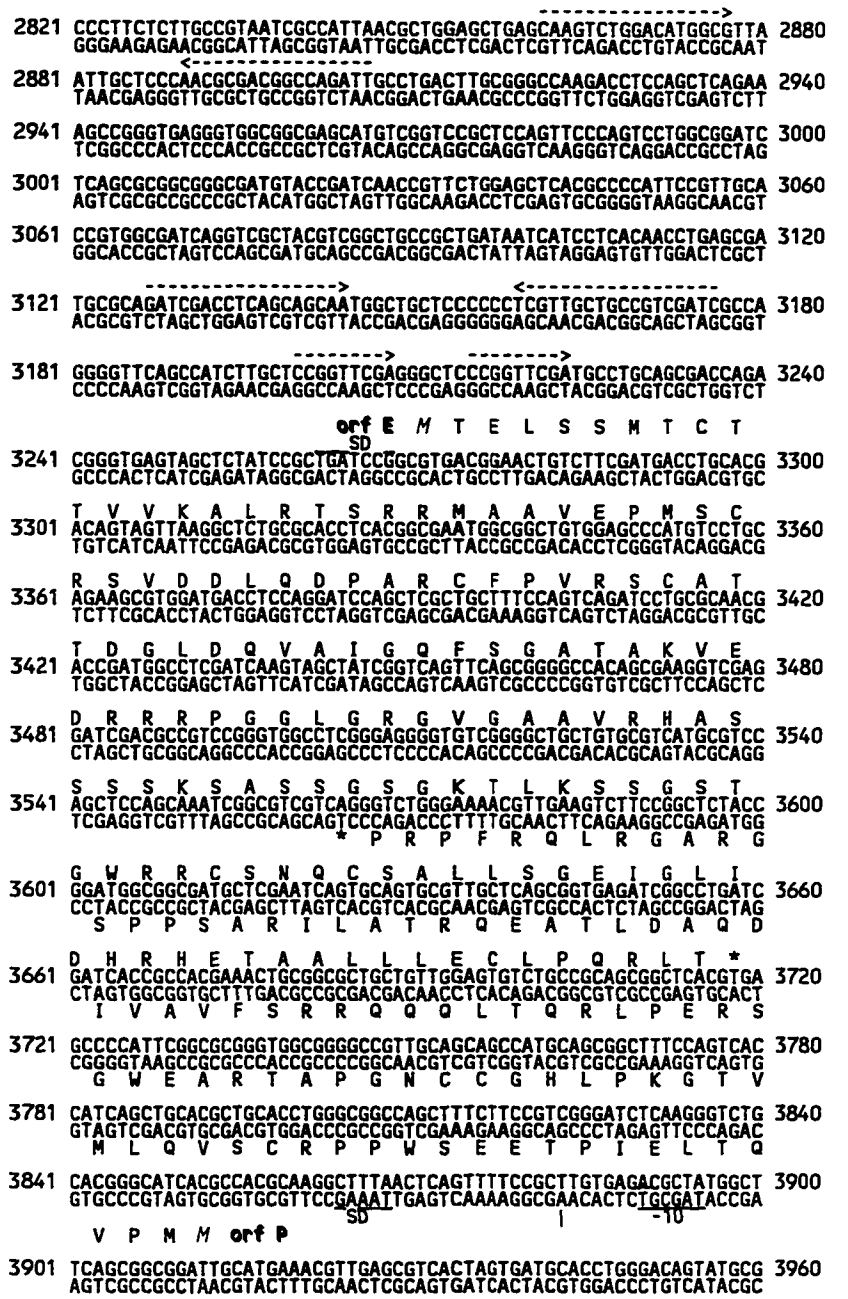
The complete nucleotide sequence of pMa025 is shown in Figure 5. Both plus and minus strands are shown. These correspond to sequences derived from the first (nested deletions) and second (confirmation) sequencing series, respectively. The plasmid is 8018 bp in length and has a 62.3% G+C content. As a result of the sequencing strategy (see below) the sequence is presumed to be without ambiguities. These sequence data have been submitted to the GenBank database under the accession number U94409.

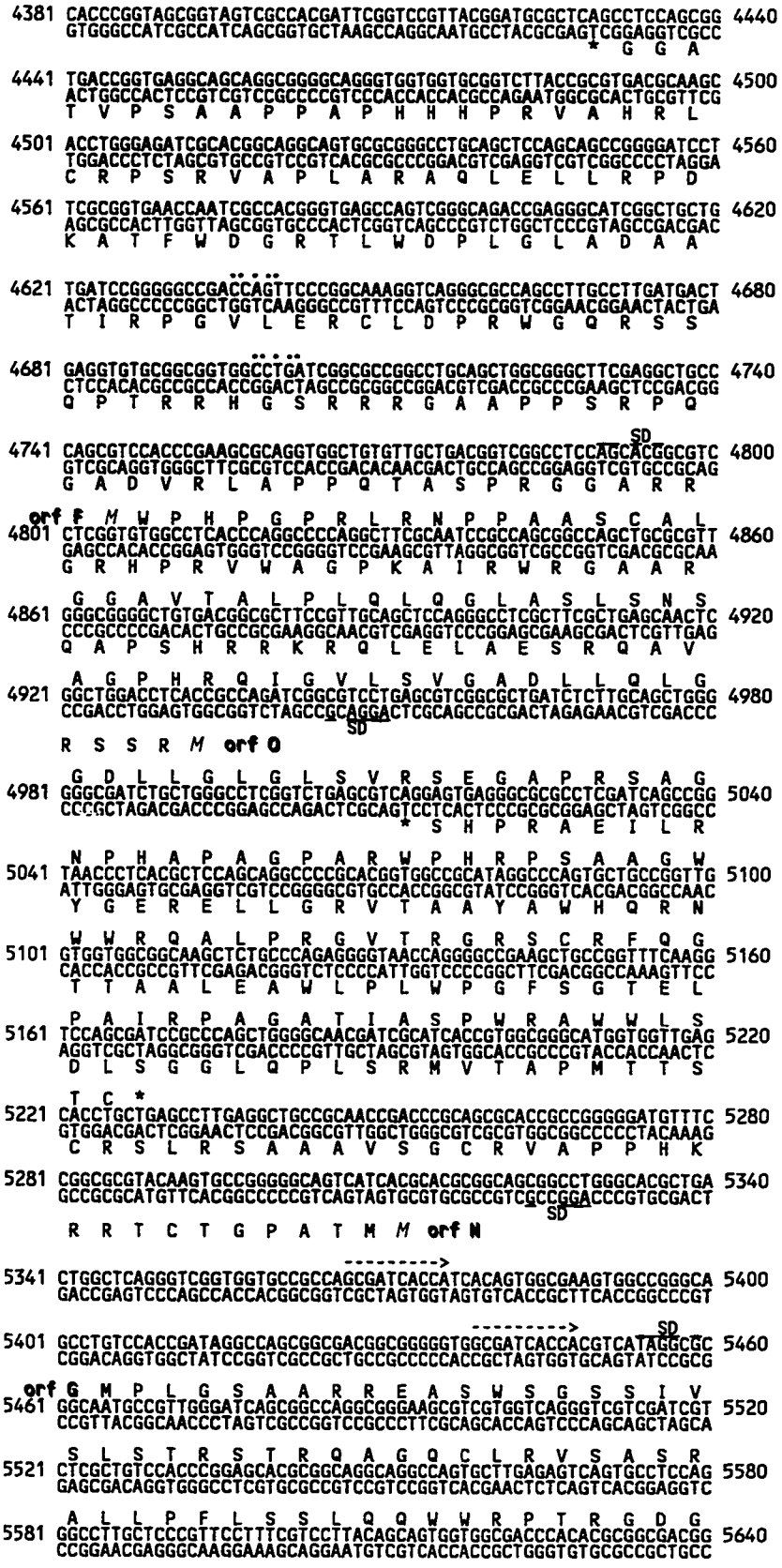
The sequence for each strand was generated from overlapping fragments, each of which was sequenced at least twice. Sequences artifacts (bands across four lanes) resulting from secondary structure were completely resolved by using either glycerol stabilized Sequenase or a TdT chase. In general, template secondary structure presented more of a problem during confirmation. The undeleted recombinant plasmid, pMaLI, was used as the template for sequencing the second strand. Thus, more opportunities for the formation of secondary structures existed in this larger template than in the smaller deletion constructs.

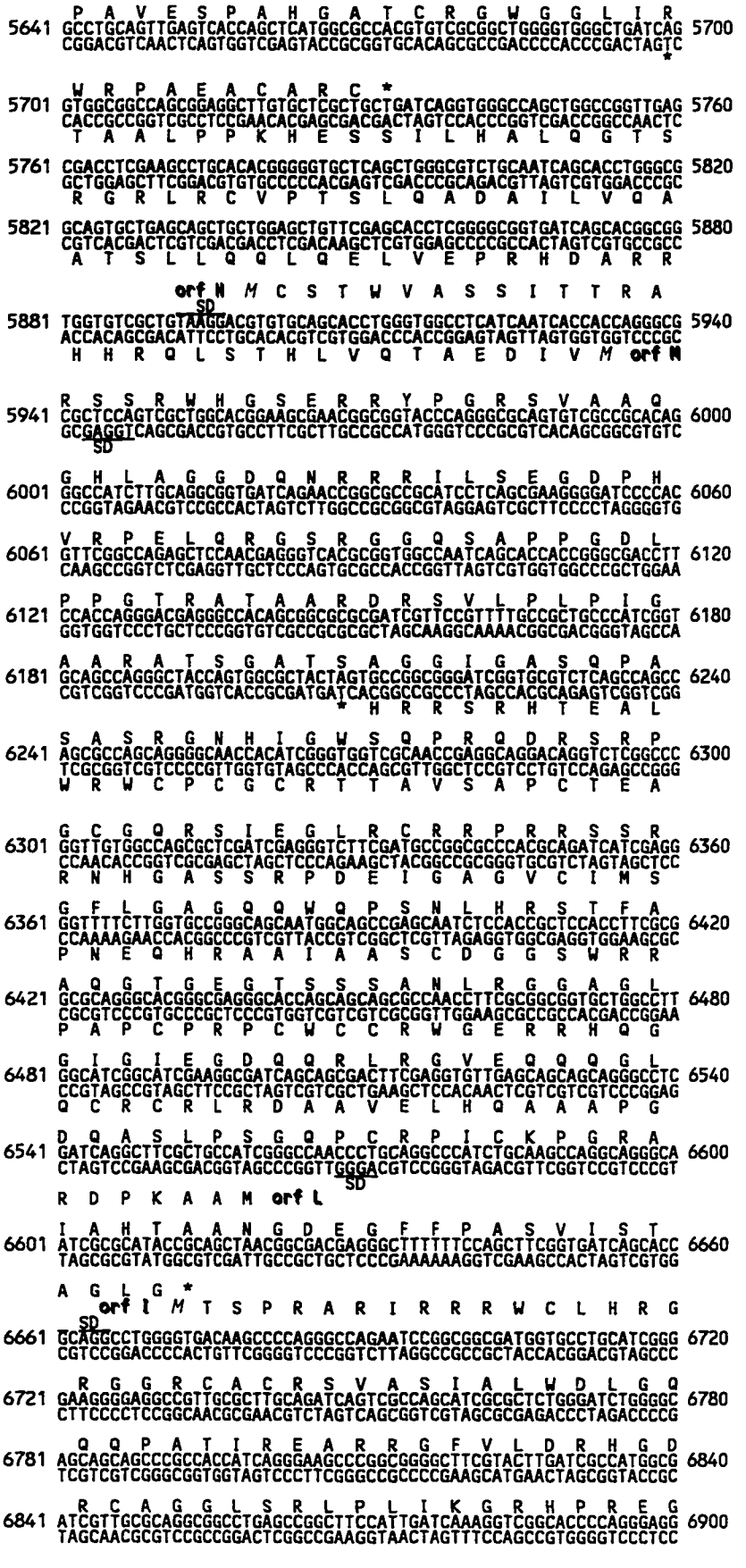
A circular restriction map, based on the sequence data, is presented in Figure 6. Only sites for some of the more common restriction enzymes are shown. This confirms and extends the early map data (Raps et al.,1991). There are 11 potential sites for Ava II; 8 for Pst I; 4 each for Ava I and BamH I; 3 each for Apo I, Hinc II and Sac I;











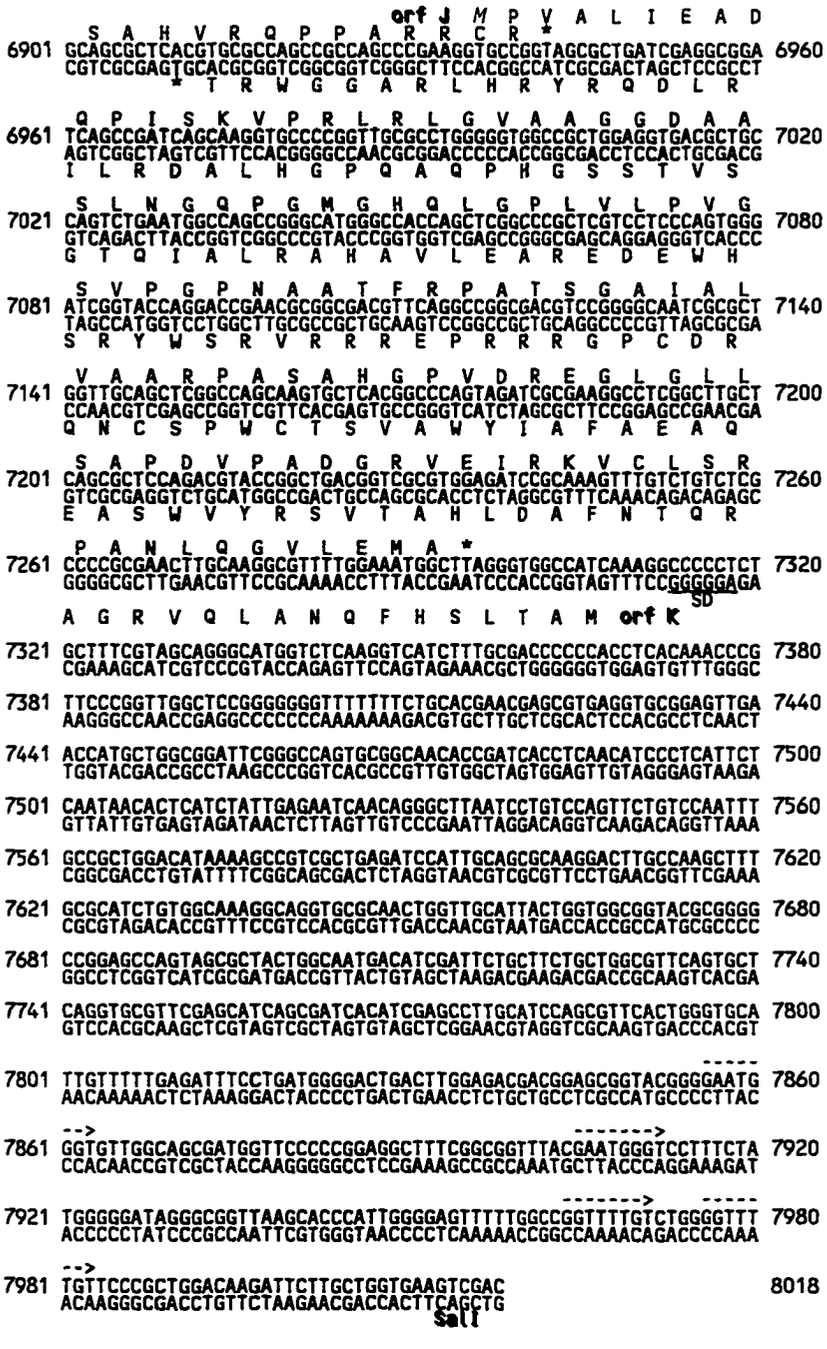


Figure 5. Complete nucletotide sequence of pMa025. The deduced amino acid sequences encoded by the ORFs are shown in standard single letter code above (translation left to right)and below (translation right to left) the sequence. Methionine residues encoded by GTG are indicated by italics. Asterisks denote stop codons. Shine-Dalgarno (SD), -10 & -35 promoter sequences are indicated by lines above the sequence. Possible transcription start sites are denoted by vertical lines (|) above the base. Direct (--->) and inverted (---> <---) repeats are marked by arrows above the sequence. The putative RepA binding sites (4636-4698)in the hypothetical *ori* are shown by dots(...) above the sequence.

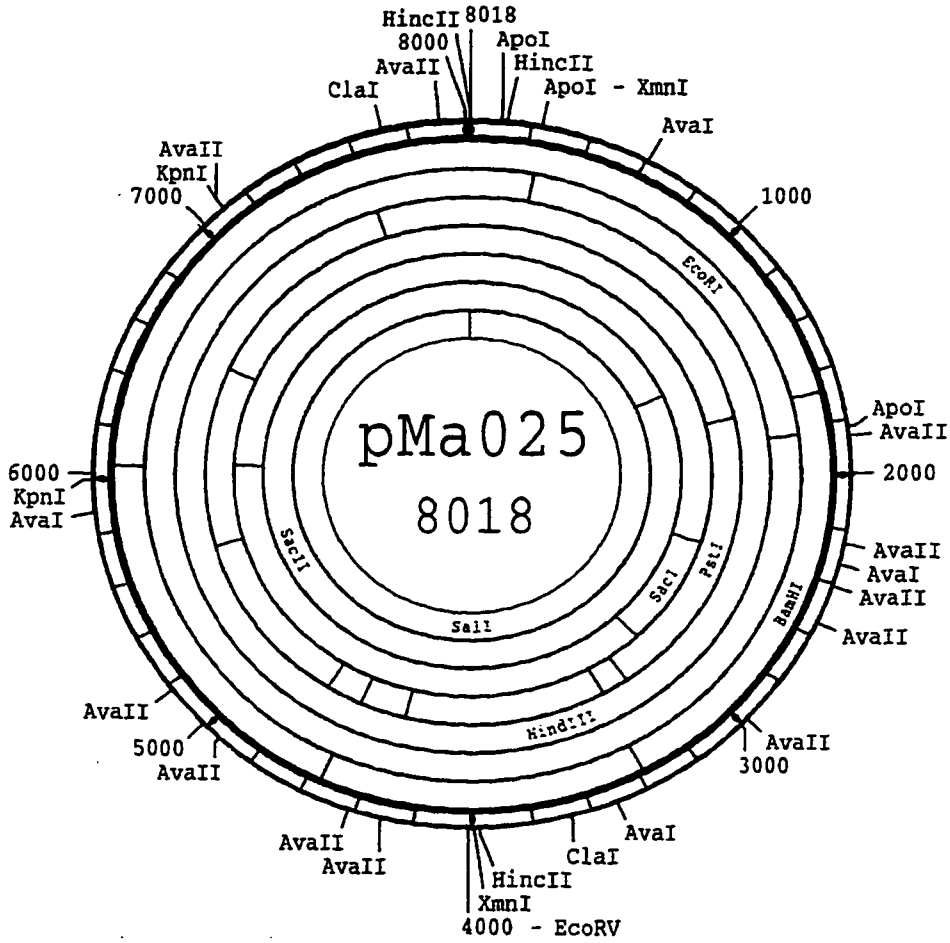


Figure 6. Restriction map of pMa025.

and 2 each for Cla I, EcoR I, Kpn I and Xmn I. Unique restriction sites exist for EcoR V. Hind III, Sac II and Sal I. There are no restriction sites for Apa I, Hpa I, Not I, Sma I, Xba I and Xho I. Studies of restriction patterns in pMa025 isolated from *Microcystis* are needed to confirm if any of the additional restriction sites (Ava II, Apo I, Hinc II) are accessible in the cyanobacterium.

B. ORFs

Thirty-two potential ORFs initiating with either (A/G)TG and coding for at least 75 amino acid residues were found. Using the previously outlined criteria (see Methods. section V., subsection C.), 19 putative ORFs (Figure 5, ORFs A-S) were identified. The nucleotide range (nt), presumptive initiation codon (underlined), number of codons and MWs of the deduced polypeptides are shown in Table 1a.

All of the putative ORFs are short; eight contain less than 100 codons. Fourteen ORFs use TGA as the termination codon: five use TAG. This reflects the compositional bias of the sequence. The triplet TGA occurs 111 and 148 times on the plus and minus strands, respectively. The trinucleotide TAG has respective frequencies of 42 and 40 on the plus and minus strands. TAA occurs infrequently: 28 and 32 times on the plus and minus strands.

When the pMa025 sequence was rescreened allowing TGA to code for the amino acid tryptophan (see Methods section C), fifteen putative ORFs (Table 1b, ORFs a-o) were found. The nucleotide range, presumptive initiation codon (underlined), number of codons and MWs of the deduced polypeptides are shown in Table 1b. In this instance. only three ORFs contained less than 100 codons. It must be noted, however, that in these

ORF	Range (nt)	Translation initiation signals	Codons[a]	MW[b]
A	62-298	TTGTTTGCGGTAACCCAGACTTATG	79	8836
B	647-1045	CTGAGCGGATCATGGGGCCTTAGTG	133	14368
C	1433-1696	GCGATGTTTACAGCAGTGCTAAGTG	88	9876
D	2517-2762	GTGGGCGATCACGTCGCTTCGAGTG	82	8650
E	3271-3720	TAGCTCTATCCGCTGATCCGGCGTG	150	15684
F	4805-5230	GGCCTCCAGCACGGCGTCCTCGGTG	142	14737
G	5465-5734	TCACCACGTCATAGGCGCGGCAATG	90	9581
H	5899-6675	GCGGTGGTGTCGCTGTAAGGACGTG	259	26833
I	6672-6944	TGATCAGCACCGCAGGCCTGGGGTG	91	10230
J	6935-7297	TGCGCCAGCCGCCAGCCCGAAGGTG	121	11993
K	7306-6908	ACGAAAGCAGAGGGGGCCTTTGATG	133	15323
L	6559-6206	ATGGGCCTGCAGGGTTGGCCCGATG	118	12942
M	5933-5697	CCAGCGACTGGAGCGCGCCCTGGTG	79	8648
N	5312-5010	TGCCCAGGCCGCTGCCGCGTGCGTG	101	10892
O	4933-4426	CGCTCAGGACGCCGATCTGGCGGTG	169	18579
P	3852-3562	CTGAGTTAAAGCCTTGCGTGGCGTG	97	10749
Q	2788-2324	GCAGCAGTTGCACCAACCGTCCATG	155	17184
R	2210-1824	TTAGTTCCGATTTAATTCTTTCATG	129	13845
S	913-683	TCCCGCATGGCGTCAAAAAACGATG	77	8485

Table 1a. Translation initiation signals for the putative coding regions of pMa025. Ribosome binding sites and start codons are indicated by lines above and below the sequences, respectively.

[a] The stop codon is included.

[b] Protein molecular weights (MW) were calculated from the deduced amino acid sequence and include the N-terminal methionine which be absent in the mature protein.

ORF	Range (nt)	Translation initiation signals	Codons[a]	MW[b]
a	1693-1995	AGACAGAAGTGGATCCAGAGGGATG	101	11272
b	2872-3099	CTGGAGCTGAGCAAGTCTGGACATG	76	8379
c	3223-4209	CGGTTCGAGGGCTCCCGGTTCGATG	329	35661
d	5101-7506	ATAGGCCCAGTGCTGCCGGTTGGTG	802	84504
e	7629-1242	ACTTGCCAAGCTTTGCGCATCTGTG	544	57747
f	1433-1729	GCGATGTTTACAGCAGTGCTAAGTG	99	10942
g	2376-2837	GGAGTGCCTCTCCATTAGGCGCATG	154	16519
h	4305-5132	AGGGGTATTTTGGTCCATTCCCGTG	276	28920
i	7801-7514	CCATCAGGAAATCTCAAAAACAATG	96	10007
j	7490-4377	TGAGTGTTATTGAGAATGAGGGATG	1038	112713
k	2669-1614	CCGGCCAGCGGAGAGAGTTGGCATG	352	383778
l	1238-60	AAATAGGAAGGTGGGGTGCTAGGTG	393	43267
m	255-7659	TTCGGCGAGGAGTCGCTCATCAATG	205	23345
n	3849-2878	AGTTAAAGCCTTGCGTGGCGTGATG	324	35710
o	2844-2446	GACTTGCTCAGCTCCAGCGTTAATG	133	38378

Table 1b. Translation initiation signals for the putative coding regions of pMa025. Ribosome binding sites and start codons are indicated by lines above and below the sequences, respectively.
[a] TGA was read as sense. The stop codon is included.
[b] Protein molecular weights (MW) were calculated from the deduced amino acid sequence and include the N-terminal methionine which be absent in the mature protein.

cases all TGA terminators were interpreted as sense, i.e. tryptophan. Although the intriguing possibility of read-through at some TGA codons exists, it is extremely unlikely that all opal termination codons would be ignored in *Microcystis*. These results will, therefore, not be discussed further.

C. Transcription & Translation Initiation Signals

Sequences matching *E. coli* -10 promoter elements (see Figure 5) were found upstream of ORFs A, B and P at positions 27, 611 and 3895, respectively. In all cases, putative transcription start sites are located 7 bases downstream of the consensus hexamer (Figure 5).

Both -10 and -35 *E.coli* consensus motifs were found upstream of ORF C at positions 1337 and 1365. The spacer is 22 bases. No transcription initiation site was located 7±1 bp downstream of the -10 element.

Twelve ORFs (63.2%) use GTG as the translational start codon (Figure 5 & Table 1). Methionine residues encoded by GTG are indicated by italics in Figure 5.

Seventeen ORFs have sequences upstream of the translation initiation codon that resemble ribosome binding sites in either *E.coli* (Shine & Dalgarno,1974) or *Synechococcus* sp. PCC6301 (Tomioka et al.,1983). These Shine-Dalgarno (SD) sequences are shown in Figure 5 and, for greater clarity, in Table 1. The spacer region between the SD sequence and the initiation codon varied from 2 to 15 bases.

D. Terminator Sequences, Direct & Inverted Repeats

No rho-independent transcription terminator sequences were found.

Five sets of direct repeats (DR) of at least 8 residues, occurring within a range of 100 bases, were found. These occurred at nucleotide positions: 1) 441-452 and 537-548; 2) 3200-3208 and 3215-3223; 3) 5367-5376 and 5439-5448; 4) 7856-7863 and 7905-7912; 5) 7964-7971 and 7976-7983. The direct repeats are marked by arrows above the sequence in Figure 5.

Imperfect inverted repeats (IR) were found at four locations: 1) 1045-1083; 2) 2863-2906; 3) 3127-3176; 4) 4333-4374. These ranges include the spacers and are indicated by arrows above the sequence in Figure 5.

E. Plasmid Replication

No sequence corresponding to the nicking cassette, CTTGATA, used by plasmids that replicate via a rolling circle mechanism was found.

Southern analysis failed to detect any homology (Wallace et al.,1995) between pMa025 and the putative *ori* (van der Plas et al., 1992) from *Synechococcus* PCC 7942 plasmid pUH24.

F. Database Searches & Sequence Comparisons

The deduced amino acid sequence from a 1700 base region (5000-6700), on the minus strand, of pMa025 showed significant similarity to three cyanobacterial proteins reputed to be involved in plasmid replication. Specifically, this region of pMa025 exhibited significant similarity, at the amino acid level, to ORF 93 from *Synechocystis* sp. strain PCC6803 genomic DNA (Kaneko et al.,1995), ORF F from *Synechococcus* sp. strain PCC7942 plasmid pUH24 (van der Plas. et al.,1992) and ORF 1 from *Synechococcus* sp. strain PCC7002 plasmid pAQ1 (Akiyama et al., unpublished). The Poisson probabilities associated with these matches are $1.8x10^{-26}$, $2.8x10^{-20}$ and $1.5x10^{-14}$, respectively. These values indicate the probability of observing a score of the same magnitude by chance with a search of the same size. ORF 93 is significantly similar to ORF 1 (Kaneko et al.,1995). The rank order of z scores (11.1, 7.4 & 5.6), determined from a shuffle of Bestfit, between this region of pMa025 and each of the three putative plasmid replication proteins paralleled the BLAST probability scores.

This 1700 base segment of pMa025 was translated in three frames (-1,-2 &-3). Each translation product was compared to each of the three presumptive plasmid replication proteins. Figure 7 shows the multiple alignment between the translation product of frame -3 and the three hypothetical replication proteins. The identity, similarity and Z scores (shuffle of Gap) are shown in Table 2. Thus, the translation product of frame -3 is significantly similar to each of three cyanobacterial hypothetical plasmid replication proteins.

```
          1                                                   50
ORF 93    MFCPYNETPR TLDMESCSGK TDGKMNHLQE WRQSCVDDQL TRLNVTPLLG
ORF 1     .......... .......... .......... .......... ..........
ORF F     ........MA FIANHLQRSQ IPTRCELIDQ LDRLPPNWRF CVVNSRKAPF
pMa025    .......... .......... .......... .......... ..........

          51                                                  100
ORF 93    DAPAQHLLYA .EALQRRNDG RVSDGLMKRY AHVASGGWWC SGIDLLSGEK
ORF 1     MTPYEYLLYS DEPALRRNDG RLRDTWLKRY AFVEHGGWWC SGIDIKTGKD
ORF F     DPGWQNNPLD HAAVAERIRS DRRVTGIGLL TGPASGGLIA VDFDGPTAHE
pMa025    .......... .......... .......... .......... ..........

          101                                                 150
ORF 93    DLWG..CFKP DRPRQDGESS KTIKYEHPPK TPIGIFALRV PLHLWERIAA
ORF 1     SLWG..CFKG DRPRKDREDK KPIKYEHPPR VATEIFTEKV DRGTWRKIAK
ORF F     ALPEGLTLTE LPPTVAYTSG KPGRHQRLYQ VPQERWAAIA TQKLHSPDGD
pMa025    .......... .......... .PDSGPGACH PRPAVLITEA GKKPSSPLAA

          151                                                 200
ORF 93    KAGVVLTEED Q..DPDQPDL GFWQWLMGHP EIPLVITEGA KKAGALLTAG
ORF 1     RHKVELPETD Q......... GFWEWVLAHP ELPIIITEGA KKAGALLTAG
ORF F     LLELRWNKLQ SVIVGQHPET GAYRWVEGCA PWEIEVAEAP PELLDAMERQ
pMa025    VCAIALPGLQ MGLQGWPDGS EA*SRPCCCS T......... ..........

          201                                                 250
ORF 93    YGAIALPGVH NGYRTPRDDQ GQRIGK.CQL IPALGKLAAP GRQILIAFDQ
ORF  1    YCAIGLPGIY NGYRTPKNDH GEPMRQLRHL IPELDLLAKN NRAIAFCFDQ
ORF  F    DRQTAPKSYK PIVSAPASDD EVAIARTMEA HVPASYADDY ESWVAVGMAL
pMa025    .PRSRC*SPS MPMPRPAPPR RLAL...LLV PSPVPCAAKV ERWRLLGCHC

          251                                                 300
ORF 93    DSKPKTIQQV NLAIQRLGYL FS...RQGCE VKVLQWEHHL GKGVDDVIAH
ORF  1    DKKPKTIKAV NGAIQTTGAL LE...KAGAK VSVITW.HQD AKGVDDLIVE
ORF  F    QSVSDALLDD WIAWSAQSSK FDGNRKLERK WASFKGSGIT IATLAKLAKE
pMa025    CPAPRKPLDD LRGRRHRRPS IERWPQPGRD LSCLGCDHPM WLPLLALAG*

          301                                                 350
ORF 93    QGSDYLQQLV GKALPLEIWK AQRLNRLTHS RGMEVEARYL ..PPLTIPAE
ORF  1    HGAKALHNRY KHRKPLAVWE MDNLTDITTQ VDLTVDQRYL DIDPRAIPKD
ORF  F    GGYRPPKLTR QEQREYQREV ANEWQRLTRS PNRLHTGGYI NGDYIPQPEQ
pMa025    DAPIPPALV. .......... .......... .......... ..........

          351                                                 400
ORF 93    EKLIALQSPK GTGKTEFLAR IVRQAQAEHR PVLVIGHRIR LVQELCHRFQ
ORF 1     AQIIFIKSAK GTGKTEWLGK IVKLAQDDCA RVLVLTHRIQ LAKELARRLD
ORF F     HRLVCLNAAM GTGKTEAIAG HLQPLIGTGI PVVLITHRRT LGAALGKRLG
pMa025    APLVALAAPM GSGKTERSRA .......... AVALVPGGRS PGGADWPPRD

          401                                                 450
ORF 93    LPYVGDLSSS PLA....... .RQRGFGLCI DSLHGHSQAQ FDPEQWQDCL
ORF 1     IDHISELDSS PTG....... .GALGMAMCI DSLHPDSQAH FNFMEWHGAH
ORF F     VPWAEECVPG S.DLRM.... ...IGLGGCL DSMHPASKLR FSAHGWQRAI
pMa025    PRWSSGRTWG SPSLRMRRRF *SPPARWPCA ATLRPGYRRS LPCQRLERAL

          451                                                 500
ORF 93    IIIDEVEQVL WHGLNSDT.C RRQRVAILRN LKQLLQHSLG GEGQIYVADA
ORF  1    IVLDEIEQVL GHALGSST.C TQDRAKILET FYNLILYALR GGKLYCSDAT
ORF  F    VVIDEVELWA EHLLSGKTEI AKHRPEVMAE IAGL....LA GGRQVIAADA
pMa025    MVIDEATQVL NTSLQRHNRR ADHRPEVLEQ LQQL....LS TAAQVLIADA

          501                                                 550
ORF 93    DLTDVSLDYL .ISLSGIPLQ PYVIRNHWQP GLEEAWPIYH FGENDPKQLV
ORF  1    DLSPISYELI KYILDGCEFK PFTILNTYKP CLEQQRDLFF YEGNDPRDLL
ORF  F    MLSDVSIALL EALTGST... AYLVSSTVQP FSGAP....V YQVGSDRELM
pMa025    QLSTPVCRLR GRSTGQL... ANLISSEWKP PLAAT....* SAHPSRDTWR

          551                                                 600
ORF 93    KQLLHHIREG GKPFVCLSA. .QKLTSAWGT RNLEAYLKKQ FPDRRILRID
ORF  1    TNLRQAIENG EKTLVFTAA. .QKTASTYST QNLESLFREK YPDKRILRID
ORF  F    AMAVEMVQAG KR..LAVFCD RQKLSQSRPS LYSAATIAQQ LSEACGQPVP
pMa025    HELVTQLQAR RRVWVATTAV RTKGTGARPW ........RH *LSSTGLPAA
```

```
          601                                                       650
ORF 93    AESLSDPHHP AHGSLTNLNQ VLADYDIVLS ..........  .SPAVETGVS
ORF  1    AESVAEPGHP AYGCIDSLNA ILPLYDIVLC ..........  .SPAVETGVS
ORF  F    VIDSEANKRG ALETLKLNPQ MFINCCPHEV ..........  FTSVIDAGLS
pMa025    CSGWTARRST TETTTLPAWP LIPTALPRL* RGDRHPRRRR  WPIGGQAARP

          651                                                       700
ORF 93    I....DLKNH FTSVWGIAQG IQTATS.VCQ SLSRIRQNIP  RYLWAASYG.
ORF  1    I....DIKDH FDSVWGMGSG VQTVNG.FCQ GLERLRDNVP  RHVWIPKFSP
ORF  F    ITQGVGDR.P FDAVLIFAGA GHIAPGAIAQ IAGRVRSNVP  RFITCPDTSN
pMa025    LRHCDGDRWR HHRP*ASQRA QAAAACVMTA PGTCTRRKHP  PAVRCGSVAA

          701                                                       750
ORF 93    .FNQVGNGAT AISKFLTAGH RLTEVNIRLL QQSDEDNLDD  LDTGFQAESL
ORF  1    HSNRIANGGY TAKAIARDQH RYAELTHKLI GEHAAEC.SG  LEDSLKP.FL
ORF  F    LM.RAGNGSI DPTEVADGLK RHQQV...IL AQLSQAAIKI  GEPVASQAYL
pMa025    AS.RLSRCS. ..TTMPATVM RSLPQ...LG GSLDLETGSF  G......PWL

          751                                                       800
ORF 93    LCWAKMAVRI NLAMVNYRES ILGLLHQEGH QLLGPPPALP  SLPVPPSDPD
ORF  1    WAYCRYAALA NRGFGSYREA IENKLLSEGY ..........  .VQKDLSEID
ORF  F    TAWAELAARR NRDGLTYAAT VEQLLRLNGF TWQDWDGAID  PFLQEVVDEL
pMa025    PLWAELAATT NRQHWAYAAT VRGLLEREGY ..........  RLIEARPHS*

          801                                                       850
ORF 93    HPLSTVAPQP LQQAIEAVRE QNYLADCQAI ADAKPLKELE  YQQLKKRLVK
ORF  1    PAL....AKD YRDELKAVKD HNYLQERVAI SKVENPDDRQ  YEKLKRQRAK
ORF  F    QRETAEALAD ASAEAILSSP DLTPKEAEAI ASADFANAEQ  LAALARHRCQ
pMa025    RSD..... .......... .......... ..........  ..........

          851                                                       900
ORF 93    NSHERQRLRR YELAQRYGIP VTADLVQKDD QNWYQKLRWH  YFLTVGRPFL
ORF  1    SETERHQERH GKLSRSYGLT VTPELVEKDD DGWYSQLQLE  YYLTVGKAFC
ORF  F    QRWG..... .........D CTRELLAADD DGLYAPLRLL  FWLSLGKDAV
pMa025    .......... .......... .......... ..........  ..........

          901                                                       950
ORF 93    GDRD.AKIAR LLLDQGQGSL FLPDFNGSQL GAIIGTYDII  GIPILLAHPQ
ORF  1    SARDRAKYDQ L...QHEGFV FKPDINRRSL SPKIHLLELL  NIHQFL...K
ORF  F    PLAMPPSSTN GAIARCSDPI LWAAVSRPKC GTCMG.....  ..........
pMa025    .......... .......... .......... ..........  ..........

          951                                                      1000
ORF 93    RELCALDEDL IALGKLALAN RDDIKTVVGI GLAKNASPIT  IVRRFLEKLG
ORF  1    PGVTFTGASL EGFKENCLRY AKPIKWILGR TITDKMSPLE  IAQALLGKLD
ORF  F    .......... .......... .......... ..........  ..........
pMa025    .......... .......... .......... ..........  ..........

          1001                                                     1050
ORF 93    FGIQLTRTKT VEKKRVRIYQ ITCPQDGREK VFQAWLQGDR  Q.....CPGS
ORF  1    RKLEYKGRFG SRDNRQRVYE AIAPNDQREK VFAHWLQRDQ  AKLGAVSNPC
ORF  F    .......... .......... .......... ..........  ..........
pMa025    .......... .......... .......... ..........  ..........

          1051                          1083
ORF 93    SEQWMAEYWK KLQSSPRAEE QHQPFVQLSL ELG (1036)
ORF  1    INRFIQEA.. .......... .......... ... (943)
ORF  F    .......... .......... .......... ... (876)
pMa025    .......... .......... .......... ... (567)
```

Figure 7. Alignment of the deduced amino acid sequences from *Microcystis aeruginosa* UV025 plasmid pMa025 (nt 5000-6700) with three putative cyanobacterial plasmid replication proteins from *Synechocystis* PCC6803 (ORF 93), *Synechococcus* PCC7942 (ORF F) and *Synechococcus* PCC7002 (ORF 1). The region of pMa025 (ORF M) encoding the putative plasmid replication protein, RepA, is indicated in bold. Stop codons in the pMa025 sequence are indicated by asterisk. Identical residues are shaded. Number of amino acid residues is indicated in parenthesis.

	% Similarity	% Identity	Z score
		pMa025	
ORF 93	45.0	25.5	8.6*
ORF F	41.2	24.6	9.2*
ORF 1	42.2	19.6	5.6*

*significant (see Doolittle, 1986)

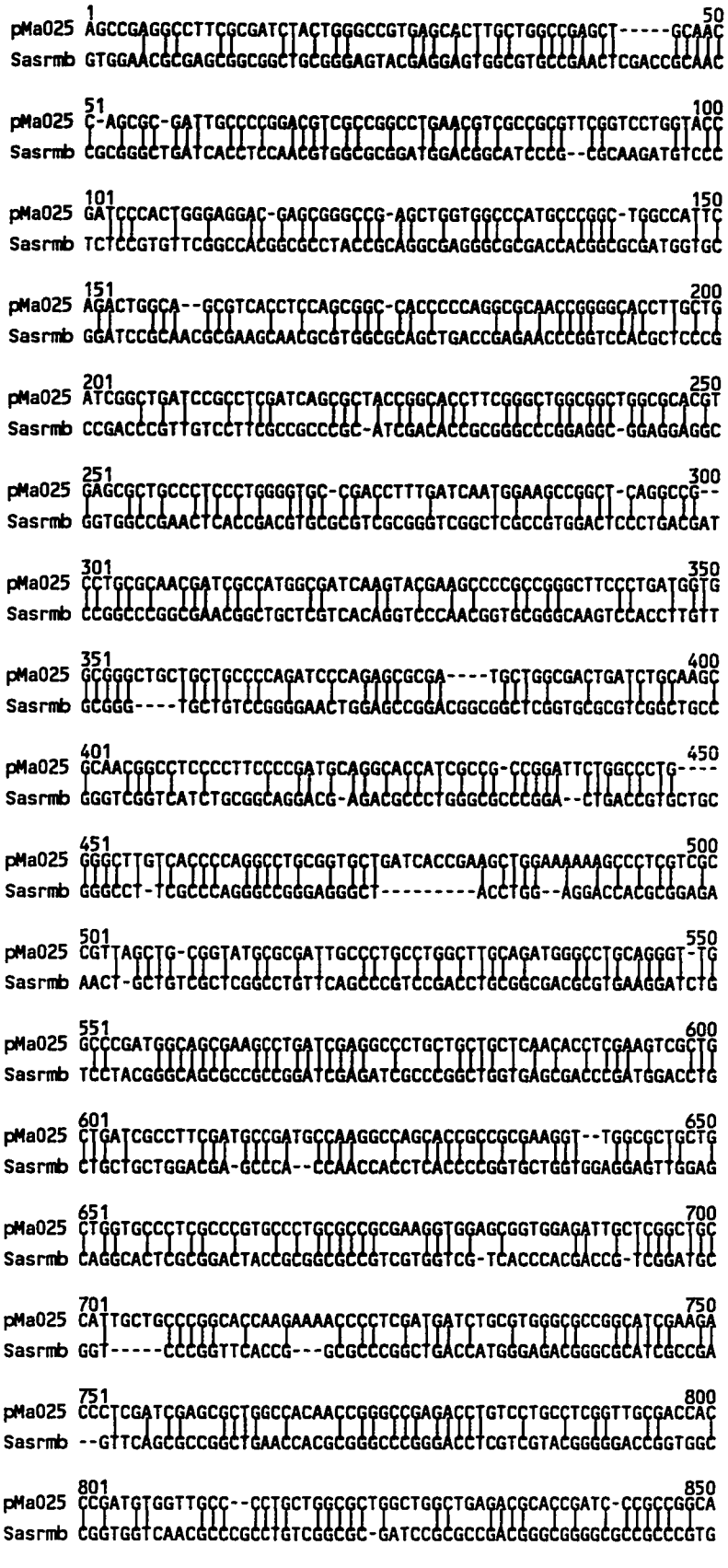
Table 2. Percent identity, similarity and z scores for the pairwise comparisons (Gap) between the deduced amino acid sequence (nt 5000-6700) of pMa025 and those of three putative cyanobacterial plasmid replication proteins.

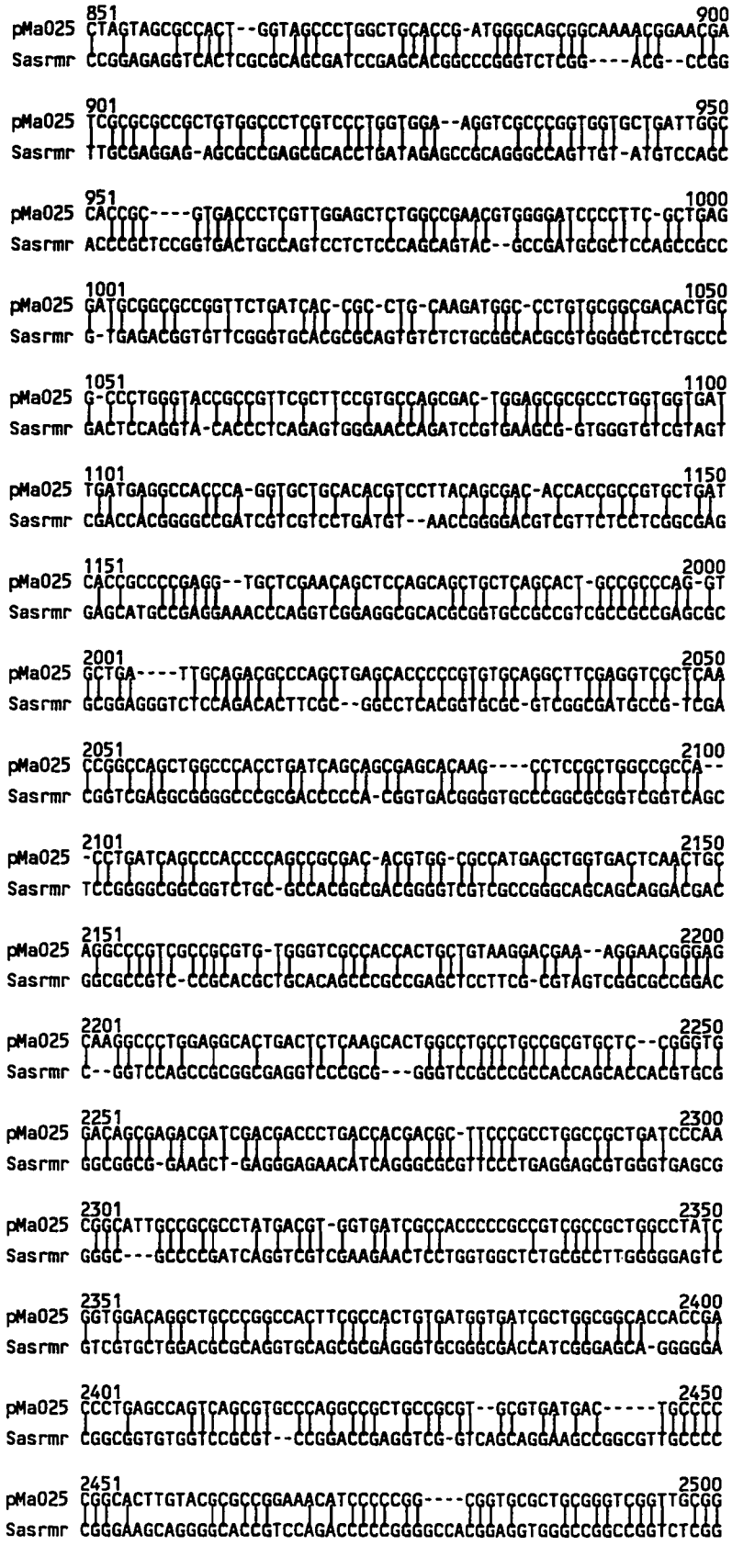
Within this 1700 base region, the deduced amino acid sequence encoded by ORF M was significantly similar ($P=2.9x10^{-7}$) to the hypothetical plasmid replication protein (repA) from *Synechococcus* PCC7942 plasmid pUH24. Thus, in *Microcystis*, ORF M corresponds to a hypothetical gene, *repA* (plasmid *rep*lication), that encodes a polypeptide (RepA) of 78 amino acids. The residues corresponding to the RepA protein are indicated in bold in Figure 7.

When read-through the TGA terminator was permitted, a presumptive ORF encoding a polypeptide of 1037 amino acids was found. This ORF encompassed ORF M and when translated in frame -3 was 98.6% identical to that of the corresponding 4kb fragment initially analyzed by BLAST (see Methods section V, subsection B). It is highly improbable that this segment of DNA constitutes an ORF: 1) read-through of all of the 14 opal termination codons in this region is extremely unlikely; 2) the encoded polypeptide is not significantly different from that generated by translation of an arbitrarily chosen region of pMa025. Thus, this line of investigation was not pursued further.

No other ORF showed significant similarity to any protein(s) in the database.

A 2.8 kb region (4390-7196) of pMa025 was significantly ($z=8.1$) similar (50.7%) to two genes, *srmB* and *srmR*, involved in the biosynthesis of macrolide antibiotics by polyketide synthase in the bacterium *Streptomyces ambofaciens* (Geistlich et al.,1992). The alignment is shown in Figure 8.





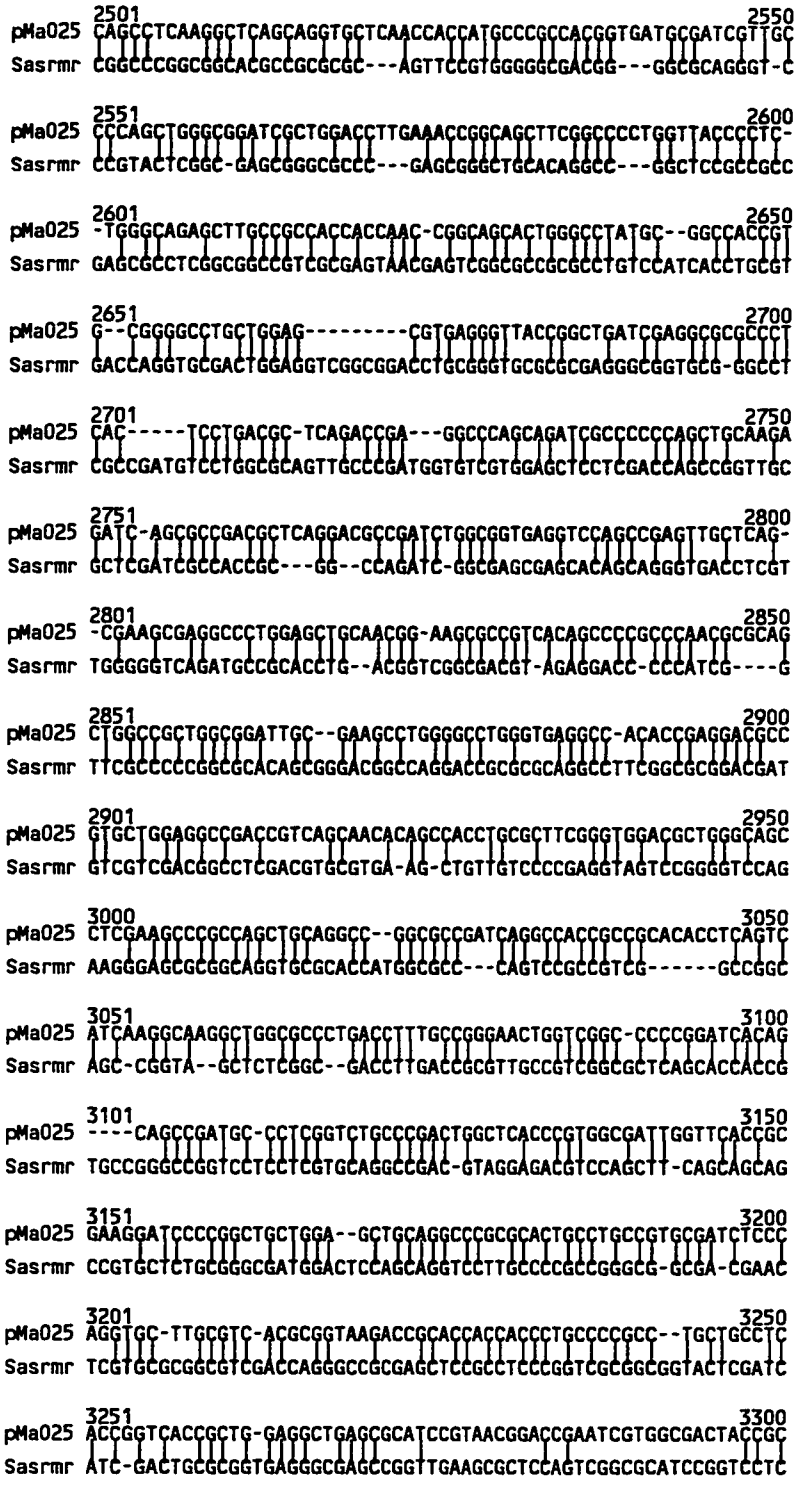


Figure 8. Comparison between pMa025 (nt 4390-7196) and *S.ambofaciens srmB* & *srmR* genes. Identical residues are indicated by vertical lines.

The deduced amino acid sequence from two regions, on the minus strand, of pMa025 showed significant similarity to polyketide synthase proteins. One region (4695-4965) was significantly ($P=4.8x10^{-6}$) similar to the product of a hypothetical polyketide synthase gene in *Streptomyces ambofaciens* (Aigle et al., unpublished). An upstream region (2517-2762) was significantly similar ($P=7.1x10^{-6}$) to the product of a gene encoding the putative ketoreductase domain of polyketide synthase in *Sorangium cellulosum* So ce26 (Schupp et al.,1995).

In addition, the deduced amino acid sequence from a region (4805-5230) on the plus strand of pMa025 showed significant ($P=4.2x10^{-9}$) similarity to the dehydratase domain encoded by a putative fatty acid synthase gene from *Mycobacterium tuberculosis* var. *bovis* BCG (Fernandes, et al.1996).

In summary, regions of pMa025 exhibiting significant similarity to three cyanobacterial plasmid replication proteins and to bacterial polyketide and fatty acid syntheases were identified.

CHAPTER II. TRANSFORMATION OF UV027

METHODS

I. Electroporation

Electroporation involves the application of a brief, high voltage pulse to a suspension of cells and DNA. As a result, the cell membrane reversibly breaks down in localized areas and becomes permeable to exogenous DNA (Knight et al.,1986). Preliminary studies were done using the recombinant plasmid, pMaLI, to determine the parameters necessary to electro-transform *Microcystis aeruginosa* UV027 to carbenicillin (CAR) resistance.

A modification of the procedure developed by Thiel & Poo (1989) was employed. *Microcystis* was grown phototropically in BG-11 medium at 28°C with shaking. Cells from 8 day cultures (OD_{630}=5.5) were used. This OD corresponds to a density of $3x10^8$ cells/ml. The cells were washed 2X in double-distilled sterile water and concentrated to a final volume of approximately $2x10^9$ cells/ml. This washing removes nucleases. Resuspension in a low ionic strength (i.e., high resistance) buffer ensures minimal current flow-through and heating of the sample. This results in greater cell viability.

385µl cells containing 1µg/ml of plasmid DNA (total vol ≈15µl in TE, pH 8.0) were added to a chilled, sterile cuvette and electroporated (Bio Rad Gene Pulser) at a field strength of 8kV/cm with a time constant of 2.2 msec. Donor DNA from pMaLI and deletions 5, 7, 9, 12, 18 and 22 was used. These deletions were chosen since preliminary sequence analysis indicated that a translation product from an≈ 4 kb region (4000-8018) of pMa025 was significantly similar to three cyanobacterial plasmid replication proteins.

Deletion 5 begins at nucleotide position 3150 in pMa025. Both pBluescript II and TE were used as controls.

After electroporation, the cells were immediately diluted 2-fold with BG-11 medium. 200µl samples were plated on each of five BG-11/1.5% agarose (Shirai et al.,1981) plates. Cells were incubated at 28°C under constant illumination.

II. Screening of Transformants

A. Antibiotic Challenge

Twenty-four hours (approx. 2 doubling times) after plating, the cells were challenged with CAR at doses of 0, 0.5, 5 and 25µg/ml, respectively. A prior experiment had indicated that UV027 was sensitive to 0.5µg/ml CAR. The agarose was partially lifted with a sterile spatula and the appropriate dose of CAR, in a total volume of 400µl sterile H_2O, was injected into the center of the Petri dish. This subjected the cells to a gradually increasing concentration of CAR (Shestakov et al.,1970; van den Hondel et al.,1980b; Kuhlemeier et al.,1981; Chauvet et al.,1983; Golden et al.,1984).

Incubation was continued 14-21 days. Dark green colonies were scored as transformants. Depending upon the outcome, 1-5 colonies were transferred to 20 ml BG-11 medium supplemented with 25µg/ml CAR. This antibiotic concentration produced the most reliable results (see Table 3).

B. Plasmid Minipreps

After 7 to 10 days of culture in liquid media plus CAR, plasmid DNA was isolated using the SpinBind Mini-Prep System (FMC). This method uses alkaline lysis of cells, followed by plasmid purification over a microporous silica membrane.

Overnight LB/CAR (100μg/ml) cultures of *E.coli* XL1-Blue cells, started from glycerol stocks of the respective deletions, were used as positive controls. In all instances three ml of culture was used for alkaline lysis. However, since initial results indicated that plasmid yield from UV027 was extremely low, the cleared lysate from subsequent experimental groups was pooled prior to column purification.

UV027 does not possess a plasmid (Raps,1990). Carbenicillin resistance is most probably due to a β-lactamase gene present on the shuttle vector(s). Recombination between homologous regions on the chromosome is not likely. Preliminary results indicated no homology between ^{32}P-labeled pMa025 and UV027 chromosomal DNA (Wallace et al.,1995). Moreover, even when regions of homology exist between a recombinant construct and the chromosome, integration only occurs when the foreign DNA is bounded on both sides by cyanobacterial DNA (Williams et al.,1983; van der Plas et al.,1990). Constructs, such as pMaLI, in which the cyanobacterial sequences are not split do not recombine with the chromosome.

Plasmid yield was monitored electrophoretically (0.8% agarose, 40V, 1.5hr, TBE). In general, cyanobacterial plasmids have very low copy numbers (Dr. W.E. Borrias, personal communication). This was confirmed by the electrophoretic profile of

the *Microcystis* plasmid preps (see Results & Figures 9-12). Thus, plasmids from transformed UV027 cells were recovered by back-transformation (Laudenbach et al., 1985; Chauvet et al.,1986; Kreps et al.,1990; Gendel et al.,1991; Mermet-Bouvier et al.,1993; Dr. J. Cobley, personal communication) of competent (Hanahan,1983) *E.coli* XL1-Blue cells .

Recovered plasmids were characterized by comparing Sal I restriction patterns to those of their respective positive controls.

RESULTS

Transformation of UV027

The number UV027 transformants and transformation efficiency across all CAR concentrations for each of the constructs tested is presented in Table 3.

Plasmid DNA isolated from UV027 cells transformed to CAR (25μg/ml) resistance with pMaLI, deletions 5, 7, 9, 12, 18, 22 and pBluescript II is shown in Figures 9 through 12. The single CAR (5μg/ml) transformant obtained with the negative control (TE) was used to start a BG-11/ CAR culture. Although no growth in liquid medium supplemented with CAR was apparent, this control was carried through the plasmid preparation procedure (see Figure 12, lane 5).

No bands corresponding to plasmid DNA isolated from the positive controls (*E.coli* glycerol stocks) were observed for any *Microcystis* prep even when the equivalent of 6 ml (Figure 11) or 12 ml (Figure 12) culture was used. However, plasmids were recovered by back- transformation of *E.coli* XL1-Blue cells with either 25μl (pMaLI, deletions 7,12,18, 22, pBluescript II) or 75μl (deletions 5, 7 & 9) of the original UV027 plasmid preps. These data are shown in Table 4. Few back-transformants were recovered from *E.coli*. This indicates that either the plasmid copy number is low in *Microcystis* and/or the incoming DNA is restricted (see Chapter I Results, Section II. A). In general, transformation is more efficient with plasmids isolated directly from cyanobacteria than with those passed through *E.coli* (Kuhlemeier et al.,1987). Restriction patterns also differ dependent upon the host. It has been suggested that some restriction sites are protected/modified in cyanobacteria (Buzby et al.,1983; Schwabe et al, 1990;

TRANSFORMATION OF UV027

Construct	Region of pMa025 (bp)	CAR (μg/ml)	Number of transformants	Transformation efficiency*
pMaLI	1-8018	0	green lawn	ND
		0.5	31	$3.1x10^2$
		5	19	$1.9x10^2$
		25	9	$9.0x10^1$
Del 5	3150-8018	0	green lawn	ND
		5	2	$2.0x10^1$
		25	7	$7.0x10^1$
		50	17	$1.7x10^2$
Del 7	4550-8018	0	green lawn	ND
		5	22	$2.2x10^2$
		25	13	$1.3x10^2$
		50	9	$9.0x10^1$
Del 9	5032-8018	0	green lawn	ND
		5	2	$2.0x10^1$
		25	7	$7.0x10^1$
		50	6	$6.0x10^1$
Del 12	5458-8018	0	green lawn	ND
		5	10	$1.0x10^2$
		25	6	$6.0x10^1$
		50	5	$5.0x10^1$
Del 18	6387-8018	0	green lawn	ND
		5	4	$4.0x10^1$
		25	6	$6.0x10^1$
		50	3	$3.0x10^1$
Del 22	7454-8018	0	green lawn	ND
		5	6	$6.0x10^1$
		25	2	$2.0x10^1$
		50	2	$2.0x10^1$
pBluescriptII	------	0	green lawn	ND
		5	2	$2.0x10^1$
		25	1	$1.0x10^1$
		50	1	$1.ox10^1$
TE	------	0	green lawn	ND
		5	1	$1.0x10^1$
		25	0	------
		50	0	------

*CFU/μg DNA; ND, not determined

Table 3. Transformatiom of UV027.

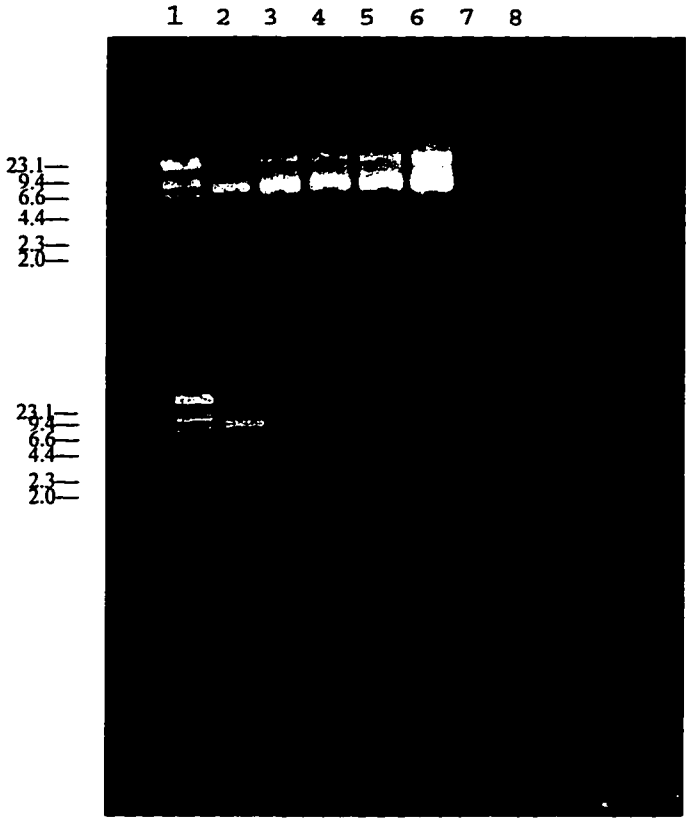


Figure 9. Agarose gel electrophoresis of plasmids isolated from transformed *E.coli* and *M.aer* UV027.
0.8% agarose was used. The gel was run at 40V, 1.5 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.
Top: Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lane 2, pMaLI; Lanes 3-6, pMaLI, *E.coli*; Lanes 7-8, pMaLI, *M.aer*.
Bottom: Lane 1, Hind III digest of lambda DNA; Lane 2, pMaLI; Lanes 3-4, pMaLI, *M.aer*; Lanes 7-8, empty.

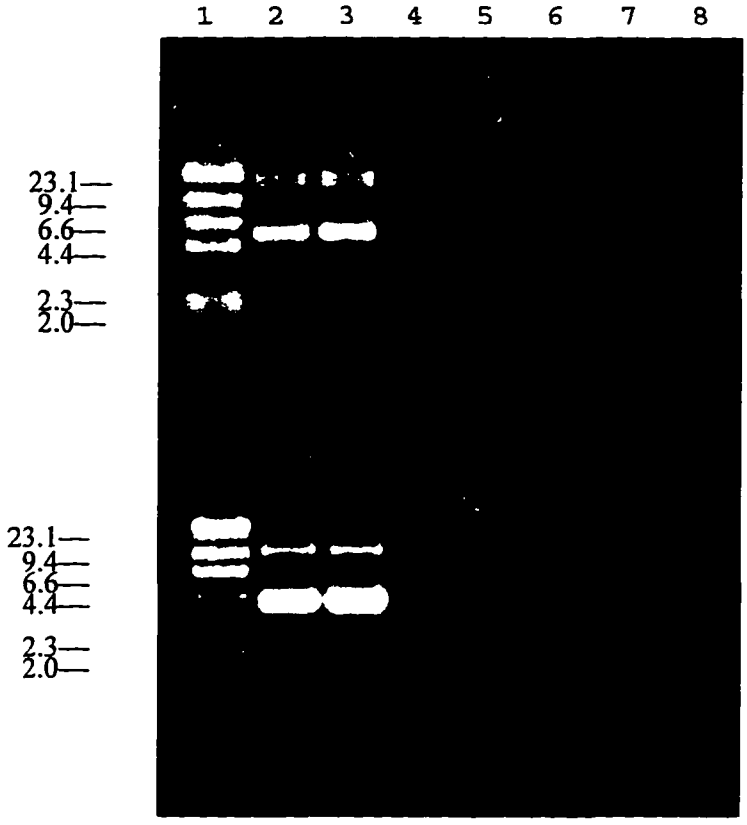


Figure 10. Plasmids recovered from *E.coli* and *M.aer* UV027.
0.8% agarose was used. The gel was run at 40V, 1.5 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.
Top: Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lanes 2-3, deletion 5, *E.coli*; Lanes 4-7, deletion 5, *M.aer*; Lane 8, empty.
Bottom: Lane 1, Hind III digest of lambda DNA; Lanes 2-3, deletion 7, *E.coli*; Lanes 4-7, deletion 7, *M.aer*; Lane 8, empty.

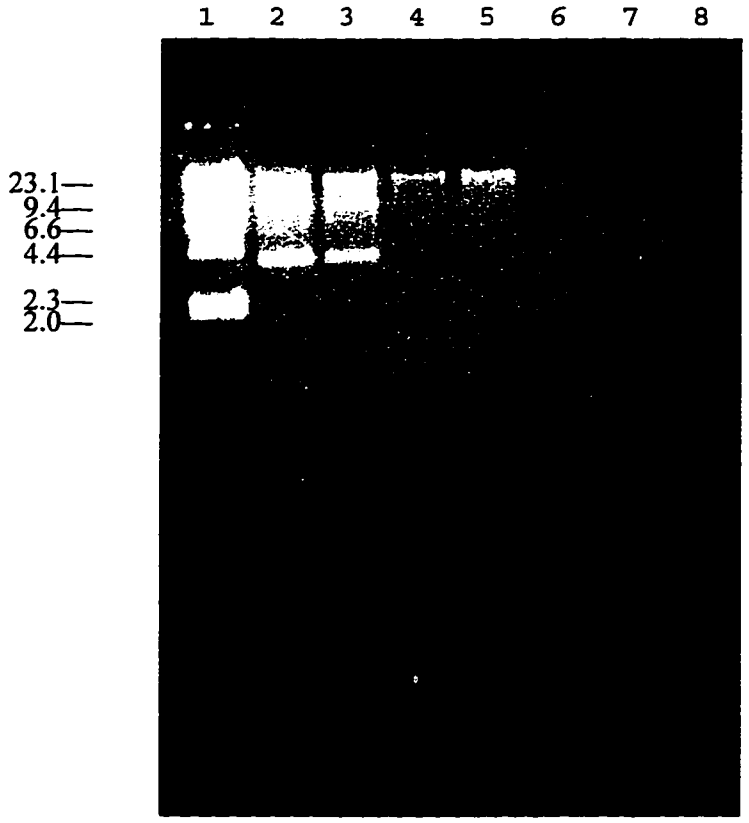


Figure 11. Plasmids recovered from *E.coli* and *M.aer* UV027.
0.8% agarose was used. The gel was run at 40V, 1.5 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.
Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lanes, 2-3, deletion 9, *E.coli*;Lanes 4-5, deletion 9, *M.aer*;
Lanes 6-8, empty.

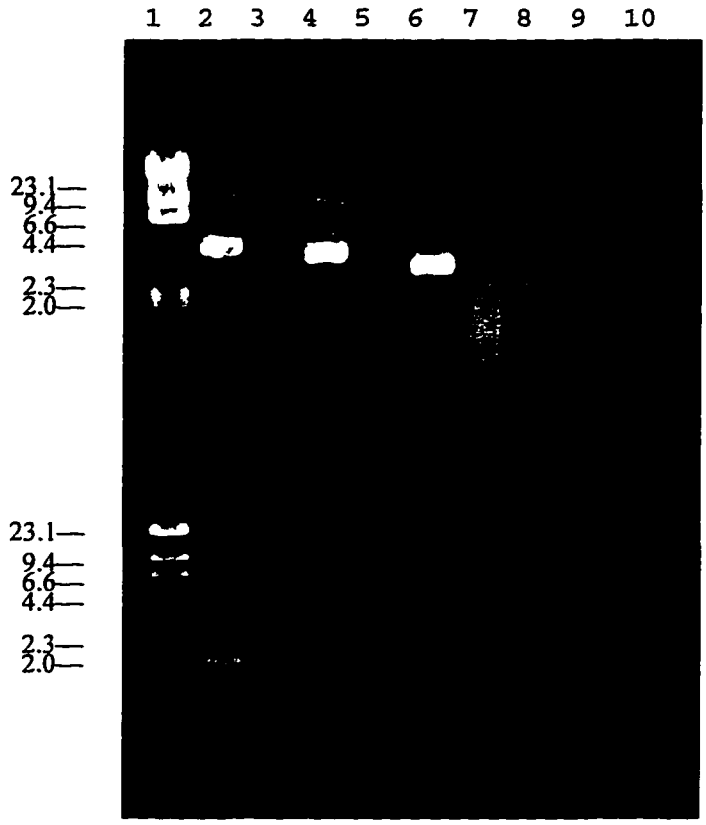


Figure 12. Plasmids recovered from *E.coli* and *M.aer* UV027.
0.8% agarose was used. The gel was run at 40V, 1.5 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.

Top: Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lane 2, deletion 7, *E.coli*; Lane 3, deletion 7, *M.aer*; Lane 4, deletion 12, *E.coli*; Lane 5, deletion 12, *M.aer*; Lane 6, deletion 18, *E.coli*; Lane 7, deletion 18, *M.aer*; Lane 8, deletion 22, *E.coli*; Lane 9, deletion 22, *M.aer*; Lane 10, empty.

Bottom: Lane 1, Hind III digest of lambda DNA; Lane 2, pBluescript II, *E.coli*; Lane 3, pBluescript II, *M.aer;* Lane 4, TE, *E.coli*; Lane 5, TE, *M.aer;* Lanes 6-10, empty.

Tominaga et al.,1993a). Indeed removal (Buzby et al,1985) or methylation (Elhai et al, 1988) of restriction sites on plasmid DNA greatly enhances transformation efficiency in cyanobacteria.

The preparation from the negative (TE) control failed to back-transform *E.coli* to CAR resistance (see Table 4). The single TE carbenicillin transformant observed in *Microcystis* (see Table 3) was most probably due to growth on a region of solid medium lacking CAR. Although the agarose underlay procedure is the method of choice for antibiotic challenge in cyanobacteria (see Method section II, subsection A), there is a potential for uneven distribution of the drug throughout the solid support. Thus, whenever possible, colonies used to start liquid cultures were picked from the center of the agarose.

The discrepancy between the number of transformants observed in UV027 (Table 3) with those obtained by back-transformation in *E.coli* (Table 4) is probably an experimental artifact. The high density of UV027 cells plated, coupled with the diffusion of β-lactamase, most probably resulted in a large number of false positives.

The electro-transformation procedure needs to be optimized. The stability of the transformants upon repeated subculturing must be determined.

The sizing patterns of Sal I restricted plasmids recovered from *E.coli* are shown in Figures 13, 14 and 15. In the majority of cases the banding patterns are identical to those of the positive controls. Only one plasmid each from deletions 7 (Fig. 15, top, lanes 6-7); 18 (Fig. 15, bottom, lanes 4-5) and 22 (Fig. 15, bottom, lanes 12-13) showed aberrant

BACK-TRANSFORMATION OF *E.COLI*		
Construct	Number of transformants	Transformation efficiency*
pMaLIE	1144	$5.7x10^5$
pMaLIM	31	ND
Del 5E	7296	$3.7x10^6$
Del 5M	1	ND
Del 7E	8384	$4.2x10^6$
Del 7M	36	ND
Del 9E	7600	$3.8x10^6$
Del 9M	3	ND
Del 12E	8192	$4.1x10^6$
Del 12M	3	ND
Del 18E	9600	$4.8x10^6$
Del 18M	3	ND
Del 22E	7680	$3.8x10^6$
Del 22M	1	ND
pBluescript IIE	6784	$3.4x10^6$
pBluescript IIM	1	ND
TE	0	ND

*CFU/μg DNA; E, plasmids isolated from *E.coli* glycerol stock
M, plasmid preps from *M.aer*; ND, not determined

Table 4. Back-transformation of *E.coli*

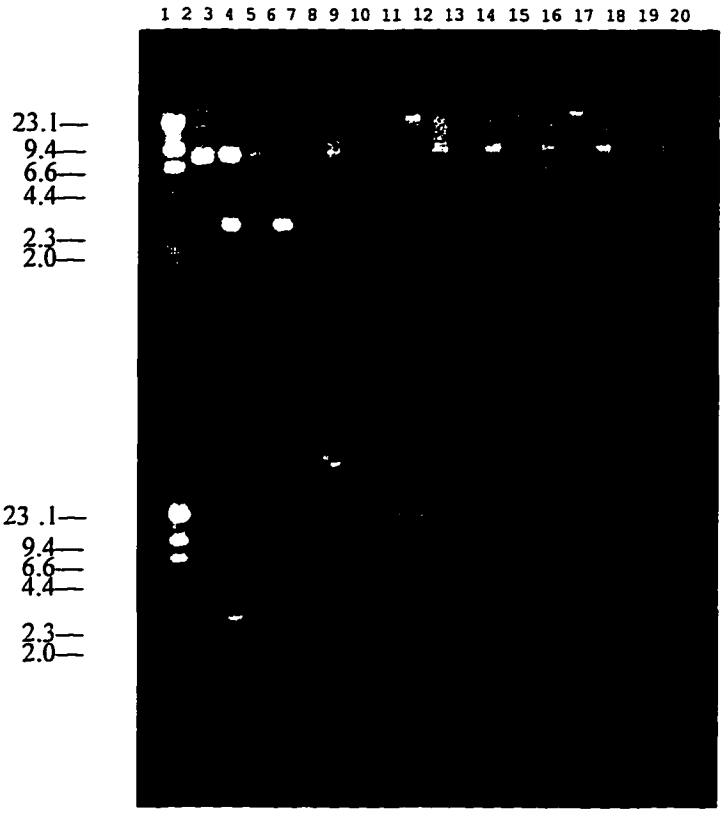


Figure 13. Agarose gel electrophoresis of plasmids recovered from *M.aer* UV027 by back-transformation of *E.coli.*

0.8% agarose was used. The gel was run at 40V, 3.0 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.

Top: Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lane 2, pMaLI; Lane 3, Sal I digest of pMaLI; Lane 4, Sal I digest of pMa025; Lane 5, Sal I digest of pBluescript II; Lanes 6,8,10,12,14,16,18, pMaLI *; Lanes 7,9,11,13,15,19, Sal I digest of pMaLI*; Lane 20, empty.

Bottom: Lane 1, Hind III digest of lambda DNA; Lane 2, Sal I digest of pMa025; Lane 3, Sal I digest of pBluescript II; Lanes 4,6,8,10,12, pMaLI*; Lanes 5,7,9,11,13, Sal I digest of pMaLI*; Lanes 14,16, pMaLI **; Lanes 15,17, Sal I digest of pMaLI**, Lanes 18-20, empty.

*Back-transformants; **_E.coli_ positive controls.

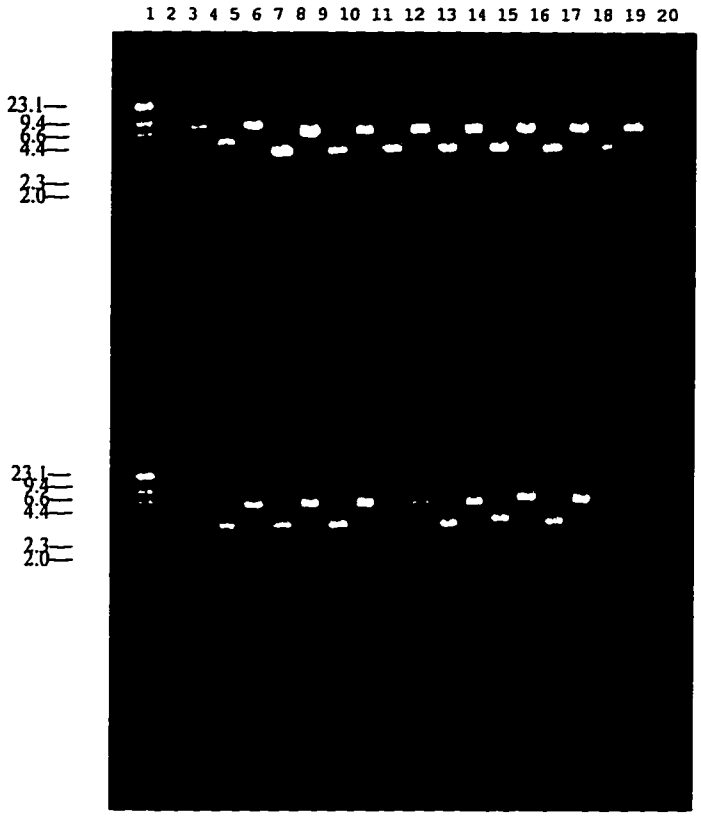


Figure 14. Agarose gel electrophoresis of plasmids recovered from *M.aer* UV027 by back-transformation of *E.coli*.

0.8% agarose was used. The gel was run at 40V, 3.0 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.

Top: Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lane 2, deletion 5**, Lane 3, Sal I digest of deletion 5**; Lane 4, deletion 5*; Lane 5 SalI digest of deletion 5*; Lane 6, deletion 7**; Lane 7, Sal I digest of deletion 7**; Lanes 8,10,12,14,16,18, deletion 7*; Lanes 9,11,13,15,17,19, Sal I digest of deletion 7*; Lane 20, empty.

Bottom: Lane 1, Hind III digest of lambda DNA; Lane 2, deletion 9**; Lane 3, Sal I digest of deletion 9**; Lanes 4,6,8, deletion 9*; Lanes 5,7,9, Sal I digest of deletion 9**; Lane 10, deletion 12**; Lane 11, Sal I digest of deletion 12**; Lanes 12,14,16, deletion 12*: Lanes 13,15,17, Sal I digest of deletion 12*; Lanes 18-20, empty.

*Back-transformants; **_E.coli_ positive controls.

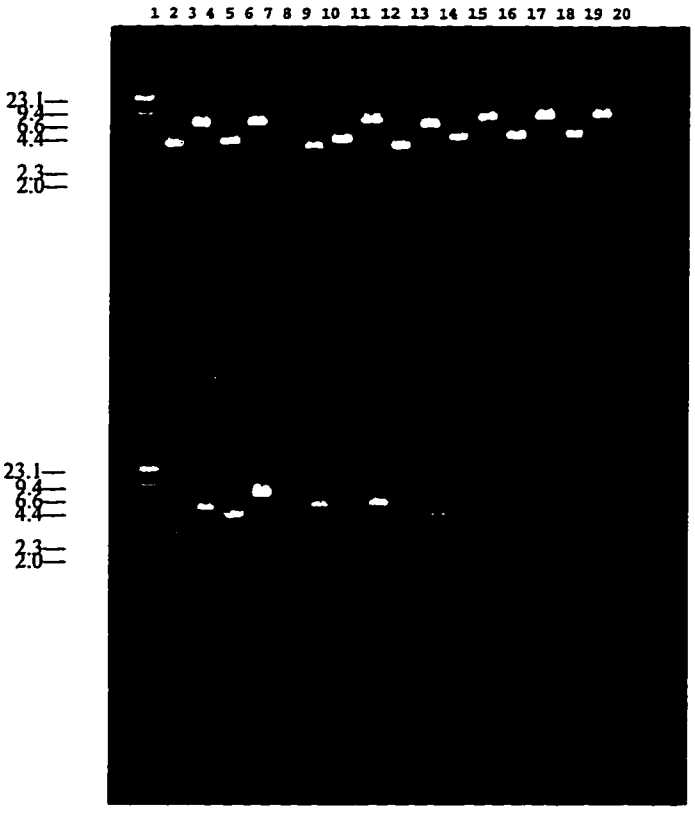


Figure 15. Agarose gel electrophoresis of plasmids recovered from *M.aer* UV027 by back-transformation of *E.coli*.
0.8% agarose was used. The gel was run at 40V, 3.0 hr. The DNA was stained with ethidium bromide and the fluorescence was photographed on a 312 nm U.V. transilluminator.
Top: Lane 1, Hind III digest of lambda DNA, fragment sizes (kb) are indicated at the left; Lane 2, deletion 7**;
Lane 3, Sal I digest of deletion 7**; Lanes 4,6,8, deletion 7*;
Lanes 5,7,9, Sal I digest of deletion 7*; Lane 10, deletion 12**;
Lane 11, Sal I digest of deletion 12**; Lanes 12,14,16, deletion 12*;
Lanes 13,15,17, Sal I digest of deletion 12*; Lanes 18-20, empty.
Bottom: Lane 1, Hind III digest of lambda DNA; Lane 2, deletion 18**;
Lane 3, Sal I digest of deletion 18*; Lanes 4,6,8, deletion 18*;
Lanes 5,7,9, Sal I digest of deletion 18*; Lane 10, deletion 22**;
Lane 11, Sal I digest of deletion 22**; Lane 12, deletion 22*;
Lane 13, Sal I digest of deletion 22*; Lane 14, pBluescript II**;
Lane 15, Sal I digest of pBluescript II**; Lane 16, pBluescript II*;
Lane 17, Sal I digest of pBluescript II*; Lanes 18-20, empty.
*Back-transformants; **E.coli* positive controls.

banding patterns. Thus, recombinant plasmids can be "shuttled", apparently unaltered, between *Microcystis aeruginosa* UV027 and *E.coli*. In addition, these recovered plasmids transformed *E.coli* XL1-Blue cells to CAR resistance with the same efficiency ($\approx 10^6$ CFU/μg DNA) as those maintained in *E.coli* (data not shown).

Comparison of the efficiency of transformationi in UV027 (Table 3), coupled with the yield (Table 4) and characterization (Figures 14 & 15) of *E.coli* back-transformants indicated that deletion 7 is the best shuttle vector. This plasmid, recovered by back-transformation to *E.coli*, transformed UV027 to CAR (25μg/ml) resistance with the same efficiency ($\approx 10^2$ CFU/μg DNA) as the original plasmid. Re-isolation via back-transformation to *E.coli* was again possible. The Sal I restriction pattern remained unaltered. This deletion construct, pMaLIND7, is being maintained in *Microcystis aeruginosa* UV027 under CAR (25μg/ml) selection pressure.

DISCUSSION

pMa025 Sequence

The *Microcystis* plasmid pMa025 is 8018 bp in length with a mol % G+C content of 62.3%. This composition is similar to the 59% value reported for the *Synechococcus* plasmid pUH24 (van der Plas et al.,1992) but decidedly different from the 41% & 42.2% values described for the *Microcystis aeruginosa* f. *aeruginosa* Kützing plasmids pMA1 & 2 (Tominaga et al., 1993,1995). However, inferences concerning genetic relatedness are difficult to make. The taxonomy of cyanobacteria is in a state of disarray (Bergey, 1989). Assignment to groups is provisional due to the sparsity of data. Indeed, the genus *Synechococcus* represents a heterogeneous group whose mol% G+C range (39-71%) is as broad as that observed for all prokaryotes. Classification of *Microcystis* is particularly problematic. These cyanobacteria have been placed temporarily, as a cluster, within the *Synechocystis* group (Bergey, 1989). The presence of identical plasmids in several strains of unicellular (van den Hondel et al.,1979) and filamentous (Felkner et al., 1988) cyanobacteria, despite differences in genome size and base composition (Herdman et al,1979), suggests interspecific transfer of these extrachromosomal elements. The existence of regions of homology between different plasmids in the same or related strains of cyanobacteria (Lau et al.,1980) suggests the presence of transposable elements. Such mobile genetic elements have been found in the filamentous cyanobacteria (Alam et al.,1991; Cai, 1991). Moreover, inter-generic plasmid transfer between *Synechococcus* and *Synechocystis* has been reported (Devilly et al.,1977; Stevens et al.,1986) and has prompted Stevens (1986) to suggest that the

taxonomic boundary between these two genera does not really exist.

Coding Regions

Nineteen presumptive ORFs were found on pMa025. Twelve (63.2%) use GTG as the translational start codon. Although the preponderance of cyanobacterial genes are reported to initiate with ATG, 50% of the putative ORFs reported for the *Synechococcus* sp. PCC7942 plasmid pUH24 (van der Plas et al.,1992) and two of the three hypothetical ORFs from *Microcystis aeruginosa* f *aeruginosa* plasmids pMA1 & 2 (Tominaga et al.,1993,1995) use GTG as the initiation codon. This is particularly striking since the compositional bias (≈ 42% G+C) of the latter two plasmids would tend to favor adenine as the first base in the start codon. Thus, *Microcystis* plasmid genes may use (A/G)TG as translation initiation signals with approximately equal frequency.

Overlapping reading frames from several cyanobacterial genes have been reported (Golden et al.,1988; van der Plas et al.,1992). ORFs H & I and I & J overlap by 3 and 10 bases, respectively (Figure 5). Therefore, the transcript from this region of pMa025 may be polycistronic.

ORFs on both strands in several regions of pMa025 were noted. Specifically, the regions 647-1045, 2788-2324, 3271-3852, 5465-6675, 4805-5230 and 6908-7306 contain putative coding sequences on both DNA strands. Interestingly, this extensive use of a restricted plasmid region for coding is similar to that reported for the two ORFs on the *Microcystis aeruginosa* plasmid, pMA2 (Tominaga et al.,1995).

Regulatory Regions

Sequences resembling consensus ribosome binding sites from *E.coli* or *Synechococcus* PCC6301 (Tomioka et al.,1983) were present upstream of the translational start codon in seventeen ORFs. Sequence motifs matching *E.coli* σ^{70} -10 and -35 promoter consensus hexamers were found upstream of several ORFs. Thus, these putative *Microcystis* genes may be readily expressed in *E.coli*. Expression and protein sequencing studies are needed to confirm the coding functions of the presumptive ORFs.

Motifs corresponding to prokaryotic rho-independent transcription termination signals were not found in pMa025. However, the imperfect inverted repeats downstream of ORFs B and O may play a role in transcription termination.

Many regulatory proteins bind to regions of dyad symmetry on DNA. The two imperfect palindromes upstream of ORFs E and Q may play a role in transcription initiation by altering the secondary structure of the DNA and, hence, it's accessibility to trans-acting factors. The direct repeats, located in the intergenic regions, may also be binding sites for trans-acting factors required for expression of *Microcystis* plasmid genes.

Plasmid Replication

The translation product of a 1700 bp region (5000-6700) of pMa025 was significantly similar to three putative cyanobacterial plasmid replication proteins encoded by ORF 93 from *Synechocystis* sp. strain PCC6803 (Kaneko et al.,1995), ORF F

from *Synechococcus* sp. strain PCC7942 plasmid pUH24 (van der Plas et al.,1992) and ORF 1 from *Synechococcus* sp. strain PCC7002 plasmid pAQ1 (Akiyama et al., unpublished). The Poisson probabilities (BLAST) associated with these matches are $1.8x10^{-26}$, $2.8x10^{-20}$ and $1.5x10^{-14}$, respectively. However, only a subset (5697-5933) of this region, ORF M, encoded a putative plasmid replication protein, RepA, that was significantly similar (BLAST, $P=2.9x10^{-7}$) to the gene product of ORF F.

The hypothetical plasmid replication proteins encoded by ORFs 93, 1 and F are significantly similar to each other. Z-scores, calculated from the pairwise comparisons (Bestfit) between ORFs 93 & 1, ORFs 93 & F and ORFs F & 1, were 77, 13 and 11, respectively. The rank order of the z-scores (11.1, 7.4 & 5.6) for the pairwise comparisons (Bestfit) between the deduced amino acid sequence from the 1700 bp region of pMa025 and the products from each of the three ORFs parallels that of the BLAST Poisson probabilities. The greatest degree of similarity is found between the translation product of this region of pMa025 and the gene product from a genomic region, ORF 93, in *Synechocystis*. It is not known to which of the four small *Synechocystis* PCC6803 plasmids (Chauvet et al.,1986) ORF 93 corresponds. These data reinforce the placement of *Microcystis*, as a cluster, within this group of cyanobacteria. However, the product, RepA, from a delimited portion (ORF M) of this 1700 bp region shows the greatest identity (37.2%) to the RepA protein encoded by ORF F in *Synechococcus* PCC7942. The putative plasmid replication proteins encoded by ORFs 93 & 1 show 28.2.8% and 29.5% identities to RepA. These percentages were calculated from the pairwise alignments (Bestfit) between ORF M and each of the three putative ORFs. Only the ORF

F vs M alignment is significant (z=11.0). This is reflected in the BLAST results. Thus, the hypothetical gene, *repA*, from the *Microcystis aeruginosa* UV025 plasmid, pMa025, encodes a putative plasmid replication protein, RepA, which is significantly similar to the hypothetical plasmid replication protein, RepA, from *Synechococcus* sp. strain PCC7942 plasmid pUH24. RepA also shows slight similarities to presumptive plasmid replication proteins from *Synechocystis* sp. strain PCC6803 and *Synechococcus* sp. strain PCC7002.

However, the *Synechococcus* PCC7002 shuttle vector, pAEQ19 (gift of Dr. D.Bryant), derived from the endogenous plasmid, pAQ1, failed to transform *Microcystis aeruginosa* UV027 to ampicillin resistance using either natural or chemical ($CaCl_2$) methods (Dr. S. Raps, personal communication). Thus, although the hypothetical plasmid replication proteins appear to be similar, other host-specific factors may limit transmission and/or stability. *Microcystis aeruginosa* UV027 is a unicellular, freshwater, obligately photoautotrophic, toxin producing cyanobacterium. *Synechococcus* sp. strain PCC7002 is unicellular, marine and facultatively photoheterophic. Thus, *Microcystis aeruginosa* UV027 and *Synechococcus* PCC7002 are phenotypically distinct microorganisms. Intergeneric plasmid transfer may not be possible between these two cyanobacteria. Studies using electro-transformation are needed to extend and confirm these findings.

All prokaryotic plasmids contain a region, required in *cis*, for plasmid replication. These origins of replication (*oris*) are usually A+T rich and contain direct and inverted repeats (Scott,1984). *Trans*-acting Rep proteins bind to sites within the origin and initiate plasmid replication. In low copy number plasmids, Rep protein concentration is rate-

limiting for initiation of replication (Scott, 1984). Other *cis*-acting elements, while not essential to plasmid replication *per se*, are important for plasmid stability.

Although several areas of pMa025 contain direct and inverted repeats, no region strictly conforming to a plasmid origin of replication was found. Moreover, low-stringency Southern hybridization failed to detect any homology (Wallace et al.,1995) between pMa025 DNA and the putative *ori* region from plasmid pUH24 from *Synechococcus* PCC7942 (van der Plas et al.,1992). However, a 483 bp region (4550-5032) of pMa025, may contain the plasmid *ori*. This region corresponds to the 5' terminus of deletion 7 and would be absent in subsequent smaller constructs that failed to transform *Microcystis*.

All the minimally effective pMa025 derived shuttle vectors (based on the number of transformants and the integrity of the recovered constructs) contain ORF M which encodes the presumptive plasmid replication protein, RepA. However, only pMaLIND7 (deletion 7) which contains both the hypothetical *ori* and the putative *repA* gene is relatively efficient.

Interestingly, deletion constructs which lack the presumptive *ori* (deletions 9 & 12) and those in which both the *ori* and the *repA* gene are absent (deletions 18 & 22) transform *Microcystis*. Plasmids can be recovered by back-transformation to *E.coli*. However, the number of back-transformants (Table 4) is low. Also, in some instances, plasmid integrity, characterized by an alteration in restriction pattern (Figure 15), was compromised. Moreover, the plasmid pBluescript II could be shuttled, albeit inefficiently, between *Microcystis aeruginosa* and *E.coli*. This may reflect the method

used for transformation. The plasmid pUC19 failed to transform *Microcystis* using either natural or chemical ($CaCl_2$) methods (Raps, 1990). Transformation of the cyanobacterium *Synechococcus* PCC7942 with the *E.coli* plasmid pBR322 has been reported (Daniell et al.,1986). In this instance transformation was much more efficient with permeaplasts than with intact cells. It may be that the cell surface alterations in permeaplasts (which lack cell walls) and in electroporated cells (which have transient membrane pores) result in enhanced plasmid uptake. ColE1 plasmids (pUC 19, pBluescript II & pBR322) do not encode replication proteins. Plasmid replication is dependent upon host factors. This suggests the intriguing possibility that the ColE1 replicon may be functional in *Microcystis*. When the pMa025 replicon is comprised or absent, host derived factors in *Microcystis* may support replication from the ColE1 *ori*. Indeed, although many large plasmids contain multiple *oris*, only one initiates replication (Scott,1984). Constructs in which the "preferred" *ori* is excised initiate replication at one of the other *oris*. Further studies are needed to determine the stability of the ColE1 replicon in *Microcystis*.

Both pMaLI and pMaLIND7 are effective shuttle vectors. The drastic reduction in the number of transformants obtained with deletion 5 suggests that a 1.4 kb (3150-4550) region of pMa025 is inhibitory. In pMaLI a stem-loop structure may form at the imperfect inverted repeat (3127-3176, Figure 5). The absence of one of the repeats in deletion 5 (3150-8018), could result in a change in secondary structure. This alteration may make binding sites accessible (the direct repeats at positions 3200-3208 & 3215-3223) to a plasmid replication protein. Binding of RepA at these site(s) would reduce its

effective concentration and result decreased plasmid replication. Indeed, two imperfect direct repeats, partially matching (56%) the nonomers at 3200 and 3215 are located at positions 4636 and 4698 (Figure 5).

In summary, a gene-transfer system has been developed for the cyanobacterium *Microcystis aeruginosa* UV027. Plasmid pMaLIND7 "shuttles" effectively between the cyanobacterium and *E.coli.*

Other Plasmid Related Functions

Several lines of evidence indicate that pMa025 may encode enzymes important in the synthesis of bioactive peptides. Polyketide and peptide synthases are multifunctional enzymes that non-ribosomally synthesize cyclic peptides. Fatty acid, polyketide and peptide synthetases employ a thio-template polymerization mechanism (Marahiel, 1992).

A 2.8 kb region (4390-7196) of pMa025 showed significant similarities to genes involved in the biosynthesis of macrolide antibiotics in the Gram-positive bacterium *Streptomyces ambofaciens* (Geistlich et al.,1992). The deduced animo acid sequence from two areas within this region (4695-4965 & 4805-5230) were also significantly similar to the products of putative polyketide and fatty-acid synthase genes in *S.ambofaciens* (Aigle, unpublished) and *Mycobacterium tuberculosis* var. *bovis* BCG (Fernandes et al.,1996), respectively. In addition, the translation product from another region (2517-2762) exhibited significant similarity to the ketoreductase domain of polyketide synthase in the Gram-negative myxobacterium *Sorangium cellulosum* So ce26 (Schupp et al.,1995). This is particularly interesting since the purple nonphotosynthetic

myxobacteria are considered to be the morphological counterparts of the photosynthetic cyanobacteria.

The cyclic hepatotoxin, microcystin, is believed to be produced non-ribosomally by peptide synthetases. These enzymes reputedly use a thio-template mechanism in the formation of peptide bonds (Arment et al.,1996). Both toxic and non-toxic strains of *Microcystis* possess genes encoding peptide synthetases (Meissner et al.,1996), suggesting a role for these enzymes in the synthesis of peptides other than microcystin.

The consensus is that plasmids are not involved in the synthesis of toxins in cyanobacteria (Vakeria et al.,1985; Bose et al.,1990). Indeed, both *Microcystis aeruginosa* UV027 and UV025 are toxic, yet only the latter possesses a plasmid. Thus, pMa025 probably does not harbor genes involved in the synthesis of microcystin. However, the plasmid may be involved in the biosynthesis of other cyclic peptides that are part of the repertoire of secondary metabolites produced by these microorganisms.

BIBLIOGRAPHY

Alam, J., Verba, J.M., Cai, Y., Martin, J.A., Weisla, L.J. & Curtis, S.E. 1991. Characterization of the IS*895* family of insertion sequences from the cyanobacterium *Anabaena* sp. strain PCC7120. *J. Bacteriol.* **173**(18):5778-5783.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.

Arment, A.R. & Carmichael, W.W. 1996. Evidence that microcystin is a thio-template product. *J. Phycol.* **32**:591-597.

Asato, Y. & Ginoza, H.S. 1973. Separation of small circular DNA molecules from the blue-green alga *Anacystis nidulans*. *Nature New Biol.* **244**:133-134.

Bagdasarian, M., Lurz, R., Rückert, B., Franklin, F.C.H., Bagdasarian, M.M., Frey, J. & Timmis, K.N. 1981. Specific-purpose plasmid cloning vectors. II Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. *Gene* **16**:237-247.

Bergey, 1989. In: Staley, J.T., Bryant, M.P., Pfennig, N. & Holt, J.G. (eds). Bergey's Manual of Systematic Bacteriology. Baltimore: Williams & Wilkins. pp. 1710-1746.

Bose, S.G. & Carmichael, W.W. 1990. Plasmid distribution among unicellular and filamentous toxic cyanobacteria. *J. Appl. Phycol.* **2**:131-136.

Brendel, V. & Trifonov, E.N. 1984. A computer algorithm for testing prokaryotic terminators. *Nuc. Acids Res.* **12**:4411-4427.

Buzby, J.S., Porter, R.D. & Stevens Jr., S.E. 1983. Plasmid transformation in *Agmenellum* quadruplicatum PR-6: construction of biphasic plasmids and characterization of their transformation properties. *J. Bacteriol.* **154**(3):1446-1450.

Buzby, J.S., Porter, R.D. & Stevens Jr., S.E. 1985. Expression of the *Escherichia coli lacZ* gene on a plasmid vector in a cyanobacterium. *Science* **230**:805-807.

Cai, Y. 1991. Characterization of the insertion sequence IS*892* and related elements from the cyanobacterium *Anabaena* sp. strain PCC7120. *J. Bacteriol.* **173**(18):5771-5777.

Carmichael,W.W. 1994. The toxins of cyanobacteria. *Scientific American* **270**: 78-86.

Caskey, C.T., Beaudet, A. & Nirenberg, M. 1968. RNA codons and protein synthesis. Dissimilar responses of mammalian and bacterial transfer RNA fractions to messenger RNA codons. *J. Mol. Biol.* **37**:99-118.

Chauvet, F., Astier, C., Vedel, F. & Joset-Espardellier, F. 1983. Transformation in the cyanobacterium *Synechococcus* R2: improvement inefficiency; role of the pUH24 plasmid. *Mol. Gen. Genet.* **191**:39-45.

Chauvet, F., De Vries, L., Van der Ende, A. & Van Arkel, G.A. 1986. A host-vector system for gene cloning in the cyanobacterium *Synechocystis* PCC6803. *Mol. Gen. Genet.* **204**:185-191.

Cobley, J.G., Zerweck, E., Reyes, R., Mody, A., Seludo-Unson, J.R., Jaeger, H., Weerasuriya, S. and Novankasattusas, S. 1993. Construction of shuttle plasmids which can be efficiently mobilized from *Escherichia coli* into the chromatically adapting cyanobacterium, *Fremyella diplosiphon. Plasmid* **30**:90-105.

Conley, P.B., Lemaux, P.G. & Grossman, A. 1988. Molecular characterization and evolution of sequences encoding light-harvesting components in the chromatically adapting cyanobacterium *Fremyella diplosiphon. J. Mol. Biol.* **199**:447-465.

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

Daniell, H., Sarojini, G. & Mc Fadden, A. 1986. Transformation of the cyanobacterium *Anacystis nidulans* 6301 with the *Escherichia coli* plasmid pBR322. *PNAS* **83**:2546-2550.

Devilly, C.I. & Houghton, J.A. 1977. A study of genetic transformation in *Gloeocapsa alpicola. J. Gen. Microbiol.* **98**:277-280.

Doolittle, R.F. 1986. Of URFs and ORFs. A Primer on How to Analyze Derived Amino Acid Sequences. Mill Valley, CA: University Science Books. pp. 12-15.

Eberhard, W.G. 1989. Why do bacterial plasmids carry some genes and not others? Plasmid **21**:167-174.

Elhai, J. & Wolk, C.P. 1988. Conjugal transfer of DNA to cyanobacteria. *Meth Enzymol.* **167**:747-754.

Eriksson, J.E., Paatero, G.I.L., Meriluto, J.A.O., Codd, G.A., Kass, G.E.N., Nicotera, P., & Orrenius, S. 1989. Rapid microfilament reorganization induced in isolated rat hepatocytes by microcystin-LR, a cyclic peptide toxin. *Exp.Cell Res.* **185**:86-100.

Essich, E., Stevens Jr., S.E. & Porter, R.D. 1990. Chromosomal transformation in the cyanobacterium *Agmenellum quadruplicatum*. *J. Bacteriol.* **172**(4):1916-1922.

Felkner, R.H. & Barnum, S.R. 1988. Plasmid content and homology of 16 strains of filamentous, nonheterocystous cyanobacteria. *Curr. Microbiol.* **17**:37-41.

Feng, D.F. & Doolittle, R.F. 1987. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.* **25**:351-360.

Fernandes, N. D. & Kolattukudy, P.E. 1996. Cloning, sequencing and characterization of a fatty acid synthase-encoding gene from *Mycobacterium tuberculosis* var. *bovis* BCG. *Gene* **170**:95-99.

Ferretti, J.J. 1971. Low-level reading of the UGA triplet in *Salmonella tryphimurium*. *J. Bacteriol.***106**(2):691-693.

Friedberg, D. & Seijffers, J. 1979. Plasmids in two cyanobacterial strains. *FEBS Lett.* **107**(1):165-168.

Friedberg, D. & Seijffers, J. 1983. A new hybrid plasmid capable of transforming *Escherichia coli* and *Anacystis nidulans*. *Gene* **22**:267-275.

Geistlich, M., Losick, R., Turner, J.R. & Rao, R.N. 1992. Characterization of a novel regulatory gene governing the expression of a polyketide synthase gene in *Streptomyces ambofaciens Mol. Microbiol.* **6**(14):2019-2029.

Gendel, S.M. 1987. Instability of Tn5 inserts in cyanobacterial cloning vectors. *J. Bacteriol.* **169**(10):4426-4430.

Gendel, S., Straus, N., Pulleyblank, D. & Williams, J. 1983a. Shuttle cloning vectors for the cyanobacterium *Anacystis nidulans*. *J. Bacteriol.* **156**(1):148-154.

Gendel, S., Straus, N., Pulleyblank, D. & Williams, J. 1983b. A novel shuttle cloning vector for the cyanobacterium *Anacystis nidulans*. *FEMS Micro. Lett.* **19**:291-294.

Gendel, S.M. & Tymeson, M. 1991. Instability of Tn5 in a plasmid cloning vector for the cyanobacterium *Anacystis nidulans* R2. *Plasmid* **25**:137-140.

Gish, W. & States, D.J. 1993. Identification of protein coding regions by database similarity search. *Nat. Genet.* **3**:266-72.

Golden, S.S. & Sherman, L.A. 1983. A hybrid plasmid is a stable cloning vector for the cyanobacterium *Anacystis nidulans* R2. *J. Bacteriol.* **155**(3):966-972.

Golden, S.S. & Sherman, L.A. 1984. Optimal conditions for genetic transformation of the cyanobacterium *Anacystis nidulans*. *J. Bacteriol.* **158**(1):36-42.

Golden, S.S. & Stearns, G.W. 1988. Nucleotide sequence and transcript analysis of three photosystem II genes from the cyanobacterium *Synechococcus* sp. PCC7942. *Gene* **67**:85-96.

Gonzalez, N., Wiggs, J. & Chamberlin, M.J. 1977. A simple procedure for the resolution of *Escherichia coli* RNA polymerase holoenzyme from core polymerase. *Arch.Biochem. Biophys.***182**:404-408.

Grigorieva, G. & Shestakov, S. 1982. Transformation in the cyanobacterium *Synechocystis* sp. 6803. *FEMS Microbiol. Lett.* **13**:367-370.

Hanahan, D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J. Molec. Biol.* **166**:557-580.

Hauman, J.H., Coetzee, W.F., Scott, W.E. 1982. Isolation of plasmids from toxic *Microcystis aeruginosa*. *South African J. Sci.* **78**:381.

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

Hirsh, D. & Gold, L. 1971. Translation of thee UGA triplet *in vitro* by tryptophan transfer RNA's. *J. Mol. Biol.* **58**:459-468.

Honkanen, R.E., Zwiller, J., Moore, R.E., Daily, S.L., Khatra, B.S., Dukelow, M. & Boynton, A.L. 1990. Characterization of microcystin-LR, a potent inhibitor of type 1 and type 2A protein phosphatases. *J. Biol. Chem.* **265**(32):19401-19404.

Hooser, S.B., Beasley, V.R., Lovell, R.A., Carmichael, W.W. & Haschek, W.M. 1989. Toxicity of microcystin LR, a cyclic heptapeptide hepatotoxin from *Microcystis aeruginosa*, to rats and mice. *Vet. Pathol.* **26**:246-252.

Hooser, S.B., Kuhlenschmidt, M.S., Dahlem, A.M., Beasley, V.R., Carmichael, W.W. & Haschek, W.M. 1991. Uptake and subcellular localization of tritiated dihydro-microcystin-LR in rat liver. *Toxicon.* **29**(6):589-601.

Houmard, J. & Tandeau de Marsac, N. 1988. Cyanobacterial genetic tools: current status. *Meth. Enzymol.* 167:808-847.

Kaneko, T., Tanaka, A., Sato, S., Kotani, H., Sazuka, T., Miyajima, N., Sugiura, M. & Tabata,S. 1995. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. Strain PCC6803. I. Sequence features in the 1Mb region from map positions 64% to 92% of the genome. *DNA Research* **2**:153-166.

Knight, D.E. & Scrutton, M.C. 1986. Gaining access to the cytosol: the technique and some applications of electroporation. *Biochem. J.* **234**:497-506.

Kreps, S., Ferino, F., Mosrin, C., Gerits, J., Mergeay, M. & Thuriaux, P. 1990. Conjugative transfer and autonomous replication of promiscuous Inc Q plasmid in the cyanobacterium *Synechocystis* PCC6803. *Mol. Gen. Genet.* **221**:129-133.

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

Kuhlemeier, C.J., Hardon, E.M., van Arkel, G.A. & van de Vate, C. 1985. Self-cloning in the cyanobacterium *Anacystis nidulans* R2: Fate of a cloned gene after reintroduction. *Plasmid* **14**:200-208.

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. & van Arkel, G.A. 1983. A host-vector system for gene cloning in the cyanobacterium *Anacystis nidulans* R2. *Plasmid* **10**:156-163.

Kuhlemeier, C.J. & van Arkel, G.A. 1987. Host-vector systems for gene cloning in cyanobacteria. *Meth. Enzymol.* **153**:199-215.

Kurokawa, M., Tominaga, H., Ashida, H., Sawa, Y. & Ochiai, H. 1994. Replication of filamentous cyanobacterial plasmids, pPF1 from *Phormidium foveolarum* and pPB1 from *Plectonema boryanum. Biosci. Biotech. Biochem.* **58**(4)796-797.

Lambert, G.R. & Carr, N.G. 1982. Rapid small-scale plasmid isolation by several methods from filamentous cyanobacteria. *Arch. Microbiol.* **133**:122-125.

Lambert, G.R., Scott, J.G. & Carr, N.G. 1984. Characterization and cloning of extra-chromosomal DNA from filamentous cyanobacteria. *FEMS Microbiol. Lett.* **21**:225-231.

Lau, R.H. & Doolittle, W.F. 1979. Covalently closed circular DNAs in closely related unicellular cyanobacteria. *J. Bacteriol.* **137**(1):648-652.

Lau, R.H., Sapienza, C. & Doolittle, W.F. 1980. Cyanobacterial plasmids: their widespread occurrence, and the existence of homology between plasmids in the same and different species. *Mol. Gen. Genet.* **178**:203-211.

Lau, R.R. & Straus, N.A. 1985. Versatile shuttle cloning vectors for the unicellular cyanobacterium *Anacystis nidulans* R2. *FEMS Microbiol. Lett.* **27**:253-256.

Laudenbach, D.E., Straus, N.A., Gendel, S. & Williams, J.P. 1983.The large endogenous plasmid of *Anacystis nidulans*: mapping, cloning and localization of the origin of replication. *Mol. Gen. Genet.* **192**:402-407.

Laudenbach, D.E., Straus, N.A. & Williams, J.P. 1985. Evidence for two distinct origins of replication in the large endogenous plasmid of *Anacystis nidulans* R2. *Mol. Gen. Genet.* **199**:300-305.

Lomax, T.L., Conley, P.B., Schilling, J. & Grossman, A.R. 1987. Isolation and characterization of light-regulated phycobilisome linker polypeptide genes and their transcription as a polycistronic mRNA. *J. Bacteriol.* **169**(6):2675-2684.

Mac Kintosh, C., Beattie, K.A., Klumpp, S., Cohen, P. & Codd, G. 1990. Cyanobacterial microcystin-LR is a potent and specific inhibitor of protein phosphatases 1& 2A from both mammals and higher plants. *FEBS* **264**(2):187-192.

Marahiel, M.A. 1992. Multidomain enzymes involved in peptide synthesis. *FEBS* **307**:40-43.

Matsunaga, T., Takeyama, H. & Nakamura, N. 1990. Characterization of cryptic plasmids from marine cyanobacteria and construction of a hybrid plasmid potentially capable of transformation of marine cyanobacterium, *Synechococcus* sp. and its transformation. *App. Biochem. Biotech.* **24/25**:151-166.

Meissner, K., Dittmann, E.& Börner,T. 1996. Toxic and non-toxic strains of the cyanobacterium Microcystis *aeruginosa* contain sequences homologous to peptide synthetase genes. *FEMS Microbiol. Letters* **135**:295-303.

Meng, B.Y., Shinozaki, K. & Sagiura, M. 1989. Genes for the ribosomal proteins S12 and S7 and the elongation factors EF-G and EF-Tu of the cyanobacterium, *Anacystis nidulans*: structural homology between 16S rRNA and S7 mRNA. *Mol. Gen. Genet.* **216**:25-30.

Mermet-Bouvier, P., Cassier-Chauvet, C., Marraccini, P. & Chauvat, F. 1993. Transfer and replication of RSF1010-derived plasmids in several cyanobacteria of the genera Synechocystis and *Synechococcus*. *Curr. Microbiol.* **27**:323-327.

Model, P. Webster, R.E. & Zinder, N.D. 1969. The UGA codon *in vitro*: chain termination and suppression. *J. Mol. Biol.* **43**:177-190.

Moore, C.H., Farron, F. Bohnert, D. & Weissman, C. 1971. Possible origin of a minor virus specific protein (A1) in Qβ particles. *Nature New Biol.*234:204-206.

Moro-Pastor, A.M., Kuritz, T., Flores, E., Herrero, A. & Wolk, C.P. 1994. Transfer of a genetic marker from a megaplasmid of *Anabaena* sp. strain PCC7120 to a different *Anabaena* strain. *J. Bacteriol.* **176**(4):1093-1098.

Murphy, R.C. & Stevens, Jr., E. 1992. Cloning and expression of the *cryIVD* gene of *Bacillus thuringiensis* subsp. *israelensis* in the cyanobacterium *Agmenellum quadruplicatum* PR-6 and its resultling larvicidal acrivity. *App. Environ. Microbiol.* **58**:1650-1655.

Murray, M.A. & Wolk, C.P. 1991. Identification and initial utilization of a portion of the smaller plasmid of *Anabaena variabilis* ATCC 29413 capable of replication in *Anabaena* sp.strain M-131. *Mol. Gen. Genet.* **227**:113-119.

Namikoshi, M., Rinehart, K.L., Sakai, R., Stotts, R.R., Dahlem, A.M., Beasley, V.R., Carmichael, W.W. & Evans, W.R. 1992. Identification of 12 hepatotoxins from a Homer Lake bloom of the cyanobacteria *Microcystis aeruginosa*, *Microcystis viridis*, and *Microcystis wesenbergii*: nine new microcystins. *J. Org. Chem.* **57**:866-872.

Needleman, S.B. & Wunsch, C.D. 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* **48**:443-453.

Nicholson, M.L., Gaasenbeek, M. & Laudenbach, D.E. 1995. Two enzymes together capable of cysteine biosynthesis are encoded on a cyanobacterial plasmid. *Mol. Gen. Genet.* **247**:623-632.

Nishiwaki-Matsushima, R., Ohta, T., Nishiwaki, S., Suganuma, M., Kohyama, K., Ishikawa, T., Carmichael, W.W. & Fujiki, H. 1992. Liver tumor promotion by the cyanobacterial cyclic peptide toxin microcystin-LR. *J. Can. Res .Clin. Oncol.* **118**:420-424.

Orkwiszewski, K.G. & Kaney, A.R. 1974. Genetic transformation of the blue-green bacterium, *Anacystis nidulans*. *Arch. Microbiol.* **98**:31-37.

Pearson, W.R. & Lipman, D.J. 1988. Improved tools for biological sequence analysis. *PNAS* **85**:2444-2448.

Perkins, D.R. & Barnum, S.R. 1992. DNA sequence analysis of a cryptic 4.2-kb plasmid from the filamentous cyanobacterium *Plectonema* sp. strain PCC6402. *Plasmid* **28**:170-176.

Porter, R.D. 1986. Transformation in cyanobacteria. *CRC Crit. Rev. Microbiol.* **13**(2):111-132.

Potts, M. 1984. Distribution of plasmids in the cyanobacteria of the LPP group. *FEMS Microbiol. Lett.* **24**:351-354.

Raps, S. 1990. Isolation and characterization of a plasmid from *Microcystis aeruginosa*. *Eastern Regional Photosyn. Conf.*, MBL, Woods Hole, Mass. p26. (abstract)

Raps, S., Wallace, M. & Scott, A. 1991. Characterization of a plasmid isolated from the cyanobacterium *M. aeruginosa*. *VII Int. Symp. Photosynth. Prokar.*, Amherst, MA, 182A. (abstract)

Reddy, K.J., Bullerjahn, G.S., Sherman, D.M. & Sherman, L.A. 1988. Cloning, nucleotide sequence, and mutagenesis of a gene (*irpA*) involved in iron-deficient growth of the cyanobacterium *Synechococcus* sp. strain PCC7942. *J. Bacteriol.* **170**(10):4466-4476.

Rebière M.C., Castets, A.M., Houmard, J. & Tandeau de Marsac, N. 1986. Plasmid distribution among unicellular and filamentous cyanobacteria: occurrence of large and mega-plasmids. *FEMS Microbiol. Lett.* **37**:269-275.

Restaino, L. & Frampton, E.W. 1975. Labeling the deoxyribonucleic acid of *Anacystis nidulans*. *J. Bacteriol.* **124**(1):155-160

Roberts, T.M. & Koths, K.E. 1976. The blue-green alga *Agmenellum quadruplicatum* contains covalently closed DNA circles. *Cell* **9**:551-557.

Roth, J. R. 1970. UGA mutations in *Salmonella typhimurium*. *J. Bacteriol.* **102**(2):467-475.

Sambrook, J.F., Fan, D.P. & Brenner, S. 1967. A strong suppressor specific for UGA. *Nature* **214**:452-453.

Sambrook, J., Fritsch, E.J. & Maniatis, F. 1989. Molecular Cloning, a Laboratory Manual, 2nd. ed. New York: Cold Spring Harbor Laboratory Press.

Sanger, F., Nicklen, S. & Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *PNAS.* **74**:5463-5467.

Schaeffer, M.R., Chiang, G.G., Cobley, J.G. & Grossman, A.R. 1993. Plasmids from two morphologically distinct cyanobacterial strains share a novel replication origin. *J. Bacteriol.* **175**(17):5701-5705.

Schmetterer, G. & Wolk, C.P. 1988. Identification of the region of cyanobacterial plasmid pDU1 necessary for replication in *Anabaena* sp. strain M-131. *Gene* **62**:101-109.

Schupp, T., Toupet, C., Cluzel, B., Neff, S., Hill, S., Beck, J.J. & Ligon, J.M. 1995. A *Sorangium cellulosum* (Myxobacterium) gene cluster for the biosynthesis of the macrolide antibiotic soraphen A: cloning, characterization, and homology to polyketide synthase genes from actinomycetes. *J. Bacteriol.* **177**(13):3673-3679.

Schwabe, W., Hübschmann, T., Meixner, M., Weihe, A. & Börner,T. 1990. Transcription and in vivo expression of a *Microcystis aeruginosa* plasmid. *Curr. Microbiol.* **20**:365-368.

Schwabe, W., Weihe, A., Börner, T., Henning, M. & Kohl, J-G. 1988. Plasmids in toxic and nontoxic strains of the cyanobacterium *Microcystis aeruginosa*. *Curr. Microbiol.* **17**:133-137.

Scott, J.R. 1984. Regulation of plasmid replication. *Microbiol. Rev.* **48**:1-23.

Sherman, L.A. & van de Putte, P. 1982. Construction of a hybrid plasmid capable of replication in the bacterium *Escherichia coli* & the cyanobacterium *Anacystis nidulans*. *J. Bacteriol.* **150**(1):410-413.

Shestakov, S.V. & Khyen, N.T. 1970. Evidence for genetic transformation in blue-green alga *Anacystis nidulans*. *Mol. Gen. Genet.* **107**:372-375.

Shine, J. & Dalgarno, L. 1974. The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *PNAS.* **71**(4):1342-1346.

Shirai, M., Matumaru, K., Ohotake, A., Takamura, Y., Aida, T. & Nakano, M. 1989. Development of a solid medium for growth and isolation of axenic *Microcystis* strains (cyanobacteria). *App. Environ. Microbiol.* **55**(10):2569-2571.

Simon, R.D. 1978. Survey of extrachromosomal DNA found in the filamentous cyanobacteria. *J. Bacteriol.* **136**(1):414-418.

Smith, T.F. & Waterman, M.S. 1981. Comparison of bio-sequences. *Adv. App. Math.* **2**:482-489.

Sode, K., Tatara, M., Ogawa, S. & Matsunaga, T. 1992. Maintenance of broad host range vector pKT230 in marine unicellular cyanobacteria. *FEMS Microbiol. Lett.* **99**:73-78.

Söll, D., Ohtsuka, E., Jones, D.S., Lohrmann, R., Hayatsu, H., Nishimura, S. & Khorana, H.G. 1965. Studies on polynucleotides, XLIX. Stimulation of the binding of aminoacyl-sRNA's to ribosomes by ribotrinucleotides and a survey of codon assignments for the 20 amino acids. *PNAS.* **54**:1378-1385.

Soltes-Rak, E., Kushner, D.J., Williams, D.D. and Coleman, J.R. 1995. Factors regulating *cryIVB* expression in the cyanobacterium *Synechococcus* PCC7942 . *Mol. Gen. Genet.* **246**:301-308.

Staden, R. 1980. A new computer method for the storage and manipulation of DNA gel reading data. *Nuc. Acids Res.* **8**:3673-3694.

Stevens Jr., S.E. & Porter, R.D. 1980. Transformation in *Agmenellum quadruplicatum.* *PNAS* **77**(10):6052-6056.

Stevens Jr., S.E. & Porter, R.D. 1986. Heterospecific transformation among cyanobacteria. *J. Bacteriol.* **167**(3):1074-1076.

Tandeau de Marsac, N., de la Torre, F. & Szulmajster, J. 1987. Expression of the larvicidal gene of *Bacillus sphaericus* 1593M in the cyanobacterium *Anacystis nidulans* R2. *Mol. Gen. Genet.* **209**:396-398

Tandeau de Marsac, N. & Houmard, J. 1987. Advances in cyanobacterial molecular genetics. In The Cyanobacteria. Fay, P & Van Baalen, C. (eds). Amsterdam: Elsevier. pp 251302.

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

Thiel, T. & Poo, H. 1989. Transformation of a filamentous cyanobacterium by electroporation. *J. Bacteriol.* **171**(10):5743-5746.

Tinoco, I., Uhlenbeck, O.C. & Levine, M.D. 1971. Estimation of secondary structure in ribonucleic acids. *Nature* **230**:362-367.

Tominaga, H., Soejima, K., Kawagishi, S., Ashida, H., Sawa, Y. & Ochiai, H. 1993a. Structural organization of a cryptic plasmid, pMA1, from *Microcystis aeruginosa* f. *aeruginosa* Kützing. *Biosci. Biotech. Biochem.* **57**(9):1503-1507.

Tominaga, H., Hayashida, Y., Hosoya, Y., Kurokawa, M., Sawa, Y. & Ochiai, H. 1993b. Characterization of a small cryptic plasmid, pPF1, from *Phormidium foveolarum* and vector construction. *Biosci. Biotech. Biochem.* **57**(10)1795-1799.

Tominaga. H., Kawagishi, S., Ashida, H., Sawa, Y. & Ochiai, H. 1995. Structure and replication of cryptic plasmids, pMA1 and pMA2, from a unicellular cyanobacterium, *Microcystis aeruginosa. Biosci. Biotech. Biochem.* **59**(7):1217-1220.

Tomioka, N. & Sugiura, M. 1983. The complete nucleotide sequence of a 16S ribosomal gene from a blue-green alga, *Anacystis nidulans. Mol. Gen. Genet.* **191**:46-50.

Toxic Cyanobacteria Web Site. 1996. In *Cyanonews* 12(2). Elhai.J (ed).Richmond: University of Richmond. p5.

Turner, P.C., Gammie, A.J., Hollinrake, K., Codd, G.A. 1990. Pneumonia associated with contact cyanobacteria. *Br. Med. J.* **300**(6737):1440-1441.

Uhlenbeck, O.C., Martin, F.H. & Doty, P. 1971. Self-complementary oligonucleotides: effects of helix defects and guanylic acid-cytidylic acid base pairs. *J. Mol. Biol.* **57**:217-229.

Vachhani, A.K., Iyer, R.K. & Tuli, R. 1992. Characterization of a small endogenous plasmid from the cyanobacterium *Plectomena boryanum. J. Biosci.* **17**(2):167-172.

Vachhani, A.K., Iyer, R.K. & Tuli, R. 1993. A mobilizable shuttle vector for the cyanobacterium *Plectonema boryanum. J. Gen. Microbiol.* **139**:569-573.

Vakeria, D., Codd, G.A., Bell, S.G., Beattie, K.A. & Priestly, I.M. 1985. Toxicity and extrachromosomal DNA in strains of the cyanobacterium *Microcystis aeruginosa*. *FEMS Microbiol. Lett.* **29**:66-72.

van den Hondel, C.A.M.J.J., Keegstra, W., Borrias, W.E. & van Arkel, G.A. 1979. Homology of plasmids in strains of unicellular bacteria. *Plasmid* **2**:323-333.

van den Hondel, C.A.M.J.J. & van Arkel, G.A. 1980a. Development of a cloning system in cyanobacteria. *Antonie van Leeuwenhock.* **46**:228-229.

van den Hondel, C.A.M.J.J., Verbeek, S., van der Ende, A., Weisbeek, P.J., Borrias, W.E. & van Arkel, G.A. 1980b. Introduction of transposon Tn901 in a plasmid of *Anacystis nidulans*: preparation for cloning in cyanobacteria. *PNAS.* **77**(3):1570-1574.

van der Plas, J., Hegeman, H., de Vrieze, G., Tuyl, M., Borrias, M. & Weisbeek, P. 1990. Genomic integration system based on pBR322 sequences for the cyanobacterium *Synechococcus* sp. PCC7942: transfer of genes encoding plastocyanin and ferredoxin. *Gene* **95**(1):39-48.

van der Plas, J., Oosterhoff-Teertstra, R., Borrias, M. & Weisbeek, P. 1992. Identification of replication and stability functions in the complete nucleotide sequence of plasmid pUH24 from the cyanobacterium *Synechococcus* sp. PCC 7942. *Molec. Microbiol.* **6**(5):653-664.

Wallace, M., Alaie-Petrillo, A. & Raps, S. 1995. Characterization of a plasmid from the cyanobacterium *Microcystis aeruginosa*. *Vth. Cyanobact. Molec. Biol. Wrkshp.*, Asilomar, CA, p106.

Walton, D.K., Gendel, S.M. & Atherly, A.G. 1992. Nucleotide sequence of the replication region of the *Nostoc* PCC7524 plasmid pDU1. *Nucl. Acids Res.* **20**(17):4660.

Walton, D.K., Gendel, S.M. & Atherly, A.G. 1993. DNA sequence and shuttle vector construction of plasmid pGL3 from *Plectonema boryanum* PCC6306. *Nucl. Acids Res.* **21**(3):746. (abstract)

Weiner, A.M. & Weber, K. 1971. Natural read-through at the UGA termination signal of QB coat protein cistron. *Nature New Biol.* **234**:206-209.

Wickrema, A.W. 1989. Complete nucleotide sequence of a small plasmid from the cyanobacterium *Plectonema boryanum*. Dissertation, p 46-68. Miami University. Oxford, Ohio.

Williams, J.G.K. & Szalay, A.A. 1983. Stable integration of foreign DNA into the chromosome of the cyanobacterium *Synechococcus* R2. *Gene* **24**:37-51.

Wolk, C.P., Vonshak, A., Kehoe, P. & Elhai, J. 1984. Construction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria. *PNAS* **81**:1561-1565.

Yang, X. & Mc Fadden, B.A. 1993. A small plasmid, pCA2.4, from the cyanobacterium *Synechocystis* sp. strain PCC6803 encodes a rep protein and replicated by a rolling circle mechanism. *J. Bacteriol.* **175**(13):3981-3991.

Yang, X. & Mc Fadden, B.A. 1994. The complete DNA sequence and replication analysis of theplasmid pCB2.4 from the cyanobacterium *Synechocystis* PCC6803. *Plasmid* **31**:131-137.

Yasui, A., Takao, M., Oikawa, A., Kiener, A., Walsh, C.T. & Eker, A.P.M. 1988. Cloning and characterization of a photolyase gene from the cyanobacterium *Anacystis nidulans. Nuc. Acids Res.* **16**(10):4447-4463.