Molecular Biology of Desiccation Tolerance in the Cyanobacterium *Nostoc commune*

by

Deborah J. Wright

Thesis submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Master of Science
in
Biochemistry

Malcolm Potts, Committee Chairman
Eugene M. Gregory
Timothy J. Larson

Feb 3, 2004
Blacksburg, VA

Keywords: *Nostoc commune*, desiccation tolerance, cyanobacteria, superoxide dismutase

Copyright 2004, Deborah J. Wright

Molecular Biology of Desiccation Tolerance in the Cyanobacterium *Nostoc commune*

by

Deborah J. Wright

Malcolm Potts, Chairman

Department of Biochemistry

Abstract

The molecular biology of desiccation tolerance was investigated in the cyanobacteria with emphasis on *Nostoc commune*. Analysis of DNA from 41 samples of desiccated *Nostoc* spp. of varied age and global distribution led to the amplification of 43 independent $tRNA^{LEU}$(UAA) group 1 intron sequences. Phylogenetic analysis of the entire data set made it possible to define the form species *Nostoc commune*.

The synthase *(spsA)* and phosphatase *(sppA)* genes required for the synthesis of sucrose were isolated from cyanobacterium *Synechocystis* sp. strain PCC 6803 and overexpressed in *E. coli* in two different vector constructions. Transformants had a marked increased capacity for desiccation tolerance. Sucrose synthesis was confirmed through thin layer chromatography (TLC) analysis of cell extracts from transformants.

Long-term stability of DNA in desiccated *Nostoc* samples was demonstrated by the ability to amplify selected gene loci from samples stored dry for decades. Successful amplification in some samples was possible only after treatment with phenacylthiazolium bromide, a reagent that disrupts covalent cross-links; indicating that the DNA was modified by cross-links that occurred between reducing sugars and the primary amines on the DNA.

Abundant superoxide dismutase was released following rehydration of desiccated field material *N. commune* CHEN after 13 years in the dry state. *sodF* mRNA was present in the dry material but was turned over within 15 min of rehydration. mRNA levels then rose and appeared to reach steady state levels after 3 hours and remained abundant after 24 hours of rehydration.

Acknowledgements

I would like to thank Dr. Malcolm Potts for hiring me in 1995 and allowing me to complete the master's degree that I began years earlier. I appreciate his confidence and encouragement. I am grateful for the camaraderie in the lab with staff members, Sue Smith in particular, since she was of great help in my first few months on the job when I had no idea what I was doing. Collaboration with undergraduate and graduate students in the lab has been a challenge and an inspiration since we always seem to learn from each other.

Thanks also to Dr. E. M. Gregory, Dr. Timothy Larson, and Dr. Thomas Keenan for allowing me to sit in on their undergraduate classes in order to catch up on the field of Biochemistry. BCHM 4115 was the most informative, and most challenging class that I have ever taken and I learned a tremendous amount.

Thanks to my husband and children who, without complaint, ate pizza, Chinese, or carry-out chicken…again. My appreciation also for understanding that Mom was not able to help much with homework since she was hibernating with her own books.

This research was made possible through the support of the Naval Research Laboratory (DARPA) and Office of Naval Research (ONR-MURI).

Table of Contents

Abstract ..... *ii*
Acknowledgements ..... *ii*
Table of Contents ..... *iv*
List of Abbreviations ..... *vi*
List of Tables ..... *vii*
List of Figures ..... *viii*
Introduction *1*
Project overview *8*
Chapter 1 ..... *10*
Characterization of form species Nostoc commune *10*
Abstract **10**
Introduction **10**
Materials and Methods **14**
Biological materials. 14
Isolation of genomic DNA. 17
Pretreatment of modified DNA. 17
Oligonucleotide primers. 18
PCR assay. 18
Subcloning and DNA sequencing. 19
Phylogenetic Analysis. 19
Sequence folding. 20
Accession numbers. 20
Results **20**
Discussion **28**
Acknowledgements **35**
References **36**
Chapter 2 ..... *40*
Cell Stabilization During Desiccation *40*
Abstract **40**
Introduction **40**
Materials and Methods **41**
Bacterial strains, plasmids, and media. 41
Cloning of *spsA/sppA*. 41
Pilot induction and electrophoresis. 43
Growth of strains and expression of *spsA* and *sppA*. 43
Cell extracts. 44
Identification of sucrose and sucrose-6-P in cell extracts. 44
Results **44**

Discussion **48**

References **49**

Chapter 3 .... ***51***

DNA and mRNA Stability in Desiccated Cells ***51***

Abstract **51**

Introduction **51**

Materials and Methods **54**

Cyanobacterial strains and growth conditions. 54
Genomic DNA purification 54
PCR amplification 55
Purification and identification of SodF. 57
Preparation of labeled probes. 57
Transcription analysis. 58

Results **59**

Discussion **64**

References **66**

Discussion ***70***

References ***73***

Appendix ***78***

Curriculum vitae ***88***

List of Abbreviations

Amp: ampicillin
AMV: avian myeloblastosis virus
BLAST: basic local alignment search tool
bp: base pair
CTAB: hexadecyltrimethylammonium bromide
dig-UTP: digoxigenin-11-2'-deoxy-uridine-5'-triphosphate, alkali-labile
dNTP: deoxynucleoside triphosphate
EDTA: ethylenediaminetetraacetic acid
EPS: extracellular polysaccharide
EtOH: ethyl alcohol
g: gravity
Kb: kilobase
kDa: kilodalton
LB: Luria-Bertani medium
MAA: microsporine amino acid
MOPS: 3-(N-morpholino) propane sulfonic acid
m/v: mass to volume ratio
nt: nucleotide
ORF: open reading frame
$P_{BAD:}$ arabinose-induced promoter
PCR: polymerase chain reaction
Pfu: plaque forming units
PTB: *N*-phenacylthiazolium bromide
ROS: reactive oxygen species
RT-PCR: reverse transcriptase - polymerase chain reaction
SDS: sodium dodecyl sulfate
SodF: Fe-containing superoxide dismutase
SpsA: sucrose phosphate synthase
T_m: melting temperature of lipid bilayer
Tris : {tris(hydroxymethyl) aminomethane} hydrochloride
UV: ultra violet
v/v: volume to volume ratio
w/v: weight to volume ratio
Wsp: water stress protein
x: times

List of Tables

Chapter 1.

Table 1.1 Origin of materials and source histories of *Nostoc* strains ………………....15-16

Table 1.2 Published sequences included in phylogenetic analysis……………………..22

Table 1.3 Structural features of *Nostoc* introns…………………………………………23

Chapter 3.

Table 3.1 Primer sequences used in PCR assays……………………………...…..….56

Table 3.2 Treatment of DNA with PTB and ability to amplify gene loci……………...60

Appendix.

Table A.1 Sequence alignment of introns in Group I of the Distance Tree……………78

Table A.2 Sequence alignment of introns in Group II of the Distance Tree…………...79

Table A.3 Sequence alignment of introns in Group III of the Distance Tree…………..81

Table A.4 Sequence alignment of introns in Group IV of the Distance Tree…………..82

List of Figures

Chapter 1

Figure 1.1 Intron PCR products…………………………………………………………...21

Figure 1.2 Structure of group 1 introns…………………………………………………24

Figure 1.3 Distance tree of conserved intron sequences………………………………..27

Chapter 2

Figure 2.1 SDS page illustrating pilot induction of P_{BAD} clones……………………..…45

Figure 2.2 SDS page illustrating SpsA and SppA production induced in *E. coli* BL21-pT7-7 clones……………………………………………………...………………46

Figure 2.3(a) TLC analysis of sucrose and sucrose-6-P production from *E. coli* pT7-7 clones……………………………………………………………………………47

Figure 2.3(b) TLC analysis of sucrose and sucrose-6-P production from *E. coli* P_{BAD} clones……………………………………………………………………………48

Chapter 3

Figure 3.1 PCR amplification of gene loci……………………………………………..61

Figure 3.2 SodF is abundant in desiccated *N. commune* CHEN/1986…………………62

Figure 3.3 Expression of *sodF* in *N. commune* ENG/1996…………………………….....63

Introduction

An actively growing cell requires a large water content to maintain cell structure and enable enzymatic activity. Water allows the cell to maintain order, plays an integral role in many reaction mechanisms in living cells, and stabilizes DNA, proteins and lipids. Removal of water from the cell is an event that the majority of organisms are unable to tolerate. Many organisms have developed a tolerance to desiccation: photosynthetic mosses and cyanobacteria, microscopic invertebrates such as nematodes and tardigrades, yeast cells and fungal spores, plant seeds and pollen. The study of these organisms may reveal the secret of their tolerance. Along with recent advances in medical research and treatment utilizing tissue engineering and genetic technologies, there is a tremendous interest in preserving human cells and cellular products such as enzymes and vaccines. The study of anhydrobiotic organisms is providing clues leading toward the ultimate goal of ambient stabilization of air-dried mammalian cells.

As a first line of defense, the membrane is of primary importance in preserving cell integrity and its proper maintenance is vital to survival. In most instances, it is not desiccation that is fatal to cells, but the event of rewetting, whereby the phospholipid membrane becomes transiently leaky. Many organisms capable of surviving desiccation accumulate a high concentration, as much as 20% of the dry cell weight, of the non-reducing sugar trehalose (reviewed by Crowe *et al*., 1998). Studies have shown that trehalose prevents damage due to dehydration by helping to maintain the membrane lipids in a liquid state by lowering the membrane transition temperature (T_m). In this 'water replacement hypothesis' trehalose and some other disaccharides are thought to intersperse between the polar headgroups of the membrane phospholipids, conserving the liquid crystalline lattice (Crowe *et al*., 1984). The strength of interaction between phospholipid bilayers and numerous sugars including inositol, glucose, sucrose, maltose, and trehalose is directly related to the sugar's ability to stabilize dry liposomes. Trehalose and maltose had the greatest interaction (Crowe *et al*., 1987). It is also thought that trehalose and other sugars protect cells through their ability to form glasses, thereby preventing fusion between neighboring membranes during drying (Crowe *et al*., 1992). When liposomes were dried in the presence of either hydroxyethyl starch or glucose, neither compound alone was able to protect the liposomes from leakage; a combination

of both, however, succeeded (Crowe *et al*., 1997). Glucose lowers the T_m through interaction with the polar headgroups whereas hydroxyethyl starch is a good glass former. Disaccharides fulfill both the requirement for lowering of T_m and for glass formation. The reduction of T_m in some organisms is insufficient to bring it below ambient storage conditions so that the membrane is in the gel phase at room temperature. These organisms, including *Saccharomyces cerevisiae*, require rehydration using warm water or alternatively, initial exposure to moisture in a vapor phase to allow the membrane to reenter the liquid crystalline state before the addition of water (Oliver *et al*., 1998). Even low concentrations of exogenously added trehalose increased the viability of *Escherichia coli* following 5 days storage in a vacuum desiccator (Welsh & Herbert, 1999). *E. coli* produces large quantities of trehalose in response to osmotic stress. Osmotic induction of trehalose in Gram-negative bacteria successfully improved desiccation tolerance, particularly when the internal moisture content of the cell was less than 10% (de Castro *et al*., 2000; Welsh & Herbert, 1999). Osmotically-stressed cells had higher survival rates after desiccation than non-stressed cells after both had been dried in the presence of external trehalose. Both intracellular and extracellular trehalose are thus required to maintain the integrity of the membrane (Tunnacliffe *et al*., 2001). Externally added trehalose or sucrose improved the viability of *E. coli* and *Bacillus thuringiensis* cells after freeze-drying. However, as the cells became transiently leaky upon entering the phase transition, disaccharide moved down the concentration gradient and into the cells and conferred protection (Leslie *et al*., 1995). The T_m of the membrane of these cells without the externally added trehalose increased 30 to 40 C^o while the T_m of those dried with trehalose remained near that of the hydrated cells. In yeast cells a trehalose transporter is expressed concurrently with trehalose synthesis, enabling the delivery of trehalose to the exterior of the membrane. The requirement for this transport to enable desiccation survival was shown by the failure of a transporter-deficient mutant to withstand drying even though it produced trehalose. A strain unable to synthesize trehalose regained its ability to survive drying upon the addition of extracellular trehalose only when the transporter was active, showing the need for the trehalose to enter the cells (Crowe *et al*., 1993). The use of a α-hemolysin pore protein to load trehalose into human fibroblasts was successful in improving the survival of cryopreserved cells (Eroglu *et al*.,

2000) and increasing membrane integrity as monitored by increased osmotolerance after convective drying. Cell viability after rehydration could not be demonstrated (Chen *et al*., 2001).

'Anhydrobiotic engineering' (de Castro, 2000) describes attempts to create desiccation tolerance in a desiccation-sensitive cell by conferring on it the ability to synthesize trehalose. The trehalose biosynthetic enzymes *otsA* and *otsB* from *E. coli* were expressed in mammalian cells to confer desiccation tolerance (Guo *et al*., 2000). However, desiccation tolerance varied dependent on the procedures employed for cell drying. In a recent trial no desiccation tolerance of mouse fibroblast or human kidney cells was observed following induction of trehalose to 80-100 mM intracellularly with or without externally added trehalose (Tunnacliffe *et al*., 2001). Attempts are being made to increase trehalose production in the cell through the inhibition of the enzyme trehalase and by creating a fusion protein of the *otsA* and *otsB* genes. A five-fold increase in trehalose production was achieved through the cloning of the complete operon containing the trehalose biosynthetic complex (reviewed in Crowe *et al*., 2001).

A new method for loading human platelets with trehalose followed by freeze-drying has been developed. This method shows great promise for the stable storage of these vital blood components whose shelf-life is only 5 days (Wolkers *et al*., 2001). Human mesenchymal stem cells incubated in 50 mM trehalose with 3% glycerol for 24 hours, air-dried, and stored under vacuum at ambient temperature for two days retained their spindle-shaped morphology and adhesive capability (Gordon *et al*., 2001). The major problem with this study is the inconsistency between replicate trials and even between neighboring wells in the same culture plate. It is apparent that the conditions for desiccation, storage, and rehydration must be carefully controlled in order to preserve cell integrity.

Other non-reducing disaccharides besides trehalose have been used to stabilize membranes during drying. A genetic engineering method utilizing a cyanobacterial *spsA* gene has been used to induce the synthesis of sucrose-6-phosphate enabling a 10,000-fold increase in the survival of *E. coli* cells following desiccation (Billi *et al*., 2000). As an initial step toward the production of stable desiccated human cells for use in solid-state sensor applications, a combination of genetic engineering and exogenously added

polysaccharide material isolated from the desiccation tolerant cyanobacterium *Nostoc commune* DRH1 enabled the survival of 10% of the human 293H kidney cells desiccated to an internal water content of 6-9% for 8 days (Bloom *et al*., 2001).

Small successes have been achieved, but it is becoming apparent from this research that loading with trehalose or sucrose alone is not sufficient. An optimized protocol leading to the maintenance of desiccated fibroblasts for 8 days at room temperature utilized the addition of a combination of trehalose and glycerol to adherent cells followed by thermal shock to promote entry of the additives into the cells. When these dried human fibroblasts were stored under vacuum, cell viability increased, indicating the possibility that it was the exposure of the cells to the reactive oxygen species (ROS) in the atmosphere that caused significant cellular damage; damage which was prevented by the elimination of air during storage (Puhlev *et al*., 2001).

DNA breakage, the accumulation of carbonyl derivatives in proteins, and lipid damage are all caused by free-radicals due to partially reduced and activated forms of O_2. This damage is exacerbated during drying when the normal repair mechanisms are inoperative. Long-term desiccation can result in damage to proteins and nucleic acids, which can be accelerated through exposure to the reactive oxygen species produced by exposure to high light intensities and UV irradiation. Condensation products between primary amines on proteins and nucleic acids and the carbonyl groups of reducing sugars result in the accumulation of covalent cross-links, producing a dark-brown pigmentation; designated as the browning or Maillard, reaction (Potts, 1999). Anhydrobiotes thus have a need for an increased efficiency of their antioxidant defenses. The importance of the free-radical scavenging systems is demonstrated in the aptly named "resurrection plants" in which the genes encoding enzymatic antioxidants such as superoxide dismutase (SOD) and glutathione reductase are up-regulated during desiccation (Reviewed in Oliver *et al*., 2001). Active SOD was measured in *Nostoc commune* CHEN/1986 after 13 years of storage in the dry state (Shirkey *et al*., 2000). In cyanobacteria, the removal of toxic oxygen species such as singlet oxygen and triplet chlorophyll as well as inhibition of lipid peroxidation was mediated by an increase of the carotenoid to chlorophyll *a* ratio (Ehling-Schultz & Scherer, 1999). The alginate found in the cell wall of *Pseudomonas aeruginosa* as well as the extracellular polysaccharide of *N. commune* restricted the

diffusion of atmospheric oxygen, offering a degree of protection from oxidative injury (Potts, 2000). Damage due to ROS can also be limited through the coordinated reduction of the metabolic activities of the cell or by the reduction of the intracellular O_2 content by oxygen-binding proteins (see Oliver *et al*., 2001 for references). Free-radical-induced damage due to exposure to fluorescent lights decreased the length of time that dried fibroblasts could be maintained in the desiccated state (Puhlev *et al*., 2001).

Members of the cyanobacteria that thrive on exposed limestone surfaces contain numerous UV-absorbing pigments that enable their survival under extreme conditions of UV exposure. One family of pigments is the microsporine amino acids (MAA) that contain chromophores that are linked to sugars such as galactose, glucose, xylose, and glucosamine (Potts, 2000). The level of protection afforded by these pigments is dependent on their cellular location; cytoplasmic pigments absorb only 10-26% of the incident photons, while the pigment located in the extracellular glycan of *Nostoc commune* absorbs 60 –70% (Böhm *et al*., 1995).

The mechanisms by which organisms survive extreme water deficit have, for the most part, remained a mystery. Only through the study of these extreme organisms may we understand the methods that enable them to survive under conditions of severe stress. There is a small group of diverse organisms, found on all continents from the Tropics to the Poles, which is able to withstand continual cycles of drying and rehydration as well as the removal of water for long periods of time. Under the extremely challenging growth conditions of deserts, intertidal marine mats, and terrestrial crusts, cyanobacteria are the dominant bacterial species (Potts, 1999). These extremophiles can remain in a dried state for centuries and regain metabolic functions rapidly and systematically upon the addition of water (Potts, 1994). A cyanobacterium, *Nostoc commune,* was selected as a model organism to study desiccation tolerance. This terrestrial microorganism is found on every continent, inhabiting nutrient-poor soils and the exposed surfaces of limestone where rainwater collects and evaporates rapidly, creating the continual cycle of desiccation and rehydration that *N. commune* tolerates easily (Potts, 1996). Desiccated colonies of these cells can typically absorb water up to 200 times their dry weight within one hour after rehydration (Hill *et al*., 1994a). Metabolic activity is recovered rapidly following desiccation, with nearly instantaneous recovery of lipid biosynthesis upon rewetting,

indicating that the enzymes required for biosynthesis, modification and turnover of lipids remain active in desiccated cells (Taranto *et al*., 1993). The synthesis of ATP, glycogen and mRNA begin immediately after the addition of water and a measurable photosynthetic rate is seen within an hour, increasing to a maximum over the following 10-24 hours. Nitrogenase activity is slower to recover, generally taking a day or more to recover fully (Potts, 1996 and Table 1 in Dodds & Gudder, 1995). Non-polar extracts of desiccated field material *N. commune* contain both sucrose [0.8 mg g^{-1} (dry weight) of cells] and trehalose [1 mg g^{-1} (dry weight) of cells]. Extractable levels of trehalose decrease after 30 minutes of rehydration, coinciding with the appearance of a putative trehalase (Hill *et al*., 1994b).

Cells of *N. commune* produce a secreted polysaccharide that forms a protective matrix to help in the regulation of water loss, aiding in survival under extreme water deficit. This polysaccharide prevents the fusion of membrane vesicles during drying (Hill *et al*., 1997). In the dry state this extracellular polysaccharide (EPS) sheath material forms a glassy mechanical solid causing the bacterial colony to have a brittle texture that initially becomes leathery upon the addition of water. The bacterial mat then achieves the consistency of a semirigid malleable gel (Potts, 1994). Cultures irradiated with UV-B yielded quantities of the EPS three times in excess of the control cultures. It was suggested that the increase in EPS not only serves to lengthen the effective pathlength of incident photons but also serves as a matrix to immobilize the increased quantities of UV-A/B-absorbing pigments which are synthesized under UV stress (Ehling-Schultz & Scherer, 1999), including MAA's and scytonemin, the yellow-brown, lipid-soluble pigment common in, and unique to, terrestrial cyanobacteria. A brown *Nostoc* species isolated in Java was resistant to high light and UV irradiation while the brown pigment and a number of other UV-absorbing pigments could not be isolated from another strain grown under a 24-hour cycle of only 4 light hours and 20 dark hours (Potts, 2000). MAA's are the water-soluble class of photoprotective pigments and can constitute up to 10% of the dry weight of desiccated colonies (Potts, 1994). The structure of the novel EPS has been elucidated. The presence of 3-O-lactyl glucuronic acid (nosturonic acid) is thought to act as a linker in the adherence of the EPS to surfaces and as a functional group enabling the immobilization of secreted proteins and both classes of UV absorbing

pigments that help protect against damage during adverse conditions (Helm *et al*., 2000). Compounds such as the MAA's that the cells synthesize and that are released upon rehydration must serve an important role for the cell since the cyclic loss of these compounds constitutes a severe drain on cellular resources (Potts, 1994.) Three very stable acidic proteins, the water stress proteins (Wsp), are also released from desiccated field material during rehydration. These proteins appear to be isoforms. Wsp are homologs of carbohydrate modifying proteins, and copurify with an associated 1,4-β-D-xylanxylanohydrolase activity that is inhibited by anti-Wsp specific antibodies (Hill, 1994). The function of the Wsp polypeptides is unknown. They were originally thought to have a structural role relating to cell stability but since the identification of ribose in the structure of the EPS are now thought to perform some modification role of the carbohydrate comprising the EPS or may participate in the transport or processing of the MAA's. The ribose moiety of the EPS is a pendant group attached to a xylopyranose unit that appears to act as a protecting group on the xylogalactoglucan backbone (Huang *et al*., 2000). Laboratory-grown strains of cyanobacteria do not ordinarily synthesize Wsp, but their synthesis can be induced in response to desiccation (Scherer & Potts, 1989, and Hill *et al*., 1994a).

No stress imposed upon a living organism elicits a single response or an invariable reaction. Cellular reserves are mobilized by a cascade as seen with the *E. coli* heat-shock sigma factor (σ^{32}) and are dependent on the immediate status of the cell and resources available to it. It is likely that these organisms that exhibit tolerance to desiccation employ some methods that are unique to their physiological composition, which simply cannot be conferred on other cells. It is in hope of detecting some ubiquitous system, however, that the study of anhydrobiotes continues. Despite an accumulation of data on physiology, structural biology and biochemistry, there is a paucity of information on molecular biology and genetics for *N. commune.* This study is an attempt to employ *N. commune* as a model system to elucidate the specific genetic adaptations that make this organism capable of colonizing inhospitable envrionments subject to continual cycles of desiccation and rehydration.

Project overview

This study involves an ancient organism that is ubiquitous in location but little studied or understood. The organism chosen for this study, *Nostoc commune* DRH1, was collected as field sample CHEN in 1986 in a desert region of Mongolia and renamed when cultured as a laboratory strain (Hill *et al*., 1994b). An attempt was made to characterize this strain by determining the scope of dispersion and fidelity of the genotype through comparison to a wide range of field samples of *Nostoc commune* and related organisms collected from all continents over a period of 150 years (Wright *et al*., 2001).

As a desiccation-tolerant cyanobacterium, *N. commune* can withstand decades or centuries in the air-dry state followed by the immediate resumption of metabolic activity upon rewetting. Several known or suspected means by which this desiccation tolerance is achieved were studied. One of these involves the stabilization of cell membranes with the nonreducing disaccharide sucrose. For this study, the sucrose phosphate synthase gene (*spsA*) was isolated from the cyanobacterium *Synechocystis* sp. strain PCC6803. Further work demonstrated that *spsA* conferred desiccation tolerance in *E. coli* (Billi *et al*., 2000). Since it is the presence of internal sucrose, not sucrose-6-phosphate, that has been shown to confer desiccation tolerance, the sucrose phosphate phosphatase gene (*sppA*) was cloned in tandem with *spsA* in both a pT7-7 and Ara-BAD *myc* His (Invitrogen) constructions. Increased quantities of sucrose were synthesized, especially for the Ara-BAD clone. This clone may be used in further studies to confer desiccation tolerance on other desiccation-sensitive organisms.

DNA stability following long-term dehydration was examined through the ability to amplify several genes from *N. commune* samples that were desiccated for up to 150 years. In many cases, successful PCR amplification of the selected genes (*sodF, phr, rrn,* and $tRNA^{LEU}$(UAA)) was dependent on pretreatment with *N*-phenacylthiazolium bromide (PTB) which was employed to break cross-links between reducing sugars and amino groups. Such cross-links are a consequence of the so-called browning or Maillard reaction (Shirkey *et al*., 2003).

Superoxide dismutase (SOD) was identified as an abundant protein in desiccated *Nostoc commune* field material following rehydration. Its mechanisms for protection

against oxidative damage may cause it to be an important factor in the resistance of prokaryotes to desiccation. To examine this possibility, a study was performed following the expression of *sodF* during rehydration of field material (Shirkey *et al*., 2000).

Chapter 1

Characterization of form species *Nostoc commune*

Abstract

The form species concept for the Cyanobacteria was evaluated using a comprehensive set of *Nostoc* samples that were collected during the past two centuries, from all continents, including regions from the Tropics to the Poles. Phylogenies were constructed based upon the conserved regions of $tRNA^{LEU}(UAA)$ group I intron DNA sequences. Thirty-four forms contained a $tRNA^{LEU}(UAA)$ intron of 284 nt. These 284-nt introns contained 200 nt of conserved sequence that in most cases shared 100% sequence identity, they had three variable regions (I, II and III) totaling 84 nt, contained no hypervariable region, and formed a discrete cluster in phylogenetic analysis. These forms represented 31 independent populations in both hemispheres and constitute examples of form species *Nostoc commune*. Multiple introns were obtained from several of the populations. Ten populations contained introns of 287 to 340 nt with an hypervariable region, 8 to 59 nt in length, located between variable regions I and II. Shared morphologies of genetically-different strains, dissimilar morphologies in strains which share identical genetic markers, incorrect naming of culture collection strains, and genetic drift in cultured strains, emphasize that the successful delineation of cyanobacterial species requires the application of multiple taxonomic criteria.

Introduction

The organism chosen for this research was identified as *Nostoc commune* when it was collected originally as field sample CHEN in 1986. The organism has a cosmopolitan habitat, occupying a niche by growing in all climates from the Poles to theTropics. Identification is usually based on colony morphology, completely independent of the actual genotype. Since there is much debate concerning the criteria that should be used for taxonomic identification, it was necessary to perform an analysis using a genetic marker to construct a phylogenetic tree to confirm the initial identification based solely on morphology.

The concept of a prokaryotic species and an understanding of the influence and consequence of environmental selection pressure on prokaryotic diversity and evolution remain intractable issues. Considerations of these issues are especially exacerbated when dealing with the Cyanobacteria (Castenholz, 1992). There are three taxonomic schemes for cyanobacteria in current use: *sensu* Geitler (Geitler, 1932), Stanier (see Rippka *et al*., 1979), and Komárek and Anagnostidis (Komárek & Anagnostidis, 1989). All three are phenetic schemes that place different emphasis on aspects of the morphology, biochemistry, physiology and DNA homologies of field communities and pure cultures. Depending on the context of usage each scheme has its own shortcomings and advantages (Whitton & Potts, 2000).

The application of molecular tools may resolve much of the controversy over cyanobacterial taxonomy, evolution and species concept (Wilmotte, 1994). Recent approaches include: DNA-DNA hybridization (Wilmotte *et al*., 1997), fingerprinting based upon PCR assay with primers from short and long tandemly repeated elements (Rasmussen & Svenning, 1998), classification of clone cultures based upon sequences from the variable regions V6, V7 and V8 of 16S rRNA (Rudi *et al*. 1997), amplified ribosomal DNA restriction analysis of the internally transcribed spacer (Scheldeman, 1999), and comparisons of the conserved structural and regulatory domains within divergent 16S rRNA-23S rRNA spacer sequences (Iteman *et al*., 2000).

The roles of gene exchange (lateral transfer), recombination and hyperbradytely (stasis; slow rate of change) in cyanobacterial evolution remain problematic (Castenholz, 1992; Rudi *et al*., 1998; Schopf, 2000). Other important questions are whether molecular studies will confirm the taxa recognized at present (from both bacteriological and classical methods), and how soon it will be possible to use molecular approaches routinely as a rapid means of characterizing cells from natural populations (Whitton & Potts, 2000). Unfortunately, the conclusions derived from such molecular studies, and the various ‘phylogenetic’ trees illustrating similarities based on sequences from particular parts of the genome, often are based on results from strains whose generic and specific names may be incorrect. In most cases these strains have been maintained in culture for decades, have lost morphological and physiological properties of determinative value, and are frequently of unknown origin. This uncertainty restricts the

value of such trees and analyses for comparing possible evolutionary relationships with the taxonomic relationships that are suggested by classical names (Whitton & Potts, 2000).

One feature of cyanobacterial populations is the frequent dominance of a single form with very characteristic morphology. Examples of such morphotypes include *Trichodesmium erythraeum, Aphanothece halophytica, Microcoleus chthonoplastes, Arthrospira sp.* and *Nostoc commune* (Garcia-Pichel *et al*., 1996; Potts, 2000; Scheleman, 1999). Because these cyanobacteria often have a global distribution and can be readily identified in diverse geographic localities, the concept of a 'form species' developed (Castenholz, 1992). The concept raises important questions about dispersal, isolation, selection, and response to environmental conditions. However, the ability to address these problems is compromised by the limitations of cyanobacterial taxonomy, and uncertainty over the spatial and temporal genetic consistency of 'form species' (Whitton & Potts, 2000). In the case of hot spring communities for example, recent studies suggest that their genetic diversity is much more complex than earlier thought (Ward & Castenholz, 2000).

The potential of group I intron analysis as a genetic marker and taxonomic tool for cyanobacteria was evaluated in several studies. Eubacterial group I introns were first identified in cyanobacteria (Kuhsel *et al*., 1990; Xu *et al*., 1990); their distribution within this group appears to be broad and includes a range of morphologically-diverse forms including *Microcystis*, *Scytonema, Gloeobacter* and *Nostoc* (Paquin *et al*., 1997; Paulsrud & Lindblad, 1998; Rudi & Jakobsen, 1999). Group I introns interrupt protein-coding regions and RNA genes in eukaryotes and bacteriophage, and are present in the anticodon loop of tRNA genes (Paquin *et al*., 1999) as well as 23S rRNA in eubacteria (Everett *et al*., 1999). The acquisition, transfer and distribution of introns are of phylogenetic significance and some controversy (e.g. Gilbert & Glynias, 1993). The presence of homologous $tRNA^{LEU}$(UAA) introns in cyanobacteria and plastids was considered evidence of an ancient origin for these sequences (Kuhsel *et al*., 1990; Xu *et al*., 1990). Subsequent studies, however, provided evidence for lateral transfer of $tRNA^{LEU}$(UAA) introns in strains of *Microcystis* and thus a more recent, possibly polyphyletic, origin was proposed (Rudi & Jakobsen, 1997). From an investigation of the diversity of *Nostoc* cyanobionts in lichens it was concluded that the $tRNA^{LEU}$(UAA)

intron can be of great value when examining cyanobacterial diversity (Paulsrud & Lindblad, 1998). A large, variable region in *Nostoc* introns, which corresponded to base 99 to 143 in *Anabaena* sp. PCC 7120 $tRNA^{LEU}$(UAA), exhibited either of two separate themes. The two classes of intron may correlate with different physiological capacities of the cyanobionts, although there were insufficient differences within the conserved regions of the intron to support or reject this proposal. More recently, these authors used $tRNA^{LEU}$(UAA) intron analysis to document geographic patterns of diversity in *Nostoc*-containing lichens (Paulsrud *et al*., 2000).

The genus *Nostoc* is one of five genera in the family Nostocaceae of subgroup 4, section A, of the oxygenic phototrophic bacteria (Holt *et al*., 1993). *Nostoc commune* is a conspicuous component of terrestrial microbial populations worldwide, especially those associated with nutrient poor soils and limestones where growths may achieve macroscopic proportions (Helm *et al*., 2000; Hill *et al*., 1994a,b; Potts, 2000; Shirkey *et al*., 2000). *N. commune* has been collected and studied for at least the past 2000 years (Potts, 1997; Potts, 2000). Numerous samples of *Nostoc* spp. are present in herbaria, many are still viable, and most are accompanied by detailed records of the time and place of collection. There are few other bacteria for which such detailed information is available and with respect to cyanobacterial diversity such collections constitute a valuable resource.

The complete genome sequence of *Anabaena* sp. PCC 7120 became available in Feb, 2000 (unannotated format; available at http://www.kazusa.or.jp/cyano/anabaena/). and that of *Nostoc punctiforme* ATCC 29133 is being annotated (available at http://genome.ornl.gov/microbial/npun/). It is important to understand the phylogenetic relationships between these, and other, forms of ecologically-relevant cyanobacteria. In this context the present study provides the most comprehensive data set to date on forms of *Nostoc*. A diverse and comprehensive set of *Nostoc* samples were used to: a) evaluate the form species concept in cyanobacteria; b) evaluate the utility of group I intron analysis in the taxonomy of *Nostoc*; c) understand the potential influence of geographic factors on intron dispersal and evolution; and d) examine phylogenetic relationships between *Anabaena* and *Nostoc* spp.

Materials and Methods

Biological materials.

Colonies of *Nostoc commune* were collected and stored desiccated, in the dark, until analysis (Table 1.1). Many of the herbarium specimens were obtained in sealed paper envelopes that had not been opened since the time of collection. Cultures of *N. commune* DRH1, *N. commune* UTEX 584, *N. punctiforme* ATCC 29133 and *Anabaena* sp. strain PCC 7120, are maintained in laboratory culture. The latter strain is considered a nomenspecies of *Nostoc* sp. PCC 6705 and 6719 (Rippka, 1988). The designation *Anabaena* sp. PCC 7120 is retained in this study (see discussion).

Table 1.1 Origin of materials and source histories

With the exception of the WH series of samples, reference numbers include a mnemonic of the place of origin and the year the sample was desiccated. Samples obtained from Wien Herbarium, Austria have the source designation WH together with the reference number provided by the herbarium. The NSW series of samples was provided by Stephen Skinner; for *N. punctiforme* ATCC 29133 see Rippka et al., (1979); for *N. commune* UTEX 584 and *N. commune* DRH1 see Potts, (2000). ND, Not Determined

Designation	Reference/strain no.	Origin	Source	Year of collection	Genebank # or Ref.
Strains (in culture)					
Nostoc punctiforme	ATCC 29133	Macrozamia Australia	ATCC	1973	
Nostoc commune	UTEX 584	UK?	T. Gibson	ND	AF204098
N. commune	DRH1	CHEN/1986	D.R. Hill	ND	AF204090
Anabaena sp.	PCC 7120	USA?	J. Elhai	ND	Paulsrud & Lindblad, 1998
Desiccated samples					
Europe					
N. commune Vaucher	MAL/1989	Malham Tarn, UK	M. Potts	1989	AF204077-AF204079
N. commune Vaucher	BER/1988	Bermatingen, Swizerland	S. Scherer	1988	AF204075
N. commune Vaucher	TAG/1988	Tägerwilen, Swizerland	S. Scherer	1988	AF204102
N. commune Vaucher	WH001/1923	Cojocna, Romania	WH3132	1923	AF204084
N. commune Vaucher	WH002/1939	Cahul, Romania	WH3132	1939	AF204092
N. commune Vaucher	WH003/1941	Caras, Romania	WH3132	1941	AF204105
N. commune	WH004/1849	Heidelberg, Germany	WH2622	1849	AF204093
N. commune	WH005/1859	Konstanz, Germany	WH2622	1859	
N. commune	WH006/1878	Germany	WH2622	1878	AF204096-AF204097
N. commune	WH007/1866	Trieste, Italy	WH2922	1866	AF204099-AF204100
N. pellucidum	WH016/1865	Sesia a Vercelli , Italy	WH670	1865	AF204088
N. alpicola Kützing	WH011/1860	ND	WH2525	1860	
The Americas					
N. commune	WH009/1880	Montevideo, Uruguay	WH2486	~1880	AF204085
N. commune var. *flagelliforme*	WH015/1878	Zacatecas, Mexico	WH2407	1878	AF204087
N. commune var. *flagelliforme*	WH008/ND	San Pedro, TX, USA	WH2525	ND	AF204106

Designation	Reference/strain no.	Origin	Source	Year of collection	Genebank # or Ref.
The Americas (continued)					
N. commune	WH013/ND	Winchester, MA, USA	WH	ND	AF204086
N. commune	WH014/1948	Lawrence Co, IN, USA	WH2009	1948	AF204095
N. commune Vaucher	TOP/1993	Topsail Is., SC, USA	S. Scherer	1993	AF204104
N. commune Vaucher	TEN/1988	Knoxville, TN, USA	W. Herndon	1988	AF204103
N. commune Vaucher	BBC/1990	Blacksburg, VA, USA	M. Potts	1990	AF204074
N. commune Vaucher	BBC/1992	Blacksburg, VA, USA	M. Potts	1992	
N. commune Vaucher	ENG/1996	Blacksburg, VA, USA	M. Potts	1992	AF204076
Antarctica					
N. commune Vaucher	MOA/1997	Mars Oasis, Alexander Is.	A. George	1997	AF204073
N. commune Vaucher	RIS/1979	Ross Ice Shelf	M. Potts	1979	AF204071-AF204072
Nostoc sp.	LBP/1995	Lake Bonney Pond, Dry Valleys	H. Paerl	1995	AF204070
Asia					
N. commune Vaucher	HUN/1982	Chiyang-Hunan. China	S. Scherer	1982	AF204091
N. commune Vaucher	HUN/1986	Hunan, China	W.-Q. Xie	1986	
N. commune Vaucher	CHEN/1986	Hunan, China	S. Scherer	1986	AF204089
N. sphaericum	SPH/1998	health store, China	Z. Huang	1998	AF204101
Australasia					
N. commune	WH012/1880	Surabaya, Java	WH447	1880	AF204094
N. commune Vaucher	MEL/1968	Victoria, Australia	VH698677	1968	AF204083
Nostoc sp.	PORT/1999	New South Wales, Australia	PMC003	1999	
N. commune Vaucher	RNP/1999	New South Wales, Australia	NSW 435425	1999	
N. commune Vaucher	ARM/1999	New South Wales, Australia	NSW 435426	1999	
N. commune var. *flagelliforme*	PARA/1979	New South Wales, Australia	NSW A3124	1979	
N. commune Vaucher	NZE/1997	South Island, New Zealand	P. Broadie	1998	AF204080
Indo-Africa					
N. commune var. lesler Bornet	WH010/1878	Columbo, Sri Lanka	WH2525	1878	
Nostoc sp.	ALD776D/1973	Aldabra Atoll, Indian Ocean	A. Donaldson	1973	AF204064-AF204066
N. commune Vaucher	ALD779D/1973	Aldabra Atoll, Indian Ocean	A. Donaldson	1973	AF204081
N. commune var. *flagelliforme*	ALD857D/1973	Aldabra Atoll, Indian Ocean	A. Donaldson	1973	AF204067-AF204069
Nostoc sp.	ALD8122/1974	Aldabra Atoll, Indian Ocean	M. Potts	1974	AF204082

Isolation of genomic DNA.

Large scale. Desiccated samples were lyophilized and ground to a powder under liquid nitrogen. The powder was rehydrated with TE (1 mM EDTA, 10 mM Tris•HCl, pH 7.5; approximately 1:10, powder:buffer). Extraction buffer (1.4 M NaCl, 20 mM EDTA, 1.5% w/v CTAB, 1% v/v β-mercaptoethanol, 100 mM Tris•HCl, pH 7.5; approximately 4 times the volume of TE added to the powder) was added and the mixture was incubated at 65°C, for 30 min. The slurry was homogenized, frozen under liquid nitrogen, and thawed at 65°C. The freeze-thaw cycle was repeated 6 times. The mixture was extracted with chloroform:isoamyl alcohol (24:1), centrifuged and the aqueous phase was mixed with an equal volume of isopropanol. DNA was recovered and stored in a minimal volume of TE buffer.

Small scale. When limited amounts of material were available the desiccated colonies were ground to a powder under liquid nitrogen. Samples were then added directly to PCR tubes and heated at 95°C for up to 4 hours. Rehydrated cells and/or the cell-free supernatant fractions were used in PCR amplifications (see below).

Pretreatment of modified DNA.

After long-term desiccation the DNA in *Nostoc* cells may become covalently modified through Maillard reactions (Shirkey *et al*, 2000) and this can contribute to subsequent failures in PCR amplification (data not shown). If PCR amplification proved negative in preliminary trials the genomic DNA was treated with *N*-phenacylthiazolium bromide (PTB) before assay. The PTB reagent was synthesized as described (Poinar *et al*., 1998; Vasan *et al*., 1996).

Desiccated colonies were ground to a powder under liquid nitrogen and resuspended in 50 mM EDTA. An equal volume of 1.5% w/v molten agarose (type 1-A; low EEO, Sigma cat. no. A-0169) was added to embed the cell material. The blocks were incubated overnight, at 50°C, with 2 mg lysozyme ml^{-1} (final concentration; Sigma grade 1) in 50 mM EDTA, and then overnight at 50°C, with 2 mg proteinase K ml^{-1} (final concentration; Sigma cat. no. P-6556) in 0.1 M EDTA, 1% w/v *N*-lauroylsarcosine, 0.05% w/v SDS and 10 mM Tris•HCl, pH 8.0. Finally, the blocks were equilibrated in 10 mM sodium phosphate buffer, pH 7.0, prior to the addition of PTB at a final concentration of 10 mM. The blocks were incubated in PTB for 10 hours

at 27°C (room temp) and then equilibrated in TAE buffer (2 mM EDTA, 40 mM Tris•acetate). The blocks were transferred to the wells of a 1% w/v agarose (type 1-A; low EEO, Sigma cat. no. A-0169) gel and the DNA was resolved at 75 mA, for 1 hour. DNA was excised from the gel, and purified further using a QIAEX II gel extraction kit (Qiagen).

Oligonucleotide primers.

Primers for intron amplification were based upon the sequence of tRNA^{LEU} (UAA) of *Anabaena* sp. PCC 7120 (Zaug *et al*., 1993). Primer LEU1 corresponds to a region of the D loop (bases 6 to 28 [$^{5'}$ TGTGGCGGAATGGTAGACGCTAC $^{3'}$]); primers LEU2 and LEU3 were based on the reverse complement of regions within the TψC and variable loops (bases 72 to 56 [$^{5'}$GACTTGAACCCACACGAC $^{3'}$]), and the TψC loop and acceptor stem (bases 82 to 67 [$^{5'}$ GGGTGGAGGGACTTGA $^{3'}$]), respectively.

PCR assay.

Assays were performed in 50 μl reactions, in pH 8.5 buffer (60 mM Tris•HCl, 15 mM $(NH_4)_2SO_4$). Reactions contained 12.5 nmoles of each dNTP (each dNTP at 250 μM), 3.5 mM Mg^{2+} and 2.5 units of *Taq* polymerase (Promega Biotec). Temperature was controlled with a Biometra thermocycler model T-3 (Labrepco). All assays began with a denaturation temperature of 95°C for 2 min, and ended with an elongation time of 10 min; in each cycle denaturation occurred at 95°C, for 1 min, and elongation was at 72°C, for 90 s. The annealing temperature (60 s) was 66°C in the first cycle and decreased 0.2 C^{o} per cycle thereafter for an additional 39 cycles.

The experiments reported here were performed over a period of approximately two years. In view of the resistance of the cells (and their DNA) to air drying, and their probable persistence in aerosols, stringent precautions were used to prevent cross contamination. Conditions for PCR assay were optimized in preliminary trials using genomic DNA from *Anabaena* sp. strain PCC 7120 and *Nostoc commune* DRH1. These DNA's served as positive controls in every assay. Negative controls were used to evaluate possible contamination of the samples, reagents and buffers with occult DNA. Contamination was identified on two occasions. An ambiguity in the published sequence of the *Anabaena* sp. strain PCC 7120 tRNA^{LEU}(UAA) intron ($^{5'}$-R-$^{3'}$) at

position 148; Genbank accession number M38692) was resolved as $^{5'}$-A-$^{3'}$. The group I intron sequence of *Nostoc punctiforme* ATCC 29133 (=*Nostoc* sp. strain PCC 73102) was obtained to compare with the published sequence for control purposes (accession numbers AFO19924, U83254).

Subcloning and DNA sequencing.

A minimum of three separate PCR assays was performed on each sample. Further assays were performed on samples that either provided more than one type of intron sequence or failed to generate a product in initial trials. Reaction mixtures were resolved in agarose gels and putative amplification products were excised from the gel, the agar was crushed in a minimal volume of water, and the DNA was allowed to diffuse passively overnight. The DNA solutions were used directly in ligation reactions with pCR2.1-TOPO (Invitrogen) that were then used to transform *E. coli* TOP10 (Invitrogen). Purified plasmid templates were subjected to DNA sequence analysis using the BigDye Terminator chemistry of Perkin-Elmer Applied Biosystems and *Taq* DNA polymerase, on an Applied Biosystems 377 Prism DNA Sequencer. For each organism, replicate samples of the colony were used; the intron DNA's which derived from three separate PCR amplifications were sequenced completely on both strands, and all six sequences for a given sample were aligned (see below). In only a few cases there were between one and three discrepancies within the aligned DNA sequence of one of the three clones of a given sample. These transversions/transitions were attributed to infidelity in the PCR assay and the consensus sequence of the other two clones was thus reported.

Phylogenetic Analysis.

An alignment was constructed using the conserved secondary structure of introns and the conserved DNA sequence corresponding to bases 1 to 98, 144 to 218, and 223 to 249 in the *Anabaena* sp. PCC 7120 intron (see Paulsrud & Lindblad, 1998). The remaining regions of the introns, referred to as variable regions I, II and III (including an especially poorly resolved "hypervariable" region), and the 34-nt extensions (Table 2) in introns ALD776DA/1973, ALD776DB/1973 and ALD8122/1973, were not used in phylogenetic analysis. DNA sequences were aligned using MEGALIGN within version 4.0 of the LaserGene software (DNASTAR).

Phylogenetic trees were constructed based upon parsimony analysis (ordinary parsimony) and distance methods using the Phylogenetic Inference Package (PHYLIP, version 3.573c) obtained from J. Felsenstein, Department of Genetics, University of Washington, Seattle. SEQBOOT was used to produce 100 data sets from bootstrap resampling (Felsenstein, 1985). Majority rule strict consensus analysis was performed with CONSENSE. In view of its lack of distinguishing morphological characteristics (in comparison to the other strains studied) *Anabaena* sp. PCC 7120 was arbitrarily designated as the outgroup form. Distance matrices were obtained with the Kimura two-parameter model using the default transition/transversion ratio (Kimura, 1980) and calculated with the DNADIST and NEIGHBOR (Saitou & Nei, 1987) programs of PHYLIP (Felsenstein, 1985). Unrooted trees were plotted in DRAWTREE or DRAWGRAM and edited in Adobe Illustrator version 9.0.

Sequence folding.

Intron sequences were analyzed with reference to the *Anabaena* sp. PCC 7120 structure (Chech *et al*., 1994). The structures of sequences that were not readily alignable were predicted using the *mfold* software version 3.1 available online at http://www.bioinfo.rpi.edu/applications/mfold/ (Zuker *et al*., 1999).

Accession numbers.

Intron sequences were deposited in Genbank under the accession numbers AF204064 to AF204106. These may be accessed at

http://www.ncbi.nlm.nih.gov/entrez/query.fcgi.

Results

PCR amplification of tRNA^{LEU}(UAA) group I introns.

Primers based upon regions flanking the intron of the tRNA^{LEU}(UAA) gene of *Anabaena* sp. strain PCC 7120 amplified consistently a single discrete fragment of DNA from the different desiccated samples in PCR assay (Fig 1.1). As expected, the two different combinations of primers generated fragments which differed in size by approximately 10 bp. Use of the primer combination that included LEUI3 (corresponding to a more conserved region of the *Anabaena* sp. strain PCC 7120 tRNA^{LEU}(UAA) intron; see Methods), resulted in more efficient amplification of the

Nostoc intron sequences (Fig. 1.1). Preliminary sequence analysis identified the products as amplified group I intron sequences with portions of the flanking $tRNA^{LEU}$(UAA) exons. The position of insertion of the intron in the different $tRNA^{LEU}$ genes was in each case between U and A within the anticodon loop (conserved $^{5'}$-GGACUU↓AA$^{3'}$). This is the same position identified for other cyanobacterial group I introns in previous studies (Paquin *et al*., 1997)

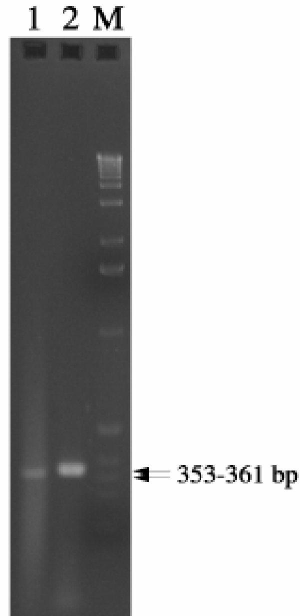


Fig. 1.1. Intron PCR products.

Sets of oligonucleotide primers based upon regions of $tRNA^{LEU}$(UAA) from Anabaena sp. PCC 7120 amplify a single product from *N. commune* WH014 desiccated for 50 years. DNA was purified using PTB reagent. Agarose gel stained with ethidium bromide. Lanes: 1, primers LEU1 and LEU2; 2, primers LEU1 and LEU3; M, 1 kbp ladder marker (LTI, inc.).

Three of the 41 desiccated samples, WH005, WH010 and WH011, failed to generate an amplification product in PCR assay in multiple trials, despite exhaustive purification of their DNA's, including treatments with PTB which was effective with other recalcitrant samples. The same genomic DNA's from WH005, WH010 and WH011 did, however, generate products of expected size when PCR assay was performed with primers designed to amplify portions of two other genes (*phr* and 23S rDNA) identified in *N. commune* DRH1 (see Table 3.2). Of the 38 desiccated samples that provided amplification products, a total of 45 independent intron sequences were obtained for analysis, *i.e.* some samples generated more than one size of fragment in repeat trials (see below). Introns were also amplified from liquid cultures of *N.*

commune DRH1, *N. commune* UTEX 584, *N. punctiforme* ATCC 29133, and *Anabaena* sp. strain PCC 7120. Published sequences from strains identified as *Anabaena* or *Nostoc* were included in the analysis (Table 1.2) to provide a total data set of 71 sequences.

Table 1.2 Published intron sequences that were included in phylogenetic analysis.

Organism	Genebank #	Reference
Nostoc sp. A	AF019912	Paulsrud and Lindblad (1998)
Nostoc sp. B	AF019913	"
Nostoc sp. C	AF019914	"
Nostoc sp. D	AF019915	"
Nostoc sp. E	AF019916	"
Nostoc sp. F	AF019917	"
Nostoc sp. G	AF019918	"
Nostoc sp. H	AF019919	"
Nostoc sp. I	AF019920	"
Nostoc sp. J	AF019921	"
Nostoc sp. K	AF019922	"
Nostoc sp. Nephroma resupinatum	AF055660	Paulsrud *et al*. (2000)
Nostoc sp. Cc1	AF095779	Costa *et al*. (1999)
Nostoc sp. Ev6a	AF095776	"
Anabaena sp. strain NIVA-CYA	AJ228705	Rudi and Jakobsen (1997)
Nostoc sp.	AJ228709	"
Nostoc flagelliforme	AJ228710	"
Nostoc commune . NIVA-CYA 308	AJ228712	"

Structure of conserved and variable regions in tRNA^{LEU}(UAA) group I introns.

Intron sequences were ordered in several provisional subgroupings according to intron size, DNA sequence homologies and presence or absence of hypervariable regions (Table 1.2). The majority (34) of the amplified *Nostoc* tRNA^{LEU}(UAA) group I intron sequences were of identical size (284 bp) and shared very high, or identical, sequence similarity (Table A-4). The sequence similarity between one given 284-bp sequence, and any one of the other thirty-three 284-bp sequences, varied from 95.8 to 100%, and it was thus possible to align these sequences without ambiguity. Sixty-seven of the introns had a conserved sequence of 200bp. In ALD776DA/1973, ALD776DB/1973, ALD8122/1973 and LBP/1995, the conserved region was 201bp. In addition, ALD776DA/1973, ALD776DB/1973 and ALD8122/1973 each contained a 34-bp extension between the first conserved region and variable region I.

Table 1.3. Structural features of *Nostoc* introns
Values are lengths in nucleotides. For further details, see Fig 1.2(b).

Designation	Intron	Conserved	Variable	Hypervariable
N. commune form species	284	200	84	—
ALD857DA/1973	298	200	65	33
ALD857DB/1973	298	200	65	33
WH9/1880	340	200	81	59
TOP/1993	340	200	81	59
ALD776DA/1973	325	201	91	33
ALD776DB/1973	325	201	91	33
ALD8122/1973	325	201	91	33
MALA/1989	300	200	66	34
MALB/1989	300	200	66	34
PARA/1979	287	200	60	27

All introns conformed to a primary structure with three conserved regions and two variable regions. The latter were subdivided into variable region I, a hypervariable region and variable region II, and variable region III (Fig. 1.2b). Variable regions III, and parts of I and II for many samples were readily aligned and secondary structure predictions suggest that many of the bases in variable region I, the hypervariable region and variable region II may be base-paired (Fig. 1.2a).

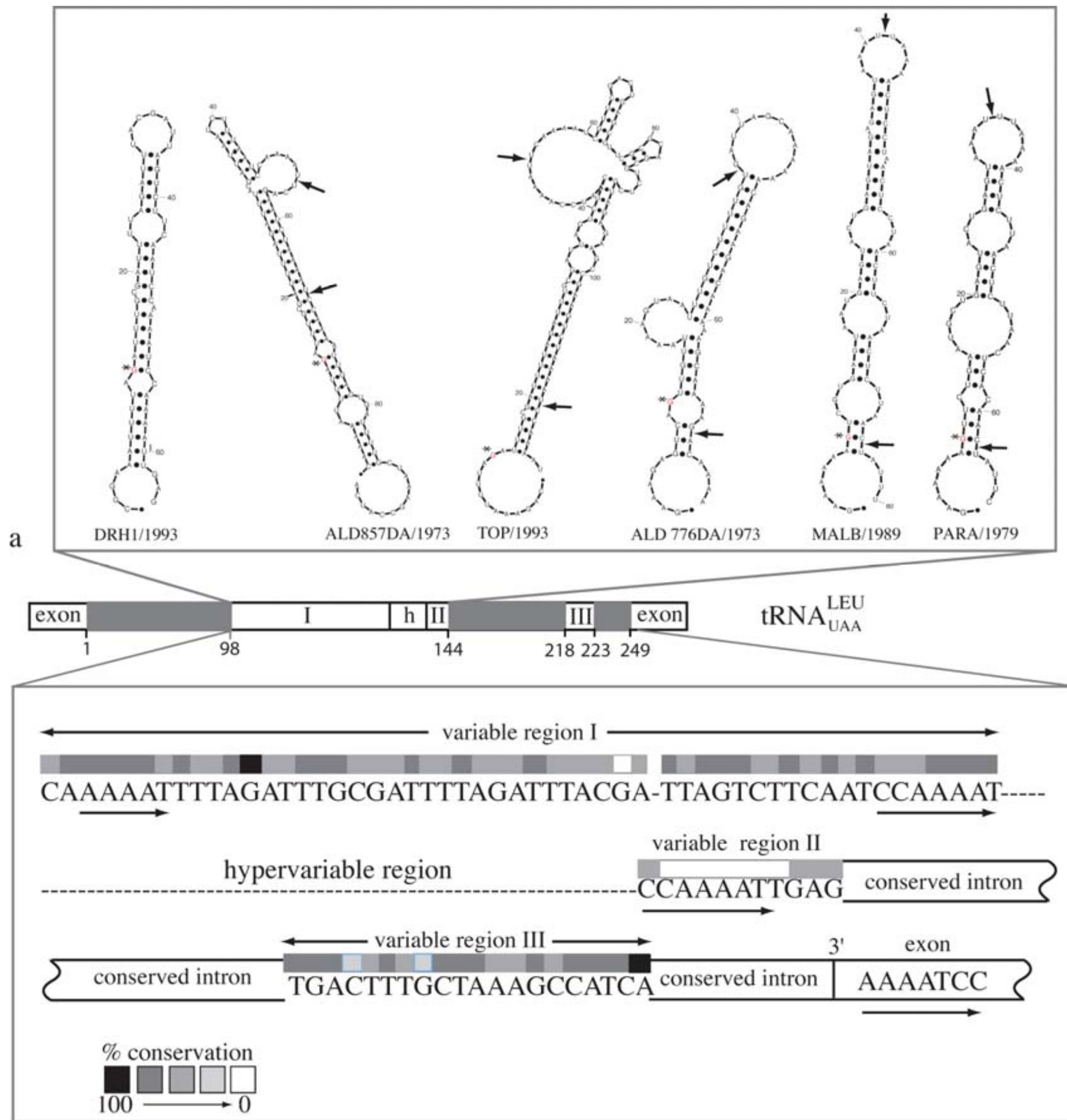


Fig. 1.2 Structure of group I introns.

(a) Secondary structure predictions of variable region 1, hypervariable region (h) and variable region II. Arrows indicate the 5' and 3' ends of the hypervariable region. The G residue conserved in region 1 of all introns analyzed is printed in red and indicated by asterisks. **(b)** Consensus sequence of the variable regions. Arrows under bases indicate sequence repeats or similarities.

Multiple introns from single samples.

Samples ALD776D, ALD857D, MAL/1989, WH006, WH007 and RIS/1979 provided multiple intron sequences. Introns ALD776DA/1973 and ALD776DB/1973 differ at two nucleotide positions (in conserved regions) and form part of group I (Fig. 1.3). They differ from ALD776DC/1973 (group IV) through 2-, 3-, 4-, 5-, 6- and 11-nt indels within the variable regions. Similarly, ALD857DA/1973 and ALD857DB/1973 differ from one another at three nucleotide positions (in conserved regions), and from ALD857DC/1973 through a 14-nt indel; these three introns cluster in group IV. MALA/1989 and MALB/1989 (group III) differ at a single nucleotide, and from MALC/1989 (group IV) through multiple indels within variable regions. RISA/1979 and RISB/1979, WH6A/1878 and WH6B/1878, and WH7A/1866 and WH7B/1866 (all in group IV) differ at six, eight and three nucleotide positions, respectively, across conserved and variable regions.

Phylogeny of *Nostoc* and *Anabaena.*

Distance trees of the conserved intron sequences were constructed using a neighbor-joining algorithm. The robustness of each tree was tested using bootstrap analysis. Maximum parsimony analysis was also used to build consensus trees.

All of the conserved sequences that corresponded with the provisional grouping of 284-bp introns (Table 1.2), with the single exception of WH2/1939, formed a single cluster characterized by little to no divergence (Fig. 1.3; cluster IV). This cluster included two samples (WH008, WH015) identified originally as *N. commune* var. *flagelliforme*, one sample (SPH/1998) identified as *N. sphaericum* and one sample (WH016) identified as *N. pellucidum*. The relative ordering of sequences in cluster IV varies upon resampling because many sequences share 100% homology and cannot be discriminated from one another. *Nostoc* samples TOP/1993, WH009 and NZE/1997 have the classic colony morphology of *N. commune* and were otherwise indistinguishable from the samples represented by cluster IV. However, they were excluded from cluster IV on the basis of maximum parsimony analysis (data not shown) and appear in cluster II with sequences from forms that enter into symbiotic associations. Four samples (WH015, WH008, PARA/1979 and ALD857D) were identified as *N. commune* var *flagelliforme*. The morphology of this form consists of bundles of hair-

like colonies. Three of the samples were included in cluster IV while *Nostoc* PARA/1979 (from Australia) grouped in cluster III with an independent isolate from China, AJ228710, which was also named as *N. commune* var *flagelliforme* (Rudi *et al*., 1998).

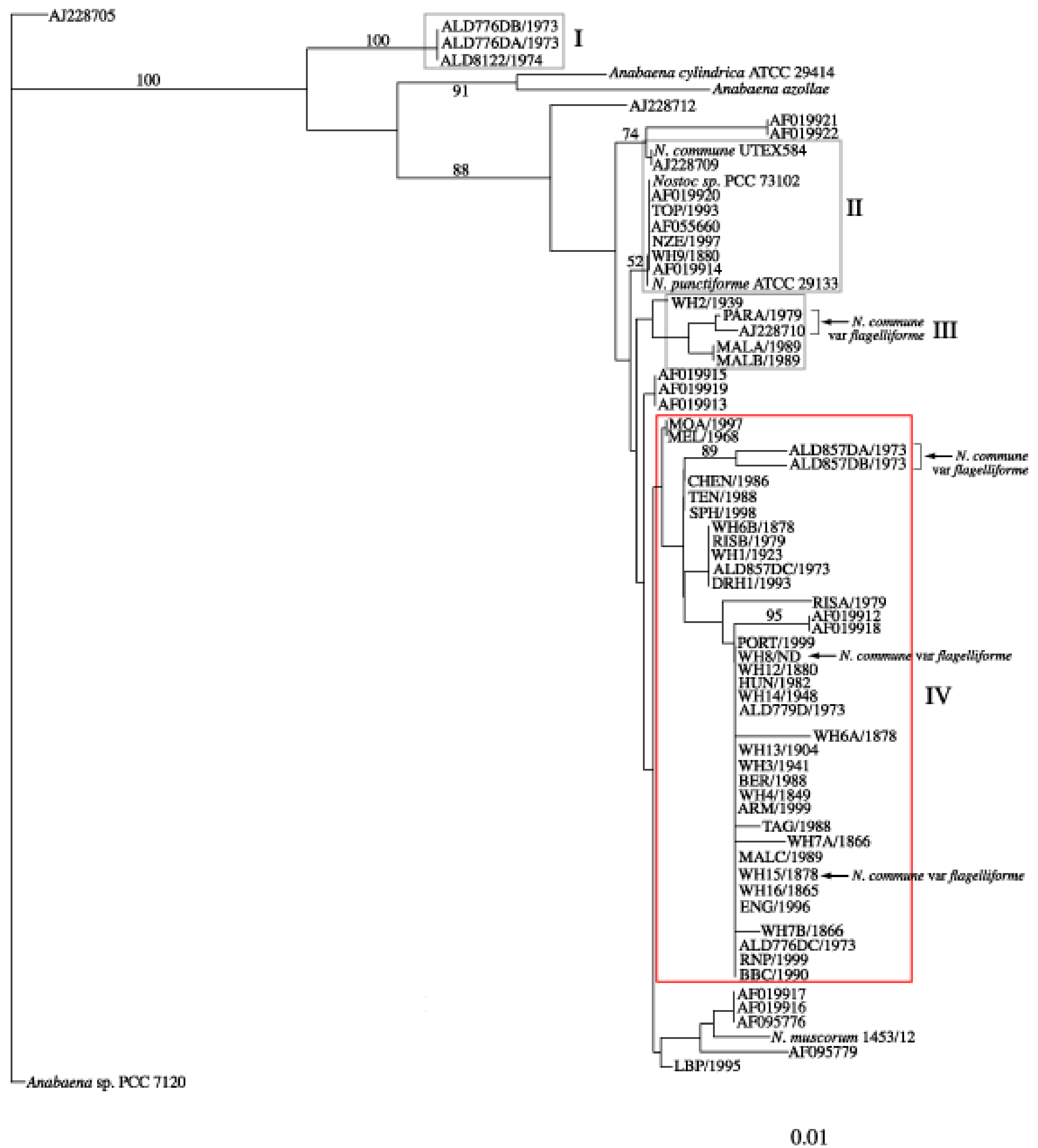


Fig. 1.3. Distance tree of conserved intron sequences.

The tree (unrooted) was inferred using distance matrices obtained with Kimura's two-parameter model (Kimura, 1980), calculated with the DNADIST and NEIGHBOR programs of PHYLIP (Felsenstein, 1985). Bootstrap values were deduced from 100 replicates using BOOTSEQ (see Methods). The distance between two strains, expressed in substitutions per nucleotide, is obtained by adding the lengths of the horizontal branches connecting them. Intron designations consist of the sample number, a letter (if necessary) to indicate an independent amplified product/year the material was desiccated. Accession numbers refer to GenBank. *Nostoc* PCC 73102 = *N. punctiforme* ATCC 29133. Boxes and labels (I-IV) demarcate major clusters.

Discussion

The growing realization of the potential breadth and extent of microbial diversity has prompted a renewed interest in the classification of prokaryotes and bacterial taxonomy. In studies that question phylogenies and evolutionary relationships the concept of a bacterial 'species' is a long-standing issue that remains a topic of controversy and debate (Pace, 1997; Stencil, 2000). A satisfactory resolution of this issue, and a clearer understanding of the genetic basis for phenotypic plasticity, will depend on the identification of defined microbial populations and their interrogation with reliable genetic markers. Cyanobacterial populations have a number of advantages for such analyses: they have a cosmopolitan distribution; they dominate the microbial populations of many extreme environments; their colonies are often conspicuous and may assume macrophytic proportions; colonies of what appear to be morphologically-identical forms occur in geographically-isolated environments; and growths, often attributed to a 'mono-specific' population, may cover many square kilometers. Cyanobacteria are believed to be ancient organisms (3.5 billion years) based upon the evidence of microfossils from Western Australia (Awramik et al, 1983; Schopf and Walter, 1983; Schopf and Packer, 1987). A recent report brings the initial research on cyanobacterial origins into question however, especially that of Schopf and Packer (1987). The new information describes the self-assembly of inorganic materials into filaments that closely resemble the microfossil evidence (García-Ruiz et al, 2003).

Of those genetic markers that may resolve clusters and groupings within taxa below the level of genus, group I intron sequences have been investigated in a broad range of eubacterial taxa (Biniszkiewicz *et al*., 1994; Paquin *et al*., 1997; Paquin *et al*., 1999; Rudi & Jakobsen, 1997; Edgell *et al*., 2000). In this study we focused on a large set of desiccated samples of free-living *Nostoc* for which detailed records of time and place of origin, and taxonomic assignment, were available. The major objectives were to test the form species concept as it pertains to cyanobacteria, to assess the utility of group I intron analysis as a determinative tool, to resolve possible relationships within the *Anabaena*-*Nostoc* lineage, and to identify possible correlations between time of collection, geographic distribution, endemic isolation and intron sequence divergence.

Form species *Nostoc commune*.

Texts such as Geitler (1932) emphasize the use of morphological features as determinative criteria for cyanobacterial taxa. Colonies of *N. commune* are cosmopolitan and occur as free-living, blackened, friable crusts when dry, and as green, lobate masses with a cartilaginous to gelatinous texture when rehydrated (see Fig. 1b in Potts, 2000). The colony morphology is sufficiently characteristic that phycologists and microbiologists assigned the name *N. commune* with confidence, the taxonomic problems of the Cyanobacteria not withstanding (see Whitton & Potts, 2000). Thirty of the samples used in this study were collected in 13 different countries, in both hemispheres, and were identified as *N. commune* by independent investigators over a period of some 200 years. Twenty-five of the forms assigned this species name each contained a 284-bp tRNA^{LEU}(UAA) group I intron that lacked a hypervariable region; collectively these sequences were characterized by little to no divergence and formed cluster IV of the distance tree (Fig. 1.3). The remaining five samples had indistinguishable colony morphology but were excluded from cluster IV in phylogenetic analysis. On the basis of these data the phylogenetic grouping represented by cluster IV was attributed provisionally to form species *N. commune*.

Reliability of morphological criteria.

The samples studied included those with designations of *Nostoc* sp. (four samples), *N. sphaericum* (one sample), *N. pellucidum* (one sample), and *N. commune* var. *flagelliforme* (four samples). In the case of *N. sphaericum* and *N. commune* var. *flagelliforme* the name designations were based, in large part, upon colony morphology. Colonies of the former cyanobacterium exist as gelatinous spheres when growing in standing water (Potts, 2000), while the desiccated thread-like and twine-like bundles of the latter are equally distinctive (the colloquial term for this cyanobacterium in China translates as "Hair-vegetable"; Potts, 1997). *N. sphaericum* (SPH/1998) and *N. commune* var. *flagelliforme* (WH015) were attributed to form species *N. commune* through intron analysis. Several of the samples of *N. commune* we have isolated in clonal culture can produce spheres or "pearls" according to the growth conditions (Hill *et al*., 1994a; unpublished data) which suggests that the use of this character as a taxonomic marker is unreliable. Whether the presence of a characteristic hair-like

morphology justifies assignment to *N. commune* var. *flagelliforme* is harder to resolve. Three of the four samples grouped within cluster IV. The intron of the most divergent of the four samples (PARA/1979; collected in 1979 from Australia) was 94.5% homologous with that of an independent isolate from China. Lateral intron transfer may be responsible for the assignment of WH015 and WH008 to cluster IV but it is also possible that a characteristic growth habit (bundles of hair-like filaments) can be achieved by genotypically-different forms of *Nostoc.* The latter possibility was discussed in regard to spherical colony morphology (above) and it can be emphasized further by noting that TOP/1993, with the classic *N. commune* lobate colony structure and cell morphology, is divergent from cluster IV on the basis of intron phylogeny.

A strain from China, identified as *N. commune* (in Rudi *et al*., 1998), contained an intron that differed at only a single nucleotide from the intron sequence of *N. commune* UTEX 584 determined in this study. *N. commune* UTEX 584 is of unknown origin and the naming of this strain is considered erroneous (Potts, 2000). Unlike form species *N. commune*, [*N. commune*] strain UTEX 584 has a conspicuous life cycle, a pronounced production of the secondary metabolite geosmin, lacks water stress protein *wsp*, and contains at least two genes (*iphP, glbN*) that are absent in *N. commune* DRH1 (Potts, 2000). The *nif* genes of [*N. commune*] UTEX 584 *and N. punctiforme* ATCC 29133 (cluster II, IIa/IIb) are also more similar to one another than to those of *Anabaena* sp. strain PCC 7120 or any other cyanobacterium (Thiel, T., pers. comm.). It appears that strains from the University of Texas Culture Collection (UTEX) were donated to the Hydrobiological Institute of Wuhan, China, from where the strains (in Rudi *et al*., 1998) were acquired (Huang Z., pers. comm.). This example emphasizes the extreme caution that must be taken when interpreting phylogenetic relationships based upon names assigned to culture collection strains (Whitton & Potts, 2000).

There is a lack of a clear understanding of the diagnostic traits that can be used to discriminate between a strain of *Anabaena*, and a strain of *Nostoc.* Based upon phylogenetic analysis of conserved regions, and irrespective of the sequence designated as the outgroup, the group I intron of *Anabaena* sp. PCC 7120 was the most divergent of the 71 sequences used in this study. In the taxonomic scheme of Stanier (see Rippka, 1988, Rippka *et al*., 1979) the production of hormogonia (motile filaments), if produced,

is considered a good discriminatory character. The maturation of an hormogonium into a mature filament was described in detail for a number of *Nostoc* species with complex life cycles (see Potts, 2000). Largely on the basis of DNA-DNA hybridization data, *Anabaena* sp. PCC 7120 is considered a nomenspecies of *Nostoc* sp. PCC 6705 and 6719, even though it produces no hormogonia. Hormogonium formation, as a negative character, is unreliable (Rippka, 1988). For example, strains derived from *Nostoc commune* TEN/1988, BBC/1990 and TAG/1988 failed to generate hormogonia in liquid culture but the parent materials did so when incubated on solid media. The complete genome sequence data of *Anabaena* sp. PCC 7120 and *Nostoc punctiforme* ATCC 29133 provide a means to resolve this taxonomic problem.

Absence of group I introns.

Three of the oldest samples, WH005 (1859), WH010 (1880) and WH011 (1860), failed to generate any amplification products in PCR assay. The samples were received with the names *N. commune*, *N. alpicola* Kützing and *N. commune* var lesler Bornet, respectively. It is unlikely that the tRNA sequences of these strains (i.e. the primer hybridization sites) were sufficiently different from those of other *Nostoc* strains to have prevented amplification using the primers based upon *Anabaena* sp. strain PCC 7120 $tRNA^{LEU}$(UAA). The latter is the most divergent strain in the study set yet the primers allowed amplification from a diverse collection of samples. Even if introns were absent in these materials PCR assay should have amplified part of the $tRNA^{LEU}$ (UAA) genes. Given the age of the samples, the quality of the DNA may have prevented amplification, although introns were recovered from even older samples where PTB reagent proved effective. The small amounts of material available prevented Southern analysis that may have resolved the issue.

Geographical correlations.

The samples used in this study are not old enough (on a geological time scale) to permit a rigorous appraisal of intron evolution in dispersed populations of the same form species. However the samples are old from the perspective of free-living populations of bacteria; many of the samples pre-date the Industrial Revolution and the onset of the widespread use of herbicides, pesticides and other xenobiotics which have poorly-understood effects on gene conversion and migration in ecosystems. The older samples

also pre-date the widespread use of air travel, and increased frequency of travel to both populated and remote regions. Availability of inocula, transport, and maintenance in aerosols are important, often overlooked, considerations in studies of microbial populations, especially those which tolerate extreme desiccation. Some cyanobacteria, and also eukaryotic algae, are sufficiently desiccation-tolerant to survive long-distance transport in aerosols even over Antarctica (Marshall & Chalmers, 1997). Viable cells of the non-heterocystous cyanobacterium *Microcoleus* were sampled from aerosols following nuclear bomb detonations (in Kraus, 1969). These are important considerations because there does seem to be some restriction on the global occurrence of cyanobacterial populations that may be circumvented by import of samples. Some species of *Synechococcus* are clearly restricted in geographic distribution. For example, all forms of thermophilic *Synechococcus* are absent from Icelandic hot springs although numerous springs exist there that appear chemically suitable (Ward & Castenholz, 2000).

The samples used for analysis included a spectrum of longitude and latitude. If there is any correlation between sequence variation within the different phylogenetic trees and geographic distributions of *Nostoc*, it is not an apparent one. For example, introns ENG/1996 and BBC/1990, from independent populations growing approximately 4 km from one another in Virginia, share 100% homology, and 98.9% homology with TEN/1988 from a population sampled from the roof of a schoolhouse several hundred km south, in Tennessee. However intron WH012, from *N. commune* collected from Surabaya, Java, (118 years previous and more than 15,000 kilometers distant), is 100% identical to ENG/1996 and BBC/1990. The significance of very small differences in sequence similarity is not clear although Paulsrud *et al*. (2000) were able to use small differences in the variable I regions of $tRNA^{LEU}$(UAA) introns to compare geographical distribution patterns of *Nostoc* cyanobionts.

TOP/1993 was excluded from form species *N. commune* in all phylogenies that were constructed. Despite the characteristic colony morphology, these populations grow on coastal sand subject to sea spray; a feature that is atypical considering the alkaline limestone environments typically colonized by *N. commune*. (Potts, 2000).

Multiple introns.

In samples that contained multiple introns, the different forms were either homologous, with only a few nucleotide substitutions e.g. MALA/1989 and MALB/1989 or, they differed more significantly and in addition had one or more indels e.g. compare ALD776DA/1973 and ALD776DB/1973 with ALD776DC/1973. In the former situation the introns clustered in the same phylogenetic grouping and these data may reflect cases of intron evolution (in colonies that are stable over time). For the latter situation, the differences were sufficient to discriminate the introns into different phylogenetic groupings which may suggest mixed populations (see Rudi and Jakobsen, 1999).

Introns in remote populations.

The question of intron stability in remote populations is an important one and it has some bearing on the controversy over ancient versus recent origins of group I introns. Aldabra (9° S) is an uninhabited, remote coral atoll situated in the extreme southwest corner of the Indian Ocean (see Potts, 2000 for refs). Visually conspicuous colonies of *Nostoc* are scattered over the limestone surface and shallow soils of the atoll, and become locally abundant in small depressions and solution-holes that fill partially with water during the wet season (WNW monsoon – lasting from December to April). Samples ALD776D and ALD8122 were collected independently, in 1973 and 1974, from pool W10 on West Island. At the time of collection these samples were identified as *Nostoc* sp. (*sensu* Geitler; Geitler, 1932). Sample ALD779D was collected from Bassin Cabri, a pool several hundred meters from W10, and was identified as *N. commune* upon collection. Sample ALD857D was collected from "platin" (flat limestone pavement) in the vicinity of W10 and Bassin Cabri, and was termed "crumbly *Nostoc*", and identified as *N. commune* +/- var. *flagelliforme* at the time of collection. The populations of these three environments are distinguished on the basis of $tRNA^{LEU}$(UAA) group I intron analysis. The group I intron sequences of ALD776D and ALD8122 are identical (with a 33-nt hypervariable region), which may suggest the dominant population in pool W10 was "stable" and monospecific at least over the course of two wet seasons and one dry season. In contrast, the intron of ALD779D has the signature of the 284-nt group I intron (no hypervariable region) of form species *N. commune*. The hypervariable regions of introns from ALD857D and ALD8122 and

ALD776D are very different. Given that there is no shortage of inoculum in any part of the atoll, the persistence of discrete populations, in very close proximity to one another, is significant.

Age correlations.

Nostoc commune strain DRH1 was brought into liquid culture in 1993 using CHEN/1986 as the primary source of inoculum (Hill *et al.*, 1994a). The intron sequences of these materials, determined in 1998, share 97.9% homology. Do the five nt changes (1 in the conserved region and 4 in the variable regions) reflect mutations derived from several years of liquid culture or, does CHEN/1986 contain more than one form (and intron)? Variations due to PCR amplification are discounted on the basis of the precautions taken to sample multiple times from each material for PCR assay and to sequence the cloned introns on both strands. For the same reason, the possibility of multiple forms of intron is unlikely; for other samples where more than one population of introns was in fact present, these were differentiated easily during PCR assay and subcloning. It is assumed, therefore, that the nt substitutions reflect genetic drift in the laboratory culture. It can be noted that this example is in contrast to the data for strains *Nostoc punctiforme* ATCC 29133 and *Nostoc* sp. PCC 73102. These strains have identical intron sequences, share a common origin, and have been maintained for many decades in separate culture collections.

Hypervariable regions and intron evolution.

There is disagreement whether tRNA^{LEU}(UAA) introns have an ancient or recent origin, and whether they are stable or mobile (Besendahl *et al*., 2000; Edgell *et al*., 2000; Paquin *et al*., 1997; Paquin *et al*., 1999; Rudi and Jakobsen 1997, 1999). In fact, current opinion is that there is no single reason that can explain adequately the current distribution of group I and II introns in bacteria (Edgell *et al*., 2000). Recent work suggests also that the evolution of the intron in the tRNA^{LEU}(UAA) gene is considerably more complex than previously thought (Rudi & Jakobsen, 1999). It is interesting that all of the forms assigned to form species *N. commune*, representing a wide range of geographically isolated communities, have a group I intron of 284 nt that has no hypervariable region (Fig. 2a). The 284-nt group I intron of cluster IV appears to have

withstood acquisition of a hypervariable region, be it from invasion by mobile elements and/or duplication events.

The data confirm and emphasize the complexity of intron evolution and suggest that an intron phylogeny alone cannot be used with confidence to infer cyanobacterial species or strain relationships. The reasons for this include, but are not limited to, the morphological similarity of genetically-different forms, morphological dissimilarities between forms that share identical genetic markers, misnaming of culture collection strains, and possible genetic drift in strains maintained in continuous culture. Examples of all of these drawbacks were encountered during the study and they may pose greater obstacles to the successful definition of other cyanobacterial species that lack characteristic morphological criteria. Note, however, that most of the samples used in this study were never brought into culture and were assigned to *N. commune* based solely on colony morphology *in situ*; significantly, the majority of these showed little to no divergence for independent genetic markers.

Summary.

Populations that constitute global form species *Nostoc commune* were identified through group I intron analysis. In many cases the phylogenetic analysis supported identifications based solely upon morphological criteria. In contrast, for several samples the analysis brought into question the reliability of certain morphological criteria. Information on the degree of genetic (genomic) diversity within the form species, and within the *Nostoc/Anabaena* lineage in general, is only fragmentary at this time. The phylogenetic tree based upon intron sequence analysis (Fig. 1.3) gives the impression of a continuum, with ill-defined borders, punctuated by clusters and groups.

Acknowledgements

I would like to thank Dr. Richard Helm for synthesizing PTB. Special thanks go to Dr. Malcolm Potts for the many hours he spent doing the phylogenetic analysis and revisions.

References

Besendahl, A., Qiu, Y.L., Lee, J.H., Palmer, J.D. & Bhattacharya, D. (2000). The cyanobacterial origin and vertical transmission of the plastid tRNA(Leu) group-I intron. *Curr. Gen.* **37**: 12-23.

Biniszkiewicz, D., Cesnaviciene, E. & Shub, D.A. (1994). Self-splicing group I intron in cyanobacterial initiator methionine tRNA: evidence for lateral transfer of introns in bacteria. *EMBO J* **13**: 4629-4635.

Castenholz, R.W. (1992). Species usage, concept, and evolution in the cyanobacteria (blue-green algae). *J. Phycol.* **28**: 737-745.

Chech, T.R., Damberger, S.H. & Gutell, R.R. (1994). Representation of the secondary and tertiary structure of group I introns. *Nature Structural Biology* **1**: 273-280.

Costa, J.L., Paulsrud, P. & Lindblad, P. (1999). Cyanobiont diversity within coralloid roots of selected cycad species. *FEMS Microbiol. Ecol.* **28**: 85-91.

Edgell, D.R., Belfort, M. & Shub, D. (2000). Barriers to intron promiscuity in bacteria. *J. Bacteriol.* **182**: 5281-5289.

Everett, K.D.E., Kahane, S., Bush, R.M. & Friedman, M.G. (1999). An unspliced group I intron in 23S rRNA links *Chlamydiales*, chloroplasts and mitochondria. *J. Bacteriol.* **181**: 4734-4740.

Felsenstein, J. (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783-791.

Garcia-Pichel, F., Prufert-Bebout, L. & Muyzer, G. (1996). Phenotypic and phylogenetic analyses show *Microcoleus chthonoplastes* to be a cosmopolitan cyanobacterium. *Appl. Environ. Microbiol.* **62**: 3284-3291.

García-Ruiz, J. M., Hyde, S. T., Carnerup, A. M., Christy, A. G., Van Kranendonk, M. J., and Welham, N. J. (2003). Self-Assembled silica-carbonate structures and detection of ancient microfossils. Science. **302**: 1194-1197.

Geitler, L. (1932). Cyanophyceae. In: *Rabenhorst's Kryptogamen-Flora*, Vol. 14. Akademische Verlagsgesellschaft, Leipzig.

Gilbert, W. & Glynias, M. (1993). On the ancient nature of introns. *Gene* **135**: 137-144.

Helm, R.F., Huang, Z., Leeson, H., Edwards, D., Peery, W. & Potts, M. (2000). Characterization of the released polysaccharide of desiccation-tolerant *Nostoc commune* DRH-1. *J. Bacteriol.* **182**: 974-982.

Hill, D.R., Peat, A. & Potts, M. (1994a). Biochemistry and structure of the glycan secreted by desiccation-tolerant *Nostoc commune* (cyanobacteria). *Protoplasma* **182**: 126-148.

Hill, D.R.H., Hladun, S., Scherer, S. & Potts, M. (1994b). Water stress proteins are secreted with UV-A/B-absorbing pigments and associate with a β-D-xylanxylanohydrolase activity. *J. Biol. Chem.* **269**: 7726-7734.

Holt, J.G., Krieg, N.R., Sneath, P.H.A., Staley, J.T., & Williams, S.T. (ed.). (1993). *Bergey's Manual of Determinative Bacteriology*, 9th ed. Williams & Wilkins Co., Baltimore, MD

Iteman, I., Rippka, R., de Marsac, N.T. & Herdman, M. (2000). Comparison of conserved structural and regulatory domains within divergent 16S rRNA-23S rRNA spacer sequences of cyanobacteria. *Microbiol.* **146**: 1275-1286.

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol. Evol.* **16**: 111-120.

Komárek, J. & Anagnostidis, K. (1989). Modern approach to the classification system of cyanophytes. 4-Nostocales. *Arch Hydrobiol Suppl 823, Algological Studies* **56**: 247-345.

Kraus, M. P. (1969). Resistance of blue-green algae to ^{60}Co gamma radiation. *Radiat. Bot.* **9**, 481-489.

Kuhsel, M.G., Strickland, R. & Palmer, J.D. (1990). An ancient group I intron shared by eubacteria and chloroplasts. *Science* **250**: 1570-1573.

Marshall, W.A. & Chalmers, M.O. (1997). Airborne dispersal of antarctic terrestrial algae and cyanobacteria. *Ecography* **20**: 585-594.

Pace, N. (1997). A molecular view of microbial diversity and the biosphere. *Science* **276**: 734-740.

Paquin, B., Kathe, S.D., Nierzwicki-Bauer, S.A. & Shub, D.A. (1997). Origin and evolution of group I introns in cyanobacterial tRNA genes. *J Bacteriol.* **179**: 6798-6806.

Paquin, B., Heinfling, A. & Shub, D.A. (1999). Sporadic distribution of $tRNA^{ARG}{}_{CCU}$ introns among α-purple bacteria: evidence for horizontal transmission and transposition of a group I intron. *J Bacteriol.* **181**: 1049-1053.

Paulsrud, P. & Lindblad, P. (1998). Sequence variation of the $tRNA^{LEU}$ intron as a marker for genetic diversity and specificity of symbiotic cyanobacteria in some lichens. *Appl Environ Microbiol.* **64**: 310-315.

Paulsrud, P., Rikkinen, J. & Lindblad, P. (2000). Spatial patterns of photobiont diversity in some *Nostoc*-containing lichens. *New Phytol.* **146**: 291-299.

Poinar, H.N., Hofreiter, M., Spaulding, W.G., Martin, P.S., Stankiewicz, B.A., Bland, H., Evershed, R.P., Possnert, G. & Pääbo, S. (1998). Molecular coproscopy: Dung and diet of the extinct ground sloth *Nothrotheriops shastensis*. *Science* **281**: 402-406.

Potts, M. (1997). Etymology of the genus name *Nostoc* (Cyanobacteria). *Int. J. Syst. Bacteriol.* **47**: 584.

Potts, M. (2000). *Nostoc*. In *The Ecology of Cyanobacteria,* pp. 465-504. Edited by B.A. Whitton, & M. Potts. Kluwer Academic Publishers, Dordrecht, Netherlands.

Rasmussen, U. & Svenning, M.V. (1998). Fingerprinting of cyanobacteria based on PCR with primers derived from short and long tandemly repeated repetitive sequences. *Appl. Environ. Microbiol.* **64**: 265-272.

Rippka, R. (1988). Recognition and identification of cyanobacteria. *Meth.. Enzymol.* **167**: 28-67.

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

Rudi, K. & Jakobsen, K.S. (1997). Cyanobacterial $tRNA^{LEU}$(UAA) group I introns have polyphyletic origin. *FEMS Microbiol. Lett.* **156**: 293-298.

Rudi, K., Skulberg, O.M., Larsen, F., & Jakobsen, K.S (1997). Strain characterization and classification of oxyphotobacteria in clone cultures on the

basis of 16S rRNA sequences from the variable regions V6, V7, and V8. *Appl. Environ. Microbiol.* **63**: 2593-2599.

Rudi, K., Skulberg, O.M., Larsen, F. & Jakobsen, K.S. (1998). Evolution of cyanobacteria by exchange of genetic material among phyletically related strains. *J. Bacteriol.* **180**: 3453-3461.

Rudi, K. & Jakobsen, K.S. (1999). Complex evolutionary patterns of $tRNA^{Leu}$ UAA group I introns in cyanobacterial radiation. *J. Bacteriol.* **181**: 3445-3451.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**: 406-425.

Scheldeman, P., Baurain, D., Bouhy, R., Scott, M., Muhling, M., Whitton, B.A., Belay, A. & Wilmotte, A. (1999). *Arthrospira* ('Spirulina') strains from four continents are resolved into only two clusters, based on amplified ribosomal DNA restriction analysis of the internally transcribed spacer. *FEMS Microbiol. Lett.* **172**: 213-222.

Schopf, J.W. (2000). The fossil record: tracing the roots of the cyanobacterial lineage. In *The Ecology of Cyanobacteria,* pp. 13-35. Edited by B.A. Whitton, & M. Potts. Kluwer Academic Publishers, Dordrecht, Netherlands.

Shirkey, B., Kovarcik, D.P., Wright, D.J., Wilmoth, G., Prickett, T.F., Helm, R.F., Gregory, E.M. & Potts, M. (2000). Active Fe-SOD and abundant *sodF* mRNA in *Nostoc commune* (Cyanobacteria) after years of desiccation. *J. Bacteriol.* **182**: 189-197.

Stencel, C. (2000). Microbial diversity; eyeing the big picture. *ASM News* **66**: 142-146.

Vasan, S., Zhang, X., Kapurniotu, A., Bernhagen, J., Teichberg, S., Basgen, J., Wagle, D., Shih, D., Terlecky, I., Bucala, R., Cerami, A., Egan, J. & Ulrich, P. (1996). An agent cleaving glucose-derived protein crosslinks *in vitro* and *in vivo. Nature* **382**: 275-278.

Ward, D.M. & Castenholz, R.W. (2000). Cyanobacteria in geothermal habitats. In *The Ecology of Cyanobacteria,* pp. 37-60. Edited by B.A. Whitton, & M. Potts. Kluwer Academic Publishers, Dordrecht, Netherlands.

Whitton, B.A. & Potts, M. (2000). Introduction to the cyanobacteria. In *The Ecology of Cyanobacteria,* pp. 1-11. Edited by B.A. Whitton, & M. Potts. Kluwer Academic Publishers, Dordrecht, Netherlands.

Wilmotte, A. (1994). Molecular evolution and taxonomy of the cyanobacteria. In *The Molecular Biology of Cyanobacteria,* pp. *1-25.* Edited by D.A. Bryant. Kluwer Academic Publishers, Dordrecht, Netherlands.

Wilmotte, A., Stam, W. & Demoulin, V. (1997). Taxonomic study of marine oscillatoriacean strains (Cyanophyceae, Cyanobacteria) with narrow trichomes. III. DNA-DNA hybridization studies and taxonomic conclusions. *Algol. Stud* **87**: 11-28.

Xie, W.-Q., Tice, D. & Potts, M. (1995). Cell water deficit regulates expression of *rpoC1C2* (RNA polymerase) at the level of mRNA in desiccation-tolerant *Nostoc commune* UTEX 584 (cyanobacteria). *FEMS Microbiol Lett* **126**: 159-164.

Xu, M.-Q., Kathe, S.D., Goodrich-Blair, H., Nierzwicki-Bauer, S.A. & Shaub, D.A. (1990). Bacterial origin of a chloroplast intron: conserved self-splicing group I introns in cyanobacteria. *Science* **250**: 1566-1570.

Zaug, A.J., McEvoy, M.M. & Cech, T.R. (1993). Self-splicing of the group I intron from *Anabaena* pre-tRNA:requirement for base-pairing of the exons in the anticodon stem. *Biochem.* **32**: 7946-7953.

Zuker, M., Mathews, D.H. & Turner, D.H. (1999). Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide. In *RNA Biochemistry and Biotechnology*, pp. 11-43. Edited by J. Barciszewski & B.F.C. Clark. NATO ASI Series, Kluwer Academic Publishers.

Chapter 2

Cell Stabilization During Desiccation

Abstract

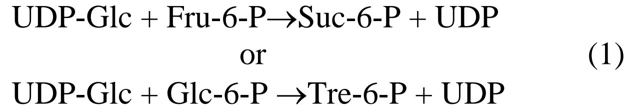
There are two major hypotheses that aim to explain desiccation tolerance, the ‘water replacement hypothesis’ and ‘glass transition theory’ (vitrification). Of the former, the synthesis of trehalose and sucrose are considered to be important. Since only plants and cyanobacteria can synthesize sucrose, the possibility of engineering sucrose synthesis in other desiccation-sensitive organisms was explored. Recombinant sucrose-6-phosphate synthase (SpsA) and sucrose-6-phosphate phosphatase (SppA) were synthesized in *Escherichia coli* using the *spsA* and *sppA* genes of the cyanobacterium *Synechosystis* sp. strain PCC 6803. Overexpression of the genes and production of free sucrose was demonstrated using two different vector constructions.

Introduction

Loss of even a small fraction of intracellular water is lethal for most cells. Nevertheless, some cells, including many microbial pathogens of humans, survive extreme desiccation, often for protracted periods (Billi & Potts, 2000, Potts, 1994). The maximal longevity of microorganisms in a metabolically inactive, desiccated state is unknown (Kennedy *et al*., 1994; Potts, 1994; and Sneath, 1962), but there have been controversial reports of ancient, yet viable bacteria in 25 million- to 40-million-year-old Dominican amber (Gerhardt, 1998; Postgate & Priest, 1995; Priest, 1995; Cano & Borucki, 1995; and Lambert *et al*., 1998).

Understanding the mechanisms that some organisms use to withstand the removal of virtually all of their water is an important problem in cell biology. The nonreducing disaccharides sucrose and trehalose protect membranes and proteins *in vitro* from dehydration damage, as described by the ‘water replacement hypothesis’ (Crowe *et al*., 1984). Survival of dehydration damage in a variety of organisms is correlated with intracellular accumulation of one of these disaccharides, and even the addition of exogenous trehalose or sucrose to cells that are sensitive to drying can increase survival (Leslie, *et al*., 1995; and Potts, 1994).

It was proposed that the ability to survive desiccation may be conferred by transfection of desiccation-sensitive cells with genes which permit synthesis of trehalose or sucrose (Crowe & Crowe, 1992). This transfection seems eminently practical since synthesis of either disaccharide requires only two steps, involves only two gene products, a synthase (reaction 1) and a phosphatase (reaction 2), and requires substrates found in all cells:



(1)

(2)

Materials and Methods

Bacterial strains, plasmids, and media.

E. coli BL21(DE3) and BL21 Star™(DE3) (Invitrogen, Carlsbad, Calif) have a chromosomal copy of the T7 RNA polymerase gene under control of the isopropyl-β-D-thiogalactopyranoside (IPTG)-inducible *lacUV5* promoter. Derivative strain BL21(DE3)(pT7-7) contains expression vector pT7-7, which has a T7 promoter that permits high-level gene expression after induction with IPTG. Strain BL21(DE3)(pSpsA) contains the sucrose-6-phosphate synthase (*spsA*) gene from the cyanobacterium *Synechocystis* sp. strain PCC 6803 cloned in pT7-7 while strain BL21(DE3)(pSpsSppA) contains the sucrose-6-phosphate phosphatase gene cloned in tandem following *spsA*. The pBAD *myc* His plasmid (Invitrogen) contains the AraC regulatory protein and allows expression to be initiated from the *ara*BAD promoter (P_{BAD}) in the presence of arabinose in the growth media. P_{BAD} clones were expressed in *E. coli* strain Top10 (Invitrogen). The strain containing vector only is identified as pBAD while the single and tandem recombinants are denoted as pSpsBAD and pSpsSppBAD, respectively.

Cloning of *spsA/sppA*.

The following two primers were used to amplify *spsA* from genomic DNA of *Synechocystis* sp. strain PCC 6803 (Curatti, 1998; Kaneko, 1996) for cloning into pT7-7 vector (or pBAD *myc* His A): $^{5'}${AGAGCGCAT}**ATG**AGCTATTCATCAAAATACA $^{3'}$ (the translation start codon is in boldface type)(The sequence in {} was replaced with

an *Xho*1 site, CTCGAGA, for subcloning into pBAD *myc* His A) and $^{5'}$ {GAGACGGTCGAC}-**TTA**AACGGGGTCTAACAACTCA $^{3'}$ (the translation stop codon is in boldface type)(The sequence in { } was replaced with a *Pst*1 site, CTGCAG, for subcloning into pBAD *myc* His A). These primers had 5'-terminal *Nde*1 and *Sal*1 sites, respectively, which aided subcloning in pT7-7. *SppA* was amplified using primers which had 5'-terminal *Sal*1 and *Cla*1 sites, respectively, to facilitate cloning in pT7-7 immediately behind *spsA*. The primers for *sppA* were $^{5'}$ {GTCGAC}AAGAAGGAGATATACAT**ATG**CGACAGTTA-TTGCTAATTTCTGA $^{3'}$ (the translation start condon is in boldface and the ribosome binding site is underlined) (sequence in { } replaced with a *Pst1* site, CTGCAG, for cloning into pBAD *myc* His A) and $^{5'}$ {ATCGAT}**TCA**GCTCAAAAAATCGAAATG-GGCGATCG $^{3'}$ (the translation stop codon is in boldface type) (sequence in { } replaced with a *Hind*111 site, AAGCTT, for cloning into pBAD *myc* His A). PCR reactions were performed in 50 µl mixtures containing pH 9.0 buffer (Promega, Madison, Wis.) for pT7-7 or pH 8.5 buffer (60 mM Tris•HCl, 15 mM $(NH_4)_2SO_4$) for P_{BAD} clones, 12.5 nmol of each deoxynucleoside triphosphate (final concentration of each deoxynucleoside triphosphate, 250 µM), 1.5 mM Mg^{2+}, and 2.5 units of *Taq* DNA polymerase (Promega). The annealing temperature used for the first cycle (60 s) was 72°C, and the annealing temperature was then decreased 0.8C° per cycle for the next nine cycles (to 65°C) and was kept constant at 65°C for the remaining 30 cycles. Each cycle included denaturation at 95°C for 1 min and elongation at 72°C for 90 s. The reactions began with denaturation at 95°C for 2 min and ended with elongation for 10 min at 72°C. DNA amplification resulted in 2,184 and 2,181-bp *spsA* products and 764-bp *sppA* products. All DNA products were purified, ligated to pCR2.1 TOPO (Invitrogen) and transformed into *E. coli* TOP10 creating pDWSPS1, pDWSPP1, pDWSpsBAD and pDWSppBAD. DNA sequence analysis was performed to confirm that all genes were complete and accurately amplified. pDWSPS1 was digested with *Nde1*and *Sal1* to remove the *spsA* fragment, which was then cloned into pT7-7 (which had been digested previously with *Nde1* and *Sal1*) to obtain pSpsA. pDWSPP1 was digested with *Sal1* and *Cla1* to release the *sppA* fragment, which was then cloned into pSpsA (which had been digested with *Sal1* and *Cla1*) to produce the tandem clone pSpsSppA. Similarly, pDWSpsBAD was digested with *Xho1* and *Pst1* and

cloned into pBAD *myc* His A to create pSpsBAD; then the *Pst1-Hind11* fragment of pDWSppBAD was ligated following *spsA* to create the tandem clone pSpsSppBAD under the control of P_{BAD}. All pT7-7 clones were expressed in *E. coli* BL21(DE3) or derivative strain BL21 STAR™ (DE3) while the P_{BAD} clones were expressed in Top10 (Invitrogen).

Pilot induction and electrophoresis.

It was necessary to determine the appropriate concentration of arabinose to add to the media for the optimal induction of the synthesis of SpsA/SppA from P_{BAD}. Cultures were grown in 10 ml LB + Amp to an OD_{600} of 0.5 and 10-fold serial dilutions of arabinose added (final concentrations ranged from 0.2% w/v to 0.0002% w/v) and incubation continued at 37°C for 4 hours. Aliquots (1 ml) were collected and cell pellets frozen. Proteins were resolved under denaturing conditions in the presence of sodium dodecyl sulfate (SDS) in 0.75 mm-thick 8% or 12% (SpsA) or 15%(w/v) (SppA) polyacrylamide gels with the use of a Mighty Small SE250/SE260 gel apparatus (Hoeffer) and visualized with Coomassie blue.

Growth of strains and expression of *spsA* and *sppA*.

Strains BL21(DE3), BL21 Star™ (DE3) (Invitrogen) and Top 10 were grown in Luria-Bertani (LB) medium, while strains BL21(DE3)(pT7-7), BL21(DE3)(pSpsA), BL21(DE3)(pSpsSppA), pBAD, pSpsBAD, and pSpsSppBAD were grown in LB medium containing ampicillin (final concentration, 100 μg ml^{-1}). For each experiment single colonies of the different strains were grown in 3 ml of LB medium or LB medium containing 100 μg of ampicillin per ml for 3 h at 37°C. The 3-ml cultures were used to inoculate 200 ml portions of fresh medium (with 100 μg of ampicillin per ml), and incubated to an optical density at 600 nm (OD_{600}) of 0.8 for pT7-7 clones or 0.5 for P_{BAD} clones. P_{BAD} induction was begun with the addition of arabinose (0.002% w/v final concentration) and continued for 5.5 h. Cells for pT7-7 induction were harvested from 200 ml of culture by low speed centrifugation, resuspended in 200 ml of M9 medium (with ampicillin, when needed), and after 0.1 mM (final concentration) IPTG was added, incubated for an additional 8-12 hours. M9 media contained (per liter) 6 g of Na_2HPO_4, 3 g of H_2PO_4, 0.5 g of NaCl, and 1 g of NH_4Cl as the basal salts. The pH of the basal medium was adjusted to 7.4 and the medium was autoclaved, after which 2 ml of $MgSO_4$

(1 M stock solution), 0.1 ml of $CaCl_2$ (1 M stock solution), and 10 ml of glucose (20% w/v, filter sterilized) were added.

Cell extracts.

Cell extracts were prepared by the method of Germer et al. (1998). Cell pellets obtained from 200-ml cultures of BL21(DE3)(pSpsA), BL21(DE3)(pSpsSppA), and BL21(DE3)(pT7-7) after induction in M9 medium (see above) and pBAD, pSpsBAD, and pSpsSppBAD after induction with arabinose were freeze-dried (5 mm of Hg, -75°C) by using a Labconco model-8 system, resuspended in 800 μl of ice-cold 15 mM trichloroacetic acid (TCA), and incubated on ice for 20 min. After centrifugation at 10,000 x *g* for 5 min, the supernatant fraction was recovered for analysis for sucrose-6-P and sucrose.

Identification of sucrose and sucrose-6-P in cell extracts.

The presence of sucrose-6-P and sucrose in cell extracts was detected by thin-layer chromatography (TLC). Two-microliter aliquots of the cell extracts from 200 ml induced cultures were separated on silica gel plates (type G/UV; Whatman International Ltd., Maidstone, Kent, England). The solvent system used was acetonitrile-water-ethyl acetate-isopropyl alcohol-acetic acid (85:20:30:30:10, v/v). The sugars were visualized by dipping the plates in a naphthorescorcinol-ethanol-sulfuric acid solution (200 mg of naphthoresorcinol, 100 ml of 95% [v/v] ethanol, 4 ml of concentrated sulfuric acid) for a few seconds and drying the plates with gentle heating. Solutions of sucrose and sucrose-6-P (3 mM each) were used as standards.

Results

Expression of *spsA* and *sppA* in *E. coli*.

In a pilot induction of P_{BAD} clones 0.002% arabinose (w/v, final concentration) was sufficient for optimal expression of *spsA* from pSpsBAD (Fig. 2.1, lanes 3 and 8, induction of two independent clones). Expression decreased at 0.0002% arabinose (Fig. 2.1, lanes 4 and 9) and was absent in the pBAD vector-only clone (Fig. 2.1, lane 10). SDS PAGE analysis of total protein extracts from cultures of BL21(DE3)(pSpsA) and BL21(DE3)(pSpsSppA) induced with 0.1 mM IPTG for 8 h revealed a prominent, approximately 75-kDa polypeptide (Fig. 2.2a, lanes 1-6). The 75-kDa peptide was not

detected in BL21(DE3)(pT7-7) (Fig. 2.2a, lane 7) under the same conditions, nor was it detected in extracts obtained prior to induction (data not shown). A second protein detected in total protein extracts from cultures of BL21DE3(pSpsSppA) (Fig. 2.2b, lanes 2-6) but not BL21DE3(pT7-7) (Fig. 2.2b, lane 7) or BL21DE3(pSpsA) (Fig. 2.2b, lane 1) had an approximate mass of 25-kDa corresponding to SppA.

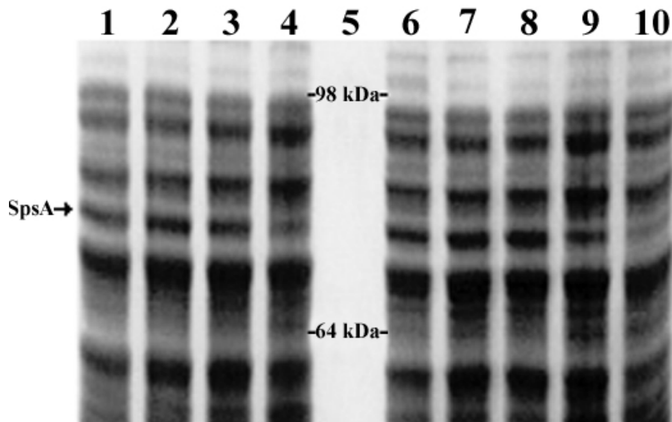
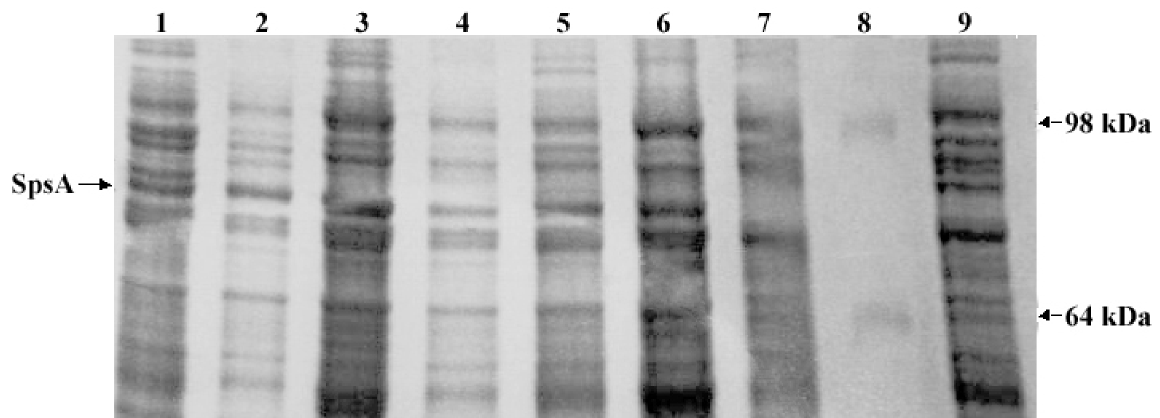


Figure 2.1. Pilot induction of P_{BAD} clones.

Induction of 2 independent pSpsSppBAD clones with decreasing concentrations of arabinose. Lanes 1 and 6, 0.2% arabinose (final concentration); lanes 2 and 7, 0.02% arabinose; lanes 3 and 8, 0.002% arabinose; and lanes 4 and 9 0.0002% arabinose. lane 5, SeeBlue pre-stained marker (Invitrogen); lane 10, pBAD, induced at 0.02% arabinose as a control.



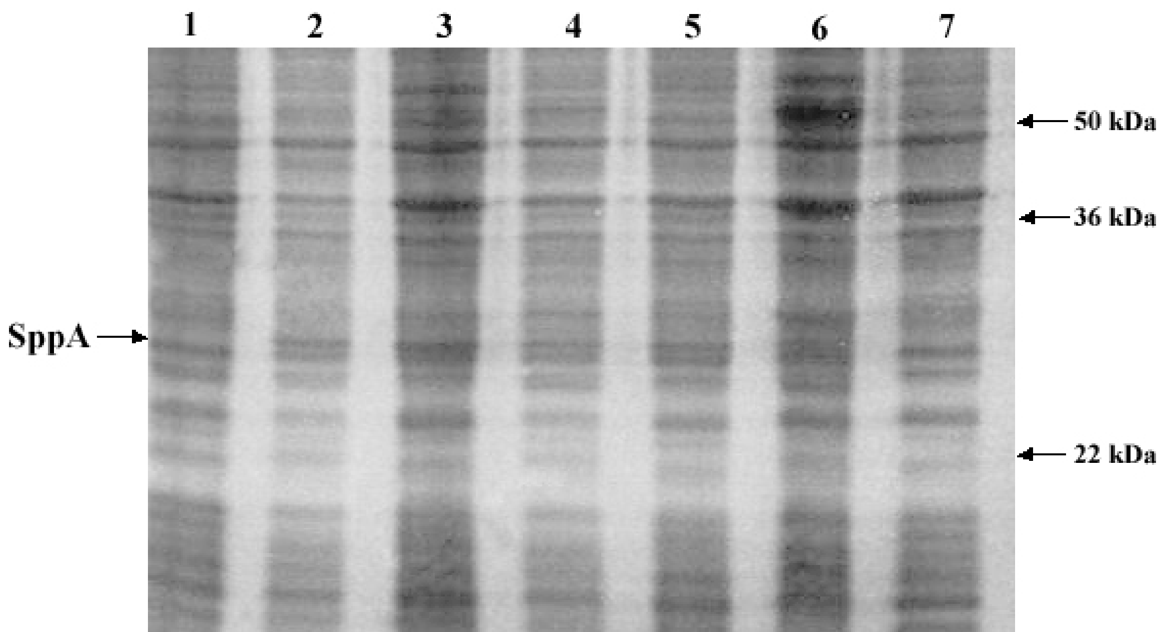


Figure 2.2 SpsA and SppA production induced in *E. coli* BL21 clones.

(a) 8% SDS PAGE gel. Lane 1, single gene clone pSpsA; lanes 2, 4, and 6, 3 independent tandem clones (pSpsSppA) expressed in BL21(DE3) or lanes 3 and 5, in BL21 STAR™(DE3); lane 8, SeeBlue marker (Invitrogen); and lane 7, pT7-7 vector-only induced in BL21(DE3) or, lane 9, in *E. coli* BL21 STAR™(DE3);. **(b)** 15% SDS PAGE gel. Lanes 1 - 7 SppA production in the same clones as in (a) lanes 1 - 7.

Synthesis of sucrose-6-P and sucrose.

The sucrose-6-P standard produced a characteristic comet-shaped band at or just above the origin on TLC plates (Fig. 2.3a, lane E, Fig. 2.3b, lane C). Cell extracts of

BL21D(E3)(pSpsA) induced for SpsA synthesis produced a faint spot (Fig. 2.3a, lane B, arrow) with an R_f value similar to that of the sucrose standard (Fig. 2.3a, lane D). When this extract was incubated with calf intestinal alkaline phosphatase prior to TLC, the intensity of the spot coincident with sucrose increased significantly concomitant with the disappearance of sucrose-6-P (Billi *et al*., 2000). TLC of extracts for three independent tandem clones (pSpsSppA) showed the production of greater quantities of sucrose (Fig. 2.3a, lanes C, F and G). TLC of the P_{BAD} clones after induction showed strong overproduction of sucrose in the tandem clone (pSpsSppBAD) (Fig. 2.3b, lane D). In the single clone (pSpsBAD), there appears to be a faint pink spot coincident with sucrose (Fig. 2.3b, lane B), but it is nearly masked by the presence of the blue signal due to arabinose. pBAD vector only, induced under the same conditions, showed no indication of sucrose (Fig. 2.3b, lane A).

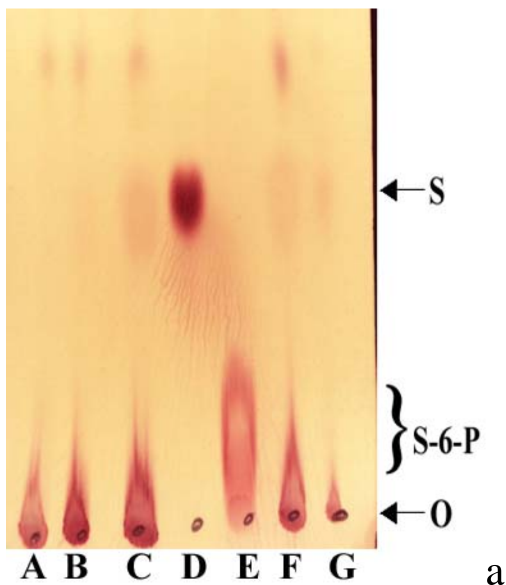
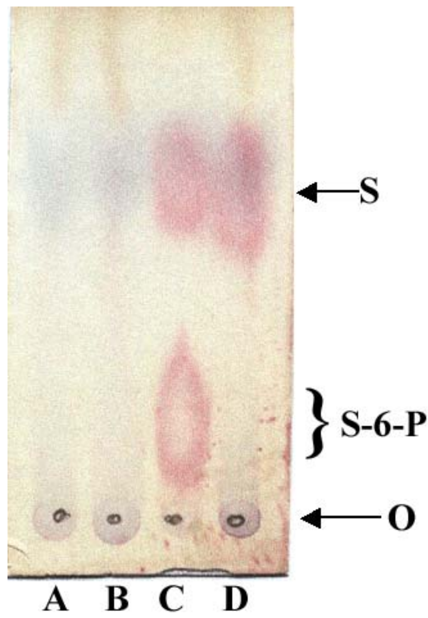


Figure 2.3. Carbohydrate analysis by thin layer chromatography.

(a) The production of sucrose 6-P and sucrose in *E. coli.* BL21(DE3)(pSpsA) and BL21(DE3)(pSpsSppA) clones, as determined by TLC analysis of cell extracts and carbohydrate standards. Lanes A, B, C, F, and G; 2 µl aliquots of 200 µl cell extracts obtained from 200-ml M9 medium-induced cultures of BL21(DE3)(pT7-7), (lane A), BL21(DE3)(pSpsA) (lane B), and three independent BL21(DE3)(pSpsSppA) clones (lanes C, F, and G); lane D, 2 µl of sucrose (3 mg ml^{-1}); and lane E, 2 µl of sucrose-6-P (3 mg ml^{-1})



b

(b). Sucrose production by the *E. coli* P_{BAD} clones induced with 0.002% (final concentration) arabinose for 5.5 h. Lane A, pBAD only; lane B, pSpsBAD; and lane D, pSpsSppBAD. Lane C, 2 µl of sucrose-6-P and sucrose standards (3 mg ml^{-1} each). S, sucrose; S-6-P, sucrose 6-phosphate; O, origin.

Discussion

Only low levels of intracellular sucrose, calculated at 2.3×10^7 molecules per cell in one experiment (Billi *et al*, 2000), are produced from clones containing the individually cloned *spsA* gene (pSpsA (pT7-7) and pSpsBAD), The tandem clone is more efficient at sucrose production, especially for the arabinose-induced clone. This probably indicates that the host phosphatase in *E. coli* is expressed at low levels, may be only marginally active or sucrose-6-P may not be an ideal substrate, allowing a buildup of sucrose-6-P with concomitant inhibition of additional sucrose-6-P synthesis. In the tandem clones pSpsSppBAD and pSpsSppA transcription was designed so that *spsA* and *sppA* are expressed equivalently, enabling the rapid dephosphorylation of sucrose-6-P with resultant accumulation of free sucrose. Billi, *et. al*. (2000), in our laboratory, demonstrated that *E. coli* BL21(DE3)(pSpsA) transformants demonstrated a 10,000-fold increase in survival compared to wild-type cells following either freeze-drying, air drying or desiccation over phosphorus pentoxide. It is possible that the tandem clone would

have achieved more significant results since it does not rely on a host non-specific phoshomonoesterase activity to release free sucrose.

References

Billi, D., and Potts, M. (2001). Desiccation. *In* K. B. Storey and J. M. Storey (ed.), *Cell and Molecular Responses to Stress*. Elsevier Science, Amsterdam, the Netherlands.

Billi, D., Wright, D. J., Helm, R. F., Prickett, T., Potts, M., and Crowe, J. H. (2000). Engineering Desiccation Tolerance in *Escherichia coli*. *Appl. Environ. Microbiol*. **66**: 1680-1684.

Cano, R., and Borucki, M. K. (1995). Revival and identification of bacterial spores in 25-million-year-old to 40-million-year-old Dominican amber. *Science*. **268**: 1060-1064.

Crowe, J. H., Crowe, L. M., and Chapman, D. (1984). Preservation of membranes in anhydrobiotic organisms: the role of trehalose. *Science* **223**: 701-703.

Crowe, J. H., and Crowe, L. M. (1992). Membrane integrity in anhydrobiotic organisms: toward a mechanism for stabilizing dry cells, p 87-103. *In* G. N. Somero, C. B. Osmond, and C. L. Bolis (ed.), *Water in Life: Comparative Analysis of Water Relationships at the Organismic, Cellular, and Molecular Level*. Springer-Verlag, Berlin, Germany.

Curatti, L., Folco, E., Desplats, P., Abratti, G., Limones, V., Herrera-Estrella, L., and Salerno, G. (1998). Sucrose-phosphate synthase from *Synechocystis* sp. strain PCC 6803: identification of the *spsA* gene and characterization of the enzyme expressed in *E. coli*. *J. Bacteriol*, **180**: 6776-6779.

Gerhardt, P. (1998). Survival of ancient bacteria desiccated within amber: believe it or not? *ASM News* **64**: 68-69.

Germer, J., Muffler, A., Hengge-Aronis, R. (1998). Trehalose is not relevant for *in vivo* activity of σ-containing RNA polymerase in *Eschericia coli*. *J. Bacteriol.* **180**: 1603-1606.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hiroswa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Osumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp strain PCC 6803. I. Sequence features in the 1Mb region from map positions 64% to 92% of the genome. *DNA Res*. **3**: 109-136.

Kennedy, M. J., Reader, S. L., and Swierczynski, L. M. (1994). Preservation records of micro-organisms: evidence of the tenacity of life**.** ***Microbiol.*** **140**: 2513-2529.

Lambert, L. H., Cox, T., Mitchell, K., Rossello-Mora, R. A., Del Cueto, C., Dodge, D. E., Orland, P., and Cano, R. J. (1998). *Staphylococcus succinus* sp. nov,.isolated from Dominican amber. *Int J. Syst. Bacteriol*. **48**: 511-518.

Leslie, S. B., Israeli, E., Lighthart, B., Crowe, J. H., and Crowe, L. M. (1995). Trehalose and sucrose protect both membranes and proteins in intact bacteria during drying. *Appl. Environ. Microbiol*. **61**: 3592-3597.

Postgate, J., and Priest, F. G. (1995). Putative Oligocene spores. *Microbiol*. **141**: 2763-2764.
Potts, M. (2001). Desiccation tolerance of prokaryotes. *Microbiol. Rev*. **58**: 755-805.
Priest, F. G. (1995). Age of bacteria from amber. *Science*. **270**: 2015.
Sneath, P. H. A. (1962). Longevity of microorganisms. *Nature*. **4842**: 643-646.

Chapter 3

DNA and mRNA Stability in Desiccated Cells

Abstract

The genomic DNA of *Nostoc commune* became covalently modified during decades of desiccation. PCR amplification of gene loci from desiccated cells required pretreatment of DNA with *N*-phenacylthiazolium bromide, a reagent that cleaves DNA- and protein-linked glycosylation end-products.

Active Fe-superoxide dismutase (SodF) was the third most abundant soluble protein in cells of *Nostoc commune* CHEN/1986 after prolonged (13 years) storage in the desiccated state. Upon rehydration, Fe-containing superoxide dismutase (Fe-SOD) was secreted from cells. The 21-kDa Fe-SOD polypeptide was purified and the N-terminus was sequenced. *sodF* mRNA was abundant and stable in cells after long-term desiccation. Upon rehydration of desiccated cells, there was a turnover of *sodF* mRNA within 15 min and then a rise in the *sodF* mRNA pool to control levels (quantity of *sodF* mRNA in cells in late logarithmic phase of growth) over approximately 24 h.

Introduction

Some cells that tolerate desiccation do so because of an ability to protect vital components of their cellular machinery from damage and/or repair them quickly upon rehydration. The latter strategy, for example, is employed by the ionizing radiation-resistant bacterium *Deinococcus radiodurans* (Battista *et al*., 1999). Prolonged desiccation of a cell may cause damage to proteins, nucleic acids, and membrane components through browning (Maillard) reactions. The modification is initiated by the spontaneous reaction of reducing sugars with the primary amino groups of proteins and nucleic acids (Papoulis *et al*., 1995). The browning reaction proceeds from reversible Schiff base and Amadori products to a class of irreversibly bound, structurally heterogeneous products referred to as advanced glycosylation end-products (AGEs) (Vasan *et al*., 1996). Accumulation of AGEs is implicated in many of the pathophysiological alterations associated with normal aging. Such Maillard products

occur in plant and animal remains and are a prominent component of ancient DNA extracts (Evershed *et al*., 1997; VanBergen *et al*., 1997; and Poinar *et al*., 1998). Other factors that contribute to DNA modification, and ultimately the killing of cells, include metal-catalyzed Haber-Weiss and Fenton reactions (Henle & Linn, 1997; and Buyuksonmez *et al*., 1999) and the presence of reactive oxygen species and free radicals. High photon flux densities and UV irradiation exacerbate the effects of reactive oxygen species in these processes. Five-membered hydantoin rings in DNA originate from oxidative decay of six-membered pyrimidines (Lindahl, 1997a; Dizdaroglu *et al*., 2002; and Höss *et al*., 1996); their presence is thought to be negatively correlated with the PCR amplification of DNA. Of the free radicals, the highly reactive hydroxyl radical (OH•) causes damage to DNA and other biological molecules (Dizdaroglu, 1992; and Breen & Murphy, 1995). This type of damage is also called 'oxidative damage to DNA' and is implicated in mutagenesis, carcinogenesis and aging (Cutler & Rodriguez, 2003). The occurrence of modified bases is problematic because DNA polymerases are unable to copy these damaged residues. The reagent *N*-phenacylthiazolium bromide (PTB) disrupts AGE crosslinks and its use made it possible to amplify DNA sequences from ancient samples (Poinar *et al*., 1998).

Mechanisms that protect cells from the effects of desiccation include those which circumvent oxidative damage (Beckman & Ames, 1998; Berlett & Stadtman, 1997; and Haslekas *et al*., 1998). Of these, the synthesis of superoxide dismutase (SOD) is an important response. SOD mediates the disproportionation of superoxide radicals to hydrogen peroxide and dioxygen; the hydrogen peroxide is then scavenged either by ascorbate peroxidase or catalase (Miyake *et al*., 1991). In many bacteria, SODs are present in the periplasmic space as well as in the cytoplasm and their synthesis is repressed under anaerobic conditions (Beaman *et al*., 1983; and Gort *et al*., 1999). Regulation of the expression of genes that encode these SODs, as well as genes involved in the synthesis of other antioxidants, may be subject to control by the alternate sigma factor RpoS (Fang *et al*., 1999; and Gort *et al*., 1999). Other mechanisms that cells use to protect against oxidative damage from desiccation include the synthesis of carotenoids and tocopherols (Ehling-Schulz *et al*., 1997; Sandman *et al*., 1998; and Stahl & Sies, 1993) and anthocyanins (in plants, Sherwin & Farrant, 1998).

Cyanobacteria may be especially prone to desiccation damage because they evolve intracellular oxygen through photosynthesis and persist in habitats subject to intense solar irradiation (Potts, 1999). In fact, many terrestrial cyanobacteria withstand the most acute long-term water deficit; how they do so is of considerable practical importance. Cyanobacteria are known to use both Fe- and Mn-containing SODs (Fe-and Mn-SODs) to scavenge superoxide radicals (Canini *et al*., 1992), but there is a paucity of information on the roles of these enzymes in response to water deficit. The *sod* genes of a number of cyanobacteria have been characterized. The genome of *Synechocystis* sp. strain PCC 6803 contains a single SOD gene (*sodB*; slr1516 [Kaneko *et al*., 1996]); in contrast, *Plectonema boryanum* UTEX 485 contains *sodB* as well as three additional Mn-SOD genes (Campbell & Laudenbach, 1995). Results from studies with *Synechococcus* sp. strain PCC 7942, in which *sodB* was inactivated, suggested that in this cyanobacterium Fe-SOD did not protect photosystem II during oxidative stress but that it did protect photosystem I. The enzyme also protected cells from the effects of chilling in the light (S. K. Herbert and D. J. Thomas, Abstr. VIth Cyanobacterial Workshop, abstr. S15, 1998).

The cyanobacterium *Nostoc commune* exhibits a marked tolerance of desiccation and was the subject of a study that sought to identify structural, physiological, and molecular mechanisms that contribute to this tolerance (Potts, 1994). It is becoming clear that these mechanisms are both numerous and diverse. The regulation of gene expression during desiccation and subsequent rehydration is of considerable interest. One of the most obvious features of gene expression in rehydrating *N. commune* is an ordered, apparently stringent, recovery of metabolic functions, beginning with respiration and followed by photosynthesis and finally nitrogen fixation (Gao *et al*., 1998; and Potts & Bowman, 1985). The control of these processes is likely to be complex, given what is known about the sensing of other environmental signals in morphologically less complex cyanobacteria (Schwartz & Grossman, 1998). Studies with field samples of *N. commune* HUN and *N. commune* UTEX 584 provided evidence that some, but not all, proteins remain stable despite extended periods of desiccation (Hill *et al*., 1994; and Scherer & Potts, 1989). Although the drying of *N. commune* UTEX 584 cells led to a rapid cessation of nitrogenase activity (Potts & Bowman, 1985; Potts *et al*., 1984), no evidence

was obtained for denaturation of at least one structural component of nitrogenase, Fe protein, which was present in cells of *N. commune* HUN following 10 years of desiccation (Peat *et al*., 1988). The intracellular ATP pool and the protein biosynthetic machinery of desiccated cells remained unperturbed for 30 min and 2 h, respectively, after rapid drying at –99.5 MPa (Angeloni & Potts, 1986; Potts, 1985; and Potts *et al*., 1984). In contrast, short-term drying led to structural changes in the pigment antenna complexes of cyanobacteria, the phycobilisomes. In the light, the phycobiliproteins were degraded, and even in the dark, short-term drying led to subtle changes in the polydispersity of the complexes when they were analyzed in sucrose gradients (Potts, 1985; Scherer & Potts, 1989). *De novo* transcription in rehydrated cells of *N. commune* UTEX 584 is directed by extant RNA polymerase holoenzyme, which maintains its stability during desiccation, at least long enough for some transcripts to accumulate to control (predrying) levels (Xie *et al*., 1995). A marked accumulation of Fe-SOD was detected in desiccated field materials of *N. commune* CHEN/1986 following rehydration.

Materials and Methods

Cyanobacterial strains and growth conditions.

Desiccated colonies of *N. commune* of different ages were collected from field localities (see Table 1.1) and stored dry in the dark until analysis. Genotypic analysis indicated these all belonged to 'form species' *N. commune* (Wright, 2001). Most herbarium samples (identified by the WH in strain name, see Tables 1.1 & 3.1) were obtained in sealed paper envelopes that were unopened since the time of collection. A liquid culture, *N. commune* DRH-1 was derived from *N. commune* CHEN/1986 and grown in liquid culture in the laboratory.

Genomic DNA purification

In the majority of cases, only limited quantities of dry field material was available, so small portions of the desiccated colonies were ground to a powder under liquid nitrogen. Samples were then added directly to PCR tubes and heated at 95°C in 100 μl of sterile distilled water for up to 4 hours. Rehydrated cells or the cell-free supernatant was used in PCR amplifications (see below). When the initial PCR trial was

unsuccessful, the DNA was pretreated with PTB (see Chapter 1) and PCR amplification repeated.

PCR amplification

The loci selected for amplification through PCR assay included the tRNA^{LEU}(UAA) group I intron (Wright *et al*., 2001), *rrn* (23S rDNA) (Wright *et al*., 2001), *phr*, and *sodF* (Shirkey *et al*., 2000) of *N commune* DRH-1. Assays were optimized with each set of primers (Table 3.1), in multiple trials, prior to experiments. Assays were performed in 50 µl reactions, in pH 9.0 (*sodF* only) buffer (supplied by the manufacturer, Promega Biotech) or in pH 8.5 (*phr*, 23S rDNA, group 1 introns) buffer [60 mM Tris•HCl, 15 mM $(NH_4)_2SO_4$]. Reactions contained 12.5 nmole of each dNTP (each dNTP at 250 µM), 1.5 (*sodF* only) or 3.5 mM Mg^{2+} (*phr*, 23S rRNA, group I introns) and 2.5 U of *Taq* polymerase (Promega). Temperature was controlled with a Biometra thermocycler model T-3. All assays began with a denaturation temperature of 95°C for 2 min and ended with an elongation time of 10 min at 72°C; in each cycle denaturation occurred at 95°C for 1 min and elongation was at 72°C for 90 s. For amplification of tRNA^{LEU} group I introns the annealing temperature (60 s) was 66°C in the first cycle and dropped 0.2 C^o per cycle thereafter for an additional 39 cycles. For amplification of *phr* and 23S rDNA the annealing temperature was 60°C (60 s) in the first cycle, dropped 0.5 C^o per cycle for the next 19 cycles (to 50°C) and was kept constant at 50°C for the remaining 19 cycles. For amplification of *sodF* the annealing temperature was 60°C (90 s) in the first cycle, dropped 1.5 C^o per cycle for the first 10 cycles (to 45°C) and was kept constant at 45°C for the remaining 30 cycles. Three independent amplifications were performed per sample.

Table 3.1. Primer sequences used in PCR assays

Primer sequence	Forward/Reverse	Target	Origin
5'-TGTGGCGGAATGGTAGACGCTAC-3'	f[a]	Bases 6-28 of $tRNA^{LEU}$ (UAA)	*Anabaena* sp. PCC7120
5'-GACTTGAACCCACACGAC-3'	r[b]	Bases 72-56 of $tRNA^{LEU}$ (UAA)	*Anabaena* sp. PCC7120
5'-GGGTGGAGGGACTTGA-3'	r[c]	Bases 67-82 of $tRNA^{LEU}$ (UAA)	*Anabaena* sp. PCC7120
5'-ACCCTATATACTACCTGCTG-3'	f	23S rRNA	*N. commune* DRH1
5'-GGTCTTCGCTGATTCACATG-3'	r	23S rRNA	*N. commune* DRH1
5'-ACATAACCACGCCCGAATGTG-3'	f	*phr*	*N. commune* DRH1
5'-AAATCACAATGACCGTAAAG-3'	r	*phr*	*N. commune* DRH1
5'-GAGTATCACTATGGCAAGCA-3'	f	*sodF*	*N. commune* DRH1
5'-CTAAAGTCAATGTAGTA-3'	r	*sodF*	*N. commune* DRH1

[a]LEU1 primer (see Table 3.2)
[b]LEU2 primer (see Table 3.2)
[c]LEU3 primer (see Table 3.2)

Purification and identification of SodF.

Desiccated colonies of *N. commune* CHEN/1986 were used to isolate preparative amounts of SodF. To achieve efficient cell lysis, colonies were frozen under liquid nitrogen and ground to a powder with a pestle and mortar. The powder was mixed with sterile distilled water, and cells were allowed to rehydrate for 1 h at room temperature to allow the release of soluble proteins. Phycobiliproteins were removed by batch adsorption with Q-Sepharose (Pharmacia) resin equilibrated in 50 mM Tris•HCl (pH7.5) buffer. After dialysis, the sample was concentrated with a 10-kDa-cutoff filter and the retentate, containing approximately 14 mg of total protein, was subjected to preparative sodium dodecyl sulfate-polyacrylamide gel electrophoresis using a BioRad Protean™ II system (10%, w/v, gels). Two major classes of soluble protein were resolved; the 27- to 39-kDa Wsp isoforms (Scherer, 1989) and a protein with a prominent band at approximately 20 kDa. Proteins were transferred to Immobilon P membrane (Millipore) via Western blotting in CAPS (3-[cyclohexylamino]-1-propanesulfonic acid) buffer at pH 10.4 in a Hoeffer transfer chamber and visualized with Coomassie brilliant blue. The 20 kDa protein, immobilized on an excised fragment of membrane, was subjected to automated Edman degradation with an Applied Biosystems model 477A protein sequencer.

Preparation of labeled probes.

The *sodF* and 23s rRNA probes were prepared from genomic DRH1 DNA. The 50-μl reaction mixes in both cases contained 12.5 nmoles each of dCTP, dATP, and dGTP (250 mM), 8 nmoles of dTTP (160 mM) and 3.5 nmoles digoxigenin-labeled dUTP (dig-dUTP) (70 mM, Roche). The "*sodF* labeling reaction" was incubated in Promega magnesium-free buffer with the addition of 1.5 mM Mg^{2+} while the "23s rRNA reaction" was incubated in pH 8.5 buffer (60 mM Tris•HCl, 15 mM $(NH_4)_2SO_4$) with 3.5 mM Mg^{2+}. Both reactions used 2.5 units of Promega *Taq* DNA polymerase and 50 pmoles of each primer. The *sodF* primers were $^{5'}$-GAGTATCACTATGGCAAGCA-$^{3'}$ and $^{5'}$-CTAAAGTCAATGTAGTAG-$^{3'}$ and the 23s rRNA primers were $^{5'}$-GTCGATTGGCGTTGTTACCA-$^{3'}$ and $^{5'}$-GTATTTAGCTGCTTCGTCTA-$^{3'}$. Amplifications for both probes were performed with a Biometra thermocycler model T-3 (Labrepco) and began with a denaturation step at 95°C for 2 min and ended with an

elongation step at 72°C for 10 min. *sodF* amplification continued with 20 cycles of a 1 min denaturation at 95°C, annealing for 1.5 min initially at 60°C with a decrease of 1.5 C^{o} per cycle and elongation at 72°C for 1.5 min followed by an additional 20 cycles with the annealing temperature held at 45°C. Amplification of 23s rRNA continued with 20 cycles beginning with denaturation at 95°C for 1 min, annealing for 1 min initially at 60°C with a decrease of 0.5 C^{o} per cycle, and elongation at 72°C for 1 min followed by an additional 20 cycles with the annealing temperature held at 50°C.

Transcription analysis.

Approximately 1 g of desiccated colonies of *N. commune* CHEN/1986 or rehydrated colonies that quickly were blotted dry, were frozen in liquid nitrogen and ground to a powder. The powder was mixed with 1.5 ml of RNA extraction buffer (0.25 M sucrose, 0.2 M NaCl, 10 mM magnesium acetate, 0.1 M Tris•HCl [pH 9]). After the addition of 1.5 ml of Tris-buffered phenol (pH> 8), the mixture was sonicated at a setting of 30 (Fisher sonic dismembrator model 300) three times for 15 s each time, with cooling on ice. Glass beads (4-mm diameter) were added (1:1 ratio), and the slurry was mixed at top speed by vortex for 10 min. The aqueous phase was recovered through centrifugation and extracted with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1) multiple times until no deposit remained at the interface. The aqueous phase was then extracted with chloroform-isoamyl alcohol (24:1), mixed with 2 volumes of ice-cold 100% ethanol, and left at –20°C for several hours. The precipitate was recovered, dissolved in 1.7 ml of diethyl pyrocarbonate-treated water, mixed with an equivalent volume of 8 M LiCl, and left overnight at 4°C. The pellet was washed with 70% ice-cold ethanol and stored in 70% ethanol at –70°C until needed.

RNA was resolved in formaldehyde agarose gels, transferred to a positively charged nylon membrane (Boehringer Mannheim GmbH) through capillary blotting in 50 mM NaOH, and cross-linked to the membrane by UV irradiation. Hybridization with DIG-dUTP-labeled PCR fragments of *sodF* or 23S ribosomal DNA (rDNA) from *N. commune* DRH1 was at 50°C overnight, with final stringency washes containing 0.1X SSC-0.1% (w/v) SDS at 60°C (Xie, 1995). DNA-RNA hybrids were visualized with anti-DIG alkaline phosphatase and the chemiluminescent detection system of Boehringer Mannheim GmbH.

Results

Stability of desiccated DNA *in vivo*.

The effect of PTB treatment on the capacity to amplify defined genetic loci was studied to obtain evidence of DNA modification during desiccation *in vivo*. Four selected genetic loci [*sodF, phr, rrn* (23S rRNA) and $tRNA^{LEU}$(UAA) group 1 intron] were amplified from colonies of *N. commune* desiccated for periods of up to 149 years (Table 3.2). Photolyase (*phr*) is an enzyme that repairs one form of damage to DNA caused by exposure to UV light. The three gene products as well as 23S rRNA are highly conserved and therefore are good indicators of the quality of the DNA. Products of the expected size were obtained from some of the different samples and their identities were confirmed through DNA sequence analysis. For the majority of the samples amplification of *phr* and *rrn* was possible only after pretreatment of the samples with PTB (Figure 3.1a). Amplification of $tRNA^{LEU}$(UAA) group I intron and *sodF* required PTB in one and two samples, respectively. With samples of *N. commune* WH002, amplification of *sodF, phr* and *rrn* was possible only after treatment of the DNA with PTB. In contrast, all loci were amplified efficiently without need of PTB using DNA from a liquid culture of *N. commune* derived from field material (Figure 3.1b).

Table 3.2. Treatment of DNA with PTB and ability to amplify gene loci

Gene Amplification Material	Age (years)	Viable	tRNA^{LEU} intron[a]		tRNA^{LEU} intron[b]		*phr*		23S rRNA		*sodF*	
			+PTB	-PTB	+PTB	-PTB	+PTB	-PTB	+PTB	-PTB	+PTB	-PTB
WH002	64	Yes	+	+	-	-	+	-	+	-	++	-
WH014	51	Yes	++	ND	+++	+	+	+	+	+	+++	+
ALD8122	25	Yes	++	+	+	+	+	+	+	+	+	+
WH010	120	-	-	-	-	-	+	+	+	-	-	-
WH012	118	-	-	-	+	-	+	+	+	-	-	-
WH001	76	ND	+	+	-	-	+	-	+	-	-	-
WH004	150	ND	+	+	-	-	+	+	+	+	+	+
WH005	140	ND	-	-	-	-	+	+	+	+	-	-
WH006	121	ND	+	-	+	-	+	-	++	+	-	-
WH011	139	ND	-	-	-	-	+	-	+	-	-	-
WH013	NK	ND	-	-	+	-	+	-	+	-	-	-
WH016	134	ND	-	-	+	-	+	+	+	+	-	-
MEL1968	31	ND	++	+	+	+	++	+	++	+	+++	++
SPH1998	1	ND	+	+	+	+	+	+	+	+	-	-
DRH1	NA	Yes	ND	+++	ND	+	ND	+	ND	+	ND	+

NA, not applicable, liquid culture
ND, not determined
+, ++, +++, relative abundance of a single amplification product
-, no amplification product
[a]Using LEU1 and LEU3 primers (see Table 3.2)
[b]Using LEU1 and LEU2 primers (see Table 3.2)

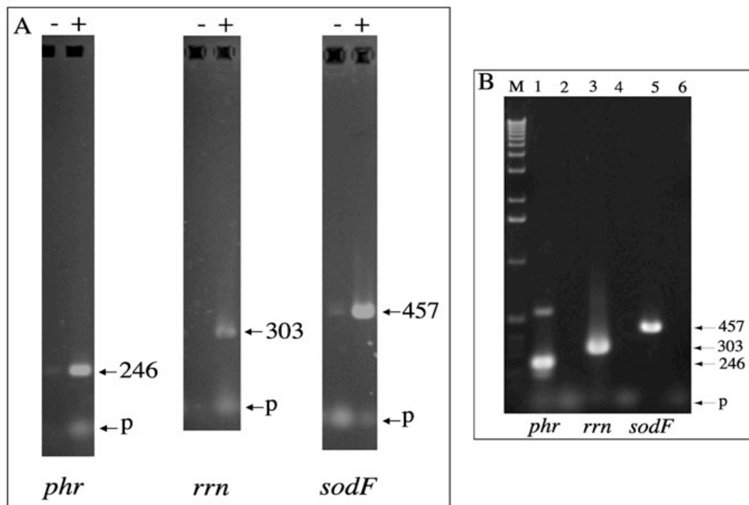


Fig. 3.1. PCR amplification of gene loci.

Numbers in bp; p, primer dimers. (A) DNA from desiccated *N. commune* WH002 (*phr* and *rrn*) or *N. commune* WH014 (*sodF*) with (+) or without (-) prior treatment with PTB. (B) DNA from liquid culture of *N. commune* with no PTB treatment. Lanes 2, 4, and 6 are control assays (no target DNA in reaction mixture).

Viability of desiccated *Nostoc* colonies.

In cases where sufficient material was available dried colonies were rehydrated on moist filter paper followed by incubation at 25°C under subdued illumination. Colonies of *N. commune* TOP/1993 (9 years dry), ENG/1996 (6 years dry), WH002 (63 years dry), and WH014 (55 years dry) all contained material that grew after rehydration. However, WH010 and WH012 (120 and 118 years old, respectively) failed to grow and did not generate two of the markers despite treatment with PTB (Table 3.2).

Identification of SodF in desiccated *N. commune* CHEN/1986.

Soluble proteins of *N. commune* CHEN/1986 were obtained from desiccated colonies after one cycle of freezing and thawing through the addition of water. Resolution of proteins with SDS-PAGE gels identified Wsp isoforms as well as a single prominent band at approximately 21 kDa (Figure 3.2). The 21-kDa protein was recovered, and the unambiguous sequence of the first 20 amino acids was determined to be AFVQDPLPFDINALEPYGMK. (*N*-formylmethinoine was apparently removed by post-translational processing.) Based on the relative amounts of Wsp proteins, which are the most abundant proteins in desiccated *N. commune* CHEN/1986 (Scherer & Potts,

1989), and phycobiliproteins, SodF was judged to be the third most abundant soluble protein in materials of *N. commune* CHEN/1986 desiccated for 13 years (Figure 3.2).

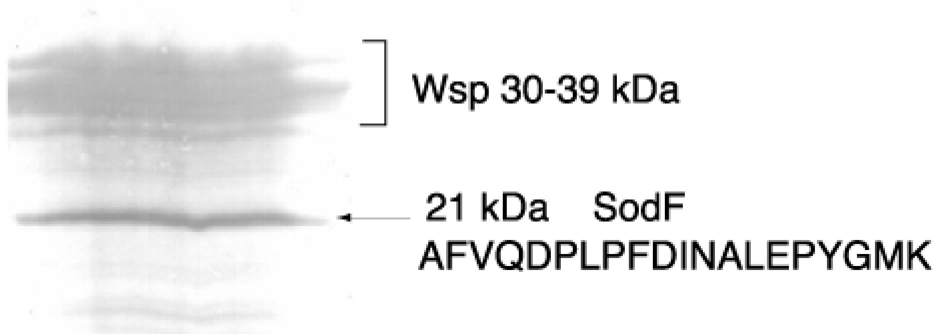


Fig. 3.2. SodF is abundant in desiccated *N. commune* CHEN/1986.
An immunoblot stained with Coomassie blue and used for N-terminal sequence analysis of SodF is shown. Wsp isoforms migrate between 30 and 39 kDa. The N-terminal sequence determined for the 21-kDa peptide (SodF) is indicated.

Expression of *sodF* in *N. commune* ENG/1996.

Probing of RNA extracts from desiccated and rehydrating materials of *N. commune* ENG/1996, as well as liquid cultures of *N. commune* DRH1 in the late logarithmic phase of growth, identified a major transcript of around 750 bases. Two minor signals between 400 and 600 bases were also detected (Figure 3.3a, lanes 2, 3, 4, and C), but their resolution was varied in repeat trials. *sodF* mRNA was abundant in cells after 3 years of desiccation (Figure 3.3a, lane 2) and remained so up to 15 min after the onset of rehydration (Figure 3.3a, lanes 3 and 4). After this time the intensity of the signals diminished and then gradually increased over the next 24 h (Figure 3.3a, lanes 5 through 9). The decrease in *sodF* mRNA after 15 min of rehydration accompanied an increase in the intensity of signals attributed to 23S rRNA that appeared to reach steady state after approx 3 h of rehydration (Figure 3.3b). Multiple bands detected when the 23S rRNA probe was used may represent precursor and processed forms of one or more copies of 23S rRNA (Angeloni & Potts, 1986). The fluctuation in the amounts of *sodF* mRNA in cells, apparently according to their degree of hydration, raised the question of the variability in mRNA content within different colonies. However, the pattern was consistent in repeat trials with different samples of cells of *N. commune* ENG/1996; most

notably, the depletion in *sodF* mRNA after 15 min of rehydration and the subsequent increase in the pool of *sodF* mRNA to control levels after approximately 24 h, were reproducible (data not shown). The patterns in the *de novo* synthesis of 23S rRNA in response to rehydration (Figure 3.3b) were consistent when different samples of the desiccated colonies of *N. commune* ENG/1996 were used (data not shown).

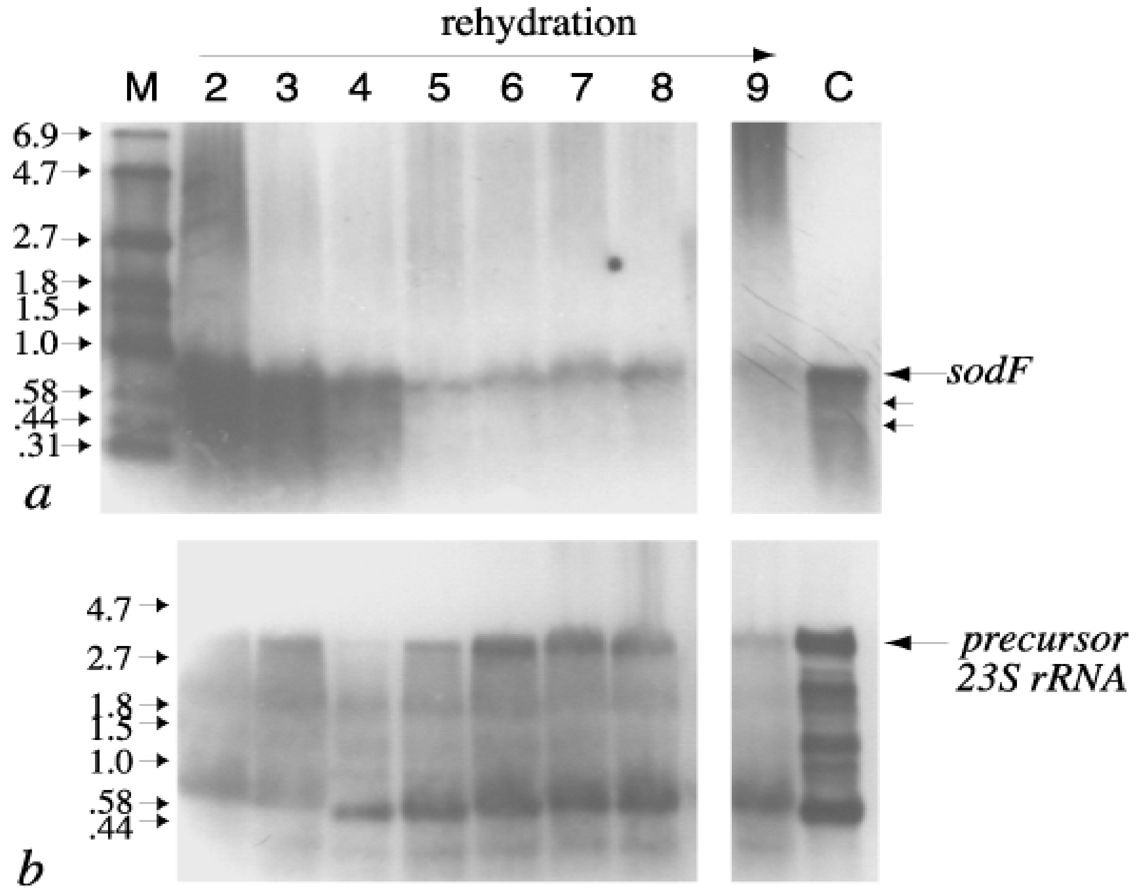


Fig 3.3. (a) *sodF* expression in *N. commune* ENG/1996 by Northern blotting.

Lane M contains RNA molecular weight markers (in kilobases). Other lanes contain mRNAs from colonies desiccated 3 years (lane 2) or rehydrated for 5 min (lane 3), 15 min (lane 4), 1 h (lane 5), 3 h (lane 6), 6 h (lane 7), 12 h (lane 8), or 24 h (lane 9). A liquid culture of *N. commune* DRH1 in late logarithmic phase of growth is shown in lane C. (b) 23sRNA in *N. commune* ENG/1996 by Northern blotting. The lanes contain identical amounts of the same samples applied in panel a.

Discussion

Nostoc commune colonizes habitats that are rarely dry for longer than 6 months; however, in those cases where sufficient material was available, it was possible to demonstrate viability and recovery of samples desiccated for at least 60 years, but not in samples desiccated for >100 years. It has been shown that purified DNA isolated from cells of *N. commune* desiccated for weeks to centuries contain a brown pigmentation that could not be removed by exhaustive solvent extraction but treatment with PTB followed by extraction did remove the coloration (Shirkey *et al*., 2003). The DNA was thus modified through covalent Maillard reactions. Clearly, DNA modification is a natural consequence of desiccation *in situ* and one that the cells can tolerate for considerable periods, but not indefinitely. To do so, the cells would presumably require one or more enzymes to remove such modification at the time of rehydration.

Oxidative DNA damage occurs in cells as a consequence of normal aerobic metabolism, by generation of oxygen-derived free radicals, from exposure to ionizing radiation and from other DNA-damaging agents (Dizdaroglu *et al*., 2002). This type of DNA damage appears to be one of the main causes of aging and gene-related diseases (Halliwell & Gutteridge, 1990; Wallace, 2002; Friedberg *et al*., 1995; Lindahl *et al*., 1997b; and Wood, 1996). It was postulated therefore, that genes involved in oxygen-scavenging mechanisms are likely to be crucial to the response of organisms to air-drying (Fang *et al*., 1999). Three biomarkers of oxidative damage (8-hydroxyguanine, 8-hydroxyadenine, and 5-hydroxyuracil), studied in samples desiccated for 1, 7, and 13 years showed that free radical damage was not a significant event during the prolonged desiccation of dormant *N. commune* cells (Shirkey *et al*., 2003), possibly due to the large quantity of active SOD that was demonstrated to be present in those cells (Shirkey *et al*., 2000). The fact that SodF remains active after such long storage may suggest that the enzyme has structural features which permit it to remain in the native state despite removal of most of the water from the cells or that it may be sequestered in a particular cell microenvironment. The amount of water in such desiccated (anhydrobiotic) cells can be as low as 0.02 g g^{-1} (dry weight) of cells in which proteins lack a monolayer coverage of water (Potts, 1994; 1999).

One of the forms of SOD is frequently located in extracellular fluids, such as synovial fluid (Marklund, 1984), or the periplasmic space of *E. coli* (Benov *et al*., 1995) and other bacteria (Barnes *et al*., 1999; Fang *et al*., 1999; and Gort *et al*., 1999), and *Norcardia asteroides* synthesizes and secretes a SOD which adheres to the outer cell membrane (Beaman *et al*., 1983). In each of these instances, SOD intercepts extracellular superoxide radicals and diminishes damage to the cell. In none of these examples, however, does the quantity of SOD released from the cells approach the magnitude of that released by rehydrating *N. commune*. This finding is all the more significant considering the time the cells were stored in a dormant state. SodF does not have a recognizable N-terminal signal sequence, and microsequencing of the protein through Edman degradation confirmed that there was no processing at the N-terminal region *in vivo*.

It is reasonable to suspect that crowding and elevated salt concentrations within desiccated cells, as well as intense solar irradiation, contribute to the production of both intra-and extracellular superoxide radicals during rehydration as cells become metabolically active in an environment laced with organic compounds. Metabolic activity exacerbates the damage acquired during desiccation, because cells quickly recover the capacity for photosynthesis and thus oxygen evolution (Scherer *et al*., 1984).

Cells are metabolically inactive following dehydration and during subsequent storage in the dry state. Transcription analysis reveals that there is turnover of *sodF* mRNA immediately upon rewetting followed by the resumption of *sodF* mRNA synthesis. Although not directly measured, *sodF* mRNA synthesis must continue or SodF must be stable as water is removed from cells to account for the large amounts in all desiccated samples analyzed. The cells after 24 h of rehydration (Figure 3.3a, lane 9) are in recovery phase, which may be equated with exponential growth. However, the system is complex. For example, in previous studies a sequential recovery of metabolic functions upon rehydration was demonstrated, beginning with respiration and followed by photosynthesis and nitrogen fixation (for references, see Potts, 1994). Nitrogen fixation is recovered after 24 h of rehydration. The fact that *sodF* mRNA is abundant after 24 h of rehydration emphasizes the possible importance of SodF during the recovery phase.

References

Angeloni, S. V., and Potts, M. (1986). Polysome turnover in immobilized cells of *Nostoc commune* (cyanobacteria) exposed to water stress. *J. Bacteriol*. **168**: 1036-1039

Barnes, A. C., Balebona, M. C., Horne, M. T., and Ellis, A. E. (1999). Superoxide dismutase and catalase in *Photobacterium damselae* subsp. *piscicida* and their roles in resistance to reactive oxygen species. *Microbiol*. 145: 483-494.

Battista, J. R., Earl, A. M., and Park, M. J. (1999). Why is *Deinococcus radiodurans* so resistant to ionizing radiation? *Trends Microbiol.* **7**: 362-365.

Beaman, B.L., Scates, S. M., Moring, S. E., Deem, R., and Misra, H. P. (1983). Purification and properties of a unique superoxide dismutase from *Norcardia asteroids . J. Biol. Chem.* **258**: 91-96.

Beckman, K. B., and Ames, B. N. (1998). The free radical theory of aging matures. *Physiol. Rev*. **78**: 547-581.

Benov, L., Chang, L. Y., Day, B., and Fridovich, I. (1995). Copper, zinc superoxide dismutase in *Escherichia coli*: periplasmic location. *Arch. Biochem. Biophys*. 319: 508-511.

Berlett, B. S., and Stadtman, E. R. (1997). Protein oxidation in aging, disease, and oxidative stress. *J. Biol, Chem*. **272**: 20313-20316.

Breen, A. P., and Murphy, J. A. (1995). Reactions of oxyl radicals with DNA. *Free Radic. Biol. Med.* **18**: 1033-1077.

Buyuksonmez, F., Hess, T. F., Crawford, R. L., Paszczynski, A. and Watts, R. J. (1999). Optimization of simultaneous chemical and biological mineralization of perchloroethylene. *Appl. Environ. Microbiol*. **65**: 2784-2788.

Campbell, W. S., and Laudenbach, D. E. (1995). Characterization of four superoxide dismutase genes from a filamentous cyanobacterium. *J. Bacteriol*. **177**: 964-972.

Canini, A., Civitareale, P. Marini, S., Grilli Caiola, M., and Rotilio, G. (1992). Purification of iron superoxide dismutase from the cyanobacteruim *Anabaena cylindrica* Lemm. and localization of the enzyme in heterocysts by immunogold labeling. *Planta* **187**: 438-444.

Cutler, R. G., and Rodriguez, H. (2003). *Critical Reviews of Oxidative Stress and Aging: Advances in Basic Science, Diagnostics and Intervention*. World Scientific Publishing, River Edge, NJ.

Dizdaroglu, M., Jaruga, P., Birincioglu, M., and Rodriguez, H. (2002). Free radical-induced damage to DNA: mechanisms and measurement. *Free Radic. Biol. Med*. **32**: 1102-1115.

Dizdaroglu, M. (1992). Oxidative damage to DNA in mammalian chromatin. *Mutat, Res*. **275**: 331-342.

Ehling-Shultz, M., Bilger, W., and Scherer, S. (1997). UV-B-induced syntheses of photoprotective pigments and extracellular polysaccharides in the terrestrial cyanobacterium *Nostoc commune*. *J. Bacteriol*. **179**: 1940-1945.

Evershed, R. P., Bland, H. A., vanBergen, P. F., Carter, J. F., Horton, M. C. and Rowley Conwy, P. A. (1997). Volatile compounds in archaeological plant remains and the Maillard reaction during decay of organic matter. *Science* **278**: 432-433.

Fang, F. C., DeGroote, M. A., Foster, J. W., Baumler, A. J., Ochsner, U., Testerman, T., Bearson, S., Giard, J. C., Xu, Y. S., Campbell, G., and Laessig, T. (1999). Virulent *Salmonella typhimurium* has two periplasmic Cu, Zn-superoxide dismutases. *Proc. Natl. Acad. Sci*. USA **96**: 7502-7507.

Friedberg, E. C., Walker, G. C., and Siede, W. (1995). *DNA Repair and Mutagenesis*. ASM Press, Washington, DC, pp. 42-43.

Gao, K., Qiu, B., Xia, J., and Yu, A. (1998). Light dependency of the photosynthetic recovery of *Nostoc flagelliforme*. *J. Appl. Phycol*. **10**: 51-53.

Gort, A. S., Ferber, D. M., and Imlay, J. A. (1999). The regulation and role of the periplasmic copper, zinc superoxide dismutase of *Eschericia coli. Mol. Microbiol.* **32**: 179-191.

Halliwell, B. and Gutteridge, J. M. (1990). Role of free radicals and catalytic metal ions in human disease: an overview. *Methods Enzymol*. **186**: 1-85.

Haslekas, C., Stacy, R. A. P., Nygaard, V., Culiáñez-Macia, F. A., and Aalen, R. B. (1998). The expression of a peroxiredoxin antioxidant gene, *AtPer1*, in *Arabidopsis thaliana* is seed-specific and related to dormancy. *Plant Mol. Biol*. **36**: 833-845.

Henle, E. S., and Linn, S. (1997). Formation, prevention and repair of DNA damage by iron/hydrogen peroxide*. J. Biol. Chem*. **272**: 19095-19098.

Hill, D. R., Kladun, S. L., Scherer, S. and Potts, M. (1994). Water stress proteins of *Nostoc commune* (cyanobacteria) are secreted with UV-A/B-absorbing pigments and associate with 1,4-β-D-xylanxylanohydrolase activity. *J. Biol. Chem*. **269**, 7726-7734.

Höss, M., Jaruga, P., Zastawny, T. H., Dizdaroglu, M., and Pääbo, S. (1996). DNA damage and DNA sequence retrieval from ancient tissues. *Nucleic Acids Res*. **24**: 1304-1307.

Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirosawa, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996). Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC 6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* **3**: 109-136.

Lindahl, T. (1997a). Facts and artifacts of ancient DNA. *Cell*. **90**: 1-3.

Lindahl, T., Karran, P., and Wood, R. D. (1997b). DNA excision repair pathways. *Curr. Opin. Genet. Dev*. **7**: 158-169.

Marklund, S. L. (1984). Extracellular superoxide dismutase in human tissue and human cell lines. *J. Clin. Investig*. **74:** 1398-1403.

Miyake, C., Michihata, F., and Asada, K. (1991). Scavenging of hydrogen peroxide in prokaryotic and eukaryotic algae: acquisition of ascorbate peroxidase during evolution of cyanobacteria. *Plant Cell Physiol*. **32**: 33-43.

Papoulis, A., Al-Abed, Y., and Bucala, R. (1995). Identification of N2-(1-carboxyethyl) guanine (CEG) as a guanine advanced glysolation end product. *Biochem*. **34**: 648-655.

Peat, A., Powell, N., and Potts, M. (1988). Ultrastructural analysis of the rehydration of desiccated *Nostoc commune* HUN (cyanobacteria) with particular reference to the immunolabelling of NifH. *Protoplasma*. **146**: 72-80.

Poinar, H. N., Hofreiter, M., Spaulding, W. G., Martin, P. S., Stankiewicz, B. A., Bland, H., Evershed, R. P., Possnert, G., and Pääbo, S. (1998). Molecular coproscopy: dung and diet of the extinct ground sloth *Nothrotheriops shastensis*, *Science*. **281**: 402-406.

Potts, M. (1994). Desiccation tolerance of prokaryotes. *Microbiol. Rev*. **58**: 755-794.

Potts, M. (1999). Mechanisms of desiccation tolerance in cyanobacteria. *Eur. J. Phycol.* **34**: 319-328.

Potts, M., and Bowman, M. A. (1985). Sensitivity of *Nostoc commune* UTEX 584 (cyanobacteria) to water stress. *Arch. Microbiol***.** **141**: 51-56.

Potts, M., and Bowman, M. A., and Morrison, N. S. (1984). Control of matric water potential (Ψ_m) in immobilized cultures of cyanobacteria. *FEMS Microbiol. Lett*. **24**: 193-196.

Sandmann, G., Kuhn, S., and Böger, P. (1998). Evaluation of structurally different carotenoids in *Escherichia coli* transformants as protectants against UV-B radiation. *Appl. Environ. Microbiol*. **64**: 1972-1974.

Scherer, S., Ernst, A., Chen, T. W., and Böger, P. (1984). Rewetting of drought-resistant blue-green algae: time course of water uptake and reappearance of respiration, photosynthesis, and nitrogen fixation. *Oecologia* (Berlin) **62**: 418-423.

Scherer, S., and Potts, M. (1989). Novel water stress protein from a desiccation-tolerant cyanobacterium. *J. Biol. Chem*. **264**: 12546-12553.

Schopf, J. W., Kudryavtsev, A. B., Agrestl, D. G., Wdowlak, T. J., and Czaja, A. D. (2002). Laser-raman imagery of Earth's earliest fossils. Nature 416: 73-76.

Schwarz, R., and Grossman, A. R. (1998). A response regulator of cyanobacteria integrates diverse environmental signals and is critical for survival under adverse extreme conditions. *Proc. Natl. Acad. Sci.* USA **85**: 11008-11013.

Sherwin, H. W., and Farrant, J. M. (1998). Protection mechanisms against excess light in the resurrection plants *Craterostigma wilmsii* and *Xerophyta viscosa*. *Plant Growth Regul*. **24:** 203-210.

Shirkey, B., Kovarcik, D. P., Wright, D. J., Wilmoth, G., Prickett, T. F., Helm, R. F., Gregory, E. M., and Potts, M. (2000). Active Fe-SOD and abundant *sodF* mRNA in *Nostoc commune* (cyanobacteria) after years of desiccation. *J Bacteriol*. **182**: 189-197.

Shirkey, B., McMaster, N., Smith, S., Wright, D., Rodriguez, H., Jaruga, P., Birincioglu, M., Helm, R., Potts, M. (2003). Genomic DNA of *Nostoc commune* (cyanobacteria) becomes covalently modified during long-term (decades) desiccation but is protected from oxidative damage and degradation. *Nucleic Acids Res*. **31**: 2995-3005.

Stahl, W., and Sies, H. (1993). Physical quenching of singlet oxygen and *cis-trans* isomerization of carotenoids. *Ann. N. Y. Acad. Sci*. **691**: 10-19.

VanBergen, P. F., Bland, H. A., Horton, M. C., and Evershed, R. P. (1997). Chemical and morphological changes in archaeological seeds and fruits during preservation by desiccation. *Biochim. Cosmochim. Acta* **61**: 1919-1930.

Vasan, S., Zhang, X., Zhang, X. N., Kapurniotu, A., Bernhagen, J., Teichberg, S., Basgen, J., Wagle, D., Shih, D., Terlecky, I., Bucala, R., Cerami, A., Egan, J., and Ulrich, P. (1996). An agent cleaving glucose-derived protein crosslinks *in vitro* and *in vivo*. *Nature* **382**: 275-278.

Wallace, S. S. (2002). Biological consequences of free radical-damaged DNA bases. *Free Radic. Biol. Med*. **33**: 1-14.

Wood, R. D. **(1996).** DNA repair in eukaryotes. *Annu. Rev. Biochem*. **65**:135-167.

Wright, D., Prickett, T., Helm, R. F., and Potts, M. (2001). Form species *Nostoc commune* (cyanobacteria). *Int. J. Syst. Evol. Microbiol*. **51**:1839-1852.

Xie, W. Q., Tice, D. & Potts, M. (1995). Cell water deficit regulates expression of *rpoC1C2* (RNA polymerase) at the level of mRNA in desiccation-tolerant *Nostoc commune* UTEX 584 (cyanobacteria). *FEMS Microbiol Lett* **126**, 159-164.

Discussion

Based on its typical colony morphology, the field material used for this study was identified as the desiccation-tolerant cyanobacterium *Nostoc commune*. Characterization of this strain by comparison to a wide range of field samples that were identified as *N. commune* or related organisms collected from all continents over a period of 150 years showed a phylogenic relationship between *N. commune* DRH1 and a group of 29 samples which amplified almost identical $tRNA^{LEU}$(UAA) group 1 intron sequences (Wright, *et. al*., 2001). While the study did show that *N. commune* strain DRH1 was a member of the 'typical' form species *N. commune,* it also illustrated that assignment of identification based on morphological characteristics alone should be done with great care.

Trehalose plays a crucial role in the stabilization of microorganisms to desiccation (Crowe *et al*, 1993). A similar role is fulfilled in plants by the disaccharide, sucrose, which can accumulate to between 14 and 20% of the total weight of seeds (Aguilera & Karel, 1997). A comparison of the sucrose content of seeds dried at low versus high humidity revealed that during slow drying, the sucrose content increased to five times the level present in seeds held at high relative humidity (128 versus 25 microgram/axis, respectively). Seeds held at high humidity did not survive desiccation (Blackman *et al*, 1992). The sucrose biosynthetic genes (*spsA* and *sppA*) were cloned from *Synechocystis* sp. 6803 and expressed in *Escherichia coli.* The hypothesis was that desiccation tolerance can be conferred on sensitive organisms through the incorporation of genes responsible for sucrose synthesis. *spsA* confers desiccation tolerance in transformed *E. coli* (Billi *et al*, 2000). The tandem *spsA/sppA* clone may be used to stabilize other organisms.

In order to survive desiccation stress, a cell must either ensure the stability of its nuclear material or employ a mechanism to repair it quickly upon the addition of water. The ionizing radiation-resistant bacterium *Deinococcus radiodurans,* for instance, suffers enormous DNA damage during desiccation but is able to re-form its genome from subgenomic fragments within three hours without mutation (Battista *et al*, 1999). The response of *N. commune* to desiccation stress is much different. *N. commune* cells desiccated for up to 60 years accumulated DNA modifications, without degradation or loss of cell viability. Many primers that failed to amplify loci from DNA extracts without

pretreatment with phenacylthiazolium bromide (PTB), a chemical which breaks cross-links between DNA and reducing sugars, were successful after PTB treatment. No significant oxidative DNA damage compared to control DNA was measured for three *Nostoc* samples dried for up to 10 years (Shirkey *et al*, 2003).

Numerous examples of repeated DNA sequences were reported in cyanobacteria (Robinson *et al*, 2000). HIP1 (for *h*ighly *i*terated *p*alindrome) is the designation for a consensus sequence first identified in *Synechococcus* PCC 6301. It was postulated that this sequence may play a role in genome plasticity and adaptation in cyanobacteria (Gupta *et al*, 1993) and may promote recombination repair following DNA damage (Robinson *et al*, 2000). The *in vivo* stability of DNA following desiccation was studied by subjecting colonies of *Nostoc* to RAPD-PCR assay following varied rehydration times. (Shirkey *et al*, 2003). Several primers gave different banding patterns dependent on the hydration time of the colonies while the HIP-derived primers were more consistent and reproducible, suggesting that these sites are readily accessible despite the degree of hydration and may promote reorganization after long-term desiccation.

Fe-superoxide dismutase was identified as the third most abundant protein in soluble extracts of *N. commune* CHEN desiccated for 13 years. The addition of water to a dried cyanobacterial cell results in the almost immediate resumption of respiration and photosynthesis (see Potts, 1994 for references). The cell must cope with reactive oxygen species generated both through intracellular oxygen synthesis as well as through exposure to UV irradiation. Methods to control the presence of these damaging reactive oxygen species must be present or rapidly mobilized. *sodF* mRNA was abundant in cells after three years of desiccation, turned over within 15 min, and gradually increased over the next 24 hours. When the extrapolysaccharide (EPS) of *N. commune* was purified and subjected to UV irradiation, significant quantities of superoxide radicals were generated (Shirkey *et al*, 2000) emphasizing the importance of the SOD activity that was identified in the extracellular fluid of rehydrating *Nostoc*.

N. commune is well adapted to survive in environments subject to multiple cycles of desiccation and rehydration. The accumulation of disaccharides aids in membrane stabilization and the production of superoxide dismutase protects against oxidative damage. These mechanisms assist in the long-term stabilization of DNA and mRNA and

the rapid recovery of metabolic activities upon rehydration. Desiccation tolerance results from coordinated cellular mechanisms involving membrane stabilization, protection from UV and free redical damage, and stabilization of nucleic acids. Further research may elucidate additional factors that may contribute to maintaining cell integrity during desiccation.

References

Aguilera, J. M., and Karel, M. (1997). Preservation of biological materials under desiccation. *Critical Rev. in Food Sci. and Nutrition*. **37**: 287-309.

Awramik, S. M., Schopf, M. R., and Walter, J. W. (1983). Filamentous fossil bacteria from the Archean of Western Australia. *Precambrian Res*. **20:** 357-374.

Battista, J. R., Earl, A. M., and Park, M. J. (1999). Why is *Deinococcus radiodurans* so resistant to ionizing radiation? *Trends Microbiol*. **9**: 362-365.

Billi, D., Wright, D. J., Helm, R. F., Prickett, T., Potts, M., Crowe, J. H. (2000). Engineering desiccation tolerance in a bacterium. *Appl. Environ. Microbiol*. **66**: 1680-1684.

Blackman, S. A., Obendorf, R. L., and Leopold, A. C. (1992). Maturation proteins and sugars in desiccation tolerance of developing soybean seeds. *Plant physiology*. **100**: 225-230.

Böhm, G. A., Pfleiderer, W., Böger, P., & Scherer, S. **(1995).** Structure of a novel oligosaccharide-mycosporine-amino acid ultraviolet A/B sunscreen pigment from the terrestrial cyanobacterium *Nostoc commune*. *J. Biol. Chem*. **270**: 8536-8539.

Bloom, F. R., Price, P., Guifang, L., Xia, J. L., Crowe, J. H., Battista, J. R., Helm, R. F., Slaughter, S., Potts, M. (2001). Engineering mammalian cells for solid-state sensor applications. *Biosen. & Bioelect*. **16**: 603-608.

Chen, T., Acker, J, P., Eroglu, A., Cheley, S., Bayley, H., Fowler, A., and Toner, M. (2001). Beneficial effect of intracellular trehalose on the membrane integrity of dried mammalian cells. *Cryobiol*. **43**: 168-181.

Crowe, J. H., Crowe, L. M., **and Chapman, D. (1984).** Preservation of membranes in anhydrobiotic organisms: the role of trehalose. *Science*. **223**: 701-703.

Crowe, J. H., Spargo, B. J., and Crowe, L. M. (1987). Preservation of dry liposomes does not require the retention of residual water. *Proc. Nat. Acad. Sci*. **84**: 1537-1540.

Crowe, J. H., Hoekstra, F. A., and Crowe, L. M. **(1992).** Anhydrobiosis. *Annu. Rev. Physiol*. **54**: 579-599.

Crowe, J. H., Crowe, L. M., Leslie, S. B., and Fisk, E. (1993). Mechanisms of stabilization of dry biomlecules in anhydrobiotic organisms. *Curr. Top. Plant Physiol*. **10**: 11-20.

Crowe, J. H., Oliver, A. E., Hoekstra, F. A., and Crowe, L. M. (1997). Stabilization of dry membranes by mixtures of hydroxyethyl starch and glucose: the role of vitrification. *Cryobiol.* **35**: 20-30.

Crowe, J. H., Carpenter, J. F., and Crowe, L. M. (1998). The role of vitrification in anhydrobiosis. *Annu. Rev. Physiol*. **60**: 73-103.

Crowe, J. H., Crowe, L. M., Oliver, A. E., Tsvetkova, N., Wolkers, W., and Tablin, F. (2001). The trehalose myth revisited: introduction to a symposium on stabilization of cells in the dry state*.* *Cryobiol*. **43**: 89-105.

de Castro, A. G. (2000a). Anhydrobiotic engineering. *Nat. Biotech*. **18**: 473.

de Castro, A. G., Bredholt, H., StrØm, A. R., and Tunnacliffe, A. (2000b). Anhydrobiotic engineering of gram-negative bacteria. *Appl. Environ. Microbiol*. **66**, 4142-4144.

Dodds, W. K., and Gudder, D. A. (1995). The ecology of *Nostoc*. *J. Phycol*. **31**: 2-18.

Ehling-Schultz, and M., Scherer, S. (1999). UV protection in cyanobacteria. *Eur. J. Phycol* **34**: 329-338.

Eroglu, A., Russo, M. J., Bieganski, R., Fowler, A., Cheley, S., Bayley, H., and Toner, M. (2000). Intracellular trehalose improves the survival of cryopreserved mammalian cells. *Nature Biotech.* **18**: 163-167.

Gordon, S. L., Oppenheimer, S. R., Mackay, A. M., Brunnabend, J., Puhlev, I., and Levine, F. (2001). Recovery of human mesenchymal stem cells following dehydration and rehydration. *Cryobiol*. **43**: 182-187.

Guo, N., Puhlev, I., Brown, DR., Mansbridge, J., and Levine, F. (2000). Trehalose expression confers desiccation tolerance on human cells. *Nature Biotech*. **18**: 168-171.

Gupta, A., Morby, A. P., Turner, J. S., Whitton, B. A., and Robinson, N. J. (1993). Deletion within the metallothionein locus of Cd-tolerant *Synechococcus* PCC 6301 involving a highly iterated palindrome (HIP1). *Mol Microbiol.* **7**: 189-195.

Helm, R. F., Huang, Z., Edwards, D., Leeson, H., Perry, W., and Potts, M. (2000). Structural characterization of the released polysaccharide of desiccation-tolerant *Nostoc commune* DRH-1. *J. Bacteriol*. **182**: 974-982.

Hill, D. R., Hladun, S. L., Scherer, S., and Potts, M. (1994a). Water stress proteins of *Nostoc commune* (cyanobacteria) are secreted with UV-A/B-absorbing pigments and associate with 1,4-β-D-xylanxylanohydrolase activity. *J. Biol. Chem*. **269**, 7726-7734.

Hill, D. R., Peat, A., and Potts, M. (1994b). Biochemistry and structure of the glycan secreted by desiccation-tolerant *Nostoc commune* (cyanobacteria). Protoplasma: **182**, 126-148.

Hill, D. R., Keenan, T. W., Helm, R. F., Potts, M., Crowe, L. M., and Crowe, J. H. (1997). Extracellular polysaccharide of *Nostoc commune* (cyanobacteria) inhibits fusion of membrane vesicles during drying. *J. Appl. Phycol.* **9**: 237-248.

Huang, Z., Prickett, T., Potts, M., and Helm, R. (2000). The use of the 2-aminobenzoic acid tag for oligosaccharide gel electrophoresis. *Carbohy. Res*. 328: 77-83.

Leslie, S. B., Israeli, E., Lighthart, B., Crowe, J. H., and Crowe, L. M. (1995). Trehalose and sucrose protect both membranes and proteins in intact bacteria during drying. *Appl and Envt. Microbiol.* **61**: 3592-3597.

Oliver, A. E., Crowe, L. M., and Crowe, J. H. (1998). Methods for dehydration-tolerance: depression of the phase transition in dry membranes and carbohydrate vitrification. *Seed Sci. Res*. **8**: 211-221.

Oliver, A. E., Leprince, O., Wolkers, W. F., Hincha, D. K., Heyer, A. G. and Crowe, J. H. (2001). Non-disaccharide-based mechanisms of protection during drying. *Cryobiol.* **43**: 151-167.

Potts, M. (1994). Desiccation tolerance of prokaryotes. *Microbiol. Rev*. **58**: 755-805.

Potts, M. (1996). The anhydrobiotic cyanobacterial cell. *Physiologia Plantarum*. 97: 788-794.

Potts, M. (1999). Mechanisms of desiccation tolerance in cyanobacteria. *Eur. J. Phycol.* **34**: 319-328.

Potts, M. (2000). Nostoc. In *The Ecology of Cyanobacteria*, *Their Diversity in Time and Space*, pp. 465-504. Whitton, B. A., and Potts, M. (eds). Kluwer Academic Publishers, Dordrecht, The Netherlands.

Puhlev, I., Guo, N., Brown, D. R., and Levine, F. (2001). Desiccation tolerance in human cells. *Cryobiol*. **42**: 207-217.

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

Robinson, N. J., Rutherford, JC., Popcock, M. R., and Cavet, J. S. (2000). Metal metabolism and toxicity: repetitive DNA. In: *The Ecology of Cyanobacteria, Their Diversity in Time and Space*. pp 443-463. Whitton, B. A., and Potts, M. (eds). Kluwer Academic Publishers, Dordrecht, The Netherlands.

Schopf, J. W., and Packer, B. M. (1987). Early Archean (3.3-billion-year-old) microfossils from Warrawoona Group, Austrailia. *Science*: 70-73.

Schopf, J. W. and Walter, M. R. (1983). In *Earth's Earliest Biosphere*, ch, 9. J. W. Schopf, (ed). Princeton University Press, Princeton, NJ.

Scherer, S., and Potts, M. (1989). Novel water stress protein from a desiccation-tolerant cyanobacterium. *J. Biol. Chem*. **264**: 12546-12553.

Shaw, E., Hill, D. R., Brittain, N., Wright, D., Täuber, U., Marand. H., Helm, R., and Potts. M. (2003). Unusual water flux in the extracellular polysaccharide of the cyanobacterium *Nostoc commune*. *Appl. and Environ. Microbiol.* **69**: 5679-5684.

Shirkey, B., Kovarcik. D. P., Wright, D. J., Wilmoth, G., Prickett. T. F., Helm, R., Gregory, E. M., and Potts, M. (2000). Active Fe-containing superoxide dismutase and abundant *sodF* mRNA in *Nostoc commune* (cyanobacteria) after years of desiccation. *J. Bacteriol*. **182**: 189-197.

Shirkey, B., McMaster, N. J., Smith, S.. C., Wright, D. J., Rodriguez, H., Jaruga, P., Birincioglu, M., Helm, R. F., and Potts, M. (2003). Genomic DNA of *Nostoc commune* (cyanobacteria) becomes covalently modified during long-term (decades) desiccation but is protected from oxidative damage and degradation. *Nucleic Acids Res*. **31**: 2995-3005.

Taranto, P. A., Keenan, T. W., and Potts, M. (1993). Rehydration induces rapid onset of lipid biosynthesis in desiccated *Nostoc commune* (Cyanobacteria). *Biochim Biophys Acta*. **1168**: 228-237.

Tunnacliffe, A., **de Castro, A. G., and Manzanera M. (2001).** Anhydrobiotic engineering of bacterial and mammalian cells: is intracellular trehalose sufficient? *Cryobiol* **43**: 124-132.

Welsh, D. T., and Herbert, R. A. (1999). Osmotically induced intracellular trehalose, but not glycine betaine accumulation promotes desiccation tolerance in *Exchericia coli*. *FEMS Microbiol. Lett.* **174**: 57-63.

Wolkers, W. F., Walker, N. J., Tablin, F., and Crowe, J. H. (2001). Human platelets loaded with trehalose survive freeze-drying. *Cryobiol*. **42**: 79-87.

Wright, D., Prickett, T., Helm, R. F., and Potts, M. (2001). Form species *Nostoc commune* (cyanobacteria). *Int. J. Sys. Evol. Microbiol.* **51**:1839-1852.

Appendix

Table A.1. Sequences and one possible alignment of introns from Group I of the Distance Tree.

```
              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5         15         25         35         45         55
ALD776DB      AGAAAACTGA GCCTTGCTGG AGAAATCCTT CAAGTGAATG CTCTCAAACT CAGGGAAACC
ALD776DA      AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
ALD8122       AGAAAACTGA GCCTTGCTGG AGAAATCCTT CAAGTGAATG CTCTCAAACT CAGGGAAACC

              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  65         75         85         95        105        115
ALD776DB      TAAATCTGGC AACAGACATG GCAATCCTGA GCCAAGCCGA AAGAGTCAGT TGTTAAAAGT
ALD776DA      TAAATCTG-- ---------- ---------- ---------- ---------- ----------
ALD8122       TAAATCTG-- ---------- ---------- ---------- ---------- ----------

              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 125        135        145        155        165        175
ALD776DB      TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAATTTTAG ATTTGCGATT
ALD776DA      --------AT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAATTTTAG ATTTGCGATT
ALD8122       --------AT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAATTTTAG ATTTGCGATT

              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 185        195        205        215        225        235
ALD776DB      TTAGATTTAC CATTAGTCTT CAATCCAAAA TCCAAAATTG AGGGAAGGTG CAGA--GACC
ALD776DA      ---AATTTGT CTTTTGTGAT TAGCAAAAAA CCGGGAGCTA CCCTAACGTA TAACCTAACC
ALD8122       TTAGATTTAC CATTAGTCTT CAATCCAAAA TCCAAAATTG AGGGAAGGTG CAGA--GACC

              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 245        255        265        275        285        295
ALD776DB      CGACGGGAGC TACCCTAACG TTA---AGCC GAGGGTAAAG GGAGAGTCC- ----------
ALD776DA      CAAAGGTTTG TATACTTGGA TTATGAAGTC GAGGGTAAAG GGAGAGTCCA ATTCTCAAAG
ALD8122       CGACGGGAGC TACCCTAACG TTA---AGCC GAGGGTAAAG GGAGAGTCCA ATTCTTAAAA

              ....|....| ....|....| ....|....| ....|....| ....|....
                 305        315        325        335        345
ALD776DB      ---------- ---------- ---------- ---------- ---------
ALD776DA      CCTTTGAAGG CAGTAGCGAA AGCTGCGGGA GAATG----- ---------
ALD8122       CCTGACTTTA CTAAAGCCAT CAGGTAGCAG TGAAAACTGC GGAAGAATG
```

Table A.2. Sequences and one possible alignment of introns from Group II of the Distance Tree.

```
              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5         15         25         35         45         55
TOP           AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
UTEX584       AAATAATTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH009         AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
NZE           AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
AF019914      AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
AF019920      AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
AF055660      AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
AJ228709      AAATAATTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
N. sp. 73102  AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
N. punctif    AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC

              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  65         75         85         95        105        115
TOP           TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA G-AAATTTTA GATTTGCGAT
UTEX584       TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCG G-AAATTTTA GATTTGCGAT
WH009         TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCG G-AAATTTTA GATTTGCGAT
NZE           TGAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA A-AAATTTTA GATTTGCGAT
AF019914      TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCGA AAAAAGTCGT GAGT------
AF019920      TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA A-AAATTTTA GATTTGTGAT
AF055660      TAAATCTGGT TACAGACATG GCAATCCTGA GCCAAGCCCA A-AAGTTTTA GATTTGCGAT
AJ228709      TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCGA AGAAAGTCCT GAGTCATGAG
N. sp. 73102  TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCG A-AAATTTTA GATTTGCGAT
N. punctif    TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCG A-AAATTTTA GATTTGCGAT

              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 125        135        145        155        165        175
TOP           T-TTAGATTT ACCATTAGTC TTCAATCCAA AAT------- ---------- ----------
UTEX584       T-TTAGATTT ACGATTAGTC TTCAATCCAA AAT------- ---------- ----------
WH009         T-TTAGATTT ACGATTAGTC TTCAATCCAA AAT------- ---------- ----------
NZE           T-TTAGATTT ACGATTAGTC TTCAATCCAA AAT------- ---------- ----------
AF019914      --------GC AGAGTAAATT -AAAACTCTT AACTCCTAAC T--------- ----------
AF019920      T-TTAGATTT ATGATTAGTC TTCAATCCAA AATTCAAAAT C--------- ----------
AF055660      T-TTAGATTT GCGATTAATC TTCAATCCAA AATTCAAAAT CGTTCGACTG AGCGAAGCCG
AJ228709      TGTTGAGTGC TGAGTAAATT TAAAACTCTT AACTCCTAAC TCCT------ ----------
N. sp. 73102  T-TTAGATTT GGGATTAGTT TTCAATCTAA AATCCAAAAC ---------- ----------
N. punctif    T-TTAGATTT GGGATTAGTT TTCAATCTAA AATCCAAAAC ---------- ----------

              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 185        195        205        215        225        235
TOP           ---------- -CCAAAATTG AG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
UTEX584       ---------- -CCAAAATTG AG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
WH009         ---------- -CCAAAATTG AG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
NZE           ---------- -CCAAAATTG AG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
AF019914      ---------- ---------A TTCGGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
AF019920      -----TAAAA TCCAAAATTG CG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
AF055660      AAGTCTAAAA TCCAAAATTG AG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
AJ228709      --------AA CTCATAACTG TTCGGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
N. sp. 73102  ---------- -CCAAAATTG AG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
N. punctif    ---------- -CCAAAATTG AG-GGAAGGT GCAGAGACCC GACGGGAGCT ACCCTAACGT
```

```
             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                245        255        265        275        285        295
TOP          TAAGCCGAGG GTAAAGGGAG AGTCCAATTC TTAAAACCTG ACTTTGCTAA AGCCATCAGG
UTEX584      TAAGCCGAGG GTAAAGGGAG AGTCCAATTC TTAAAACCTG ACTTTGCTAA AGCCATCAGG
WH009        TAAGCCGAGG GTAAAGGGAG AGTCCAATTC TTAAAACCTG ACTTTGCTAA AGCCATCAGG
NZE          TAAGCCGAGG GTAAAGGGAG AGTCCAATTC TTAAAACCTG ACTTTACTAA AGCCATCAGG
AF019914     TAAGTCGAGG GTAAAGGGAG AGTCCAATTC TCAAAACCTG ATTTGGCTAT TGCTATCAGG
AF019920     TAAGTCGAGG GTAAAGGGAG AGTCCAATTC TCAAAACCTG GCCTTGCTAA AGCCATCAGG
AF055660     TAAGTCGAGG GTAAAGGGAG AGTCCAATTC TCAAAATCTG ATCTGGCTAT TGTCATCAGG
AJ228709     TAAGCCGAGG GTAAAGGGAG AGTCCAATTC TCAAAATCCG AGTTGGCTAT TGTCATCAGG
N. sp. 73102 AAAGTCGAGG GTAAAGGGAG AGTCCAATTC TCAAAATCTG ATTTGGCTAT TGCCATCAGG
N. punctif   AAAGTCGAGG GTAAAGGGAG AGTCCAATTC TCAAAATCTG ATTTGGCTAT TGCCATCAGG

             ....|....| ....|....| ....|
                305        315        325
TOP          TAGCAGTGAA AACTGCGGAA GAATG
UTEX584      TAGCAGTGAA AACTGCGGAA GAATG
WH009        TAGCAGTGAA AACTGCGGAA GAATG
NZE          TAGCAGTGAA AACTGCGGAA GAATG
AF019914     TAGCAGTGAA AACTGCGGGA GGATG
AF019920     TAGCAGTGAA AACTGCGGGA GAATG
AF055660     TAGCAGTGAA AACTGCGGGA GAATG
AJ228709     TAGCAGTGAA AACTGCGGGA GAATG
N. sp. 73102 TAGCAGTGAA AACTGCGGGA GGATG
N. punctif   TAGCAGTGAA AACTGCGGGA GGATG
```

Table A.3. Sequences and one possible alignment of introns from Group III of the Distance Tree.

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    5         15         25         35         45         55
WH002           AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
AJ228710        AGACAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
MALB            AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
MALA            AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                    65         75         85         95        105        115
WH002           TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA A-AAATTTTA GATTTGCGAT
AJ228710        TAAATCTGGT GACAGATATG GCAATCCTGA GCCAAGCCGA AGAAAGTGCT GAGTACTGAG
MALB            TAAATCTGGT AACAGACATG GCAATCCTGA GCCAAGCCGA AAAAAGTCGT GAGTAATGAG
MALA            TAAATCTGTT TACAGATATG GCAATCCTGA GCCAAGCCCG G-AAATTTTA GAATCAAGAT

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   125        135        145        155        165        175
WH002           TTTAGATTTA CCATTAGTCT TCAATCCAAA ATCCAAAATT GAGGGAAGGT GCAGAGACCC
AJ228710        TAAAGTT--- ---AAAACTC TTAACTCACA ACT----ATT C--GGAAGGT GCAGAGACTC
MALB            TACTG----- ----AG--TA GTGAGTAA-- ATT----ATT T--GGAAGGT GCAGAGACTC
MALA            TTTAGATTTA CCATTAGTCT TCAATCCAAA ATCCAAAATT GAGGGAAGGT GCAGAGACCC

                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   185        195        205        215        225        235
WH002           GACGGGAGCT ACCCTAACGT TAAGCCGAGG GTAAAGGGAG AGTCCAATTC TTAAAACCTG
AJ228710        GACGGGAGCT ACCCTAACGT TAAGTCGAGG GTAAAGGGAG AGTCCAATTC TCAAAATCTG
MALB            GACGGGAGCT ACCCTAACGT GAAGACGAGG GTAAAGGGAG AGTCCAATTC TCAAAACCTG
MALA            GACGGGAGCT ACCCTAACGT TAAGCCGAGG GTAAAGGGAG AGTCCAATTC TTAAAACCTG

                ....|....| ....|....| ....|....| ....|....| ....|
                   245        255        265        275        285
WH002           ACTTTACTAA AGCCATCAGG TAGCAGTGAA AACTGCGGAA GAATG
AJ228710        AAGCTTCTGA AGTCATCAGG TAGCAGTGAA AGCTGCGGGA GAATG
MALB            ACGTTGGTAA AACTATCAGG TAGCAGTGAA AGCTGCGGGA GAATG
MALA            ACTTTACTAA AGCCATCAGG TAGCAGTGAA AACTGCGGAA GAATG
```

Table A.4. Sequences and one possible alignment of introns from Group IV of the Distance Tree.

```
             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  5         15         25         35         45         55
TEN          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
SPH          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH006B       AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGAAAG CTCTCAAACT CAGGGAAACC
HUN          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
DRH1         AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGAAAG CTCTCAAACT CAGGGAAACC
CHEN         AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH001        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGAAAG CTCTCAAACT CAGGGAAACC
MEL          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
MOA          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
RISB         AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGAAAG CTCTCAAACT CAGGGAAACC
RISA         AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGCGAAAG CTCTCAAACT CAGGGAAACC
ALD857DC     AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGAAAG CTCTCAAACT CAGGGAAACC
ALD857DB     AGAAAACTGA GCCTTGAAGG AGAAATCCCT CAAGTAGAAG CTCTCAAACT CAGGGAAACC
ALD857DA     AGAAAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
AF019912     AGATAACTGA GCCTTGGAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
AF019918     AGATAACTGA GCCTTGGAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH003        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH008        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
TAG          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH007B       AGATAACTGA GCCTTGAGGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH007A       AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH014        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH012        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH004        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH016        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH015        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
WH013        AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
ALD779D      AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
MALC         AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
ENG          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
BER          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
BBC          AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
ALD776DC     AGAAAACTGA GCCTTGCTGG AGAAATCCTT CAAGTGAATG CTCTCAAACT CAGGGAAACC
WH006A       AGATAACTGA GCCTTGAAGG AGAAATCCCT CAAGTGGAAG CTCTCAAACT CAGGGAAACC
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                     65         75         85         95        105        115
TEN             TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
SPH             TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
WH006B          TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
HUN             TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCG GAAAT-TTTA GATTTGCGAT
DRH1            TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
CHEN            TAAATCTGAT GACAGACATG GCAATCCTGA GCCAGGCCCA AAAAT-TTTA GATTTGCGAT
WH001           TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA GAAAT-TTTA GATTTGCGAT
MEL             TAAATCTGGC AACAGACATG GCAATCCTGA GCCAAGCCGA AAGAG-TCAG TTGTTAAAAG
MOA             TAAATCTGGT TACAGACATG GCAATCCTGA GCCAAGCCCA GAAAT-TTTA GATTTGCGAT
RISB            TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCG GAAAT-TTTA GATTTGCGAT
RISA            TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
ALD857DC        TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCG GAAAT-TTTA GATTTGCGAT
ALD857DB        TAAATTTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
ALD857DA        TAAATTTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
AF019912        TAAATCTGAC AACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGTGAT
AF019918        TAAATCTGAC AACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGTGAT
WH003           TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA GAAAT-TTTA GATTTGCGAT
WH008           TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA GAAAT-TTTA GATTTGCGAT
TAG             TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA GAAAT-TTTA GATTTGCGAT
WH007B          TAAATCTGAT GACAGACATG GCAATCCTGG GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
WH007A          TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCGA AAAAG-TCCT GAGTCATGAG
WH014           TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
WH012           TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
WH004           TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
WH016           TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
WH015           TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
WH013           TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCCA GAAAT-TTTA GATTTGCGAT
ALD779D         TAAATCTGGT GACAGACATG GCAATCCTGA GCCAAGCCGA AGAA------ -AGTCCTGAG
MALC            TAAATCTGGT AACAGACATG GCAATCCTGA GCCAAGCCGA AAAAAGTCGT GAGTAATGAG
ENG             TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
BER             TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
BBC             TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
ALD776DC        TAAATCTGGC AACAGATATG GCAATCCTGA GCCAAGCCGA AAGAG-TCAG TTGTTAAAAG
WH006A          TAAATCTGAT GACAGACATG GCAATCCTGA GCCAAGCCCA AAAAT-TTTA GATTTGCGAT
```

```
             ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                  125        135        145        155        165        175
TEN          TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
SPH          TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH006B       TTTAGATTTA CCATTAGTCT TTAATCC--- ---------- -AAAATCCAA AATTGAGGGA
HUN          TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
DRH1         TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
CHEN         TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH001        TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTCAGGGA
MEL          T---AATTTG TCTTTTGTG- ---ATTA--- --------CA AAAAACCGGG AGCTACCCTA
MOA          TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
RISB         TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
RISA         TTTAGATTTG CGATTGGTCT TCAATCC--- ---------- -AAAATCCAA AATTGCGGGA
ALD857DC     TTTAGATTTA CGATTAATCT TTAATCC--- ---------- -AAAATCCAA AATTGAGGGA
ALD857DB     TTTAGATTTA CGATTAATCT TTAATCC--- ---------- -AAAATCCAA AATTGAGGGA
ALD857DA     T---AATTTG TCTTTTGTG- ---ATTAG-- --------CA AAAAACCGGG AGCTACCCTA
AF019912     TTTAGATTTA CGATTAGTCT TCAATCCAAA ATTCAAAATC TAAAATCCAA AATTGCGGGA
AF019918     TTTAGATTTA CGATTAGTCT TCAATCCGAA ATTCAAAATC TAAAATCCAA AATTGCGGGA
WH003        TTTAGATTTG CGATTAGTCT ACAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH008        TTTAGATTTG CGATTAGTCT ACAATCC--- ---------- -AAAATCCAA AATTGAGGGA
TAG          TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH007B       TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH007A       TGTTGAGTGC TGAGTAAATT ---GTTC--- ---------- ---------- -------GGA
WH014        TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH012        TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH004        TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATTCAA AATTTAGGGA
WH016        TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH015        TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH013        TTTAGATTTG CGATTAGTCT ACAATCC--- ---------- -AAAATCCAA AATTGAGGGA
ALD779D      TGCTGAGTAG TGAGTGAAT- ---ATTC--- ---------- ---------- -------GGA
MALC         TACTGAGTAG AGAGTAAATT ---ATTT--- ---------- ---------- -------GGA
ENG          TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
BER          TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
BBC          TTTAGATTTA CGATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
ALD776DC     TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
WH006A       TTTAGATTTA CCATTAGTCT TCAATCC--- ---------- -AAAATCCAA AATTGAGGGA
```

```
              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                 185        195        205        215        225        235
TEN           AGGTGCAGA- -GACCCGACG GGAGCTACCC AAACGTTA-- A-GCCGAGGG TAAAGGGAGA
SPH           AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH006B        AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGGGGG TAAAGGGAGA
HUN           AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
DRH1          AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
CHEN          AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH001         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
MEL           ACGTATAACC TAACCCAAAG GTTTGTATAC TTG--TTATG AAGTCGAGGG TAAAGGGAGA
MOA           AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
RISB          AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
RISA          AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- AAGTCGAGGG TAAAGGGAGA
ALD857DC      AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
ALD857DB      AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
ALD857DA      ACGTATAACC TAACCCAAAG GTTTGTATAC TTGGATTATG AAGTCGAGGG TAAAGGGAGA
AF019912      AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GTCGAGGG TAAAGGGAGA
AF019918      AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GTCGAGGG TAAAGGGAGA
WH003         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GTCGAGGG TAAAGGGAGA
WH008         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GTCGAGGG TAAAGGGAGA
TAG           AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH007B        AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH007A        AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH014         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH012         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH004         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GGCGAGGG TAAAGGGAGA
WH016         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH015         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH013         AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GTCGAGGG TAAAGGGAGA
ALD779D       AGGTGCAGA- -GACCCGACG GGAGCTA--C TAACGTTA-- A-GTCGAGGG TAAAGGGA-A
MALC          AGGTGCAGA- -GACTCGACG GGAGCTACCC TAACGTGA-- A-GACGAGGG TAAAGGGAGA
ENG           AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
BER           AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
BBC           AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
ALD776DC      AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
WH006A        AGGTGCAGA- -GACCCGACG GGAGCTACCC TAACGTTA-- A-GCCGAGGG TAAAGGGAGA
```

```
              ....|....| ....|....| ....|....| ....|....| ....|....| ....|....|
                   245        255        265        275        285        295
TEN           GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
SPH           GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH006B        GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
HUN           GTCCAATTCT TAAAACCTGA CTTTGCTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
DRH1          GTCCAATTCT TAAAACCTGA TTTTGCTAAA GTCATCAGGT AGCAGTGAAA ACTGCGGAAG
CHEN          GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH001         GTCCAATTCT TAAAACCTGA CTTTGCTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGGAG
MEL           GTCCAATTCT CAAAGCCT-- -----TTG-- ---A--AGGC AGTAGCGAAA GCTGCGGGAG
MOA           GTCCAATTCT TAAAACCTGA CTTTGCTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
RISB          GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGTGGAAG
RISA          GTCCAATTCT TAAAACCTGA TCTAATTAAA GCGATCAGGT AGCAGTGAAA ACTGCGGGAG
ALD857DC      GTCCAATTCT TAAAACCTGA TTTTGCTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
ALD857DB      GTCCAATTCT TAAAACCTGA TTTTGCTAAA GCCATCAGGT AGCAGTGGAA ACTGCGGAAG
ALD857DA      GTCCAATTCT CAAAGCCT-- -----TTG-- ---A--AGAC AGTAGCGAAA GCTGCGGGAG
AF019912      GTCCAATTCT TAAAACCTGA GCTGGCTATT GCCATTAGGT AGCAGTGAAA ACTGCGGAAG
AF019918      GTCCAATTCT TAAAACCTGA GCTGGCTATT GCCACTAGGT AGCAGTGAAA ACTGCGGAAG
WH003         GTCCAATTCT CAAAACCTGA T---GCTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGGAG
WH008         GTCCAATTCT CAAAACCTGA T---GCTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGGAG
TAG           GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH007B        GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH007A        GTCCAATTCT CAAAATCCGA GTTGGCTATT GTCATCAGGT AGCAGTGAAA ACTGCGGGAG
WH014         GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH012         GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH004         GTCCAATTCT TAAAACCTGA CTTTGCTAAA GCCATCAGGT AGCAGTGAAA GCTGCGGGAG
WH016         GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH015         GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH013         GTCCAATTCT CAAAACCTGA T---GCTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGGAG
ALD779D       GTCCAATTCT CAAAACCTGA TCTGGCTATT GTCATCAGGT AGCAGTGAAA ACTGCGGGAG
MALC          GTCCAATTCT CAAAACCTGA CGTTGGTAAA ACTATCAGGT AGCAGTGAAA GCTGCGGGAG
ENG           GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
BER           GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
BBC           GTCCAATTCT TAAAACCTGA TTTTGCTAAA GCTGTCAGGT AGCAGTGAAA ACTGCGGGAG
ALD776DC      GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
WH006A        GTCCAATTCT TAAAACCTGA CTTTACTAAA GCCATCAGGT AGCAGTGAAA ACTGCGGAAG
```

```
               ....

TEN            AATG
SPH            AATG
WH006B         AATG
HUN            AATG
DRH1           AATG
CHEN           AATG
WH001          AATG
MEL            AATG
MOA            AATG
RISB           AATG
RISA           AATG
ALD857DC       AATG
ALD857DB       AATG
ALD857DA       AATG
AF019912       AATG
AF019918       AATG
WH003          GATG
WH008          GATG
TAG            AATG
WH007B         AATG
WH007A         AATG
WH014          AATG
WH012          AATG
WH004          AATG
WH016          AATG
WH015          AATG
WH013          GATG
ALD779D        AATG
MALC           AATG
ENG            AATG
BER            AATG
BBC            AATG
ALD776DC       AATG
WH006A         AATG
```

Curriculum vitae

Deborah J. Wright

Virginia Polytechnic Institute and State University

205 Engel Hall, Blacksburg, Virginia 24061

Education

B. S. Summa cum laude with Honors in Chemistry—Georgetown College, Biology and Chemistry double major, Georgetown, Kentucky 40324

22 hours graduate coursework in pharmacology—Indiana University-Purdue University at Indianapolis (IUPUI)

M. S. Biochemistry—Virginia Polytechnic Institute and State University, Blacksburg, Virginia, 24061

Work Experience

Virginia Tech, Blacksburg, Virginia. Laboratory Specialist Senior. 11/95 – present.

Assisted Dr. Malcolm Potts in molecular biology research on the cyanobacteria *Nostoc commune.*

National Sun Industries, Goodland, Kansas 10/92 – 6/94

Quality control of sunflower, soybean and canola oils and meal for livestock feed.

Harrison and Associates, Midland, Texas 5/89-3/91

Performed confirmation of drug positives by mass spectrometer at drug testing facility.

Eli Lilly & Co., Indianapolis, Indiana 10/76—5/80

Assisted Dr. John Parli in drug metabolism research.

Honors and Awards

Member, Phi Lambda Upsilon Chemistry Honor Society

Member, Gamma Sigma Delta National Agricultural Honor Society

Publications

Shirkey, B., McMaster, N. J., Smith, S. C., **Wright, D**. **J**., Rodriguez, H., Jaruga, P., Birincioglu, M., Helm, R. F., and Potts, M. (2003). Genomic DNA of *Nostoc commune* (cyanobacteria) becomes covalently modified during long-term (decades) desiccation but is protected from oxidative damage and degradation. *Nucleic Acids Res*. **31**: 2995-3005.

Shaw, E., Hill, D. R., Brittain, N., **Wright, D. J.,** Täuber, U., Marand. H., Helm, R. F., and Potts. M. (2003). Unusual water flux in the extracellular polysaccharide of the cyanobacterium *Nostoc commune*. *Appl. Environ Microbiol.* **69**: 5679-5684.

Wright, D., Prickett, T., Helm, R. F., and Potts, M. (2001). Form species *Nostoc commune* (cyanobacteria). *Int. J. Sys. Evol. Microbiol.* **51**: 1839-1852.

Billi, D., **Wright, D. J.,** Helm, R. F., Prickett, T., Potts, M., and Crowe, J. H. (2000). Engineering desiccation tolerance in a bacterium. *Appl. Environ. Microbiol*. **66**: 1680-1684.

Shirkey, B., Kovarcik. D. P., **Wright, D. J**., Wilmoth, G., Prickett. T. F., Helm, R. F., Gregory, E. M., and Potts, M. (2000). Active Fe-containing superoxide dismutase and abundant *sodF* mRNA in *Nostoc commune* (cyanobacteria) after years of desiccation. *J. Bacteriol*. **182**: 189-197.