MOLECULAR BIOLOGY AND BIOTECHNOLOGY OF CYANOBACTERIAL NITROGEN FIXATION

G. S. VENKATARAMAN
Division of Microbiology, Indian Agricultural Research Institute, New Delhi 110012, India.

BIOLOGICAL nitrogen fixation is restricted to certain prokaryotic diazotrophs like bacteria and blue-green algae (Cyanobacteria), which are either feee-living or associated with plants. This biological source accounts for 40–45% of the total nitrogen fixed annually in the biosphere[1,2]. The physiology and biochemistry of nitrogen fixation including the enzymology of nitrogenase have been well documented[3–5]. In this report, I shall briefly summarize some of the recent developments in the molecular biology of the nitrogen-fixing cyanobacterial systems. The main objective is to convince the molecular biologists that this ancient group of organisms, occupying the twilight zone in evolution should be of interest to them and also to impress upon the traditional algologists that cyanobacteria are no more 'biological enigma', but have begun to acquire a truly 'molecular biology' of their own.

Their resemblance to photosynthetic bacteria in the prokaryotic nature and their close proximity to eukaryotic algae in their oxygen-evolving photosynthetic machinery have for long made cyanobacteria the wards of both botanists and phycologists. The major molecular biological interest is largely due to their evolutionary significance, autotrophic nutrition, cellular differentiation and a coupling of photosynthetic and nitrogen-fixing machineries. We now have the skill and background information to manipulate the cyanobacterial systems for biotechnological purposes. Cyanobacterial mutants for antibiotic resistance, nitrogenase absence, oxygen sensitivities, heterotrophy and altered heterocysts are now available and many strains can be grown clonally with ease[6–8]. There is also some evidence for mating, transformation and possibilities for transduction in this system[9–11].

Cyanobacterial genome: Ultrastructural and cytochemical studies have shown that the centrally located nucleoplasm is not bounded by a membrane and contains loosely organized DNA fibrils[12,13]. A histone-like binding protein is present[14], although cytochemically identifiable histones may be absent[15–17]. This protein is very similar to the histone-like protein HU_E of *E. coli*[18]. DNA of *Anabaena cylindrica* can also be isolated as an RNA- and protein containing 'nucleoid' analogous to that of *E. coli*[17,19]. Base composition values show a wide range among chroococcalean (35–71 mol% GC) and Oscillatorian (40–67 mol%) members, but a narrow range among pleurocapsalean (39–47 mol%) and heterocystous (38–47 mol%) forms[10]. Modified bases, 5-methylcytosine and 6-methylaminopurine, occur[20,21] as in many bacterial DNAs, indicating the operation of sequence-specific restriction-modification systems[22,23]. Specific restriction nucleases are known in several cyanobacterial strains such as *Anabaena variabilis* (*Ava I* and *Ava* II)[24].

The cyanobacterial genome complexity is comparable to that of *E. coli*. The genome sizes have been found to fall into four groups with averages of 2.2, 3.6, 5.0 and 7.4×10^9 daltons, which correspond to multiples of 2, 3, 4 and 6 times a basic unit of 1.2×10^9 daltons[25,26]. The chroococcalean members fall in the first (dimeric) or second (trimeric) group; the chroococcales which can fix nitrogen anaerobically fall in the second, while the aerobic fixers in the third. Pleurocapsales show trimeric genomes and heterocystous forms show tetrameric or hexameric genomes[25]. These genome sizes suggest a possible evolution of nutritionally and structurally complex cyanobacteria from simple non-nitrogen fixing unicellular forms[25]. Duplication or quadruplication of genome seems to have an evolutionary significance as suggested for modern bacteria from a mycoplasm-like ancestor[27–29]. There seems to be a correlation be-

tween genome size and organizational complexity. In eukaryotes, much DNA does not code for protein and may play only an evolutionary role. It will be interesting to know whether complex prokaryotes also carry large amounts of non-informational DNA![10]

Ribonucleic acids: RNA polymerase activity was first demonstrated in *Anacystis nidulans* as early as 1967[30,31]. Using the *E. coli* terminology, the subunits were designated as β, β', σ and α[32]. The cyanobacterial enzyme differs from that of *E. coli* in that (i) σ subunit is in stoichiometric amounts and the enzyme cannot be separated into σ and 'core' ($\beta'\beta\alpha_2$) on phosphocellulose, (ii) rifampicin-resistant initiation complexes can be formed at 0°C and (iii) *in vitro* reconstitution of urea-denatured subunits is very slow and dependent upon the presence of σ[10,33]. The enzyme has been shown to be rifampicin-sensitive and α-amantin insensitive and cyanobacterial transcription is generally rifampicin- and streptolydigin-sensitive *in vivo*[34-36]. It will be interesting to see if these polymerases undergo subunit modification and/or substitution when transcriptional specificity is expected to change as during chromatic adaptation and heterocyst and spore differentiation.

A unique feature of cyanobacteria is the presence of 'conditionally stable' RNAS which were first detected in *Anacystis nidulans*[37]. These novel RNAS are synthesized both in light and dark, but in the light, like mRNA, they turn over rapidly. They, however, appear completely stable in darkness. They can also accumulate in light, if photosynthetic electron flow is blocked by DCMU. Possibilities of their being stabilized messengers for proteins required upon restoration of growth conditions or their being involved in controlling gene expression in non-growing cells have been suggested[10]. These RNAS contain many oligonucleotide sequences absent from rRNA and an understanding of their nature and function is essential.

Control of gene expression: The obligately photoautotrophic tendencies of cyanobacteria have been generally attributed to their comparative stability and inability to alter, unlike *E. coli*, their enzyme levels in response to environmental changes and manipulations. It is rather difficult to visualize such an insensitivity at molecular level. As Doolittle[10] remarks, "... there are many assimilatory pathways in which cyanobacteria *do* regulate gene expression. I find it hard to believe that they could not have evolved controls on gene expression in other pathways, were it to their selective advantage. Furthermore, there appears to be no peculiarity of the cyanobacterial genome or the machinery for its expression which could stand in the way of the evolution of such controls. In bacteria, genetic regulatory mechanisms are clearly of multiple and independent origins."

A simple regulatory system involves a direct chemical relationship between regulatory effector molecules and the regulated genetic system. In a complex regulatory system, the regulator effector serves a symbol of general nutritional status and regulates a domain of chemically and genetically distinct systems related only in their biological role in maintaining that status[10]. In photoautotrophs, one would expect that alterations in the availability of any one of the requirements such as CO_2, nitrogen, phosphorus etc might affect in multiple domains extending beyond the pathways directly involved in their utilization. The regulation of heterocysts is a pointer in this direction. In heterocystous cyanobacteria, removal of combined nitrogen, regulatory effector, results first in the degradation of phycobiliprotein in all cells of the filament and subsequently in the induction of nitrogenase and heterocyst formation. In *Anabaena* 7120, Fleming and Haselkorn[38,39] have shown that (i) proteins present in the vegetative cells are degraded in heterocysts by proteases which are induced by nitrogen starvation, (ii) proteins synthesized early, late or continuously in heterocysts, but not in vegetative cells and *vice versa* and (iii) proteins synthesized in both vegetative cells and proheterocysts, but enjoying continued synthesis only in the latter. Similarly, glutamine, the major export product of functioning heterocysts, seems to be the likely negative effector for nitrogenase and heterocyst differentiation. A formal understanding of this

model system is beginning to emerge[40-43], and for both academic and practical reasons, our major research thrust should be to understand the molecular events during nitrogenase induction and heterocyst differentiation. It will be useful to identify by Southern[44] hybridization, mRNAS specific to differentiating cells and clone in *E. coli* DNA fragments from which they are transcribed. Cloned DNAS can initially serve as probes for heterocyst-specific transcription and later as templates to duplicate heterocyst transcription patterns *in vitro*.

Cyanobacteria may also be persuaded to convert light energy into chemical energy. Cyanobacterial nitrogenase also acts as an ATP-dependent hydrogenase and splits water, liberating molecular hydrogen[45,46].

Organization and regulation of nif genes: Most of what is known about the genetics and regulation of nitrogen fixation in prokaryotes was initially established in *Klebsiella pneumoniae*[47]. In this bacterium, a cluster of 17 *nif* genes, localized on the chromosome and organized in 7 or 8 transcription units has been identified[48-53]. Information about *nif* genes and their regulation in cyanobacteria has largely come from *Anabaena* 7120[54]. In this strain, homology was found with *K. pneumoniae nif* HDK and with the *nif* VS region[54,55]. Unlike *Klebsiella*, *Azotobacter* and *Rhodopseudomonas*, where *nif* H, D and K are closely linked and cotranscribed, in *Anabaena* 7120, *nif* H and *nif* D are closely linked whereas *nif* K is separated by 11 kb[56,58], (figure 1). An additional *nif* H copy has also been detected in *Anabaena* 7120[54,55]. Steven J. Robinson of the University of Massachusetts has now sequenced the extra *nif* H gene and found that it codes for a protein nearly identical in aminoacid sequence to the first copy. The extra *nif* H is transcribed to yield a 1.9 kb message under anaerobic conditions. The function of this second dinitrogenase reductase is not known, but Robinson suggests that it might be capable of accepting electrons from a different donor than is used by the first gene's product.

During heterocyst differentiation in *Anabaena* 7120, the DNA located between *nif* D and *nif* K is

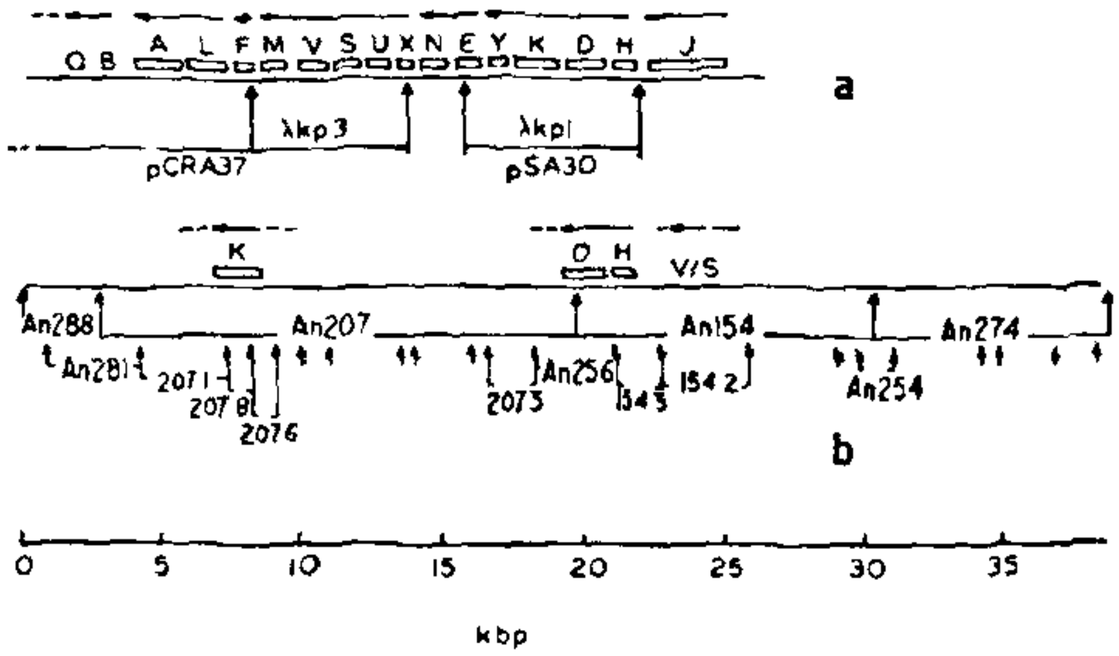

Figure 1. Physical map of *nif* genes of *Klebsiella* (a) and *Anabaena* 7120 (b)[41].

excised in the form of a circle Goldon *et al*[68] (figure 2). Excision is accomplished by conservative recombination between directly repeated identical 11 bp sequences. One copy of the direct repeat is located 335 bp upstream of the start of the *nif* K coding region. The other copy is located about 85 bp inside the open reading frame of the *nif* D gene, determined for vegetative cell DNA. As a consequence of the excision, the open reading frame of *nif* D is fused to 5-flanking sequence of *nif* K which replaces the carboxy terminal 26 amino acids with 43 new amino acids. The operon created by the fusion produces a transcript nearly 5 kb long containing sequences of all three genes, *nif* H, D and K (Goldon *et al*[68]). Thus, under nitrogen fixing conditions, the *nif* HDK genes in *Anabaena* 7120 become contiguous in the heterocyst. No *ntr* related control of nitrogen fixation has been detected in this strain. The *gln*A cloned gene expresses in *E. coli* independently of *ntr*C[59,60]. The nucleotide sequence of *gln*A has been determined and two starting points for initiation of transcription have been mapped[61]. One has a sequence similar to the *E. coli* consensus promotor and is used under conditions of nitrogen excess, whereas the second one resembles the *Anabaena nif* H promotor and is used under conditions of nitrogen deficiency[61]. These results were interpreted by Turner *et al*[61] to indicate control of *nif* expression through a modified RNA polymerase. In contrast to the filamentous *Anabaena* 7120, the *nif* HDK genes have been found to be clustered in the unicellular *Gloeothece*[62].

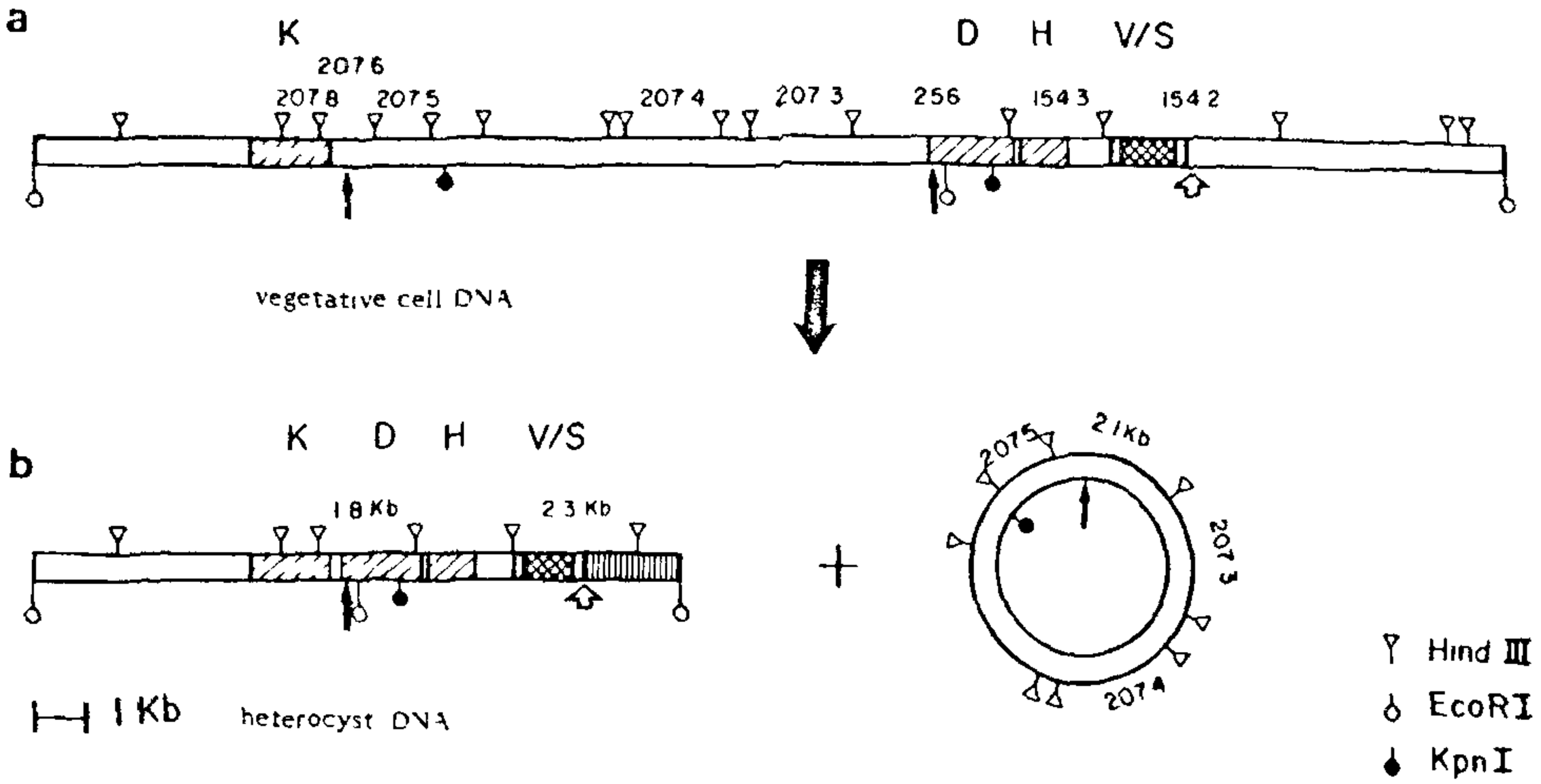

Figure 2. Organization of *nif* genes in *Anabaena* vegetative cell (a) and heterocyst DNA (b). In heterocysts, the DNA is rearranged such that *nif* K and D are adjacent and the intervening DNA is excised and is found as a circle[68].

Plasmid biology: Recently substantial progress has been made in plasmid biology of cyanobacteria. These extrachromosomal elements of DNA have been detected in all the major typological groups such as unicellular, filamentous heterocystous and non-heterocystous forms[63-65]. Construction of genetically marked cyanobacterial plasmids has also been achieved by introducing *amp*^{r} encoding transposon Tn901 from *E. coli* into the cyanobacterial plasmid pUH24 from *Anacystis nidulans*[66] and this is the first example of expression of an *E. coli* gene in a cyanobacterial cell. The construction of a recombinant plasmid consisting of pUH24 in which the 5Md part of *E. coli* pR146 is inserted offers the possibility of further constructing a plasmid carrying *amp*^{r} and *strep*^{r} genes.

The number of cyanobacterial species in which proper genetic analysis is possible has increased dramatically in the past year, thanks to the development in Peter C. Wolk's laboratory in Michigan State University, USA, of a system for introducing shuttle plasmids into cyanobacteria by conjugation from *E. coli*. Three elements are required for this: (i) a shuttle vector capable of replication in *E. coli* and in *Anabaena*, carrying drug resistance markers that are expressed in *Anabaena* and that do not contain too many *Ava* I and *Ava* II sites, (ii) a colicin K or colicin D plasmid capable of mobilizing the shuttle vector in *trans* and (iii) an IncP plasmid such as RP4 to provide a wide host range pilus for transfer. A beautiful application of the conjugation system has been perfected in *Aphanocapsa* 6174 by G. Bullerjahn of the University of Missouri-Columbia. He has constructed a plasmid containing the colEl *ori*, RP4 *tra* functions and Tn501, a 7.9 kbp transposon that confers resistance to mercuric ion. Joseph Thomas and his group[67] at BARC, Bombay have shown an irreplaceable requirement for Na^+ by *Anabaena torulosa* under nitrogen-fixing conditions. The salt-tolerance of this strain appears to be mediated by Na^+ exclusion linked to ATPase. It is hoped that the gene transfer mechanisms can be used to study salt tolerance in this agriculturally important system.

Cyanobacterial nitrogen fixation is quantitatively important and by using modern tools of molecular biology and genetic engineering, can be exploited at ecological and agronomical levels. For long, much attention has been given to a

holistic approach, concentrating on various aspects of taxonomy and physiology of many species. Although this is certainly philosophically laudable, it has put this system at least two decades behind what was achieved with *E. coli*. Molecular biological tools and techniques are now available and are being increasingly employed to unravel the control mechanisms and biotechnologically domesticate these organisms. This requires collaborative efforts of many laboratories and this has now begun.

4 February 1985

1. Hardy, R. W. F. and Havelka, U. D., *Science*, 1975, **188**, 633.
2. Burris, R. H., In: *Nitrogen fixation: Free-living systems and chemical models.* (eds) W. E. Newton and W. H. Orme-Johnson, Univ. Park Press, Baltimore, 1980, p. 7.
3. Mortenson, L. E. and Thorneley, R. N. F., *Annu. Rev. Biochem.*, 1979, **48**, 387.
4. Brill, W. J., *Microbiol. Rev.*, 1980, **44**, 449.
5. Eady, R., In: *Current perspectives in nitrogen fixation* (eds) A. H. Gibson and W. J. Newton, Aust. Acad. Sci., Canberra, 1981, p. 172.
6. Currier, T. C., Haury, J. F. and Wolk, C. P., *J. Bacteriol.*, 1977, **129**, 1556.
7. Gotto, J. W., Tabita, F. R. and Van Baalen, C., *J. Bacteriol.*, 1979, **140**, 327.
8. Stewart, W. D. P. and Singh, H. N., *Biochem. Biophys. Res. Commun.*, 1975, **62**, 62.
9. Delaney, S. F., Herdman, M. and Carr, N. C., In: *Genetics of the algae* (ed.) B. A. Lewin, Univ. Calif. Press, 1976, p. 7.
10. Doolittle, W. F., *Adv. Micro. Physiol.*, 1979, **20**, 1.
11. Van Baalen, C., In: *The biology of blue-green algae* (eds) N. C. Carr and B. A. Whitton, Blackwell, Oxford, 1973, p. 201.
12. Leach, C. K. and Herdman, M., In: *The biology of blue-green algae* (eds) N. C. Carr and B. A. Whitton, Blackwell, Oxford, 1973, p. 186.
13. Wolk, C. P., *Bacteriol. Rev.*, 1973, **37**, 32.
14. Haselkorn, R. and Rouviere-Yaniv, J., *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 1917.
15. Biswas, B. B., *Cytologia*, 1957, **22**, 90.
16. De, D. N. and Ghosh, S. N., *J. Histochem. Cytochem.*, 1965, **13**, 298.
17. Makino, F. and Tsuzuki, J., *Nature (London)*, 1971, **231**, 446.
18. Searcy, D. G., Stein, D. B. and Green, G. R., *Biosystems*, 1978, **10**, 19.
19. Worcel, A. and Burgi, F., *J. Mol. Biol.*, 1972, **72**, 127.
20. Pakhomova, M. V., *Doklady Biochem.*, 1974, **214**, 71.
21. Pakhomova, M. V., Zaitseva, G. N. and Belozerskii, A. N., *Doklady Biochem.*, 1968, **182**, 227.
22. Arber, W., *Progress in nucleic acid research and molecular biology*, 1974, **4**, 1.
23. Roberts, R. J., *CRC Crit. Rev. Biochem.*, 1976, **3**, 123.
24. Murray, K., Hughes, S. G., Brown, J. S. and Bruce, S. A., *Biochem. J.*, 1976, **159**, 317.
25. Herdman, M., In: *Proc. 2nd Int. Symp. on Photosynthetic Prokaryotes* (eds) G. A. Codd and W. D. P. Stewart, Univ. Dundee, 1976, p. 229.
26. Herdman, M., Javier, M., Rippka, K. and Stanier, R. Y., *J. gen. Microbiol.*, 1979, **111**, 73.
27. Wallace, D. C. and Morowitz, H. J., *Chromosoma*, 1973, **40**, 121.
28. Hopwood, D. A., *Bacteriol. Rev.*, 1967, **31**, 373.
29. Zipkas, D. and Riley, M., *Proc. Natl. Acad. Sci. USA*, 1975, **72**, 1354.
30. Capesius, I. and Richter, G., *Zeit. Naturf.*, 1967, **228**, 876.
31. Von der Helm, K. and Zillig, W., *Hoppe-Seyler's Zeit. Physiol. Chemie*, 1967, **348**, 902.
32. Herzfield, F. and Zillig, W., *Eur. J. Biochem.*, 1971, **24**, 242.
33. Herzfield, F. and Kiper, M., *Eur. J. Biochem.*, 1976, **62**, 189.
34. Rodriguez-Lopez, M., Munoz, M. L. and Vasquez, D., *Fed. Europ. Biochem. Soc. Lett.*, 1970, **9**, 171.
35. Bogorad, L., *Annu. Rev. Plant Physiol.*, 1975, **26**, 369.
36. Doolittle, W. F., *J. Bacteriol.*, 1972, **111**, 316.
37. Singer, R. A. and Doolittle, W. F., *J. Bacteriol.*, 1978, **12**, 44.
38. Fleming, H. and Haselkorn, R., *Proc. Natl. Acad. Sci. USA*, 1973, **70**, 2727.
39. Fleming, H. and Haselkorn, R., *Cell*, 1974, **3**, 159.
40. Wilcox, M., Mitchison, G. J. and Smith, R. J., *J. Cell Sci.*, 1973, **13**, 637.
41. Haselkorn, R., Robinson, S. J. and Rice, D., In: *Structure and function of plant genomes* (eds) O. Ciferri and L. Dure, Plenum Press, N.Y., 1983, p. 285.
42. Wolk, C. P., In: *Spores VI* (eds) P. Gerhardt, H. Sadoff and R. Costilow, Amer. Soc. Microbiol., Washington, 1975, p. 85.

43. Haselkorn, R., *Annu. Rev. Plant Physiol.*, 1978, **29**, 219.
44. Southern, E. M., *J. Mol. Biol.*, 1975, **94**, 51.
45. Peterson, R. B. and Burris, R. H., *Arch. Mikrobiol.*, 1978, **116**, 125.
46. Lambert, G. R. and Smith, G. D., *Fed. Europ. Biochem. Soc. Lett.*, 1977, **83**, 159.
47. Robson, R., Kennedy, C. and Postgate, J. R., *Can. J. Microbiol.*, 1983, **29**, 954.
48. McNeil, T., McNeil, D., Roberts, G. P., Supianio, M. A. and Brill, W. J., *J. Bacteriol.*, 1978, **136**, 253.
49. Riedel, G. E., Ausubel, F. M. and Cannon, F. G., *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 2866.
50. Merrick, M., Filser, M., Dixon, R., Elmerich, C., Sibold, L. and Houmard, J., *J. Gen. Microbiol.*, 1980, **117**, 509.
51. Puhler, A. and Klipp, W., In: *Biology of inorganic nitrogen and sulphur* (eds) W. Bothe and A. Trebst, Springer Verlag, Berlin, 1981, p. 276.
52. Sibold, L., *Mol. Gen. Genet.*, 1982, **186**, 569.
53. Bevnon, J., Cannon, M., Buchanan-Wollaston, C. and Cannon, F., *Cell*, 1983, **34**, 665.
54. Mazur, B. J., Rice, D. and Haselkorn, R., *Proc. Natl. Acad. Sci. USA*, 1980, **77**, 186.
55. Rice, D., Mazur, B. J. and Haselkorn, R., *J. Biol. Chem.*, 1982, **257**, 13157.
56. Maverech, M., Rice, D. and Haselkorn, R., *Proc. Natl. Acad. Sci. USA*, 1980, 77, 6476.
57. Lammers, P. J. and Haselkorn, R., *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 4723.
58. Mazur, B. J. and Chiu, C. F., *Proc. Natl. Acad. Sci. USA*, 1982, **79**, 6782.
59. Fisher, R., Tuli, R. and Haselkorn, R., *Proc. Natl. Acad. Sci. USA*, 1981, **78**, 3393.
60. Tuli, R., Fisher, R. and Haselkorn, R., *Gene*, 1982, **19**, 109.
61. Tumer, N. E., Robinson, S. J. and Haselkorn, R., *Nature* (*London*), 1983, **306**, 337.
62. Kallas, T., Rebiere, M. C., Rippka, R. and Tandeau de Marsae, N., *J. Bact.*, 1983, **155**, 427.
63. Lau, R. H., Sapienza, C. and Doolittle, W. F., In: *Plasmids: Environmental effects and maintenance mechanisms* (eds) K. R. Rozee and C. Stuttard, Academic Press, N.Y., 1980.
64. Lau, R. H., Sapienza, C. and Doolittle, W. F., *Mol. Gen. Genet.*, 1980, **178**, 203.
65. Simon, R. D., *J. Bact.*, 1978, **136**, 414.
66. Van den Hendel, C. A. M. J. J., Verbeek, S., Van der Ende, A., Weisbeck, P. J., Borrias, W. E. and Van Arkel, G. A., *Proc. Natl. Acad. Sci. USA*, 1980, **77**, 1570.
67. Apte, S. K. and Thomas, J., *Curr. Microbiol.*, 1980, **3**, 291.
68 Goldon, J. W., Robinson, S. J. and Haselkorn, R., *Nature* (*London*), 1985, **314**, 419.

NEWS

WILD WEATHER

. . . "Most meteorologists, being the conservative sort and preferring to analyze a long record, have been reticent about recent weather. But now three government researchers have offered an estimate of just how crazy the weather has been. . . . Three consecutive winters of the past eight—when averaged over the 48 states—were much colder than normal. Less memorable perhaps but still significant were three winters much warmer than normal. Such a combination of six very abnormal winters in 8 years should not be expected to recur for more than 1,000 years . . . according to [Thomas Karl, Robert Livezey, and Edward Epstein (Natl. Oceanic & Atmospheric Admin.)]. The variability from winter to winter in the late 1970's and early 1980's may be extreme, but the stability of winter temperatures during the preceding 20 years is no less striking, according to Karl. By the standards that set off six of the eight recent winters as extremely abnormal, all twenty winters between 1955-56 and 1974-75 were unexceptional, often by comfortable margins. The unbroken string of relative stability of winter temperatures is at least as unusual as the recent period of high variability. . . ."

[(Richard A. Kerr in *Science* 227(4686):506, 1 Feb 85) Reproduced with permission from Press Digest, *Current Contents*®, No. 11, March 18, 1985, p. 14. (Published by the Institute for Scientific Information®, Philadelphia, PA, USA.)]