Chloroplast promoters from higher plants

S.D.Kung and C.M.Lin

Department of Biological Sciences, University of Maryland Baltimore County (UMBC), Catonsville, MD 21228, USA

Received ³ May 1985; Revised and Accepted 2 October 1985

ABSTRACT

This survey compiles 60 chloroplast promoter sequences from higher plants published to date and compares them with these sequences from procaryotic systems. The current evidence demonstrates that structurally defined chloroplast promoters are, in most cases, functionally active in initiating gene expression in chloroplasts.

INTRODUCTION

The transcriptional cycle consists of three main steps: initiation, elongation and termination. The initiation of transcription is a crucial stage at which gene expression can be regulated by the interaction of RNA polymerase with the promoters. The promoter is defined as a region of DNA involved in the binding of RNA polymerase to initiate transcription. In procaryotic systems, with the notable exception of the nif genes in Klebsiella (1), promoters consist of two regions of conserved sequences, located about 10 and 35 bp upstream of the transcription startpoint and separated by an optimal distance of 17 bp (2,3). These characteristics, together with some secondary sequence conservation in the region upstream of the primary promoter (4) , are major determinants of promoter activity.

After analyzing the DNA sequence of 112 well-defined promoters, a consensus sequence of procaryotic promoters was established (2,3,4). The AT-rich sequence around the -35 region has often been referred to as the recognition site and the sequence around the -10 region as the Pribnow box (5). Within the -35 region (TTGACA), the trimer TTG is strongly conserved, appearing at a frequency of 82, 84, and 79% for each base respectively (3). At the -10 region (TATAAT), the first TA pair and the last T are also highly preserved.

Generally, there is one promoter per gene. Sharing of promoter by more than one gene also occurs. On the other hand, there are genes which have more than one promoter (6-10). A working model for the function and regulation of multiple promoters of stable RNA has been proposed (11,12).

TABLE I: COMPILATION OF CHLOROPLAST PROMOTERS FROM HIGHER PLANTS

 $\ddot{}$

CHLOROPLAST PROMOTERS

The chloroplast is probably procaryotic in origin and therefore possesses the procaryotic type machinery for protein synthesis (13-18). Many chloroplast genes and their 5'-flanking regions have been sequenced and studied (19-60). There are conserved promoter sequences in front of many chloroplast genes analyzed to date. Table I lists two groups of chloroplast promoters. The inclusion of promoters in the first group (A) was based on both structural and functional considerations. Chloroplast promoters must resemble those of procaryotic promoter in structure and satisfy one of the functional criteria. (a) These structures are protected by $B.$ coli RNA polymerase against DNase digestion. (b) These structures are determined by S1 mapping to be located in the 5' flanking region of a gene. And (c) these structures are active in initiating gene expression in either heterologous or homologous system. All the promoters included in the second group (B) are defined structurally only. For example, a promoter can be identified by locating DNA sequences matching the structure of procaryotic consensus sequence promoters proximal to a chloroplast gene. As can be seen from Table I, most chloroplast promoters identified thus far fall into the second group (B). The alignment of the promoters in Table 1 was based on the format and considerations given by Hawley & McClure (3) to maximize the homology of both the -35 and -10 regions. In order to align promoters with different spacing, two breaks were placed in the sequences; one immediately after the -35 and one immediately before the -10 regions. In addition to the -35, -10 regions and the spacer, 20 more bases were included at each end in most cases when they are available to have a total of about 70-80 bases for each promoter.

In both groups (A & B) these promoters contain two conserved hexamers separated by a short stretch of about 17 bp. The first hexamer is TTGACA resembling the -35 region of procaryotic promoter sequences, in which the trimer TTG is highly conserved. In the second hexamer, the conserved sequence of TATAAT is identical to the -10 region (5). Within this hexamer, three bases (TA---T) are highly conserved. The first TA pair was present in all but few cases, in which the pair was either AA or TT (Table I). Similar to the procaryotic promoters, the last "T" in this hexamer is also highly conserved, appearing in all but few of the 60 promoters compiled (Table I). The -35 and -10 regions are separated by 11-24 bases which is very close to the allowed procaryotic spacing of 15-21 bp (2,3). Overall, the distribution of bases in each position in the -35 and -10 regions of these chloroplast promoters is statistically similar to that of procaryotic promoters (Table II).

As in *B. coli*, secondary or tertiary promoters in the region upstream from the primary promoter sites for stable chloroplast RNA genes were also found (Table I). For example, there are two promoter sites reported for many tRNA genes $(24,25)$. In most cases, the secondary promoters have -35 and -10 regions separated by 11-22 bp, and in a few cases only the individual isolated -10 regions were identified (30). The existence of multiple promoter sites was reported for maize and duckweed rRNA genes (26,41). Conversely, sharing of promoter by more than one gene also exists. The rRNAs (23), some tRNAs (51) and the β and ϵ (32) genes are co-transcribed.

The functional assay of chloroplast promoters (38,47, 52) in the pKOl system established a certain structure-function relationship of the sequences tested. Using plasmids with sequential deletions in the 5'-flanking region of the psbA gene, and a homologous chloroplast extract, Link (57) demonstrated

POSITION	DISTRIBUTION												
						TTGACA				TATA AT			
A						$0 \t5 \t2 \t35 \t7 \t30$				1 55 14 36 40 1			
T						58 49 1 4 13 12				59 5 27 7 7 56			
G		$0 \t2 \t51 \t6 \t14 \t7$									9 12 2 1		
c		$0 \t2 \t4 \t13 \t24 \t9$									$10 \t 5 \t 11 \t 2$		
Chloroplast $(5)^*$ 100 84 88 60 41 52										98 92 45 60 67 93			
Procaryotes $(5)^{**}$ 82 84 79 64 54 45										81 95 44 59 51 96			

TABLE II Distribution of Bases at each Position in Promoters of Chloroplast Genes

Calculated from Table I

** Calculated from reference 3

that the upstream region containing the sequences of TTGACA and TATACT, which resemble the procaryotic -35 and -10 regions, is required for efficient in vitro transcription. Employing a similar technique Chua and his co-workers (48) have demonstrated that increasing the distance between the -35 and -10 regions of maize rbcL promoter from 18 pb to 20 bp with an AT base pair insertion reduced the level of transcription drastically. It is clear that the -35, the -10 regions and the distance between them are involved in the modulation of the efficiency of chloroplast promoters. The evidence obtained from the functional assays in the homologous system demonstrated that some structurally well-defined chloroplast promoters are indeed functionally active promoters (52, 60). Therefore, this survey can be considered as a preview of the nature and properties of chloroplast promoters from higher plants.

ACKNOWLEDGEMENTS

The author wishes to thank the Fulbright Foundation for the Award, Dr. J.R. Ellis of Warwick University and Dr. J.C. Gray of Cambridge University, United Kingdom, for providing laboratory space during my sabbatical leave and helpful discussion in preparing this manuscript. The skill and patience of Karen Sweeney in typing this manuscript is deeply appreciated.

REFERENCES
1. Benyan

- Benyan, J., Cannon, M., Buchanan-Wollaston, V. and Cannon, F. (1983) Cell 34, 665-671
- 2. Rosenberg, M. and Court D. (1979) Ann. Rev. Genet. 13, 319-353
- 3. Hawley, D.K. and McClure, W.R. (1983) Nucleic Acids Res. 11, 2237-2255
- 4. Travers, A.A. (1984) Nucleic Acids Res. 6: 2605-2618
- 5. Pribnow, D. (1975) Proc. Natl. Acad. Sci. 72, 784-787
6. Glass. R.E. (1983) Gene Function. Croom Helen Ltd. Lo
- 6. Glass, R.E. (1983) Gene Function. Croom Helen Ltd. London
7. Nomura, M. and Morga, E. (1977) Ann. Rev. Genet. 11, 297-
- 7. Nomura, M. and Morga, E. (1977) Ann. Rev. Genet. 11, 297-347
- 8. Miura, A., Kruege, J.H., Itsh, S. deBoer, H.A. and Nomura M. (1981) Cell 25, 773-782
- 9. deBoer, H.A., Gilbert, S.F. and Nomura, M. (1979) Cell 17: 225-234
- 10. Boros, I., Csordas-Toth, E., Kiss, A., Kiss, I., Torok, I., Udvardy, A., Udvardy, K. and Venetianer, P. (1983) Biochim. Biophy. Acta 739, 173-180
- 11. Travers, A.A., Lamond, A.T., Mace, H.A.F. and Bergan, M.L. (1983) Cell 35, 265-273
- 12. Shepherd, N.S., Churchward, G. and Bremer, H. (1980) J. Bact 141, 1098-1108
- 13. Kung, S.D. (1977) Ann. Rev. Plant Physiol. 28, 401-437
- 14. Whitfeld, P.R. and Bottomley, W. (1983) Ann. Rev. Plant Physiol. 34, 279-310
- 15. Hartley, M.R., Wheeler, A. and Ellis, R.J. (1975) J. Mol. Biol. 91, 67-73
- 16. Gatenby, A.A., Castleton, J.A. and Saul, M.W. (1981) Nature 291, 117-121
- 17. Erion, J.L., Tarnowski, J., Weissbach, H. and Brot, N. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 3459-3463
- 18. Zhu, Y.S., Lovett P.S., Williams, D.M. and Kung, S.D. (1984) Theor. Appl. Genet. 67, 333-336
- 19. Zurawski, G. Perrot, B., Bottomley, W. and Whitfeld, P.R. (1981) Nucleic Acids Res. 9, 3251-3270
- 20. Crossland, L.D., Rodermel, S.R. and Bogorad, L. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 4060-4064
- 21. Shinozaki, K. and Sugiura, M. (1982) Gene 20, 91-102
- 22. Takaiwa, F. and Sugiura, M. (1980) Mol. Gen. Genet. 180, 1-4
- 23. Tohdoh, N., Shinozaki, K. and Sugiura, M. (1981) Nucleic Acids Res. 9, 5399-5406
- 24. Deno, H., Kato, A., Shinozaki, K. and Sugiura, K. (1982) Nucleic Acids Res. 10, 7511-7520
- 25. Steinmetz, A.A., Krebbers, E.T., Schwarz, Z., Gubbins, E.J. and Bogorad, L. (1983) J. Biol. Chem. 258, 5503-5511
- 26. Schwarz, Z., Kossel, H., Schwarz, E. and Bogorad, L. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 4748-4752
- 27. Kato, A., Shimada, H., Kusuda, M. and Sugiura, M. (1981) Nucleic Acid Res. 9, 5601-5607
- 28. Sugita, M., Kats, A., Shimada, H. and Sugiura, M. (1984) Mol. Gen. Genet. 194, 200-205
- 29. Krebbers, E., Steinmetz, A. and Bogorad, L. (1984) Plant Mole. Biol. 3, 13-20
- 30. Deno, H. and Sugiura, M. (1984) Proc. Natl. Acad. Sci. U.S.A. 81, 405-408
- 31. Zurawski, G., Bohnert, H.J. Whitfeld, P.R. and Bottomley, W. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 7699-7703
- 32. Zurawski, G., Bottomley, W. and Whitfeld, P.R. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 6260-6264
- 33. Shinozaki, K., Deno, H., Kato, A. and Sugiura, M. (1983) Gene 24, 147-155
- 34. Steimetz, A., Gubbins, E.J. and Bogorad, L. (1982) Nucleic Acids Res. 10, 3027-3037
- 35. Morrio, J. and Herrmann, R.G. (1984) Nucleic Acids Res. 12, 2837-2850
- 36. Takaiwa, F. and Sugiura, M. (1982) Nucleic Acids Res. 10, 2665-2676
- 37. Koch, W., Edwards K. and Kossel, H. (1981) Cell 25, 203-213
- 38. Kong, X.F., Lovett, P.S. and Kung, S.D. (1984) Gene 31, 23-30
- 39. Krebbers, E.T., Larrinua, I., McIntosh, L. and Bogorad, L. (1982) Nucleic Acids Res. 10, 4985-5002
- 40. Schwarz, Z., Jolly, S.O., Steinmetz, A.A. and Bogorad, L. (1981) Proc.

Natl. Acad. Sci. U.S.A. 78, 3423-3427

- 41. Keus, R.J.A., Dekker, A.F., van Roon, M.A. and Groot, G.S.P. (1983) Nucleic Acids Res. 11, 6465-6474
- 42. Sugita, M. and Sugiura, M. (1984) Mol. Gen. Genet. 195, 308-313
- 43. Keus, R.J.A., Roovers, D.J., Dekker, A.F. and Groot, G.S.P. (1983) Nucleic Acids Res. 11, 3405-3410
- 44. Spielmann, A. and Stutz, E. (1983) Nucleic Acids Res. 11, 7157-7167
- 45. Link, G. and Langridge, U. (1984) Nucleic Acids 12, 945-958
- 46. Ohme, M., Kamogashira, T., Shinozaki, K. and Sugiura, M. (1984) Nucleic Acids Res. 12, 6741-6749
- 47. Kung, S.D., Kong, X.F. and Gray, J.C. (in preparation)
- 48. Orozco, E.M. Jr., Hanley-Bowdoin, L. Poulsen, C. and Chua, N.H. (1984). Twenty-second Harden Conf. p. L27.
- 49. Goloubinoff, P., Edelman, M. and Hallick, R.B. (1984) Nucleic Acid Res. 24, 9489-9496
- 50. Shinozaki, K., Sun, C.R. and Sugiura, M. (1984) Mol. Gen. Genet. 197: 363-367
- 51. Ohme, M., Kamogashira, T., Shinozaki, K. and Sugiura, M. (1985) Nucleic Acids Res. 13: 1045-1056
- 52. Hanley-Bowdoin, L., Orozco, E.M. Jr. and Chua, N.H. (1985) Photosynthetic Apparatus, Cold Spring Harbor Press (in press)
- 53. Deno, H., Shinozai, K., and Sugiura, M. (184) Gene 32: 195-201
- 54. Kuntz, M., Weil, J.H., and Steinmetz, A. (1984) Nucleic Acids Res. 12: 5037-5047
- 55. Rasmussen, O.F., Bookjans, G., Stummann, B.M. an Henningsen, K.W. (1984) Plant Mol. Biol. 3: 191-199
- 56. Lin, C.M., Liu, C.C. and Kung, S.D. (1985) Plant Mole. Biol. (in press)
- 57. Link, G. (1984) EMBO J. 3: 1697-1704
- 58. Mullet, J.E., Orozeo, E.M., Jr. and Chua, N.-H (1985) Plant Mole. Biol. 4: 39-54.
- 59. Erion, J.L. (1985) Plant Mole. Biol. 4: 169-179.
- 60. Orozco, E.M., Jr., Mullet, J.E. and Chua, N.H. (1985) Nucleic Acids Res. 13: 1283-1302