Theor Appl Genet (1987) 73:357–363

Nucleotide sequence of the chloroplast gene responsible for triazine resistance in canola

M. Reith * and N. A. Straus

Department of Botany, University of Toronto, Toronto, Ontario M5S 1A1, Canada

Received June 3, 1986; Accepted July 25, 1986
Communicated by R. Hagemann

Summary. The nucleotide sequence for the *psbA* gene from a triazine resistant cultivar of *B. napus* (cv 'Triton') has been determined. This gene encodes an open reading frame of 353 amino acids that is highly homologous to other higher plant *psbA* genes at both the nucleotide and amino acid levels. As has been found for other triazine resistant *psbA* genes, the 'Triton' *psbA* contains an A to G nucleotide change which results in a serine to glycine amino acid substitution at position 264. The *B. napus psbA* gene also has a G insertion at position −9 resulting in a ribosome binding site sequence (AGGA) just before the initial methionine and suggesting that the entire open reading frame is translated. A large (72 bp) insertion is also found upstream of the *B. napus psbA* gene which resembles a similar insertion in the mustard *psbA*. The "uncloneable" nature of the entire gene is further investigated through reconstruction experiments and the implications discussed.

Key words: Triazine resistance – *psbA gene* – Nucleotide sequence – Canola

Introduction

The chloroplast *psbA* gene encodes a thylakoid protein of M_r 32,000 (Zurawski et al. 1982) that has been extensively characterized in many higher plants and algae (Hoffman-Falk et al. 1982; Delepelaire 1983, 1984; Mattoo et al. 1981). This photosystem II protein is abundantly synthesized but does not accumulate due to a rapid turnover rate (Mattoo et al. 1984; Wettern et al. 1983). It has been shown to be synthesized as a M_r 33,500–35,000 precursor (Grebanier et al. 1978; Reisfeld et al. 1982) which is processed at the carboxy terminus (Marder et al. 1984). Recent evidence suggests that this protein, with a bound plastoquinone molecule, functions as the secondary electron acceptor on the reducing side of the PS II reaction center (Velthuys 1981; Vermaas et al. 1983; Hirschberg et al. 1984). This protein has been designated the Q_B protein (Hirschberg et al. 1984).

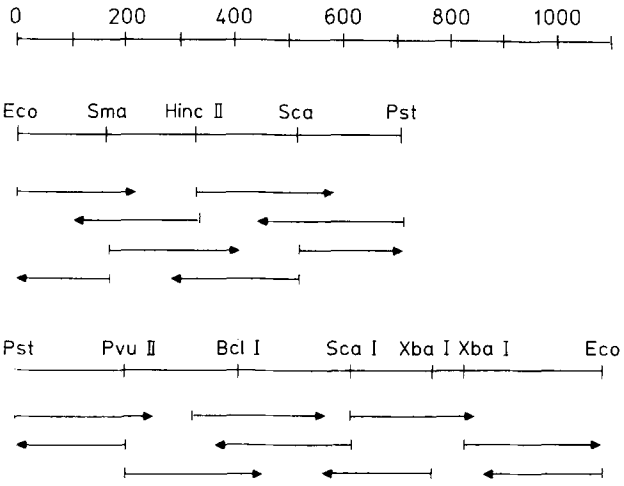

Fig. 1. Restriction map and sequencing strategy of the *B. napus* chloroplast DNA region containing the *psbA* gene. *Upper portion:* map of the pBNEP0.7 insert. Lower portion: map of the pBNEP1.1 insert. *Arrows* below each map indicate the direction and extent of DNA regions analyzed by the chain termination sequencing procedure. A size scale calibrated in base pairs is shown at the upper margin

* Present address: Division of Life Sciences, Scarborough Campus, University of Toronto, Scarborough, Ontario M1C 1A4, Canada

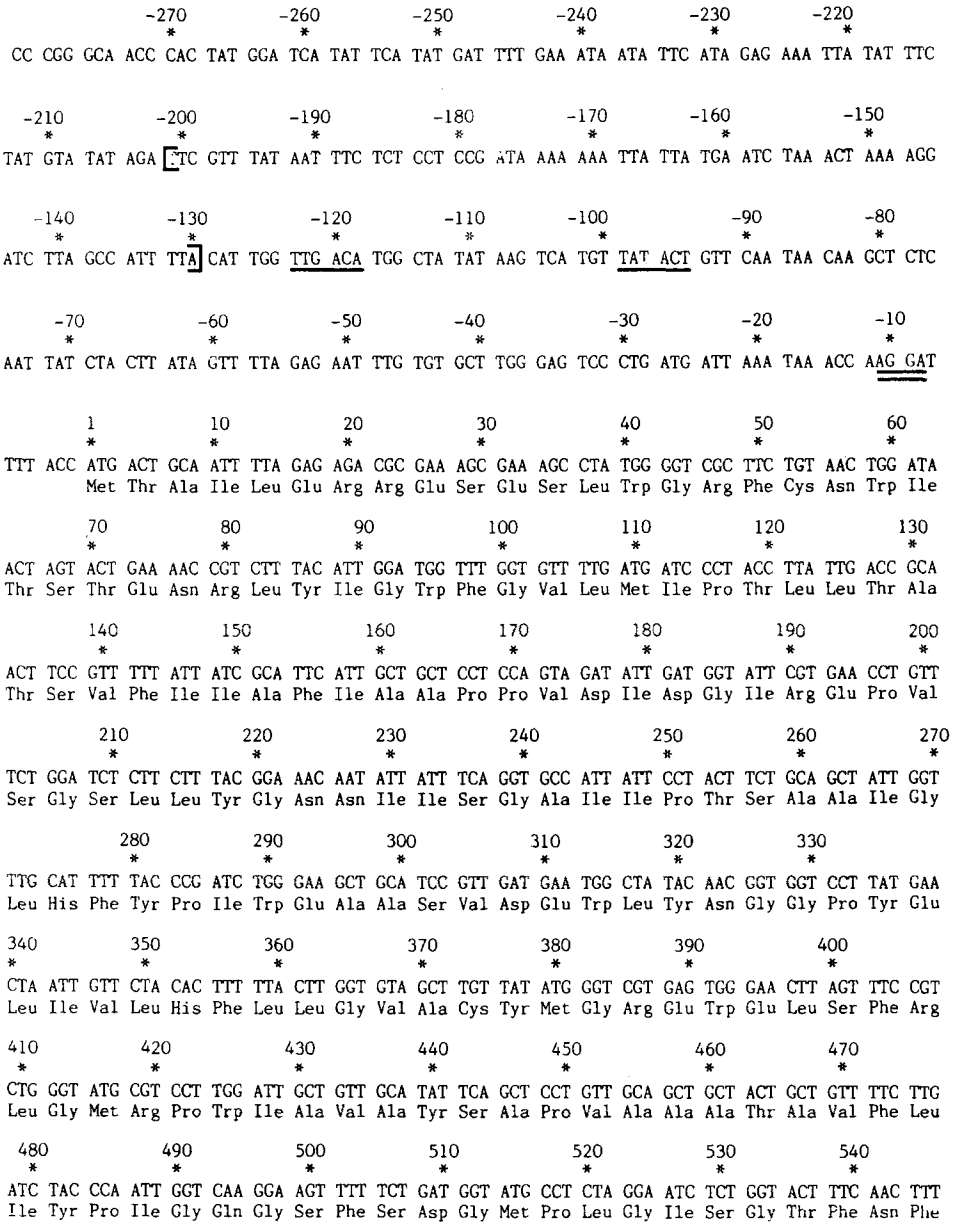

Fig. 2. Nucleotide sequence of the *B. napus psbA* gene. The glycine residue responsible for triazine resistance is boxed. Promoter regions are underlined. A ribosome binding site sequence is double underlined. The transcription termination structure is shown by the two arrows. The 72 bp insertion is bracketed

The *psbA* gene has now been sequenced from a number of higher plants and algae. Comparison of the deduced amino acid sequences from spinach, *Nicotiana debneyi* (Zurawski et al. 1982), *Amaranthus hybridus* (Hirschberg and McIntosh 1983), soybean (Spielmann and Stutz 1983), mustard (Link and Langridge 1984) and *N. tabaccum* (Sugita and Sugiura 1984) indicates an exceptionally high degree of conservation, with no more than three amino acid differences between these proteins. Algal (Karabin et al. 1984; Keller and Stutz 1984; Erickson et al. 1984) and cyanobacterial (Curtis and Haselkorn 1984; Mulligan et al. 1984; Golden and Haselkorn 1985) *psbA* genes show a similar conservation, having 87 to 93% homology with the higher plant genes.

The Q_B protein has been implicated as the primary target of several classes of herbicides (Pfister et al. 1981; Steinback et al. 1981). Presumably these herbicides displace the bound quinone molecule and thus block electron transport. Recently, several studies have correlated changes in the *psbA* gene with herbicide resistance. In herbicide resistant biotypes of *Amaranthus hybridus* (Hirschberg and McIntosh 1983) and *Solanum nigrum* (Hirschberg et al. 1984), a single A to G nucleotide replacement, which results in a serine to glycine amino acid change at amino acid 264 has been detected. In *Chlamydomonas reinhardtii* (Erickson et al. 1984; Erickson et al. 1985) and *Anacystis nidulans* R2 (Golden and Haselkorn 1985), a similar single base change at amino acid 264 (re-

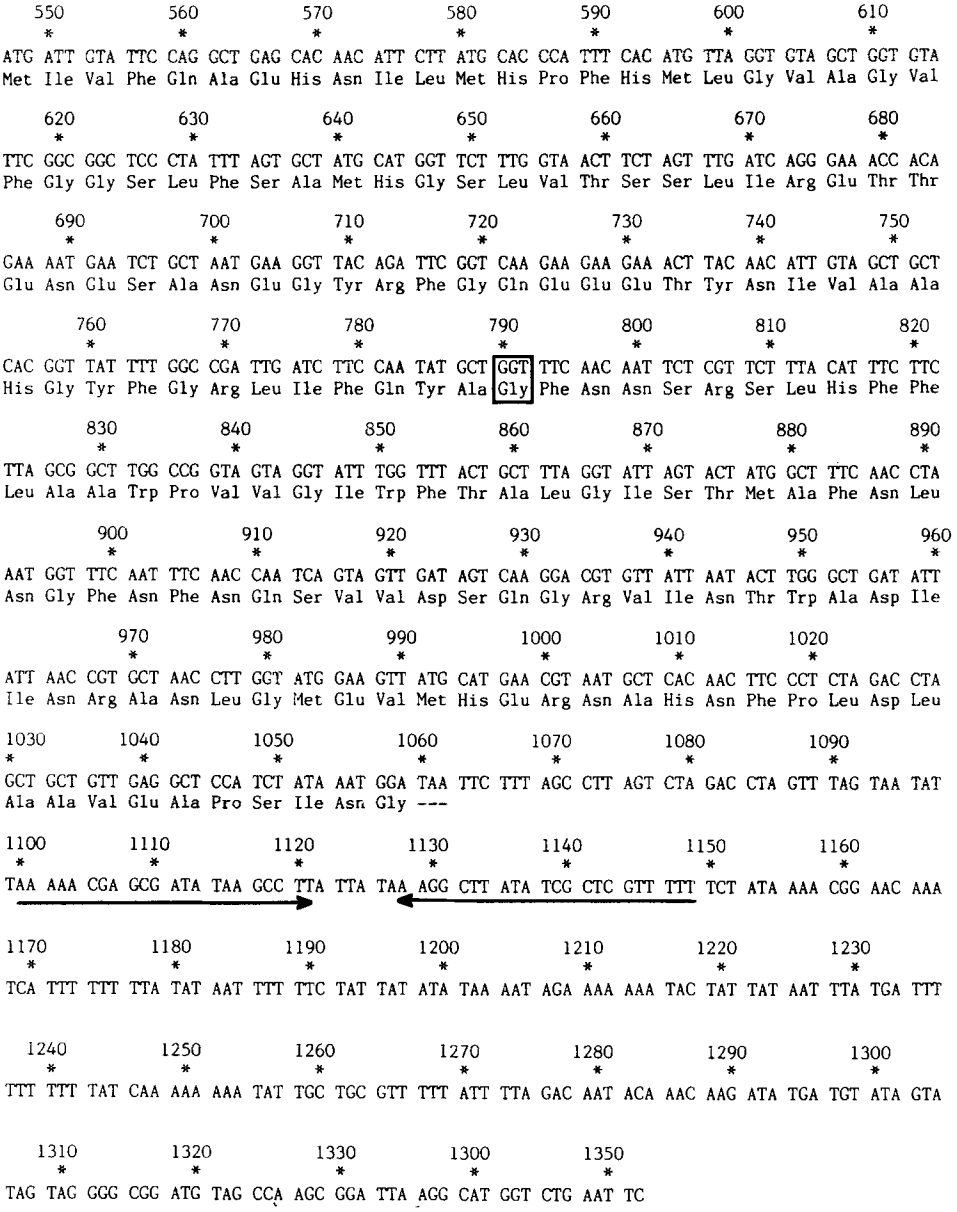

Fig. 2 (continued)

sulting in a serine to alanine conversion) is associated with resistance to triazine and diuron herbicides. In addition, in *Chlamydomonas,* two other mutations in *psbA* have been described at amino acids 219 and 255 which result in increased resistance to diuron and atrazine respectively (Erickson et al. 1985). By transforming wild-type *Anacystis nidulans* to herbicide resistance with small DNA fragments which carry the mutation at amino acid 264, Golden and Haselkorn (1985) have demonstrated that the mutation is indeed responsible for herbicide resistance.

In the last 20 years, triazine resistant biotypes of a number of weed species have developed in fields regularly treated with atrazine. A triazine resistant biotype of *Brassica campestris* (bird's rape) (Maltais and Bouchard 1978) has been used to introduce triazine resistance into agriculturally important *Brassica* cultivars (Beversdorf et al. 1980). The 'Triton' cultivar of *Brassica napus* (oilseed rape) is a triazine resistant strain of canola that is low in both erucic acid and thioglucosinolate. The chloroplast genome of this cultivar has been extensively mapped and several genes (including the *psbA* gene) located on the map (Xiao et al. 1986). In this communication, we report the nucleotide sequence of the *psbA* gene from *B. napus* (cv 'Triton') and demonstrate that this gene also contains the serine to glycine amino acid change at position 264.

Methods

Plasmids pBNP4 and pBNP8 have been previously shown to contain the entire *psbA* gene of *B. napus* (cv 'Triton') (Xiao et al. 1986). Plasmid pBNEP0.7 was constructed by subcloning a 0.7 kb EcoR I/Pst I fragment from pBNP4 into the plasmid pUC9 (Viera and Messing 1982). Likewise, pBNEP1.1 contains a 1.1 kb EcoR I/Pst I subclone from pBNP8. Plasmids were screened and isolated from *E. coli* strain JM103 as described by Ko et al. (1983). Both EcoR I/Pst I fragments were cloned into M13mp18 and mp19 (Norrander et al. 1983) and a variety of subclones were generated for DNA sequence analysis. The sequential deletion method (Dale et al. 1985) was used to produce subclones for sequencing certain regions of these fragments. Single stranded M13 DNA was isolated according to Dale et al (1985). Sequencing reactions were carried out using the dideoxy chain termination method (Sanger et al. 1977) with ^{35}S-dATP (Amersham) and a universal primer (PL Biochemicals). Sequencing reactions were electrophoresed on 0.4 mm thick 7 M urea/8% polyacrylamide gels and the results analyzed using a series of computer programs created by J. Pustell (Pustell and Kafatos 1984).

Results and discussion

Using heterologous hybridization with a *psbA* probe from mung bean, the *psbA* gene of *B. napus* (cv 'Triton') was mapped previously (Xiao et al. 1986). This gene is located in the large single copy region of the chloroplast genome, near the end of one of the inverted repeats. The *B. napus psbA* gene is contained within two Pst I fragments, Pst4 (15.9 kb) and Pst8 (2.2 kb). Subclones of these fragments were constructed in order to facilitate sequence analysis. Detailed physical maps of pBNEP0.7 and pBNEP1.1 are shown in Fig. 1, along with the sequencing strategy used for the analysis of these fragments.

The nucleotide sequence of the region beginning with the Sma I site of the pBNEP0.7 insert and ending at the EcoRI site of pBNEP1.1 is presented in Fig. 2. This sequence contains an open reading frame of 1,062 nucleotides coding for a 353 amino acid protein with a calculated molecular weight of approximately 39,000 d. Comparison of this sequence with those published for *psbA* genes from other higher plants reveals that the nucleotide homology among these genes is greater than 95%. Relative to the nucleotide sequence of the *psbA* gene from a closely related plant, white mustard (Link and Langridge 1984), the *B. napus* gene contains only ten nucleotide changes, resulting in two amino acid changes.

The first of these amino acid changes is the replacement of serine with glycine at amino acid 264 which is the result of a single A to G nucleotide change. An autoradiograph of a sequencing gel containing this portion of the sequence is shown in Fig. 3. This change is identical to those found in the triazine resistant

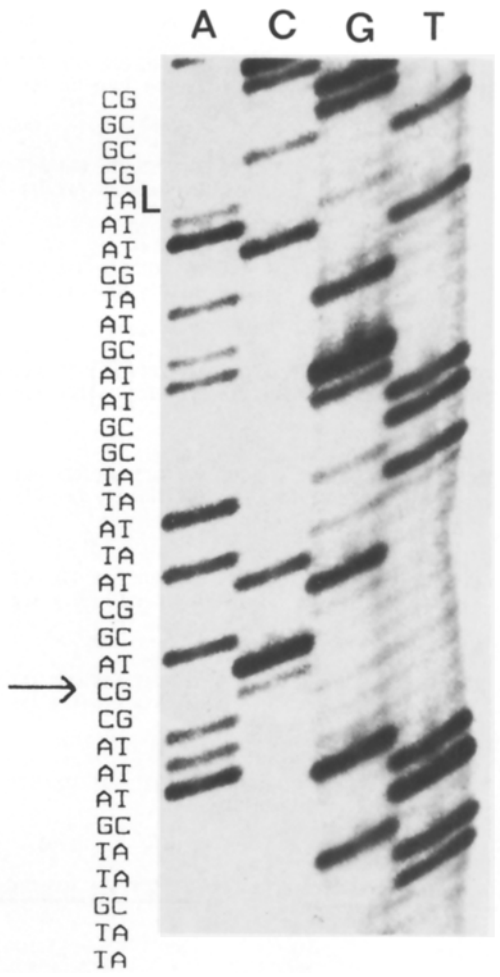

Fig. 3. Autoradiograph from a sequencing gel showing the A to G base change (*arrow*) which results in a GGT codon (glycine) at amino acid 264 and is responsible for triazine resistance

biotypes of *A. hybridus* and *S. nigrum* (Hirschberg and McIntosh 1983; Hirschberg et al. 1984).

The other amino acid difference between *B. napus* and mustard occurs at amino acid 348, near the carboxy terminus of the protein. Here, a serine is found in mustard, while an alanine is present in *B. napus* as well as spinach and tobacco. The mustard gene also contains two other amino acid changes in the carboxy portion of the protein when compared to spinach and tobacco (Link and Langridge 1984), both of which are present in *B. napus*. As the precursor protein is apparently processed at the carboxy end (Marder et al. 1984), it is likely that these differences are not present in the mature peptide.

The other eight nucleotide differences between mustard and *B. napus* are all silent with respect to amino acid sequence. One of these changes is somewhat unusual, however, in that it occurs in the first position

of the codon. Amino acid 106 (leucine) is encoded by TTA in mustard but CTA in *B. napus*.

Flanking sequences

An examination of the sequences flanking the *B. napus psbA* gene reveals the presence of a number of typical prokaryotic control regions. Centered at positions −120 and −96 are the sequences TTGACA and TATACT, respectively. Both of these sequences are highly homologous to the consensus *E. coli* promoter sequence (Rosenberg and Court 1979). The position of these regions suggests that transcription is initiated from approximately position −89, as has been demonstrated in spinach, tobacco and mustard (Zurawski et al. 1982; Link and Langridge 1984). Additionally, the sequence TATATAA (positions −113 to −107) which resembles the 'TATA' box sequence of the eukaryotic RNA polymerase II promoter (Breathnach and Chambon 1981) is present. Link (1984) has shown that this sequence promotes low levels of *psbA* transcription in vitro and has proposed that in etioplasts, in which low levels of *psbA* mRNA are found, this promoter, rather than the typical prokaryotic one, is active.

Centered at position −10 is the sequence AGGA (Fig. 2) which is homologous to the 3′ end of higher plant 16S rRNAs (Tohdoh and Sugiura 1982) and is an apparent ribosome binding site. This sequence is unique among *psbA* genes sequenced to date and arises due to the insertion of a G residue at position −9 (Fig. 4). The absence of an obvious ribosome binding site sequence before the first methionine of the open reading frame in other *psbA* genes has led some authors (Hirschberg and McIntosh 1983) to suggest that translation initiates at the second methionine (amino acid 37) with ribosome binding occuring at the sequence GGTG (positions 100 to 103). The size of the putative protein produced from this start site (317 amino acids, calculated molecular weight of 34,600 d) is in better agreement with that determined for the precursor protein on SDS polyacrylamide gels. Support for this hypothesis comes from dipeptide initiation studies which demonstrate that in pea, tobacco and maize, initiation begins much more frequently at the second methionine start site (Cohen et al. 1984). However, none of these plants contain the AGGA sequence found in *B. napus*. It would be interesting to determine how this sequence affects translation initiation in *B. napus*.

On the 3′ side of the *B. napus psbA* gene is a sequence (positions 1,100–1,149) capable of forming a large stem and loop structure. This structure would consist of a 22 base pair stem and a six base loop, with the stem ending in a run of six Ts. These characteristics are typical of a prokaryotic transcription termination signal (Rosenberg and Court 1979) and similar sequences have been found in all *psbA* genes sequenced to date.

Just 5′ to the −35 region of the *psbA* promoter is a relatively large insertion which is unique to the *B. napus* and mustard sequences (Link and Langridge 1984). In *B. napus* this insertion is 72 base pairs in length, two bp longer than the mustard insertion which also differs at five nucleotide positions. These insertions are capable of forming several stem and loop structures as demonstrated by Link and Langridge (1984). Since the expression of the *psbA* gene in mustard and *B. napus* is not apparently affected by these insertions, one must conclude that the approximately forty base pairs 5′ to the transcriptional start site contain all the appropriate regulatory sequences for this gene. This is

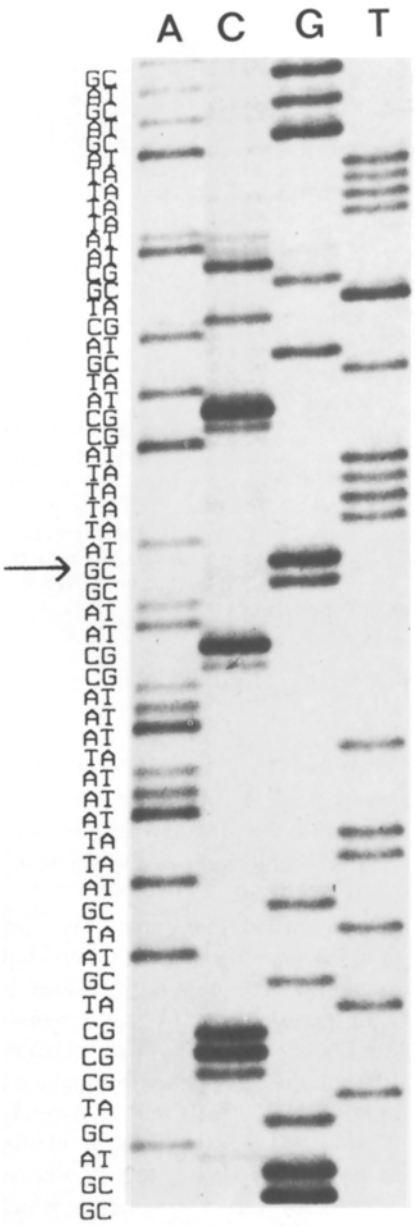

Fig. 4. Autoradiograph of a sequencing gel showing the insertion of a G at position −9 (*arrow*) which generates a typical ribosome binding site (AGGA)

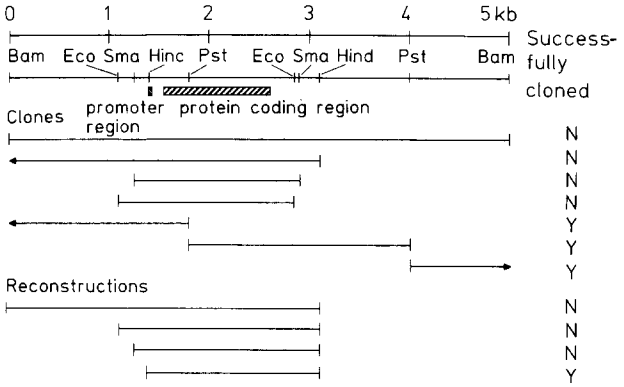

Fig. 5. Summary of attempts at cloning the *B. napus psbA* gene as a single restriction fragment. Cloning was attempted by isolating chloroplast DNA fragments from low melting temperature agarose gels and cloning into pBR322 or pUC9. Reconstructions were attempted by first cloning a fragment of pBNP4 (Bam/Pst, Eco/Pst, Sma/Pst or HincII/Pst) into the polylinker of pUC9. The Pst/HindIII fragment of pBNP8 was then inserted into each of these clones

consistent with the proposal of Link (1984) which suggests that a promoter switch between the eukaryotic promoter (TATA box) and the prokaryotic promoter (−35 and −10 regions) is the mechanism by which this gene is developmentally regulated.

We have recently described the apparently "uncloneable" nature of the *B. napus psbA* gene (Xiao et al. 1986). Other investigators have noted difficulties in cloning fragments with the *psbA* gene from *N. otophora* (Zhu et al. 1982) and *Vicia faba* (K. Ko, personal communication), although the *psbA* genes from spinach and *N. debneyi* (Zurawski et al. 1982) and *A. hybridus* (Hirschberg and McIntosh 1983) have been cloned successfully in their entirety. To date, the cloning of BamHI, Hind III, EcoRI or Sma I fragments which contain the complete *B. napus psbA* gene has been unsuccessful (Fig. 5). Consequently, reconstructions of the entire gene from the two Pst I clones (pBNP4 and pBNP8) were attempted to define the lethal nature of this gene. It is evident from these experiments that the entire *psbA* coding region can be cloned only if the promoter region is disrupted in the process. The high degree of homology of the *psbA* promoter with the *E. coli* consensus promoter indicates that it is likely to be very highly transcribed when cloned into *E. coli*. Perhaps the presence of a typical prokaryotic ribosome binding site also results in increased translational efficiency in *E. coli*. Thus, in our hands, the destruction of the promoter region is necessary to clone the *B. napus psbA* gene, suggesting that the expression of this gene in *E. coli* is lethal to the bacterium.

Acknowledgements. We would like to thank B. Wong and Z. Ko for expert technical assistance. This research was supported by grants from the National Science and Engineering Research Council (NSERC) and a University of Toronto Connaught Special Research Grant to N.A.S.

References

Beversdorf WD, Weiss-Lerman J, Erickson LR, Souza-Machado VS (1980) Transfer of cytoplasmically-inherited triazine resistance from bird's rape to cultivated oilseed rape (*Brassica campestris* and *B. napus*). Can J Genet Cytol 22: 167–172

Breathnach R, Chambon P (1981) Organization and expression of eucaryotic split genes coding for proteins. Annu Rev Biochem 50:349–383

Cohen BN, Coleman T, Schmitt JJ, Weissbach H (1984) Expression and characterization of the translation start site of the *psbA* gene product (Q_B protein) from higher plants. Nucleic Acids Res 12:6221–6234

Curtis S, Haselkorn R (1984) Isolation, sequence and expression of two members of the 32 KD thylakoid membrane protein gene family from the cyanobacterium *Anabaena* 7120. Plant Mol Biol 3:249–258

Dale RMK, McClure BA, Houchins JP (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: Application to sequencing the corn mitochondrial 18 S rDNA. Plasmid 13:31–40

Delepelaire P (1983) Characterization of additional thylakoid membrane polypeptides synthesized inside the chloroplast in *Chlamydomonas reinhardii*. Photochem Photobiophys 6: 279–292

Delepelaire P (1984) Partial characterization of the biosynthesis and integration of the photosystem II reaction centers in the thylakoid membrane of *Chlamydomonas reinhardii*. EMBO J 3:701–706

Erickson JM, Rahire M, Rochaix JD (1984) *Chlamydomonas reinhardii* gene for the 32,000 mol. wt. protein of photosystem II contains four large introns and is located entirely within the chloroplast inverted repeat. EMBO J 3:2753–2762

Erickson JM, Rahire M, Rochaix JD, Mets L (1985) Herbicide resistance and cross-resistance: changes at three distinct sites in the herbicide-binding protein. Science 228:204–207

Golden SS, Haselkorn R (1985) Mutation to herbicide resistance maps within the *psbA* gene of *Anacystis nidulans* R2. Science 229:1104–1107

Grebanier A, Coen D, Rich A, Bogorad L (1978) Membrane proteins synthesized but not processed by isolated maize chloroplasts. J Cell Biol 78:734–756

Hirschberg J, McIntosh L (1983) Molecular basis of herbicide resistance in *Amaranthus hybridus*. Science 222:1346–1349

Hirschberg J, Bleecker A, Kyle DJ, McIntosh L, Arntzen CJ (1984) The molecular basis of triazine-herbicide resistance in higher plant chloroplasts. Z Naturforsch 39c:412–419

Hoffman-Falk H, Mattoo AK, Marder JB, Edelman M, Ellis RJ (1982) General occurrence and structural similarity of the rapidly synthesized 32,000 dalton protein of the chloroplast membrane. J Biol Chem 257:4583–4587

Karabin G, Farley M, Hallick RB (1984) Chloroplast gene for M_R 32,000 polypeptide of photosystem II in *Euglena gracilis* is interrupted by four introns with conserved boundary sequences. Nucleic Acids Res 12:5801–5812

Keller M, Stutz E (1984) Structure of the *Euglena gracilis* chloroplast gene (*psbA*) coding for the 32 KDa protein of photosystem II. FEBS Lett 175:173–177

Ko K, Straus NA, Williams JP (1983) Mapping the chloroplast DNA of *Vicia faba*. Curr Genet 7:255–263

Link G (1984) DNA sequence requirements for the accurate transcription of a protein-coding plastid gene in a plastid in vitro system from mustard (*Sinapis alba* L.). EMBO J 3: 1697–1704

Link G, Langridge U (1984) Structure of the chloroplast gene for the precursor of the M_r 32,000 photosystem II protein from mustard (*Sinapis alba* L.). Nucleic Acids Res 12: 945–958

Maltais B, Bouchard CJ (1978) Une moutarde des oiseaux (*Brassica rapa* L.) resistante a l'atrazine. Phytoprotection 59:117–119

Marder JB, Goloubinoff P, Edelman M (1984) Molecular architecture of the rapidly metabolized 32 kilodalton protein of photosystem II. Indications for COOH terminal processing of a chloroplast membrane polypeptide. J Biol Chem 259:3900–3908

Mattoo A, Pick U, Hoffman-Falk H, Edelman M (181) The rapidly metabolized 32,000 dalton polypeptide of the chloroplast is the proteinaceous shield regulating photosystem II electron transport and mediating diuron herbicide sensitivity. Proc Natl Acad Sci USA 78:1572–1576

Mattoo A, Hoffman-Falk H, Marder JB, Edelman M (1981) Regulation of protein metabolism – coupling of photosynthetic electron transport to in vivo degradation of the rapidly metabolized 32 kilodalton protein of the chloroplast membranes. Proc Natl Acad Sci USA 81:1380–1384

Mulligan B, Schultes N, Chen L, Bogorad L (1984) Nucleotide sequence of a multiple copy gene for the B protein of photosystem II of a cyanobacterium. Proc Natl Acad Sci USA 81:2693–2697

Norrander J, Kempe T, Messing J (1983) Construction of improved M13 vectors using oligonucleotide directed mutagenesis. Gene 26:101–106

Pfister K, Steinback KE, Gardner G, Arntzen CJ (1981) Photoaffinity labeling of an herbicide receptor protein in chloroplast membranes. Proc Natl Acad Sci USA 78: 981–985

Pustell J, Kafatos FC (1984) A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis and homology determinations. Nucleic Acids Res 12:643–655

Reisfeld A, Mattoo A, Edelman M (1982) Processing of a chloroplast translated membrane protein in vivo. Analysis of the rapidly synthesized 32,000 dalton shield protein and its precursor in *Spirodela oligorrhiza*. Eur J Biochem 124: 125–129

Rosenberg M, Court D (1979) Regulatory sequences involved in the promotion and termination of RNA transcription. Annu Rev Genet 13:319–353

Sanger F, Nicklen S, Carlson AR (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74:5463–5467

Spielmann A, Stutz E (1983) Nucleotide sequence of soybean chloroplast DNA regions which contain the *psbA* and *trnH* genes and cover the ends of the large single copy region and one end of the inverted repeats. Nucleic Acids Res 11: 7157–7167

Steinback K, McIntosh L, Bogorad L, Arntzen CJ (1981) Identification of the triazine receptor protein as a chloroplast gene product. Proc Natl Acad Sci USA 78:7463–7467

Sugita M, Sugiura M (1984) Nucleotide sequence and transcription of the gene for the 32,000 dalton thylakoid membrane protein from *Nicotiana tabacum*. Mol Gen Genet 195:308–313

Tohdoh N, Sugiura M (1982) The complete nucleotide sequence of a 16S ribosomal RNA gene from tobacco chloroplasts. Gene 17:213–218

Velthuys BR (1981) Electron-dependent competition between plastoquinone and inhibitors for binding to photosystem II. FEBS Lett 126:277–281

Vermaas W, Arntzen C, Gu L-Q, Yu C-A (1983) Interactions of herbicides and azidoquinones at a photosystem II binding site in the thylakoid membrane. Biochim Biophys Acta 723:266–275

Vieira J, Messing J (1982) The pUC plasmids: an M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene 19:259–268

Wettern M, Owens JC, Ohad I (1983) Role of thylakoid polypeptide phosphorylation and turnover in the assembly and function of photosystem II. Methods Enzymol 97: 554–567

Xiao W, Reith M, Erickson LR, Williams JP, Straus N (1986) Mapping the chloroplast genome of triazine resistant canola. Theor Appl Genet 71:716–723

Zhu YS, Duvall EJ, Lovett PS, Kung SD (1982) *Nicotiana* chloroplast genome. 5. Construction, mapping and expression of a clone library of *N. otophora* chloroplast DNA. Mol Gen Genet 187:61–66

Zurawski G, Bohnert HJ, Whitfeld PR, Bottomley W (1982) Nucleotide sequence of the gene for the M_R 32,000 thylakoid membrane protein from *Spinacia oleracea* and *Nicotiana debneyi* predicts a totally conserved primary translation product of M_R 38,950. Proc Natl Acad Sci USA 79:7699–7703