Volume 13 Number 11 1985 Nucleic Acids Research

Conservation and variability of wheat α/β-gliadin genes

Martin Sumner-Smith[1], J.Antoni Rafalski[2], Takashi Sugiyama, Mario Stoll[3] and Dieter Söll

Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06511, USA

Received 28 March 1985; Accepted 6 May 1985

ABSTRACT

We have sequenced two genomic clones for wheat α/ß-gliadin storage protein genes. Comparison with a known sequence reveals close homology between the three and confirms the previously suspected evolutionary relatedness of members of this gliadin family. The coding region can be divided into six domains. Two unusual structures were found within this region: (i) The P-boxes which are composed of 12 codons, six of which are for proline, that are tandemly repeated four or five times; and (ii) Two polyglutamine stretches which consist of 18-22 tandemly repeated glutamine codons in one case, and 7-28 in the second. Analysis of the P-box structures revealed that certain mutations were probably present in the hypothetical ancestral α/ß-gliadin gene prior to gene multiplication. None of the genes have introns. All of the genes appear to contain typical eukaryotic promoters and also possess the double polyadenylation signal of plants.

INTRODUCTION

During wheat seed development the predominant protein synthesis is of two groups of proteins, totalling more than 50 members, which are thought to provide a stored source of nitrogen for future germination. These storage proteins, the gliadins and glutenins, have been the subject of extensive study.

Originally, the gliadins were classified according to their electrophoretic mobility in starch gels in aluminum lactate (1). Recently, they have been reclassified according to the size, amino acid composition and N-terminal sequences of purified species, into the predominant sulfur-rich α/ß- and γ-gliadins, and the less abundant sulfur-poor omega-gliadins (2,3). Gliadins are very rich in glutamine (approx. 35%) and proline (approx. 15%).

On the basis of the apparent homology between the N-terminal sequences of purified members of each gliadin class, it is thought that the gliadins are the products of several multigene families (3-5). These families presumably arose by the repeated duplication of a few ancestral genes. Since all of the multigene families seem to be present in each of the ancestral genomes which have contributed to modern hexaploid wheat (6-8), multiplication of the original

gliadin genes must have occurred in some ancestor common to the diploid strains.

Three gliadin gene loci have been identified by genetic means, two map to the short arm of homoeologous chromosome group 1 and one to group 6 (9-14). Individual genes within these loci are tightly linked (14-17).

Wheat storage proteins represent a convenient system to study both the coordinate expression of several gene families during development and also the evolution of these families. We have previously presented the complete sequence of an α/β-gliadin gene and its flanks (18). Here we report the sequence of two additional genomic clones, discuss a domain structure for gliadin proteins and attempt insights into the evolution of their genes.

MATERIALS AND METHODS

Materials: Klenow fragment of E. coli DNA polymerase I was a gift of Nancy Templeton, Yale University. Other materials were obtained commercially.

General: Handling and analysis of nucleic acids, including restriction enzyme digestions, agarose electrophoresis and elution, Southern blots, ligations, plasmid and phage DNA isolation were by established methods (19). Bacterial transformation was by the method of Hanahan (20).

Gliadin Clones: The gliadin genes selected here were selected from a Wheat (cv. Yamhill) partial EcoR1 library in Charon 32 (courtesy of Drs. J. Slightom and M. Murray, Agrigenetics Corp., Madison, WI) and recloned into pBR325 (21) in both orientations. The complete sequence of pW8233 has already been described (18). Restriction maps were determined for pW8142 and pW1215 (now shown) and the gliadin gene localized by Southern hybridization. pW8142 contains a 7.7 kb fragment, within which is a 3231 bp EcoRI-HindIII subfragment carrying the gene. pW1215 has a 9.8 kb fragment within which a 3043 bp HindIII-HindIII subfragment contains the gene. The two subfragments were completely sequenced. In addition, the ends of the flanking subfragments were sequenced. Data not shown in Fig. 1 were submitted to GenBankTm.

Sequencing: Both strands of the gliadin gene containing subfragments of pW8142 and pW1215 were determined by the dideoxy method using M13mp8, M13mp9, M13mp10 and M13mp11 and DNA fragments generated by a variation of the Bal31-deletion method (22,23). Sequence data were compiled and analyzed using the programs described by Larson and Messing (24) and by Sege et al. (25).

Nuclease S1 Mapping of the 5'-End of α/β-Gliadin mRNA: For all three genes examined, complementary probes (coding strand) were made by universal oligonucleotide (17-mer) primed synthesis on appropriate Bal31-deleted templates cloned in M13. Sequencing conditions were used except that dideoxynucleotides

were omitted. The resulting partially double-stranded molecules were cleaved with PstI (after nucleotide 653, see Fig. 1). Restricted probes were recovered after phenol extraction, denatured by boiling in 80% formamide and hybridized overnight with an excess (10 µg) of wheat endosperm polyA+ RNA (courtesy of K. Scheets, Kansas State University) according to the conditions (80% formamide, 0.4 M NaCl, 0.04 M Pipes pH 6.4, 1 mM EDTA) of Weaver and Weissmann (29) at 53° under paraffin oil. Controls contained no polyA+ RNA. After hybridization, samples were added to 200 µl S1 buffer (0.25 M NaCl, 30 mM sodium acetate pH 4.6, 1 mM $ZnSO_4$, 20 µg/µl denatured salmon sperm DNA) and incubated at 30° M. DNA was recovered by ethanol precipitation and run on an 8% sequencing gel alongside a set of sequencing reactions as a length reference.

RESULTS AND DISCUSSION

Gene Sequences: The sequences of the three genes and their immediate flanks are shown in Fig. 1. The predicted N-terminal amino acid sequences and compositions are consistent with those of α/ß-gliadin genes (3). The genes are clearly related, but show mutational differences at a number of sites as well as changes which could have arisen by insertions and deletions that preserve the ready frames.

The 5'-flanks of the genes are homologous for approximately 600 bp upstream of the ATG start codon. The sequences diverge 20-30 bp upstream of the HindIII site (Position 1 in Fig. 1). Clones pW8142 and pW1215 share homologous 3'-flank for at least 1600 bases (data not shown), but these are not related to the 3'-flank of pW8233 beginning at the position 1680. Comparison with 3'-noncoding regions of barley B1 hordein cDNA clones (30) revealed a close homology that extends downstream from the translation stop codon to the second polyadenylation signal. The spacing of this polyadenylation signal and of the stop codon is also conserved between the hordein clones and pW8233/pW8142 (except for a two bp deletion). This homology complements the observed similarity in the coding sequences of hordein and gliadin genes (see below and refs. 30-32). Close homology of zein genes has been reported (26, 27).

The ends of the mRNA: All three genes possess a typical eukaryotic RNA polymerase II promoter sequence (TATAAAA/TA) 104 bases upstream of the ATG start codon. We determined the 5'-end of the α/ß-gliadin mRNA species by nuclease S1 protection studies (Fig. 2). Subtracting 1-2 bases to account for the putative 5'-cap, we estimate that transcription in vivo begins 30 bases downstream from the end of the TATA-box at the indicated A (Fig. 1). The 3'-flank contains two potential polyadenylation signals (AATAAA/T and AATAAA)

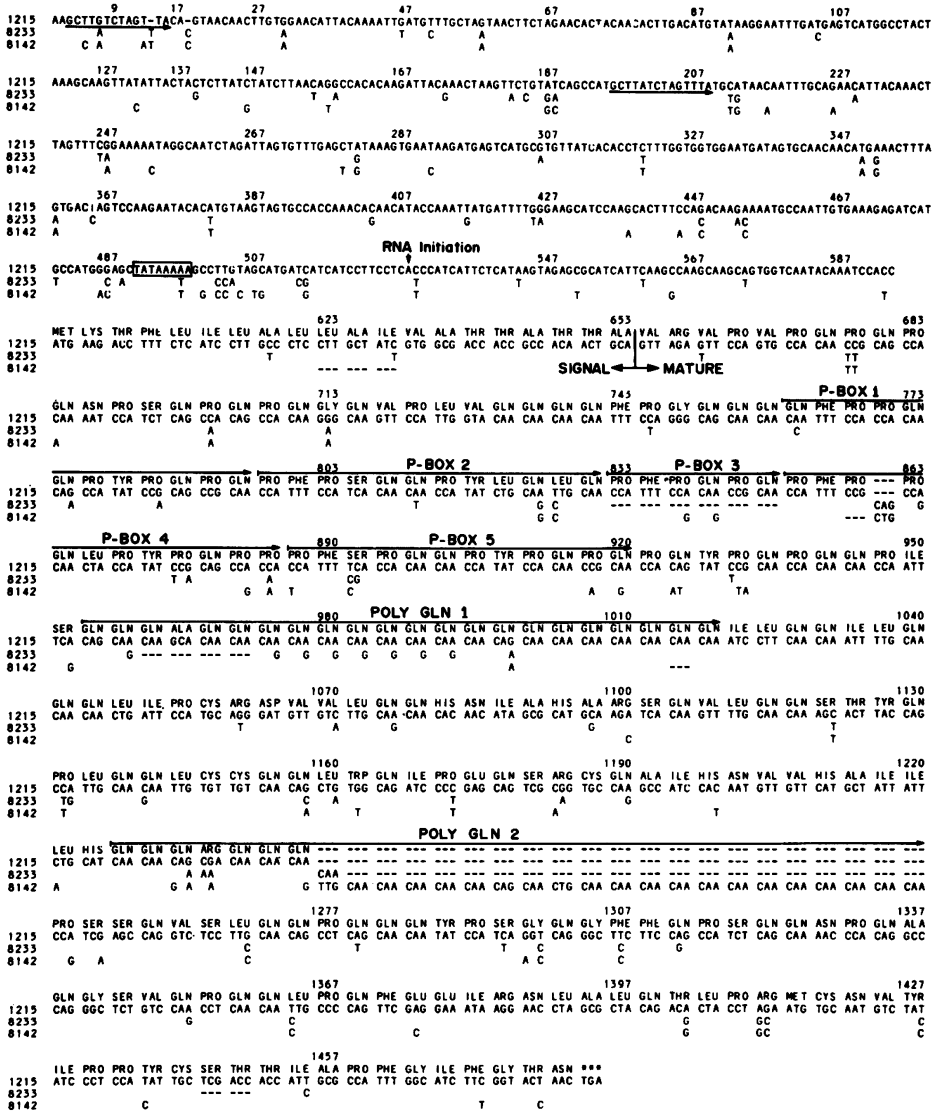


Figure 1. The DNA sequence and the derived protein sequence of three α/ß-gliadin genes. The sequenced region of the clone pW1215 is presented (upper line); bases in clones pW8233 and pW8142 which differ from those are indicated below; deleted bases are indicated by a dash. The numbering is

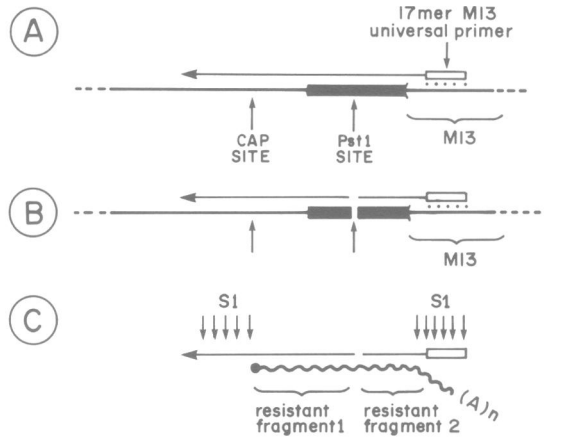
based on the pW1215 sequence. The predicted amino acid sequence of the product of pW1215 is indicated. The position of the P-boxes (see text) is indicated as are the polyglutamine stretches. The direct repeat in the 5'-flank (see text) is indicated by arrows. The TATA box and polyadenylation signals are boxed. The RNA initiation (arrow) and the polyadenylation (*) site are given.

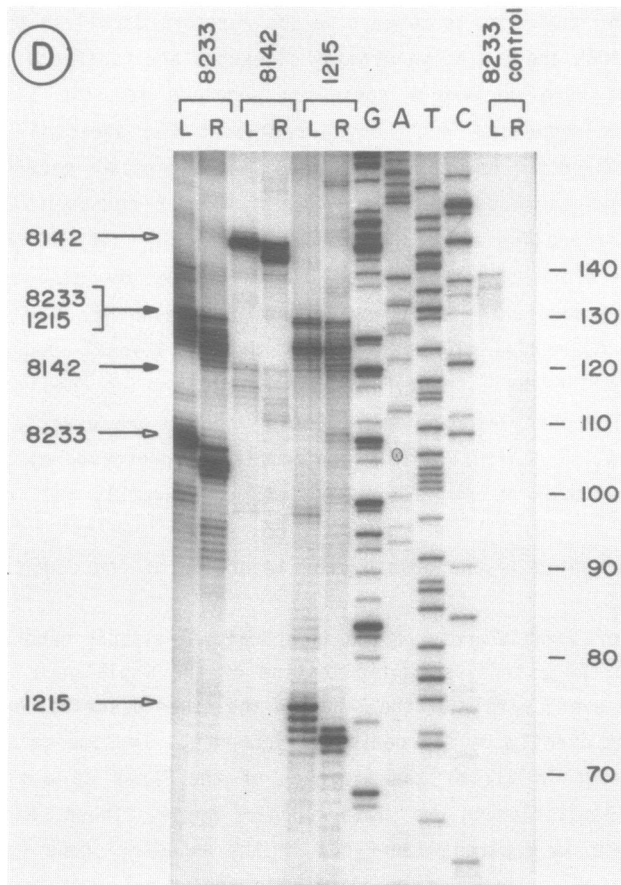
(Fig. 1); sequencing of four cDNA clones shown that the poly A tail starts at position 1625 (Fig. 1) (18).

Domain Structure of the Coding Region: Fig. 3 shows a generalized structure of wheat α/ß-gliadin genes derived from the sequences of the three genomic clones. The coding region can be divided into six domains: a signal peptide, a region of nine dodecapeptide repeats, five of which show very close homology (the P-boxes), two polyglutamine stretches and two regions of unique sequence. A similar structure has recently been proposed (31); however, since we have a larger number of sequences to compare we observe more detail in this structure. The repeat regions and the polyglutamine stretches are further discussed below.

P-box: We have derived a consensus sequence for the 12 codon repeat (Fig. 4). This sequence, which might represent the ancestral gene, is the sequence from which the fewest base changes (mutations) are necessary to arrive at the actually observed sequences. Six of the 12 codons in the consensus sequence are for proline and represent the greatest density of these codons in the genes; therefore the designation P-box. There are also four glutamine codons and one codon each for tyrosine and phenylalanine. While all but one each of the proline and glutamine codons show mutational changes in one or more P-box examples, both the tyrosine and phenylalanine codons are unchanged in any box. The phenylalanine and tyrosine codons are exactly half the box size apart (i.e. six codons). This periodicity is preserved by three deletion variants of the third P-box, i.e. by deletion of exactly half of the box in two cases (pW8142, pW1215) and the entire box in one (pW8233). The periodicity is disrupted slightly in one case by the insertion of one codon in the fourth P-box of pW8233.

The P-box presumably arose in the ancestral α/ß-gliadin gene and was multiplied prior to the extensive multiplication of the whole gene. Base changes are present in every sample of the P-box in the genes described here; no single box corresponds exactly to the consensus sequence. In some cases these mutations are present in all of the examples of the P-box at a given position, e.g. the A to T mutation in the sixth codon of Box 4. These 'early' mutations presumably arose in a particular P-box in the ancestral gene (or at least in an ancestor to all of the genes described here), and were preserved during


A
17mer M13
universal primer
CAP
SITE
Pst1
SITE
M13
B
M13
C
S1
S1
(A)n
resistant
fragment1
resistant
fragment 2


D
8233
8142
1215
8233
control
L R L R L R G A T C L R
8142
8233
1215
8142
8233
1215
140
130
120
110
100
90
80
70

gene multiplication. It might be argued that these mutations could have arisen in one gene of the repeated family and been spread to the others by unequal crossing over; however such a process would be more likely to transfer a given mutation to other P-boxes within the same gene, and there is little evidence for this.

Recently, the sequences of two complete α/ß-gliadin cDNA clones (pGliA-42 and pCH1941) from different cultivars of T. aestivum were elucidated (31,33). Comparison of these sequences in the P-box region reveals that the seven early mutations in the five boxes are present as expected (Fig. 4). Further, amino acid sequence analysis (31) of a mixed α-gliadin fraction confirms that the four non-silent 'early' mutations are present in the five or more polypeptides which make up that fraction.

The significance of the P-box organization is unknown at this time. Possibly these peptides may confer on the proteins a structure important for their function in vivo. A γ-gliadin cDNA sequence containing the 5'-half of the mature coding region shows 14 repeats of a 7 codon sequence (28). In the case of corn, certain zein genes show a 20 codon sequence tandemly repeated nine times that make up the bulk of the final polypeptide (26,27,34). It has been suggested that each example of this repeat is able to assume an α-helical structure and that the nine resulting helices are able to stack side-by-side (35).

The Polyglutamine Stretches: A second interesting feature is the presence of two long polyglutamine stretches (Figs. 1 and 4). The first one is found near the center of each gene and consists of from 18-22 codons. In two of the genes this stretch is composed of a single CAG codon followed by either 20

Figure 2. Nuclease S1 mapping of the 5'-end of α/ß-gliadin mRNA. DNA complementary to mRNA was synthesized from appropriate BAL-31 deletion clones in M13 from each of the genes (A, top legend) and cleaved by PstI to produce uniform 3'-end (B). Two fragments of this DNA were protected from digestion by nuclease S1 (reaction times 5 min [lanes L] and 40 min [lanes R]) after hybridization to mRNA (C,D): 1.) The first fragment extends from the 5'-end of mRNA to the Pst site (D, filled arrows). The length of this fragment establishes the location of the cap site, in basepairs, upstream from the PstI site (position 653). 2.) The fragment extends from the PstI site to the end of the deleted clone (D, open arrows). The length of this fragment was different for each clone, because the deletions used were different. The sizes of the fragments were determined by comparison to sequencing reactions performed on a known clone run in adjacent lanes; calculated lengths are given in the right hand column of D. The size of the 5'-fragment from pW8142 was nine bases smaller than those for the other two genes because of a three codon deletion in the signal peptide of that gene. Control reactions, lacking mRNA were performed for all three clones - that for the pW8233 subclone is shown. The origin of the faint bands in L is unclear. For further details see Experimental Section.

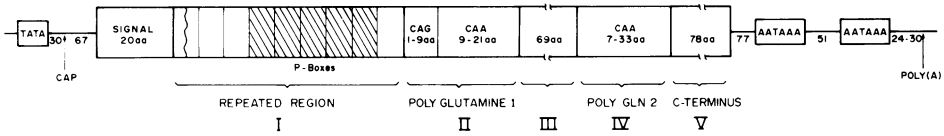


Figure 3. General structure of α/ß-gliadin genes. The generalized structure is derived from the three genomic clones described here as well as from the cDNA clones described by Kasarda et al. (31) and Proffitt et al. (30). The TATA box, CAP site, polyadenylation signals and site are found the indicated number of bases away from the coding sequence shown as an enlarged box. The signal sequence precedes five regions in the mature polypeptide defined by analysis of the DNA sequences. The first region (I) consists of a series of 9, typically dodecapeptide repeats. Five of these repeats (crosshatched) appeared to be more closely related to each other (P-boxes, Fig. 4). The first repeat, which is partial, is preceded by a three codon stretch having no obvious relationship to the rest of the region. Two polyglutamine stretches (II and IV) separate two regions of non-repeated sequence (V).

(pW8142) or 21 (pW1215) CAA codons, whereas in the other the stretch consists of nine CAG codons followed by nine CAA codons (pW8233). These stretches are actually disrupted by a putative mutant GCA codon at the fourth position in both pW8142 and pW1215, as well as a silent A to G mutation in the fifteenth position of pW1215. It therefore appears that the polyglutamine stretch was initially generated by multiplication of a CAA codon and subsequently in some genes (e.g. pW8233) by multiplication of a CAG codon found immediately in front of the CAA stretch. The latter event might therefore be more recent. Alternatively, the poly CAG stretch may have been reduced to a single codon in some genes. Another possible mechanism to account for the CAG to CAA transition derives from the fact that CG and CNG sequences in wheat DNA are over 80% methylated to m5C (36). Should 5-methylcytosine suffer spontaneous deamination to an appreciable extent (37), CAG would tend to be converted to CAA. Deamination of C in the first position of CAG would result in a nonsense codon. Thus, a selective change of CAG to CAA could be explained.

pW8142 shows a second polyglutamine stretch later in the gene consisting of 28 codons, 3 of which are mutated away from glutamine. A similar but shorter stretch is located at the corresponding position (base 1227 in pW1215) in the other two genes described here, which consists of 7 (pW1215) or 8 (pW8233) codons, one of which is mutated away from glutamine. In pGliA-42 this stretch has 33 glutamine codons (31).

Evolutionary Implications: Since the α/ß-gliadin genes in the three diploid genomes which contribute to the hexaploid genome of modern wheat are present

Consensus Sequence	Pro CCA	Phe TTT	Pro CCA	Pro CCA	Gln CAA	Gln CAA	Pro CCA	Tyr TAT	Pro CCG	Gln CAG	Pro CCG	Gln CAA	Gene
Box 1	-A-	---	---	---	---	--g	---	---	---	---	---	---	pW1215
	---	---	---	---	---	---	---	---	--a	---	---	---	pW8233
	-A-	---	---	---	---	--g	---	---	---	---	---	---	pW8142
	-A-	---	---	---	---	--g	---	---	--a	---	---	---	pGliA-42
	---	---	---	---	---	---	---	---	--a	---	---	---	pCH1941
Box 2	---	---	---	T--	---	---	---	---	-T-	--a	TT-	---	pW1215
	---	---	---	T--	---	-T-	---	---	-T-	---	-T-	---	pW8233
	---	---	---	T--	---	---	---	---	-T-	---	-T-	---	pW8142
	--t	---	---	T--	---	---	---	---	-T-	---	-T-	---	pGliA-42
	---	---	---	T--	---	---	---	---	-Ta	---	-T-	---	pCH1941
Box 3	---	---	---							--a	---	---	pW1215
													pW8233
	---	---							---	---	---	---	pW8142
	---	---							---	---	---	---	pGliA-42
													pCH1941
Box 4	---	---	--g	---	---	-T-	---	---	---	---	--a	-C-	pW1215
	---	---	--g*--g		---	-T-	---	---	T-a	---	--a	---	pW8233
	---	---	-Tg	---	---	-T-	---	---	---	---	---	---	pW8142
	---	---	--g	---	---	-T-	---	---	---	---	---	---	pGliA-42
	---	---	--g*--g		---	-T-	---	---	T--	---	--a	---	pCH1941
Box 5	---	---	T--	---	---	---	---	---	--a	--a	---	---	pW1215
	---	---	-G-	---	---	---	---	---	--a	--a	---	---	pW8233
	T--	---	---	---	---	---	---	---	--a	--a	-A-	-G-	pW8142
	T--	---	---	---	---	---	---	---	--a	--a	-A-	---	pGliA-42
	---	---	-G-	---	---	---	---	---	--a	--a	---	---	pCH1941

Figure 4. Putative mutations in the P-boxes. The P-boxes within each gene are compared to the consensus sequence derived from all of the boxes (upper line). The sequential position of the five boxes indicated in Fig. 1 is shown. Putative mutations, i.e. bases which differ from the consensus sequence are indicated; silent mutations are shown in lower-case characters, whereas mutations which change the coded amino acid are shown in upper-case letters. Bases which are preserved are indicated by a horizontal line. Deletions are indicated by a blank. * indicates insertion of a CAG codon. Data from T. aestivum cDNA clones pGliA-42 (31) and pCH1941 (33) are also included.

in multiple copies (6-14), the multiplication of that gene is likely to have occurred before these species diverged. This hypothesis could be confirmed if the mutations which we believe occurred in the ancestral gene are found in α/ß-gliadins derived from each diploid genome. At present we do not know from which genome the genes which we have sequenced have been derived. Analysis of DNA from the diploid ancestors of wheat is needed.

The high number of mutations in the P-boxes implies that they are evolutionarily ancient structures; however, we cannot yet derive an estimate of the rate of accumulation of mutations in the structure. The gliadin genes of wheat share homology with various hordein genes from barley (30,32). Thus, we would expect that the P-box structure and some of the earliest mutations in it might be preserved in barley hordeins. If so, it may then be possible to calibrate the rate of mutant accumulation using the time of divergence of wheat ancestors and barley as a standard, and thus calculate the evolutionary age of various features of the genes.

It is possible to make a crude estimate of the relative period of time that the repeated P-box structure existed prior to the multiplication of the ancestral gene, compared to the time since that event first occurred. Since the number of P-box mutations common to all genes is approximately equal to the number of mutations unique to any one gene, one can speculate that the repeated P-box structure must have existed for about as long before gene multiplication began as after.

Possible Regulatory Sequences: The expression of gliadin genes is coordinately regulated during seed development (38). It is therefore likely that these genes share common target sequences for the developmental regulatory mechanism(s). Such sequences have been demonstrated in some animal and viral genomes, and typically are present in multiple copies in the flanking regions of the regulated genes (39). If this generalization is true for plants we would expect the sites to be in the large stretches of conserved sequence in the 5'- and/or 3'-flanking regions of these genes. Several repeated sequences were found in the common 5'-flanks. One of these, which is present twice in the 5'-flank (Fig. 1), shows a surprising homology (up to 73%) to a 19 bp consensus sequence flanking the ovalbumin and related genes and identified as a probable binding site for chicken oviduct progesterone receptor (40). There are no common shorter sequences in the non-homologous 3'-flanks of pW8142/pW1215 and pW8233 (data not shown).

ACKNOWLEDGEMENTS

We are indebted to Jerry Slightom and Michael Murray for the gift of a wheat genomic library and to Nancy Templeton for E. coli DNA polymerase. We are grateful to Thomas Okita and Ralph Quatrano for making their sequence data available to us before publication.

M.S.-S. was a Fellow of the Medical Research Council of Canada; A.R. was

supported by E.I. du Pont de Nemours & Co.; T.S. was supported by a grant from the National Institutes of Health.

[1]Present address: Allelix, Inc., 6850 Goreway Drive, Mississauga, Ontario L4V 1P1, Canada
[2]Present address: Central Research and Development Department, Experimental Station, E.I. du Pont de Nemours & Co., Wilmington, DE 19898, USA
[3]Permanent address: Division Citogenetica, Instituto de Investigaciones Biologicas C. Estable, Montevideo, Uruguay

REFERENCE

1. Woychik, J.H., Boundy, J.A., and Dimler, R.J. (1961) Arch. Biochem. Biophys. **94**, 477-482.
2. Kasarda, D.D., Autran, J.-C., Lew, E.J.-L., Nimmo, C.C., and Shewry, P.R. (1983) Biochim. Biophys. Acta **747**, 138-150.
3. Bietz, J.A., Heubner, F.R., Sanderson, J.E., and Wall, J.S. (1977) Cereal Chem. **54**, 1070-1083.
4. Kasarda, D.D., DaRosa, D.A., and Ohms, J.I. (1974) Biochim. Biophys. Acta **351**, 290-294.
5. Shewry, P.R., Autran, J.C., Nimmo, C.C., Lew, E.J.-L., and Kasarda, D.D. (1980) Nature **286**, 520-522.
6. Lawrence, G.J. and Shepherd, K.W. (1981) Theor. Appl. Genet. **60**, 333-337.
7. Wrigley, C.W., Lawrence, G.J., and Shepherd, K.W. (1982) Aust. J. Plant Physiol. **9**, 15-30.
8. Galili, G. and Feldman, M. (1983) Theor. Appl. Genet. **66**, 77-86.
9. Wrigley, C.W. and Shepherd, K.W. (1973) Ann. NY Acad. Sci. **209**, 154-162.
10. Kasarda, D.D., Bernardin, J.E., and Qualset, C.O. (1976) Proc. Natl. Acad. Sci. USA **73**, 3646-3650.
11. Brown, J.W.S., Kemble, R.J., Law, C.N., and Flavell, R.B. (1979) Genetics **93**, 189-200.
12. Brown, J.W.S., Law, C.N., Worland, A.J., and Flavell, R.B. (1981) Theor. Appl. Genet. **59**, 361-371.
13. Galili, G. and Feldman, M. (1983) Theor. Appl. Genet. **64**, 97-101.
14. Galili, G. and Feldman, M. (1984) Mol. Gen. Genet. **193**, 293-298.
15. Damidaux, R., Autran, J.-C., Grignac, P., and Feillet, P. (1980) C.R. Acad. Sci. (Paris) **291D**, 585-588.
16. Sozinov, A.A. and Poyerelga, F.A. (1980) Ann. Technol. Agric. **29**, 229-245.
17. Payne, P.I., Jackson, E.A., Holt, L.M., and Law, C.N. (1984) Theor. Appl. Genet. **67**, 235-243.
18. Rafalski, J.A., Scheets, K., Metzler, M., Peterson, D.M., Hedgcoth, C., and Söll, D.G. (1984) EMBO J. **3**, 1409-1415.
19. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) Molecular Cloning - A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
20. Hanahan, D. (1983) J. Mol. Biol. **166**, 557-580.
21. Bolivar, F. (1978) Gene **4**, 121-136.
22. Guo, L.-H., Yang, R.C.A., and Wu, R. (1983) Nucl. Acids Res. **11**, 5521-5540.
23. Poncz, M., Solowiejczyk, D., Ballantine, M., Schwartz, E., and Surrey, S. (1982) Proc. Natl. Acad. Sci. USA **79**, 4298-4302.
24. Larson, R. and Messing, J. (1982) Nucl. Acids Res. **10**, 39-49.
25. Sege, R.D., Söll, D., Ruddle, F.H., and Queen, C. (1981) Nucl. Acids Res. **9**, 437-444.
26. Geraghty, D., Peifer, M.A., Rubenstein, I., and Messing, J. (1981) Nucl. Acids Res. **9**, 5163-5174.

27. Messing, J., Geraghty, D., Heidecker, G., Hu, N.-T., Kridl, J., and Rubenstein, I. (1983) In: Genetic Engineering of Plants: An Agricultural Perspective (Kosuge, T., Meredith, C.P., Hollaender, A., eds.) pp. 211-227 (Plenum Press, NY).
28. Scheets, K., Rafalski, J.A., Hedgcoth, C., and Söll, D.G. (1985) Plant Sci. Letters, in press.
29. Weaver, R.F. and Weissmann, C.H. (1979) Nucl. Acids Res. **7**, 1175-1193.
30. Rasmussen, S.K., Hopp, E., and Brandt, A. (1983) Carlsberg Res. Commun. **48**, 187-199.
31. Kasarda, D.D., Okita, T.W., Bernardin, J.E., Baecker, P.A., Nimmo, C.C., Lew, E.J.-L., Dietler, M.D., and Greene, F.C. (1984) Proc. Natl. Acad. Sci. USA **81**, 4712-4716.
32. Forde, B.G., Kreis, M., Bahramian, M.B., Matthews, J.A., Miflin, B.J., Thompson, R.D., Bartels, D., and Flavell, R.B. (1981) Nucl. Acids Res. **9**, 6689-6707.
33. Proffitt, J.H., Chakerian, R.L., Sheehy, R.E., and Quatrano, R.S., manuscript submitted.
34. Pedersen, K., Devereux, J., Wilson, D.R., Sheldon, E., and Larkins, B.A. (1982) Cell **29**, 1015-1026.
35. Argos, P., Pedersen, K., Marks, M.D., and Larkins, B.A. (1982) J. Biol. Chem. **257**, 9984-9990.
36. Gruenbaum, Y., Naveh-Many, Cedar, T., and Razin, A. (1981) Nature **292**, 860-862.
37. Lindahl, T. and Nyberg, B. (1974) Biochemistry **13**, 3405-3410.
38. Kasarda, D.D., Bernadin, J.E., and Nimmo, C.C. (1976) In: Advances in Cereal Science and Technology 1, Pomeranz, Y., ed., Am. Assoc. of Cer. Chemists, St. Paul, MN.
39. Davidson, E.H., Jacobs, H.T., and Britten, R.J. (1983) Nature **301**, 468-470.
40. Mulvihill, E.R., LePennec, J.-P., and Chambon, P. (1982) Cell **28**, 621-632.