14. M. Fromm, L. P. Taylor, V. Walbot, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 5824 (1985).
15. B. Reiss, R. Sprengel, H. Will, H. Schaller, *Gene* **30**, 211 (1984).
16. I. J. Mettler, A. R. Lachmansingh, J. J. Detmer, D. A. Pierce, D. Mascarenhas, unpublished data.
17. C. A. Rhodes, R. L. Phillips, C. E. Green, *Can. J. Genet. Cytol.* **28**, 374 (1986).
18. S. Edallo, C. Zucchinali, M. Perenzin, F. Salamini, *Maydica* **26**, 39 (1981).
19. W. Werr, W. B. Frommer, C. Maas, P. Starlinger, *EMBO J.* **4**, 1373 (1985).
20. S. Dellaporta, J. Wood, J. Hicks, *Plant Mol. Biol.* **1** (1983).
21. G. Schroder and J. Schroder, *Mol. Gen. Genet.* **185**, 51 (1982).
22. E. Southern, *J. Mol. Biol.* **98**, 503 (1975).
23. We thank our co-workers at Stauffer Chemical Co., especially R. Lachmansingh, K. Lowe, P. Olson, and L. Pomeroy, for their many contributions to this research. We also appreciate the assistance of D. Gray with the illustrations.

8 December 1987; accepted 29 February 1988

Insecticidal Activity and Lectin Homology of Arcelin Seed Protein

THOMAS C. OSBORN, DANNY C. ALEXANDER,* SAMUEL S. M. SUN,† CESAR CARDONA, FREDRICK A. BLISS

Arcelin, a major seed protein discovered in wild beans (*Phaseolus vulgaris*), has toxic effects on an important bean bruchid pest, *Zabrotes subfasciatus*. Transfer of the arcelin-1 allele to bean cultivars and addition of purified arcelin to artificial seeds results in high levels of insect resistance. The nucleotide and derived amino acid sequences of the arcelin-1 complementary DNA are very similar to those of genes encoding the bean seed lectin, phytohemagglutinin. The gene or genes encoding arcelin may have evolved from a phytohemagglutinin gene or genes resulting in an effective mechanism for resistance to bean bruchids.

PLANTS HAVE EVOLVED WITH VARIous mechanisms to protect their seeds from insect predators. Proteins, which are major components of legume seeds, represent potential antibiosis factors that could affect predation (*1–3*). Seeds of common bean, *Phaseolus vulgaris* L., contain a carbohydrate-binding lectin protein called phytohemagglutinin (PHA). Although the function of PHA has not been demonstrated conclusively (*4*), Janzen *et al.* (*2*) suggested that a major part of its adaptive significance is to protect bean seeds from insect predators. That conclusion was based on the toxic effects of PHA on the cowpea weevil (*Callosobruchus maculatus* F.) when PHA was incorporated into artificial cowpea seeds.

Although PHA may protect bean seeds from predation by some insects, it is ineffective against the two most important bruchid pests of bean, the bean weevil, *Acanthoscelides obtectus* (Say), and the Mexican bean weevil, *Zabrotes subfasciatus* (Boheman). Most bean cultivars contain PHA, but no high levels of resistance have been found among cultivated materials (*5*). Among wild beans, however, accessions with high levels of resistance to these bruchid species have been identified (*6*). These wild accessions also contain a major seed protein, named arcelin, which has not been detected in seeds of bean cultivars (*7*). Four arcelin variants have been identified in wild beans. Accessions containing arcelin-2, -3, or -4 are resistant to the two bruchid species, but arcelin-1–containing accessions have not been identified as resistant, probably because the arcelin-1 allele occurs at low frequencies in wild accessions containing this variant (*6, 7*). In earlier studies, we found that genes controlling arcelin and PHA expression are tightly linked (*7*) and that arcelin-1 has several properties in common with PHA (*8*). In this study we report on insecticidal activity of the arcelin-1 protein in backcross-derived bean lines and in artificial seeds. We also report on the cloning and sequencing of a complementary DNA (cDNA) for arcelin-1 and on comparisons of arcelin and lectin sequences.

Although the presence of arcelin is correlated with bruchid resistance in wild beans, factors other than arcelin protein might confer the resistance property. To test whether resistance is associated with the genetic transfer of arcelin, we introduced the arcelin-1 allele from the wild line UW325 (*9*) into the bean cultivar Sanilac by two generations of backcrossing followed by two selfing generations. The expression of arcelin is controlled by a single Mendelian gene, and the presence of arcelin is dominant to its absence (*7, 9*). Seeds of backcross lines were tested for resistance to *Z. subfasciatus* (Table 1). On the basis of days until adult emergence and percentage emergence of adults, all arcelin-1–containing lines showed high levels of resistance. Lines lacking arcelin-1 were fully susceptible compared to the check cultivar, and lines segregating for arcelin-1 had intermediate levels of resistance. These results demonstrate that the arcelin-1 variant is associated with high levels of resistance to *Z. subfasciatus*. They also indicate that resistance is associated with the genetic transfer of arcelin-1 expression.

Analogous sets of backcross lines were developed from different cultivated bean types (for example, Pinto and black-seeded types) as recurrent parents. When these lines

Table 1. Levels of resistance to *Z. subfasciatus* in 'Sanilac' backcross–derived lines with arcelin-1 (Arc^1/Arc^1), without arcelin-1 (arc/arc), and segregating for arcelin-1 (Arc^1/arc). Lines were screened for resistance as described previously (*5*). Values represent the mean (±SEM) of two replicates containing 50 seeds each, each replication infested with seven insect pairs.

Line or cultivar	Arcelin genotype	Days until adult emergence	Percentage emergence
Backcross line			
3	Arc^1/Arc^1	53.0 (±0.7)	2.5 (±0.3)
5	Arc^1/Arc^1	47.8 (±3.2)	2.1 (±0.2)
4	Arc^1/arc	33.2 (±2.3)	20.9 (±5.6)
7	Arc^1/arc	37.2*	38.7*
8	Arc^1/arc	38.1*	34.6*
9	Arc^1/arc	35.4*	30.2*
1	arc/arc	34.2 (±0.2)	89.5 (±3.8)
2	arc/arc	34.7 (±0.1)	76.3 (±1.5)
6	arc/arc	34.4 (±0.3)	93.8 (±6.2)
Susceptible cultivar			
Calima	arc/arc	34.0 (±0.3)	92.9 (±3.9)

*Values based on one observation.

T. C. Osborn, Department of Agronomy, University of Wisconsin, Madison, WI 53706; formerly at ARCO Plant Cell Research Institute, 6560 Trinity Court, Dublin, CA 94568.
D. C. Alexander and S. S. M. Sun, ARCO Plant Cell Research Institute, 6560 Trinity Court, Dublin, CA 94568.
C. Cardona, Centro Internacional de Agricultura Tropical, AA 6713, Cali, Colombia.
F. A. Bliss, Department of Horticulture, University of Wisconsin, Madison, WI 53706.

*Present address: Calgene, 1920 5th Street, Davis, CA 95616.
†Present address: Department of Plant Molecular Physiology, University of Hawaii, Honolulu, HI 96822.

Table 2. Levels of resistance to *Z. subfaciatus* in intact and artificial bean seeds with and without the addition of purified arcelin-1. Artificial seeds were prepared and screened for resistance with the system devised for cowpea weevil (*3*). Values represent the mean (±SEM) of four replicates containing five seeds each.

Material screened	Days until adult emergence	Percentage emergence
Intact seed		
L12-56	32.2 (±0.3)	100.0 (±0)
Sanilac	31.3 (±0.3)	95.9 (±2.5)
SARC1–7	50.3 (±2.0)	7.4 (±3.7)
Calima	31.5 (±0.3)	93.0 (±2.9)
Artificial seed		
Control		
L12-56	38.4 (±0.7)	74.7 (±9.4)
Sanilac	37.8 (±1.0)	86.1 (±3.0)
SARC1–7	53.8 (±1.2)	18.4 (±7.3)
Calima	37.9 (±0.3)	87.7 (±6.8)
Experimental		
Sanilac + 2.5% arcelin-1	38.9 (±0.8)	76.1 (±5.9)
Sanilac + 5.0% arcelin-1	44.7 (±0.8)	76.1 (±9.8)
Sanilac + 10.0% arcelin-1	53.4 (±1.8)	18.4 (±8.8)

were tested for resistance to *Z. subfasciatus,* similar results were obtained. The 'Sanilac' backcross lines also were tested for resistance to *A. obtectus,* but in this case only low levels of resistance were associated with the presence of arcelin-1 (*10*). The resistance of arcelin-1–containing bean lines to *Z. subfasciatus* is due to larval antibiosis (up to 97% mortality of first instar larvae). This could be caused by the arcelin protein or by some factor that is linked genetically to arcelin expression.

To determine whether arcelin protein is the factor conferring bruchid resistance, we produced artificial bean seeds containing different levels of purified arcelin-1 and tested these seeds for resistance to *Z. subfasciatus* (Table 2). In this test, we included intact seeds and artificial seeds from four cultivated bean lines as controls. 'Calima', 'Sanilac', which contains PHA, and L12-56, a backcross line near-isogenic to 'Sanilac' but which is PHA-deficient (*11*), were all susceptible to bruchid infestation, whereas SARC1-7, a 'Sanilac' backcross–derived line homozygous for the presence of arcelin-1, was resistant. These results indicate that the presence or absence of PHA does not affect bruchid development, but as noted before, arcelin-1 is associated with a high level of resistance. After soaking and removal of the seed coat, seeds of these lines were ground into flour, reconstituted as control artificial seeds (*3*) and tested for bruchid resistance. Although the absolute values for days until adult emergence and percentage emergence were different than those for intact seeds, resistant and susceptible responses were easily distinguished.

Experimental artificial seeds consisting of 'Sanilac' flour to which purified arcelin-1 (*8*) was added at three different levels were tested for resistance. The highest level of arcelin-1 (10% w/w) represents the approximate concentration of arcelin in seed of SARC1-7 (*12*), and the lower levels, 5 and 2.5% w/w, represent approximately one-half and one-quarter, respectively, of the arcelin concentration present in SARC1-7. At the lowest level tested, arcelin had no significant antibiosis effect on larvae. At the intermediate arcelin level, there was a significant increase in the number of days until adult emergence but not significant effect on percentage emergence. However, the response of insects to the highest level of arcelin was nearly identical to that of insects reared on control artificial seeds of SARC1-7 for both measures of resistance. This indicates that the presence of arcelin-1 in bean seeds confers resistance to *Z. subfasciatus*. The dosage response of larvae as measured by days until adult emergence was nearly linear over the range of arcelin levels tested. For percentage emergence, a significant dosage response was observed only at the highest arcelin level, indicating that high levels are needed to affect this parameter.

Since PHA was shown to have insecticidal activity on the cowpea weevil (*2*), we were interested in comparing the molecular properties of arcelin and PHA. Biochemically, arcelin and PHA are related proteins (*8*). They have similar amino acid compositions, and their deglycosylated molecular weights are almost identical. They are related antigenically, and arcelin behaves as a lectin in that it agglutinates pronase-treated erythrocytes from some animal species. To compare the amino acid and nucleotide sequences of these proteins and their genes, we cloned and sequenced a cDNA encoding arcelin-1 and compared this sequence to the published sequences of lectin genes.

Messenger RNA (mRNA) was isolated from developing seeds of SARC1-7 (*13*) and used to construct a cDNA library in the pARC7 cDNA cloning vector (*14*). Candidate clones for arcelin-1 were selected by differential hybridization of colony lifts

```
1 C12ATGAATGCATAC ATG GCT TCC TCC AAC TTA CTC ACC CTA GCC CTC TTC CTT GTG  66
                  MET Ala Ser Ser Asn Leu Leu Thr Leu Ala Leu Phe Leu Val

67 CTT CTC ACC CAC GCA AAC TCA AGC AAC GAC GCC TCC TTC AAC GTC GAG ACG TTC 120
   Leu Leu Thr His Ala Asn Ser Ser Asn Asp Ala Ser Phe Asn Val Glu Thr Phe

121 AAC AAA ACC AAC CTC ATC CTC CAA GGC GAT GCC ACC GTC TCA TCC GAA GGC CAC 172
    Asn Lys Thr Asn Leu Ile Leu Gln Gly Asp Ala Thr Val Ser Ser Glu Gly His

173 TTA CTA CTA ACC AAT GTT AAA GGC AAC GAA GAG GAC TCT ATG GGC CGC GCC TTC 228
    Leu Leu Leu Thr Asn Val Lys Gly Asn Glu Glu Asp Ser Met Gly Arg Ala Phe

229 TAC TCC GCC CCC ATC CAA ATC AAT GAC AGA ACC ATC GAC AAC CTC GCC AGC TTC 282
    Tyr Ser Ala Pro Ile Gln Ile Asn Asp Arg Thr Ile Asp Asn Leu Ala Ser Phe

283 TCC ACC AAC TTC ACA TTC CGT ATC AAC GCT AAG AAC ATT GAA AAT TCC GCC TAT 336
    Ser Thr Asn Phe Thr Phe Arg Ile Asn Ala Lys Asn Ile Glu Asn Ser Ala Tyr

337 GGC CTT GCC TTT GCT CTC GTC CCC GTC GGC TCT CGG CCC AAA CTT AAA GGC CGT 390
    Gly Leu Ala Phe Ala Leu Val Pro Val Gly Ser Arg Pro Lys Leu Lys Gly Arg

391 TAT CTA GGT CTT TTC AAC ACA ACC AAC TAT GAC CGC GAC GCC CAT ACT GTG GCT 444
    Tyr Leu Gly Leu Phe Asn Thr Thr Asn Tyr Asp Arg Asp Ala His Thr Val Ala

445 GTG GTG TTC GAC ACC GTC AGC AAC CGT ATT GAA ATC GAC GTG AAC TCC ATC CGG 498
    Val Val Phe Asp Thr Val Ser Asn Arg Ile Glu Ile Asp Val Asn Ser Ile Arg

499 CCT ATC GCA ACG GAG TCT TGC AAT TTC GGC CAC AAC AAC GGA GAA AAG GCC GAG 552
    Pro Ile Ala Thr Glu Ser Cys Asn Phe Gly His Asn Asn Gly Glu Lys Ala Glu

553 GTT CGG ATC ACC TAT GAC TCC CCC AAG AAC GAC TTG AGG GTT TCT CTG CTT TAC 606
    Val Arg Ile Thr Tyr Asp Ser Pro Lys Asn Asp Leu Arg Val Ser Leu Leu Tyr

607 CCT TCT TCG GAA GAA AAG TGC CAC GTC TCT GCC ACA GTG CCG CTG GAG AAA GAA 660
    Pro Ser Ser Glu Glu Lys Cys His Val Ser Ala Thr Val Pro Leu Glu Lys Glu

661 GTT GAG GAC TGG GTG AGC GTT GGG TTC TCT GCC ACC TCA GGG TCG AAA AAA GAG 714
    Val Glu Asp Trp Val Ser Val Gly Phe Ser Ala Thr Ser Gly Ser Lys Lys Glu

715 ACC ACT GAA ACG CAC AAC GTC CTC TCT TGG TCT TTT TCT TCC AAC TTC ATC AAT 768
    Thr Thr Glu Thr His Asn Val Leu Ser Trp Ser Phe Ser Ser Asn Phe Ile Asn

768 TTT AAG GGC AAA AAA TCT GAA CGT TCC AAC ATC CTC CTC AAC AAG ATC CTC TAG 822
    Phe Lys Gly Lys Lys Ser Glu Arg Ser Asn Ile Leu Leu Asn Lys Ile Leu  .

823 ACTCCCAAAGCCAGCTTCACTGTGACAGTAAAACCTTCCTTATACGCTAATAATGTTCATCTGTCACACAA 893

894 ACTCCAATAAATAAAATGGGAGCAATAAATAAAATGGGAGCTCATATATTTACAC                948
```

Fig. 1. Nucleotide sequence and derived amino acid sequence of arcelin-1 cDNA, pAR1-11. The 265–amino acid open reading frame is shown; the third ATG of the three potential initiation sites at the 5′ end (overlined) is presumed to be the initiation codon by analogy to the PHA genes (*18*). The 47-residue amino-terminal sequence of the mature protein, as determined by degradation sequencing (*17*), is underscored.

(*15*). The nucleotide sequence (*16*) of candidate clone pAR1-11 contained three potential initiation sites near the 5′ end of the clone, two of which (the first and third) were in the same large open reading frame encoding 269 amino acids (Fig. 1). The second ATG had a short open reading frame for 11 amino acids. Initiation at the third ATG would yield a 265–amino acid polypeptide. Comparison of this derived amino acid sequence to the partial amino acid sequence of purified arcelin-1 protein demonstrated that pAR1-11 encodes arcelin-1. The mature protein sequence, which begins at codon 22 of the 265–amino acid open reading frame, matches the predicted sequence exactly through 47 of 48 amino acids determined by degradation sequencing (*17*); amino acid residue 12 yielded a blank by protein sequencing, presumably because of glycosylation of this predicted asparagine residue. The hydrophobic nature of the 21-residue peptide not found in the mature protein suggests that this region serves as a signal peptide.

Fig. 2. Comparisons of nucleotide-derived amino acid sequences of arcelin-1, "lectin-like" protein (*19*), PHA-L (*18*), and PHA-E (*18*). The complete amino acid sequence is presented only for arcelin-1. Dots shown in the sequence lines represent amino acid identity with arcelin-1, and dashes indicate gaps introduced to maximize sequence identity. The arrow at the amino-terminal region indicates the start of the mature arcelin-1 protein. Sequence comparisons were made using Intelligenetics version 5.0 computer software.

```
                1                         25                         50
                *                         *                          *
Arcelin-1       MASSNLLTLALFLVLLTHANSSNDASFNVETFNKTNLILQGDATVSSEGH
Lectin-like     ....K..S.....A..S....ATET..IIDA.................---
PHA-L           .......S.............ASQT..SFQR..E......R......K.Q
PHA-E           ....-KFFTV.............IY..FQR..E......R..S...S.Q

               51                         75                         100
                *                         *                          *
Arcelin-1       LLLTNVKGN---EEDSMGRAFYSAPIQINDRTIDNLASFSTNFTFRINAK
Lectin-like     ---NGNLQLSYNSY...S.........R.S.TG.V...D....MN.RTH
PHA-L           .R....ND.GEPTLS.L...........W.N.TGAV.ASP.S...N.DVP
PHA-E           .R....N..GEPRVG.L...........W.N.TGTV...A.S...N.QVP

               101                        125                        150
                *                         *                          *
Arcelin-1       NIENSAYGLAFALVPVGSRPKLKGRYLGLFNTTNYDRDAHTVAVVFDT--
Lectin-like     RQA...V..D.V....--Q.ES..----------------D..T.E...--
PHA-L           .NSGP.D....V.L....Q..D..GL.....NYK..SN......E...LY
PHA-E           .NAGP.D...........Q..D..GF....--DGSNSNF.....E...LY

               151                        175                        200
                *                         *                          *
Arcelin-1       ------VSNRI IDVNSIRPIATESCNFGHNNGEKAEVRITYDSPKNDLR
Lectin-like     ------FLS..S....-NND.KSVPWDVHDYD.QN.......N.STKVFS
PHA-L           NVHWDPKPRH.G......--KSIKTTTWDFVK..N...L.....STKL.V
PHA-E           NKDWDPTERH.G.......--SIKTTRWDFV...N...L.....STNL.V

               201                        225                        250
                *                         *                          *
Arcelin-1       VSLLYPSSEEKCHVSATVPLEKEVEDWVSVGFSATSGSKKETTETHNVLS
Lectin-like     ...SNPSTGKSNN..T..E.....Y............AYQWSY...D...
PHA-L           A..V...LKTSFI..D..D.KSVLPE..I...T..T.IT.GNV..NDI..
PHA-E           A..V...QKTSFI..D..D.KSVLPE.........T.IN.GNV..ND...

               251                        275
                *                         *
Arcelin-1       WSF-SSNFINFKGKKSERSNILLNKIL
Lectin-like     ...-......L.DQ.......V.....
PHA-L           ...A.KLSDGTTSEALNLA.FA..Q..
PHA-E           ...A.KLSDGTTSEGLNLA.LV.....
```

At the nucleotide level, the arcelin-1 coding sequence is very similar to those of the two genes encoding PHA (*18*). There is approximately 78% identity between the protein coding sequence of pAR1-11 and those of pdlec1 (PHA-E) and pdlec2 (PHA-L). A cDNA clone, pPVL134 (*19*), encoding a lectin-like protein has even higher percentage identity (81%) to the pAR1-11 coding sequence. High degrees of sequence similarity occur at the amino acid level as well (Fig. 2). The derived amino acid sequence of arcelin-1 was 58 to 61% identical with the amino acid sequences derived from these three other lectin genes.

The high degree of similarity between arcelin and PHA in nucleotide and amino acid sequences is evidence for an evolutionary relation between these genes. PHA occurs in approximately 90% of bean cultivars and wild bean accessions (*20*), whereas, arcelin occurs in only 10% of wild bean lines and it has not been observed in bean cultivars (*7*). On the basis of their sequence similarity, the frequencies at which these seed proteins occur, and the tight linkage between arcelin and PHA genes (*7*), we hypothesize that arcelin and PHA are encoded by homologous genes and that one or more arcelin genes arose by tandem duplication and divergence of a PHA gene or genes or by divergence of one or more of the existing members of the PHA gene family.

This evolutionary step may be the most recent event in an evolutionary process involving lectin genes. Low levels of lectins have been detected in stems, leaves, and roots of bean plants (*21*); in the seed of a PHA-deficient bean cultivar, a small amount of lectin has been detected that differs biochemically from PHA (*22*). These lectins may represent a class of ubiquitous tissue lectins that are expressed at low levels in the plant and are encoded by a lectin gene or genes other than those encoding PHA or arcelin. In fact, a cDNA for a lectin-like protein that differs from PHA and arcelin has been isolated from common bean (*19*). Genes encoding PHA may have arisen from this lectin gene by duplication and divergence. Expression of PHA genes at high levels in the seed could have provided a selective advantage by conferring resistance to some insect predators (*2*). The appearance of a new lectin seed protein, arcelin, that confers resistance to other insect predators may represent a similar type of event in lectin gene evolution.

The mechanism of insecticidal action of arcelin is unknown, but it may resemble that of PHA which is a carbohydrate-binding lectin. Before heat denaturation, PHA is toxic to mammals because it binds to and disrupts epithelial cells in the intestine (*23*). An analogous mechanism has been proposed for the toxic effect of PHA on the cowpea weevil (*24*). Since arcelin appears to have some carbohydrate-binding activity (*8*), this also may be the mode of insecticidal action for arcelin.

Two possible explanations for insecticidal activity in arcelin that PHA does not show are as follows. First, arcelin may have some unique biochemical characteristic that is responsible for insecticidal activity on *Z. subfasciatus,* such as a different carbohydrate-binding specificity. Second, insecticidal activity may be related to the high level of arcelin expression in the seed. Arcelin-1 represents about 10% of the total bean seed weight, whereas the maximum level of PHA found in beans is only 3% of the total seed weight (*25*). Since high levels of arcelin-1 are required for insecticidal action in artificial seeds, bean seeds with levels of PHA higher than those found in nature also may be toxic to bean bruchids.

The occurrence of arcelin in bean seeds appears to be a very effective means of controlling important storage pests in common bean. The utility of this form of resistance depends not only on its effectiveness as an insect antibiosis factor, but also on its nutritional effects in mammalian diets. Initial studies indicate that arcelin-1 has no adverse effects on rat growth and metabolism when supplied as a diet of cooked beans (*26*). Since arcelin is controlled by a single Mendelian gene, this trait can be easily transferred from wild accessions into bean cultivars by backcross breeding.

REFERENCES AND NOTES

1. S. W. Applebaum and Y. Birk, in *Insect and Mite Nutrition,* J. R. Rodriguez, Ed. (North-Holland, Amsterdam, 1972), pp. 629–636; D. H. Janzen, H. B. Juster, E. A. Bell, *Phytochemistry* **16**, 223 (1977).
2. D. H. Janzen, H. B. Juster, I. E. Liener, *Science* **192**, 795 (1976).
3. R. E. Shade, L. L. Murdock, D. E. Foard, M. A. Pomeroy, *Environ. Entomol.* **15**, 1286 (1986).
4. M. E. Etzler, in *The Lectins: Properties, Functions, and Applications in Biology and Medicine,* I. E. Liener, N. Sharon, I. J. Goldstein, Eds. (Academic Press, New York, 1986), pp. 371–435.
5. A. V. Schoonhoven and C. Cardona, *J. Econ. Entomol.* **75**, 567 (1982).
6. ———, J. Valor, *ibid.* **76**, 1255 (1983).

7. T. C. Osborn, T. Blake, P. Gepts, F. A. Bliss, *Theor. Appl. Genet.* 71, 847 (1986).
8. T. C. Osborn, M. Burow, F. A. Bliss, *Plant Physiol.* 86, 399 (1988).
9. J. Romero Andreas, B. S. Yandell, F. A. Bliss, *Theor. Appl. Genet.* 72, 123 (1986).
10. Resistance of 'Sanilac' backcross–derived lines to *A. obtectus* was tested as described previously (6) with two replicates of 50 seeds each. The means (±SEM) of backcross lines for numbers of days until adult emergence and for percentages emergence were 39.6 (±0.3) and 31.5 (±1.0), respectively, for arcelin-1–containing lines, and 35.6 (±0.4) and 51.7 (±2.9), respectively, for backcross lines without arcelin-1.
11. T. C. Osborn and F. A. Bliss, *J. Am. Soc. Hortic. Sci.* 110, 484 (1985).
12. The approximate amount of arcelin-1 in seeds of SARC1-7 was determined by extracting bean flour with dissociation buffer or dissolving purified arcelin-1 in the same buffer, then separating seed proteins by SDS–polyacrylamide gel electrophoresis (7), and scanning the Coomassie blue–stained gel with a densitometer. By comparisons with known amounts of purified arcelin run on the same gel, the arcelin protein was estimated to be 10% (w/w) of SARC1-7 bean flour.
13. Developing bean seeds were harvested 13 to 19 days after flowering and mRNA was isolated by the procedure described by T. C. Hall *et al.* [*Proc. Natl. Acad. Sci. U.S.A.* 75, 3196 (1978)] except that the sucrose gradient centrifugations were omitted.
14. D. C. Alexander *et al.*, *Gene* 31, 79 (1984).
15. Filters were prepared according to F. Taub and E. B. Thompson [*Anal. Biochem.* 126, 222 (1982)]. Filters were probed sequentially with three different probes. The probes, which were ^{32}P-labeled cDNA first strands made from mRNA fractions of developing seeds, were from (i) L12-56, a PHA-deficient bean line with phaseolin as the only major seed protein, (ii) the cultivar Sanilac, which contains PHA as well as phaseolin, and (iii) SARC1-7, which contains phaseolin, PHA, and arcelin-1. Arcelin cDNA candidates were selected as those colonies that were heavily labeled with the SARC1-7 cDNA, but not by the other two cDNAs, after high-stringency washing.
16. F. Sanger, S. Nicklen, A. R. Coulson, *Proc. Natl. Acad. Sci. U.S.A.* 74, 5463 (1977).
17. P. Edman and G. Begg, *Eur. J. Biochem.* 1, 80 (1967).
18. L. M. Hoffman and D. D. Donaldson, *EMBO J.* 4, 883 (1985).
19. L. M. Hoffman, Y. Ma, R. F. Barker, *Nucleic Acids Res.* 10, 7819 (1982).
20. J. W. S. Brown, T. C. Osborn, F. A. Bliss, T. C. Hall, *Theor. Appl. Genet.* 62, 263 (1982); O. Brücher, *Proc. Trop. Region Am. Soc. Hortic. Sci.* 12, 68 (1968).
21. C. A. K. Borrebaeck and B. Mattiasson, *Physiol. Plant.* 58, 29 (1983).
22. A. Pusztai, G. Grant, J. C. Stewart, *Biochim. Biophys. Acta* 671, 146 (1981).
23. T. P. King, A. Pusztai, E. M. W. Clarke, *J. Comp. Pathol.* 90, 585 (1980); *ibid.* 92, 357 (1982).
24. A. M. R. Gatehouse, F. M. Dewey, J. Dove, K. A. Fenton, A. Pusztai, *J. Sci. Food Agric.* 35, 373 (1984).
25. T. C. Osborn, J. W. S. Brown, F. A. Bliss, *Theor. Appl. Genet.* 70, 22 (1985).
26. G. Hosfield, personal communication.
27. The authors thank C. Posso and R. Vogelzang for technical assistance and G. Colwell and M. Babcock for DNA sequencing.

28 December 1987; accepted 11 February 1988

2-Aminopurine Selectively Inhibits the Induction of β-Interferon, c-*fos*, and c-*myc* Gene Expression

KAI ZINN,* ANDREW KELLER, LISA-ANNE WHITTEMORE, TOM MANIATIS

The protein kinase inhibitor 2-aminopurine (2AP) blocks the induction of the human β-interferon gene by virus or poly(I)-poly(C) at the level of transcription. This inhibition is specific, since 2AP does not inhibit induction of either the *hsp*70 heat-shock gene by high temperature or the metallothionein gene by cadmium or dexamethasone. However, 2AP does block the induction of the c-*fos* and c-*myc* proto-oncogenes by serum growth factors or virus, suggesting that a protein kinase may be involved in the regulation of these genes, as well as of the β-interferon gene. However, different factors must be required for the induction of these three genes, since they are not coordinately regulated by the same inducers in most of the cell lines examined.

THE HUMAN β-INTERFERON (β-IFN) gene is induced by viral infection or by treatment of cells with synthetic double-stranded RNAs (dsRNAs) such as poly(I)-poly(C). The viral induction signal is thought to be dsRNA generated during infection (see *1* for review). Although the mechanism by which the presence of dsRNA leads to β-IFN gene activation is not understood, the DNA sequences required for induction have been extensively characterized (*2, 3*). A 40-bp region known as the IRE (interferon gene regulatory element) is sufficient for induction in mouse C127 cells. The IRE is a transcription enhancer composed of two distinct positive regulatory elements and an adjacent negative regulatory sequence (*4*). A second negative regulatory sequence is located 5′ to the IRE (*2*). Genomic deoxyribonuclease protection (footprinting) experiments show that proteins are bound to both negative regulatory sequences before induction. After induction, these proteins dissociate and other proteins bind to the positive regulatory sequences (*5*). In vitro DNA binding experiments have identified three different factors that specifically bind to the positive regulatory domains of the IRE (*6*). Two of these factors are detected in nuclear extracts from uninduced cells, whereas the third is detected exclusively in extracts from cells induced with virus or poly(I)-poly(C) (*6*). Since protein synthesis is not required for induction, the inactivation of the repressor and the inducible DNA binding activity may be the consequence of post-translational modifications of preexisting proteins. One possible mechanism by which transcription factors or repressors could be modified is phosphorylation. To examine the possibility that a protein kinase might be involved in the pathway of β-IFN induction, we studied the effect of the kinase inhibitor 2-aminopurine (2AP) (*7*) on β-IFN messenger RNA (mRNA) transcription in vivo.

Human MG63 osteosarcoma cells were induced with poly(I)-poly(C) in the presence or absence of 10 m*M* 2AP. RNA was prepared 4 to 6 hours after induction and analyzed for β-IFN and γ-actin mRNA by quantitative ribonuclease (RNase) mapping (*2*). As shown in Fig. 1, 2AP caused a dramatic (>100-fold) inhibition of the induction of β-IFN mRNA, but had no effect on the level of γ-actin mRNA. A similar inhibitory effect of 2AP was observed with Sendai virus induction. No cytotoxicity due to 2AP was apparent during the induction period. However, after incubation for 12

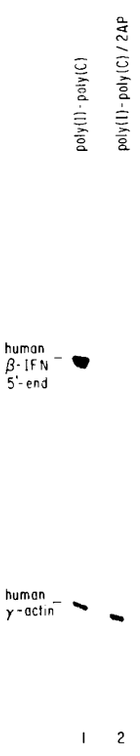

Fig. 1. Effects of 2AP on induction of the β-IFN gene in human MG63 cells. Total cellular RNA (10 μg per lane) was hybridized to a mixture of labeled hybridization probes specific for the 5′ end of human β-IFN mRNA (*2*) and the coding region of a human γ-actin gene (*12*) and analyzed by RNase mapping (*12*). Correctly initiated β-IFN mRNA protects a 277-nucleotide RNA fragment, and γ-actin mRNA protects a 145-nucleotide RNA fragment. Cells were induced with poly(I)-poly(C) as described elsewhere (*12*) in the absence (lane 1) or presence (lane 2) of 10 m*M* 2AP. Longer exposures of this and other autoradiograms indicate that 2AP inhibits induction of β-IFN mRNA by greater than 100-fold.

Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, MA 02138.

*Present address: Department of Biochemistry, University of California, Berkeley, CA 94720.