9-1993

The Primary Structure of a Fungal Chitin Deacetylase Reveals the Function for Two Bacterial Gene Products.

Dimitris Kafetzopoulos
University of Crete

George Thireos
Institute of Molecular Biology and Biotechnology

John N. Vournakis
Dartmouth College

Vassilis Bouriotis
University of Crete

Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 8005–8008, September 1993
Biochemistry

The primary structure of a fungal chitin deacetylase reveals the function for two bacterial gene products

(*Mucor rouxii*/chitosan/NodB/peptidoglycan deacetylase)

DIMITRIS KAFETZOPOULOS*†, GEORGE THIREOS*, JOHN N. VOURNAKIS‡, AND VASSILIS BOURIOTIS*†§

*Institute of Molecular Biology and Biotechnology, P.O. Box 1515, Heraklion 711 10, Crete, Greece; †Department of Biology, Division of Applied Biology and Biotechnology, University of Crete, P.O. Box 1470, Heraklion 711 10, Crete, Greece; and ‡Department of Biology, Dartmouth College, Hanover, NH 03755

Communicated by Fotis C. Kafatos, May 4, 1993 (received for review December 12, 1992)

ABSTRACT **Chitin deacetylase (EC 3.5.1.41) hydrolyzes the *N*-acetamido groups of *N*-acetyl-D-glucosamine residues in chitin. A cDNA to the *Mucor rouxii* mRNA encoding chitin deacetylase was isolated, characterized, and sequenced. Protein sequence comparisons revealed significant similarities of the fungal chitin deacetylase to rhizobial nodB proteins and to an uncharacterized protein encoded by a *Bacillus stearothermophilus* open reading frame. These data suggest the functional homology of these evolutionarily distant proteins. NodB is a chitooligosaccharide deacetylase essential for the biosynthesis of the bacterial nodulation signals, termed Nod factors. The observed similarity of chitin deacetylase to the *B. stearothermophilus* gene product suggests that this gene encodes a polysaccharide deacetylase.**

Chitin exists in the cell wall of several Zygomycetes species in its deacetylated form, referred to as chitosan (1). A chitosan layer has also been identified in the spore wall of *Saccharomyces cerevisiae* (2). The biosynthesis of chitosan in fungi proceeds by the coordinated action of chitin synthase (EC 2.4.1.16) and chitin deacetylase (EC 3.5.1.41). The former enzyme synthesizes chitin by polymerization of *N*-acetylglucosaminyl residues from UDP-N-acetylglucosamine, whereas the latter hydrolyzes the *N*-acetamido groups on the nascent chitin chains (3, 4).

Chitosan is a biopolymer with unique properties favorable for a broad variety of industrial and biomedical applications (5–7). Presently, chitosan is produced by the thermochemical deacetylation of crab chitin. To develop an alternative enzymatic process for chitosan production, we have initiated a study of fungal chitin deacetylases.

Chitin deacetylase from the fungus *Mucor rouxii* has been purified to homogeneity (8). The enzyme is an acidic glycoprotein of ≈75 kDa with ≈30% (wt/wt) carbohydrate content. Further biochemical characterization revealed that the enzyme has a very stringent specificity for β1-4-linked *N*-acetylglucosamine homopolymers, requires at least four residues (chitotetraose) for catalysis, and can achieve extensive deacetylation on chitin polymers.

We report here the cloning of a cDNA that encodes *M. rouxii* chitin deacetylase,¶ and we present the striking protein sequence similarities between chitin deacetylase, the rhizobial nodB proteins, and a *Bacillus stearothermophilus* open reading frame located next to a gene encoding class I Fe-S fumarase. These similarities suggest the functional homology of these proteins.

MATERIALS AND METHODS

Materials. *M. rouxii* (ATCC 24905) was obtained from the American Type Culture Collection. The cDNA synthesis kit, the λZAP vector, and the packaging extracts were bought from Stratagene. The Sequenase 2.0 kit was purchased from United States Biochemical. Immobilon-P and nitrocellulose membranes were purchased from Millipore. Restriction enzymes were obtained from Minotech. Synthetic oligonucleotides were obtained from the Institute of Molecular Biology and Biotechnology Microchemistry Facility.

Growth of *M. rouxii* Mycelia. *M. rouxii* was grown with vigorous shaking at 28°C in a medium containing 0.3% yeast extract, 1% peptone, and 2% glucose; the pH was adjusted to 4.5 with H_2SO_4 (1). The medium was inoculated with 10^8 spores per liter, and the mycelia were harvested at early growth phase by filtration.

Enzyme Purification. Chitin deacetylase was purified to homogeneity from mycelial extracts of the fungus *M. rouxii* as described (8).

Protein Sequencing. The amino-terminal sequence of chitin deacetylase was determined after electroblotting a purified enzyme preparation onto a poly(vinylidene difluoride) membrane. Sequence analysis was performed by the Harvard Microchemistry Facility using methods and instrumentation described (9).

Isolation of poly(A)$^+$ RNA. *M. rouxii* mycelia were extracted by grinding using liquid nitrogen, and total RNA was subsequently purified by the guanidinium thiocyanate method of Chirwin *et al.* (10). Further purification by ultracentrifugation was performed as described by Glisin *et al.* (11). Poly(A)$^+$ RNA was isolated by three passes through an oligo(dT)-cellulose column according to the method of Aviv and Leder (12).

Construction of the cDNA Library. The oligo(dT)-primed *M. rouxii* cDNA library was constructed using Stratagene's cDNA synthesis and cloning kit. Approximately 5 μg of poly(A)$^+$ RNA was used in the cDNA synthesis, and the double-stranded cDNA was directionally cloned into the *Xho* I–*Eco*RI sites of the λZAP vector (13). The constructed cDNA library consisted of ≈2.5×10^5 recombinant phages.

Isolation and Sequencing of Chitin Deacetylase cDNA Clones. Based on the amino-terminal protein sequence analysis data, four degenerate oligonucleotides were designed: 1A, GA(C/T)ACITCIGCIAA(C/T)TA(C/T)TGGC; 1B, GA(C/T)ACIAG(C/T)GCIAA(C/T)TA(C/T)TGGC; 2A, TT(C/T)ACITCICA(A/G)ATIAA(C/T)CC; 2B, TT(C/T)ACIAG(C/T)CA(A/G)ATIAA(C/T)CC. Oligonucleotides 1A and 2A, which gave positive signals when tested by Southern blot analysis of the *Bam*HI-digested library DNA, were initially combined in the cDNA library screening and subsequently alternated in the rescreening. The hybridization buffer used in Southern blot analysis and library screening was as described (14) with the addition of sonicated and

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: bv., biovar.

§To whom reprint requests should be sent at the * address.

¶The nucleotide sequence data reported in this paper has been deposited in the GenBank data base (accession no. Z19109).

```
                                                                                5'  TCTACATCCAAAAAGAATTGAAACG    25

    ATGCAAATCAAGACATTCGCCCTTTCAGCT GCAATTGCACAAGTTGCTACTCTTGCTTTA GCCGACACCTCCGCAAATTACTGGCAATCA TTTACTTCTCAAATTAATCCCAAGAACATC   145
-21 MetGlnIleLysThrPheAlaLeuSerAla AlaIleAlaGlnValAlaThrLeuAlaLeu AlaAsnThrSerAlaAsnTyrTrpGlnSer PheThrSerGlnIleAsnProLysAsnIle

    TCCATTCCCTCTATTGAGCAAACTTCATCC ATTGACCCCACTCAAGAATGTGCTTACTAC ACTCCTGATGCTTCATTGTTCACATTCAAC GCTTCCGAATGGCCCTCTATCTGGGAAGTC   265
 20 SerIleProSerIleGluGlnThrSerSer IleAspProThrGlnGluCysAlaTyrTyr ThrProAspAlaSerLeuPheThrPheAsn AlaSerGluTrpProSerIleTrpGluVal

    GCTACTACCAATGGTATGAATGAGTCTGCC GAGTTCCTCAGTGTCTACAATTCTATTGAC TGGACCAAGGCACCCAATATTTCTGTGCGT ACCCTTGACGCTAACGGCAACTTGGATACC   385
 60 AlaThrThrAsnGlyMetAsnGluSerAla GluPheLeuSerValTyrAsnSerIleAsp TrpThrLysAlaProAsnIleSerValArg ThrLeuAspAlaAsnGlyAsnLeuAspThr

    ACTGGTTACAATACTGCTACTGACCCTGAT TGTTGGTGGACAGCTACCACATGTACCTCT CCCAAGATTTCTGATATCAATGACGATATC TCCAAGTGTCCTGAACCCGAGACTTGGGGT   505
100 ThrGlyTyrAsnThrAlaThrAspProAsp CysTrpTrpThrAlaThrThrCysThrSer ProLysIleSerAspIleAsnAspAspIle SerLysCysProGluProGluThrTrpGly

    TTGACTTACGATGATGGACCTAACTGCTCT CACAACGCTTTCTATGACTACCTTCAAGAG CAAAAGTTGAAGGCCTCCATGTTTTATATT GGTTCCAATGTTGTTGACTGGCCATACGGT   625
140 LeuThrTyrAspAspGlyProAsnCysSer HisAsnAlaPheTyrAspTyrLeuGlnGlu GlnLysLeuLysAlaSerMetPheTyrIle GlySerAsnValValAspTrpProTyrGly

    GCTATGCGTGGTGTTGTTGATGGCCATCAC ATTGCATCCCACACATGGTCTCACCCTCAA ATGACAACCAAGACCAATCAAGAGGTCCTT GCTGAATTCTATTATACTCAAAAGGCCATC   745
180 AlaMetArgGlyValValAspGlyHisHis IleAlaSerHisThrTrpSerHisProGln MetThrThrLysThrAsnGlnGluValLeu AlaGluPheTyrTyrThrGlnLysAlaIle

    AAGCTCGCTACTGGTTTGACCCCTCGTTAC TGGCGTCCTCCTTATGGTGATATCGATGAT CGTGTTCGTTGGATTGCCTCTCAATTAGGT TTAACTGCTGTTATTTGGAACCTCGATACT   865
220 LysLeuAlaThrGlyLeuThrProArgTyr TrpArgProProTyrGlyAspIleAspAsp ArgValArgTrpIleAlaSerGlnLeuGly LeuThrAlaValIleTrpAsnLeuAspThr

    GATGATTGGTCTGCTGGTGTCACTACTACT GTCGAAGCTGTTGAGCAAAGTTATTCCGAT TATATTGCTATGGGTACCAATGGTACTTTT GCCAACAGTGGTAACATTGTATTGACCCAT   985
260 AspAspTrpSerAlaGlyValThrThrThr ValGluAlaValGluGlnSerTyrSerAsp TyrIleAlaMetGlyThrAsnGlyThrPhe AlaAsnSerGlyAsnIleValLeuThrHis

    GAAATCAACACAACTATGAGTCTCGCTGTC GAGAACTTGCCCAAGATCATTTCTGCCTAT AAACAAGTCATCGATGTCGCTACCTGTTAC AACATTTCTCACCCTTACTTTGAAGACTAC  1105
300 GluIleAsnThrThrMetSerLeuAlaVal GluAsnLeuProLysIleIleSerAlaTyr LysGlnValIleAspValAlaThrCysTyr AsnIleSerHisProTyrPheGluAspTyr

    GAATGGACCAATGTCTTGAACGGCACAAAA TCTTCTGCTACCGCCAGTGGATCTGCTACT TCTGCTAGTGCTTCTGGAGGCGCTACTACC GCTGCCGCTCATATCCAAGCTTCTACTAGC  1225
340 GluTrpThrAsnValLeuAsnGlyThrLys SerSerAlaThrAlaSerGlySerAlaThr SerAlaSerAlaSerGlyGlyAlaThrThr AlaAlaAlaHisIleGlnAlaSerThrSer

    GGCGCCATGTCTGTCCTTCCCAACCTCGCC TTGATCTCTGCCTTCATTGCTACCCTGTTG TTTTAGTCAAACATCGTCATATATCACCTT CCTGTCATAATTTATAATAGTAAAACATCA  1345
380 GlyAlaMetSerValLeuProAsnLeuAla LeuIleSerAlaPheIleAlaThrLeuLeu PheEnd

    TATTTAGATTTTTCTACATCTTAAAAAAAA 3'                                                                                              1375
```

FIG. 1. Nucleotide and deduced amino acid sequence of the *M. rouxii* chitin deacetylase cDNA clone. Amino acid residues are numbered on the left from the first amino acid (indicated by a double underline) of the mature protein. Nucleotide positions are numbered on the right. Underlined asparagine residues indicate the predicted N-glycosylation sites.

denatured calf thymus DNA at a concentration of 100 μg/ml. Hybridizations were performed at 37–42°C for 16–24 hr. The isolated positive clones were subsequently used for the *in vivo* excision of the recombinant pBluescript (Stratagene), performed according to manufacturer's instructions. Double-stranded DNA sequencing was carried out by the dideoxy chain-termination method using specific primers.

Protein Sequence Comparisons. All protein sequence comparisons and the evaluation of the observed homologies were done using the Genetics Computer Group software (15) and the Schartz and Dayhoff data matrix for amino acid comparisons (16). The deduced amino acid sequences of chitin deacetylase and the *B. stearothermophilus* gene (GenBank accession no. L05611) were used in multiple sequence alignments (PILEUP algorithm) with all available nodB protein sequences: *Rhizobium leguminosarum* biovar (bv.) *trifolii* (17), *R. leguminosarum* bv. *viciae* (18), *Rhizobium fredii* (19), *Rhizobium meliloti* (20), *R. leguminosarum* bv. *phaseoli* (21), *Bradyrhizobium* sp. (*Parasponia*) (22), and *Azorhizobium caulinodans* (23).

RESULTS AND DISCUSSION

We have recently reported the purification and characterization of chitin deacetylase from the fungus *M. rouxii* (6). In this report we present the isolation of a cDNA clone encoding the enzyme, its sequencing, and derived amino acid sequence comparisons with known genes.

Two different enzyme preparations were used for the determination of the amino-terminal sequence of the protein. This analysis provided with high confidence a sequence of 17 amino acids: Asp-Thr-Ser-Ala-Asn-Tyr-Trp-Gln-Ser-Phe-

Table 1. Matrix of the pairwise identity and similarity scores between deacetylases

	M.r. CDA	*B.st.* PDA	*A.c.* nodB	*R.l.p.* nodB	*B.sp.* nodB	*R.m.* nodB	*R.f.* nodB	*R.l.v.* nodB	*R.l.t.* nodB
M.r. CDA	100	22 (35)	21 (40)	23 (38)	25 (37)	24 (35)	25 (38)	29 (42)	27 (42)
B.st. PDA		100	22 (39)	29 (46)	33 (47)	31 (48)	29 (48)	31 (47)	32 (47)
A.c. nodB			100	36 (53)	39 (56)	40 (55)	41 (58)	41 (59)	40 (57)
R.l.p. nodB				100	58 (70)	59 (69)	63 (74)	60 (73)	54 (68)
B.sp. nodB					100	58 (69)	66 (76)	65 (75)	56 (68)
R.m. nodB						100	64 (76)	66 (74)	61 (70)
R.f. nodB							100	66 (78)	60 (74)
R.l.v. nodB								100	71 (80)
R.l.t. nodB									100

The numbers show the percentage of identity and similarity (in parentheses) between deacetylases from various organisms. The calculation of the percentage of identity and similarity in the pairwise comparison of the sequences was performed with the DISTANCES algorithm of the Genetics Computer Group software package (15). Similarities were defined according to Dayhoff's table (16), setting the threshold for a match to 0.6. *M.r., M. rouxii; B.st., B. stearothermophilus; A.c., Azorhizobium caulinodans; R.l.p., R. leguminosarum* bv. *phaseoli; B.sp., B.* sp. (*Parasponia*); *R.m., R. meliloti; R.f., R. fredii; R.l.v., R. leguminosarum* bv. *viciae*; *R.l.t., R. leguminosarum* biovar *trifolii*; CDA, chitin deacetylase; PDA, assumed polysaccharide deacetylase.

```
Mr CDA   116   TCTSPKISDI NDDISKCPEP ETWGLTYDDG PN-CSHNAFY DYLQEQKLKA SMFYIGSNVV
Rlt nodB   1   MKRRAYISEV PFDEAGSDD- RNIYLTFDDG PNPHCTGQIL DVLAEHRVPA TFFVLGGHVK
Bst PDA   47   KEWNELLNKY GAFYLGDPSK KEIYLTFDNG YENGYTSKIL DVLKKHDVHA TFFVTGHYLK

Mr CDA   175   DWPYGAMRGV VDGHHI-ASH TWSHPQMTTK TNQEVLAEFY YTQKAIKLAT G-LTPRYWRP
Rlt nodB  60   DHPDLVRRVA AEG-HLVANH TMTHPDLTAC DSEAIEREIK ETNEAIVSAC PQVAVQHIR-
Bst PDA  107   TAPDLVKRMV KEG-HIVGNH SWSHPDMTTI SADKIKKELD AVSDKVKELT GQEGTVYVRP

Mr CDA   233   PYGDIDDRVR WIASQLGLTA VIWNLDTDDW S-AGVTT--- TVEAVEQSYS DYIAMGTNGT
Rlt nodB 118   RYGAWNADVL SRSMNAGLRP VHWSIDPRDW SRPGVDS--- IVDAVLAA-- ----------
Bst PDA  166   PRGIFSERTL ALSEKYGYRN IFWSLAFVDW KVNEQKGWRY AYDNIINQ-- ----------

Mr CDA   289   FANSGNIVLT HEINTTMSLA VENLPKI--- -ISAYKQVID VATCYNISHP YFED
Rlt nodB 163   -ARPGAIVLL HDGCPPDEIG NCKLTGLRDQ TLSALLAIIP ALHSRGFSLR SLPQ
Bst PDA  214   -IHPGAIILL H--------- ----TVSKDN A-DALDEAIK TLKQQGYTFK SLDD
```

FIG. 2. Comparison of chitin deacetylase from *M. rouxii* (*Mr* CDA), nodB protein from *R. leguminosarum* bv. *trifolii* (*Rlt* nodB), and the deduced amino acid sequence of the presumed polysaccharide deacetylase from *B. stearothermophilus* (*Bst* PDA). Invariant residues among all nodB proteins examined are indicated by underlining. Starting amino acid numbers for each protein sequence are given on the left. Vertical lines indicate identities, and colons indicate conservative substitutions, according to the following amino acid similarity groups: V, I, L, M; D, E, N, Q; F, Y, W, H; A, S, T; K, R; P, G.

Thr-Ser-Gln-Ile-Asn-Pro-Lys. On the basis of this information, synthetic oligonucleotide probes (described in the *Materials and Methods*) were designed and used for the screening of a *M. rouxii* cDNA library. Among the positive clones, the one with the longest insert was subsequently used for the determination of the DNA primary structure (Fig. 1).

A search of the European Molecular Biology Laboratory data bank, using the deduced amino acid sequence of chitin deacetylase, revealed striking sequence similarities with the various rhizobial nodB proteins and a *B. stearothermophilus* protein encoded by an uncharacterized open reading frame located next to a fumarase gene (*fumA*). To highlight the most significant similarities, we performed a multiple sequence alignment using the deduced amino acid sequences of these genes. Furthermore, we calculated the pairwise identity and similarity scores between them, as an indication of their relative distances (Table 1).

The nodB protein sequences (≈215 amino acids long) are well conserved among all rhizobial species examined. The *B. stearothermophilus* sequence is somewhat longer (265 amino acids) and exhibits similarities extending from position number 53 toward its carboxyl terminus. Chitin deacetylase is a much longer polypeptide (400 amino acids), and the region that displays homology to the other proteins is located in the central part of the molecule. The amino-terminal (amino acids 1–121) and the carboxyl-terminal (approximately amino acids 325–400) regions do not display any significant similarity to any other known protein sequence. The sequence comparison between the fungal chitin deacetylase and the bacterial proteins is shown in Fig. 2. Invariant residues in all nodB sequences available in the data base are indicated by underlining. It is noteworthy that these invariant nodB residues are also extensively conserved in both chitin deacetylase and the *Bacillus* sequence. In contrast, six of nine predicted N-glycosylation sites in the chitin deacetylase sequence (Fig. 1) are found in the nonconserved regions of the protein, suggesting that these glycosylated domains may be important for the specialized function of the enzyme in cell wall biosynthesis.

The biochemical pathway for Nod factor synthesis is presently under intense investigation. Nod factors from different rhizobial species share a common basic structure: they are all *N*-acetylglucosamine oligomers, with an N-acyl substitution at the nonreducing end residue (24–28). This N-acyl substitution is required for the biological activity of all Nod factors, since *N*-acetyglucosamine oligomers fail to elicit any nodulation-specific responses in host plants (25). Furthermore, the similarity of nodC protein to chitin synthase is indicative that *N*-acetylglucosamine oligomers may serve as precursors in Nod factor synthesis (29–31). Given the observed similarity of chitin deacetylase to nodB proteins and the above information, it is reasonable to assume that *nodB* encodes a specific deacetylase for the nonreducing terminal residue of the precursors. This specific deacetylation may in turn provide the necessary free amino group for a subsequent N-acylation. While this paper was under review, the above suggested role of nodB protein in Nod factor biosynthesis was demonstrated biochemically by John *et al.* (32).

Some lysozyme-resistant strains of *Bacillus cereus* have been found to contain glucosamine residues in their cell wall peptidoglycan (33), and a specific enzyme activity that catalyzes the deacetylation of *N*-acetylglucosamine residues in peptidoglycan has been identified (34). The carbohydrate part of the peptidoglycan consists of alternating residues of *N*-acetylglucosamine and its lactic acid ether, *N*-acetylmuramic acid, linked by β1-4 glycosidic bonds. On the other hand, *Bacillus* species have been found to produce chitinases and chitosanases degrading chitinous substrates (35). In this process a specific deacetylation of chitin chains may be involved. The observed similarity of chitin deacetylase to the *B. stearothermophilus* gene product, together with the structural similarity of peptidoglycan to chitin, suggests that this gene encodes a peptidoglycan deacetylase, if not a chitin deacetylase.

The fungal chitin deacetylase, the rhizobial nodB proteins, and the deacetylase from *B. stearothermophilus* define a group of proteins with structural and functional homology, although evolutionarily distant. The fact that *N*-acetylglucosamine-6-phosphate deacetylase does not exhibit any significant sequence similarity to this group suggests that the protein conservation relates to both the hydrolysis of the acetamido groups on *N*-acetylglucosamine residues and the polymer structure (i.e., the type and configuration of the glycosidic bond).

The similarities that we have described highlight the catalytic domain in chitin deacetylase sequence and can direct the design of an enzyme with improved efficiency in the deacetylation of chitinous substrates.

We thank W. S. Lane for protein sequencing, A. Athanasiadis and Y. Papanikolaou for their help in sequence comparisons, C. Louis

and C. Savakis for many helpful suggestions, and F. C. Kafatos for critical reading of the manuscript. The chitin deacetylase technology is protected by patents issued to the Institute of Molecular Biology and Biotechnology.

1. Bartnicki-Garcia, S. & Nickerson, W. J. (1962) *Biochim. Biophys. Acta* **58,** 102–119.
2. Briza, P., Ellinger, A., Winkler, G. & Breitenbach, M. (1988) *J. Biol. Chem.* **263,** 11569–11574.
3. Araki, Y. & Ito, E. (1975) *Eur. J. Biochem.* **55,** 71–78.
4. Davis, L. L. & Bartnicki-Garcia, S. (1984) *Biochemistry* **23,** 1065–1073.
5. Rha, C. K., Rodriguez-Sanchez, D. & Kienzle-Sterzer, C. (1984) in *Biotechnology of Marine Polysaccharides*, eds. Colwell, R. R., Pariser, E. R. & Sinskey, A. J. (Hemisphere, Washington, DC), pp. 283–311.
6. Knorr, D. (1984) in *Biotechnology of Marine Polysaccharides*, eds. Colwell, R. R., Pariser, E. R. & Sinskey, A. J. (Hemisphere, Washington, DC), pp. 313–332.
7. Hirano, S. (1989) in *Chitin and Chitosan*, eds. Skjak-Braek, G., Anthonsen, T. & Sandford, P. (Elsevier, Essex, UK), pp. 37–43.
8. Kafetzopoulos, D., Martinou, A. & Bouriotis, V. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 2564–2568.
9. Lane, W. S., Galat, A., Harding, M. W. & Scheiber, S. L. (1991) *J. Protein Chem.* **10,** 151–160.
10. Chirwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18,** 5294–5299.
11. Glisin, V., Crkvenjakov, R. & Byus, C. (1974) *Biochemistry* **13,** 2633–2637.
12. Aviv, H. & Leder, P. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1408–1412.
13. Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. (1988) *Nucleic Acids Res.* **16,** 7583–7600.
14. Church, G. M. & Gilbert, W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 1991–1995.
15. Devereux, J., Haeberli, P. & Smithies, O. (1984) *Nucleic Acids Res.* **12,** 9509–9524.
16. Schartz, R. M. & Dayhoff, M. O. (1979) in *Atlas of Protein Sequence and Structure*, ed. Dayhoff, M. O. (Natl. Biomed. Res. Fnd., Washington, DC), pp. 353–358.
17. Schofield, P. R. & Watson, J. M. (1986) *Nucleic Acids Res.* **14,** 2891–2903.
18. Rossen, L., Jonhston, A. W. B. & Downie, J. C. (1984) *Nucleic Acids Res.* **12,** 387–395.
19. Krishnan, H. B. & Pueppke, S. G. (1991) *Mol. Pl.-Microbe Interact.* **4,** 512–520.
20. Toeroek, I., Kondorosi, E., Stepkowski, T., Postfai, J. & Kondorosi, A. (1984) *Nucleic Acids Res.* **12,** 9509–9524.
21. Velazquez, M. V., Davalos, A., Las Penas, A., Sanchez, F. & Quinto, C. (1991) *J. Bacteriol.* **173,** 1250–1258.
22. Scott, K. F. (1986) *Nucleic Acids Res.* **14,** 2905–2919.
23. Goethals, K., Gao, M., Toemekpe, K., Van Mondagu, M. & Holsters, M. (1989) *Mol. Gen. Genet.* **219,** 289–298.
24. Lerouge, P., Roch, P., Faucher, C., Maillet, F., Truchet, G., Promé, J. C. & Dénarié, J. (1990) *Nature (London)* **344,** 781–784.
25. Spaink, H. C., Sheeley, J. M., VanBrussel, A. A., Glushken, J., York, W. S., Tak, T., Geiger, O., Kennedy, E. P., Reinhold, V. N. & Lugtenberg, B. J. J. (1991) *Nature (London)* **354,** 125–130.
26. Roche, P., Debelé, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Dénarié, J. & Promé, J. C. (1991) *Cell* **67,** 1131–1143.
27. Truchet, G., Roch, P., Lerouge, P., Vasse, J., Camut, S., de Billy, J., Promé, J. C. & Dénarié, J. (1991) *Nature (London)* **351,** 670–673.
28. Schulze, M., Quiclet-Sire, B., Kondorosi, E., Virelizier, H., Glushka, J. N., Endre, G., Géro, S. & Kondorosi, A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 192–196.
29. Bulawa, C. E. (1992) *Mol. Cell. Biol.* **12,** 1764–1776.
30. Atkinson, E. M. & Long, S. R. (1992) *Mol. Pl.-Microbe Interact.* **5,** 439–442.
31. Debellé, F., Rosenberg, C. & Dénarié, J. (1992) *Mol. Pl.-Microbe Interact.* **5,** 443–446.
32. John, M., Röhrig, H., Schmidt, J., Wieneke, U. & Shell, J. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 625–629.
33. Araki, Y., Nakatani, T., Nakayama, K. & Ito, E. (1972) *J. Biol. Chem.* **247,** 6312–6322.
34. Araki, Y., Fukuoka, S., Oba, S. & Ito, E. (1971) *Biochem. Biophys. Res. Commun.* **45,** 751–758.
35. Shimahara, K., Takiguchi, Y. & Nagano, Y. (1992) in *Advances in Chitin and Chitosan*, eds. Brine, C. J., Sandford, P. A. & Zikakis, J. P. (Elsevier, Essex, UK), pp. 364–371.