New *Rhizobium leguminosarum* Flavonoid-Induced Proteins Revealed by Proteome Analysis of Differentially Displayed Proteins

Nelson Guerreiro, John W. Redmond, Barry G. Rolfe, and Michael A. Djordjevic

Plant-Microbe Interaction Group, Research School of Biological Sciences, Australian National University,
P.O. Box 475, Canberra City, Australia, 2601

Received 8 November 1996. Accepted 3 February 1997.

Proteome analysis was used to establish the first two-dimensional protein map of *Rhizobium. R. leguminosarum* bv. *trifolii* strain ANU843 was grown in defined medium in the presence and absence of the flavonoid 7,4'-dihydroxyflavone. Over 1,700 constitutive proteins were resolved, representing about 30% of the estimated genomic output. Proteome analysis of flavonoid-treated cells was done to reveal differentially displayed proteins. The results showed that while the global expression pattern of proteins was largely unaltered by the treatment, four inducible proteins were observed. The four inducible proteins and 20 constitutively expressed proteins were subjected to sequence analysis to provide internal standards for the construction of a two-dimensional *Rhizobium* protein data base. The identity of 12 proteins, including NodE and NodB, was established. NodE was present throughout the growth of the cells but was diminished in amount in stationary phase cells whereas NodB was not detected in the later stages of growth. Two of the induced proteins sequenced did not match any known nodulation gene product, with one of these being present in mid-late log and stationary phase cells and possessing four consecutive His residues at the N terminus. N-terminal sequencing was successful with 100 to 200 fmol of protein. Proteome analysis provides a sensitive new tool to examine plant-microbe interactions.

Additional keywords: amino acid composition (AAC) analysis, nodulation genes, *Rhizobium* ChvE homologue, two-dimensional gel electrophoresis.

Rhizobia (*Rhizobium*, *Bradyrhizobium*, and *Azorhizobium*) infect plants in the family Leguminosae (legumes) and induce nitrogen-fixing root nodules. A central role in promoting the infection process is provided by plant-secreted flavonoids (Redmond et al. 1986). In the early stages of legume-*Rhizobium* symbiosis, rhizobia are chemotactic toward flavonoids and other molecules released in the plant root exudates (including sugars) and move toward localized sites on the legume root (Gaworzewska and Carlilie 1982; Djordjevic et al. 1987; Cae-tano-Anolles et al. 1988). In addition, a series of nodulation

genes are rapidly induced by these plant signals (Peters et al. 1986; Redmond et al. 1986; Djordjevic et al. 1987) via a mechanism involving the NodD family of transcription activators. Because of the central role played by flavonoids in rhizosphere colonization and infection, it is likely that rhizobia may have coupled the expression of other rhizosphere- or infection-related genes to the presence of these molecules.

To investigate the effects of flavonoids upon *Rhizobium* gene expression, we have used the technique of proteome analysis (Wilkins et al. 1996a) to evaluate the changes in *Rhizobium* gene expression in response to the presence of such flavonoids. Proteome analysis exploits the high resolving power of two-dimensional gel electrophoresis (2-DE) for the global analysis of cellular gene expression (Jungblut and Wittmann-Liebold 1995; Kamo et al. 1995; Pasquali et al. 1996; Wilkins et al. 1996b), in conjunction with a range of post separation techniques (Jungblut et al. 1996; Wheeler et al. 1996; Yan et al. 1996) for the identification of individual proteins. With the availability of immobilized pH gradients (IPG) (Bjellqvist et al. 1982; Görg et al. 1988), several thousands of proteins can be resolved on a single gel to provide reproducible patterns with sufficient resolution for the parallel study of the response of many genes or proteins to different biological conditions. Comparisons can be made (Celis et al. 1994; Knecht et al. 1994; Graumann et al. 1996; van Bogelen et al. 1996) between the generated protein maps for characterization of what we term differentially displayed proteins (DDPs), which are induced or suppressed under different conditions.

Rhizobium is a genus currently under-represented in the available data bases, as is the case with many agriculturally important bacteria. Most of the genes identified to date are associated with infection, polysaccharide production, or nitrogen fixation (van Rhijn and Vanderleyden 1995; Gray and Rolfe 1990; Fischer 1994). With increasing interest in the complex regulatory changes that occur in *Rhizobium*-legume interactions under different environments, together with the ever-increasing availability of mutants, it is appropriate to characterize the *Rhizobium* gene expression via proteome analysis. In organisms with extensive data bases, amino acid composition (AAC) analysis combined with a knowledge of the molecular mass and isoelectric point have proved useful parameters for rapid protein screening and identification (Jungblut et al. 1994; Galat et al. 1995; Wilkins et al. 1996a). More generally, improvements in the sensitivity of N-terminal sequencing

Corresponding author: Michael A. Djordjevic; Fax: 61 6 2490754;
E-mail: michael@rsbs-central.anu.edu.au

and the development of sequencing-stable membranes for the immobilization of proteins have enabled sequencing of proteins in the low picomole range (Matsudaira 1987).

In this paper we present the first 2-DE protein map for a *Rhizobium* species. By using N-terminal sequencing, we have characterized 20 constitutive proteins, several of which appear to be novel. These proteins provide the basis of a 2-DE *Rhizobium* protein data base. N-terminal sequencing proved to be a reliable, rapid, and sensitive identification tool for unknown proteins present at 100-fmol levels. In addition, we have examined the response of *R. leguminosarum* bv. *trifolii* strain ANU843 to growth in the presence of the flavonoid 7,4'-dihydroxyflavone (DHF). The addition of DHF does not result in global changes in the levels of expressed proteins, but four specifically induced proteins were identified, including NodE, NodB, and a phase-dependent protein with a four-histidine repeat motif at the N terminus. The *Rhizobium* protein map provides a rapid and sensitive link between genomic sequences and expressed protein and speeds up the cellular characterization of new open reading frames (ORFs), while offering a global approach to the study of genome organization and function in rhizobia under different biological conditions.

RESULTS

Rhizobia show growth phase-dependent expression of nodulation genes.

Before proteome analysis was undertaken, the *nod* gene expression of strain ANU843 in the presence and absence of the DHF inducer was evaluated throughout its growth cycle. These experiments were conducted to establish a cell harvesting time that favors maximal *nod* gene expression. ANU843 derivatives with translational or transcriptional *lacZ* fusions to *nodA* and *nodE*, respectively, were evaluated for β -galactosidase activity after a single addition of 2 μ M DHF to lag phase cells (Djordjevic et al. 1987). Figure 1 shows the expression of *nodA* and *nodE* was maximal in cells in early log phase at 6 to 8 h after the addition of DHF. The β -galactosidase activity decreased during mid-log phase and reduced to near basal levels by the end of log phase. In addition, the presence of DHF reduced markedly the growth of cells containing the *nodA::lacZ* fusion but did not severely effect the growth of the cells containing the *nodE::lacZ* fusion (Fig. 1) or the wild-type strain ANU843 (results not shown).

Analysis of proteins affected by flavonoid addition.

The proteome for strain ANU843 cells grown in defined (BIII) medium in the presence and absence of DHF is shown in Figures 2, 3, 4, and 5. Excellent reproducibility and resolution of protein spots after silver staining were obtained. Image analysis revealed over 1,700 protein spots within pH 4 to 7 (Figs. 2 and 3) and pH 3 to 10 ranges (Figs. 4 and 5).

Early log phase cells were exposed to DHF for a period of 6 h before harvesting and the proteins subsequently extracted for 2-DE analysis. Total protein expression patterns obtained after DHF exposure are represented in Figures 3 and 5. Image analysis was used to superimpose the images of the flavonoid-treated and nontreated cells to identify DDPs. Subtraction of one image from the other confirmed that the addition of DHF did not result in a drastic change in gene expression at the protein level. However, three flavonoid-induced proteins were

identified at this stage of growth. These protein differences are marked with arrows and located within the enlarged framed areas of Figure 6.

The effect of growth phase was also assessed by harvesting cells at mid-late log (corresponding to 1×10^9 cells/ml) and stationary phase. Little difference was observed in the protein patterns of these cells compared with early log phase cells (data not shown), except that an additional induced protein corresponding to spot 165 (Fig. 6) was found in both mid-late log and stationary phase cells.

N-terminal sequencing of constitutive and induced proteins generates reliable data at 100 fmol.

A two-dimensional data base for strain ANU843 was initiated by subjecting a total of 20 constitutive protein spots and the four flavonoid-induced protein spots to N-terminal sequencing. Reliable sequence data to 10 or more residues was obtained for proteins present in the 100 to 200 fmol range, which is equivalent to 3.5 ng for a 35-kDa protein. This result demonstrates the remarkable purity of proteins isolated from the polyvinylidene difluoride membranes as well as the sensitivity of N-terminal sequencing. For example, Figure 7 shows the chromatograms of the first 10 residues obtained from the N-terminal sequence of spot 70, which share homology to as-

Fig. 1. Strain ANU843 *nod* gene expression throughout its growth cycle, assessed by assaying for β -galactosidase activity in the presence and absence of 7,4'-dihydroxyflavone (DHF). Cells containing (A) a *nodA::lacZ* translational fusion or (B) a *nodE::lacZ* transcriptional fusion were inoculated to an OD₆₀₀ of 0.01 in BIII medium and exposed to a single addition of DHF (2 μ M final). Cells were taken at various time points during their growth cycle and the level of *nod* gene induction was measured as units of β -galactosidase activity.

partate aminotransferase of *R. meliloti*. The high repetitive yields and stable reproducible chromatograms evident for every cycle allowed for unambiguous interpretation of the sequence at these low levels. To our knowledge, this is the first report of N-terminal sequencing with these low amounts of materials.

Sequence alignment searches of the SWISS-PROT and GenBank data bases revealed that 10 constitutive proteins exhibited high sequence homology to previously identified proteins (Table 1), with the majority of the matched proteins being from species other than rhizobia. Some of the *R. leguminosarum* bv. *trifolii* proteins corresponded to previously hypothetical proteins known only from DNA sequence analysis of ORFs. For example, the N-terminal sequence of spot 144 has a 100% match to the previously hypothetical protein of *R. leguminosarum* bv. *viciae* specified by a gene in the ORF 240 5' region (Colonna-Romano et al. 1990) and shows homology to acyl carrier proteins. Similarly, spot 78 matches the previously hypothetical ChvE protein of *A. tumefaciens* that is required for virulence gene induction and function as a periplasmic receptor protein involved in chemotaxis and uptake of sugars (Huang et al. 1990). Furthermore, spots 144 and 78 contained inferred leader sequences, since

the first residue was 15 and 25 residues, respectively, from the methionine initiation site predicted by DNA sequencing, therefore suggesting that the proteins at spots 144 and 78 are most likely periplasmic or membrane-located (Broome-Smith et al. 1994). Nine proteins yielded N-terminal sequences with no significant homology to previously described proteins, indicating they are likely to be novel. Furthermore, the abundance of some of these unknown proteins suggests that they play important functional roles in general cell metabolism. No sequence information could be obtained for one protein, presumably because it was N-terminally blocked.

Table 2 contains data on the flavonoid-induced proteins, including those characterized by N-terminal sequencing. Of the flavonoid-induced proteins, spots 162 and 163 were identified as NodE and NodB, respectively. NodE (and spot 164) was present in log phase cells, and still present in stationary phase cells, but in lower amounts, even though stationary phase cells showed negligible *nodE* expression (Fig. 1). Consistent with Figure 1, NodB was present in early log phase cells but diminished to nondetectable amounts in late log and stationary phases. Sequencing of the protein spot 165 revealed an N terminus with four consecutive histidines from residues 3 to 6

Fig. 2. Two-dimensional gel electrophoresis (2-DE) reference protein map of total *Rhizobium leguminosarum* bv. *trifolii* (ANU843) proteins after silver staining. Isoelectric focusing in the first dimension was carried out on linear pH 4 to 7, 18-cm immobilized pH gradient (IPG) strips loaded with 100 µg of total cellular protein. The numbered proteins illustrate those that were randomly selected for N-terminal sequencing and correspond to those listed in Table 1. M.m. = molecular mass.

(Table 2) but no significant homology to previously described proteins was found. Although spot 164 was present in amounts too low for confident sequencing, none of the different permutations possible with the sequence obtained for this protein matched any known Nod proteins.

In *Rhizobium*, protein identification with AAC analysis is currently unreliable.

The AAC of 88 randomly chosen constitutive proteins was also analyzed. The AAC together with the molecular mass and isoelectric point parameters of individual proteins were subjected to cross-species matching against all SWISS-PROT data base entries. Single-species matching was not practical because *Rhizobium* is under-represented in the available protein data bases. The computer-generated score patterns that resulted enabled only three of 88 proteins (aspartate aminotransferase and flagellin of *R. meliloti* and a prokaryotic 50S ribosomal protein) to be identified with confidence. The remaining 85 data base searches resulted in score patterns showing little or no difference between the top-ranked and the next highest ranked proteins but the correct protein was, in some cases, among those with a high ranking.

DISCUSSION

The examination of *nod* gene expression during the growth cycle revealed that early log phase cells responded strongly to a single addition of DHF. By stationary phase the expression of *nodA* and *nodE* was reduced almost to background levels. This strong response of late lag/early log phase cells to DHF is consistent with previously published results (Djordjevic et al 1987). The expression of *nodE* remained elevated during log phase whereas *nodA* expression showed a continual decline from the peak at early log phase. This is consistent with NodE being present in cells throughout the growth cycle and NodB being absent in late log and stationary phases. These results also suggest that NodE is turned over slowly after the decline in *nodE* gene expression but NodB does not persist after *nodA* expression is down-regulated. NodE is an inner-membrane protein, containing at least two nonpolar transbilayer alpha-helices (Spaink et al. 1989). There is evidence to suggest that NodE is responsible for alterations to the degree of unsaturation of the fatty acyl group of the ANU843 lipo chito-oligosaccharides (Bloemberg et al. 1995; Philip-Hollingsworth et al. 1995), and these results suggest that NodE should function throughout the growth of

Fig. 3. Two-dimensional gel electrophoresis protein map of total protein from *Rhizobium leguminosarum* bv. *trifolii* (ANU843) in response to 7,4'-dihydroxyflavone (DHF). Isoelectric focusing in the first dimension was carried out on linear pH 4 to 7, 18-cm immobilized pH gradient (IPG) strips loaded with 100 µg of total cellular protein. Comparisons of DHF-treated and nontreated protein maps revealed differences located within the framed areas, with numbered arrows pointing to the induced spots. M.m. = molecular mass.

ANU843 cells. It is clear that the addition of DHF markedly slows the growth of the cells containing the *nodA::lacZ* fusion but has less effect upon cells containing the *nodE::lacZ* fusion. This is possibly due to polar effects upon *nodBCIJT* expression, which lie on the same transcript as *nodA* (Innes et al. 1985; Djordjevic et al. 1987). NodI, NodJ, and NodT are thought to form a membrane-associated transporter complex (Vasquez et al. 1993; Rivilla et al. 1995; Spaink et al. 1995) and perturbation of these functions may have pleiotropic effects that extend beyond Nod factor secretion (McKay and Djordjevic 1993).

The data presented establish the first standard 2-DE protein reference map for *Rhizobium* bacteria and will be submitted to the SWISS-2DPAGE data base. The size of the *R. leguminosarum* bv. *trifolii* genome is unknown but is estimated to be no larger than that reported for *R. meliloti* (6.5 Mbp) (Sobral et al. 1991), which would code for about 5,000 average sized proteins, each of 330 amino acids. Considering that over 1,700 distinct proteins were detected, we calculate about 30% of the *R. leguminosarum* bv. *trifolii* coding capacity was visualized on one 2-DE gel.

The addition of flavonoid to ANU843 cells did not result in global changes in gene expression. Of the known nodulation proteins, only NodE and NodB were identified. The two other

induced proteins of 20.9 and 13.2 kDa showed no significant homology to the known nodulation proteins, with the latter being phase-dependent and containing a histidine-rich region at the N terminus. We are yet to establish whether these observed protein changes are Sym plasmid- or NodD-dependent.

The protein differences identified on the protein maps represent a minimum since 12 nodulation genes are known to be induced in *R. leguminosarum* bv. *trifolii* (Schofield and Watson 1986; Djordjevic et al. 1987; Spaink et al. 1989; Djordjevic and Weinman 1991; Lewis-Henderson and Djordjevic 1991). Whereas proteome analysis possesses high resolution power and sensitivity, there may be limitations in the analysis of total cellular protein. The inability to detect some proteins may indicate that they (i) are present in relatively low amounts, (ii) are not soluble, (iii) co-migrate with more abundant proteins, or (iv) are of very low molecular masses and are not resolved in the second dimension. Most Nod proteins in rhizobia are present in very low amounts (Schlaman et al. 1989), making it difficult to characterize them biochemically. In the past, the characterization of several Nod proteins has relied on coupled transcription/translation systems and the expression of cloned genes and gene fusions in *E. coli* (Dusha et al. 1986; Schmidt et al. 1986; Schmidt et al. 1988; Spaink et

Fig. 4. Two-dimensional gel electrophoresis (2-DE) separation of total *Rhizobium leguminosarum* bv. *trifolii* (ANU843) proteins in the pH 3 to 10 range. ANU843 proteins were separated by 2-DE with nonlinear pH 3 to 10, 18-cm immobilized pH gradient (IPG) strips in the first dimension loaded with 100 µg of total cellular protein. M.m. = molecular mass.

al. 1989). Given that silver staining can detect as low as 1 to 2 ng (or 142 pg/mm²) of standard protein (N. Guerreiro and M. Djordjevic, *unpublished data*; Rabilloud 1992), our inability to detect a large range of known Nod proteins is most likely due to their very low relative abundance in cells. Therefore, 2-DE analysis of cellular subfractions having an increased relative amount of low abundant proteins is likely to be a more promising approach, in particular to examine some Nod proteins (for example NodC, NodI, NodJ, NodT) that are thought to be membrane-located, likely to show poor solubility, and are present in low abundance. Increasing the resolving power of the first dimension with narrow isoelectric point range IPG separations (1 pH unit), together with a high sample load, should also provide a possible solution to the problem encountered with characterization of low-abundance and poorly resolved proteins such as spot 164. Of the known Nod proteins, only NodF (9.9 kDa) would be too small to resolve with the two-dimensional gel system used in this study. NodF protein is translationally coupled with NodE (Schofield and Watson 1986) and would therefore be expected to be present in similar quantities.

Twenty constitutive proteins were sequenced and assigned

to the *R. leguminosarum* bv. *trifolii* reference map and 88 others were subjected to AAC analysis. Ten of these were identified on the basis of their sequence homology, with two matching currently hypothetical proteins and several matching proteins found in other prokaryotes but not previously reported in rhizobia. Of those proteins identified, the *Rhizobium* ChvE homologue and the potential acyl carrier protein are of significance to *Rhizobium*-legume interactions. ChvE is involved in chemotaxis and sugar uptake in *Agrobacterium* (Huang et al. 1990). To our knowledge this is the first report of this protein in *Rhizobium* and it may reflect a general conservation of functions seen between *Agrobacterium* and *Rhizobium* that extends to β -glucan and polysaccharide synthesis (Gray and Rolfe 1990). The potential acyl carrier protein could be involved in Nod factor biosynthesis since *nodF* mutants retain an acyl group attached to the oligo chitin chain (Ardourel et al. 1994).

Success in identifying proteins has been achieved in the genetically well characterized *E. coli* with AAC analysis (Wilkins et al. 1996a), but only three of 88 unknown *Rhizobium* proteins were confidently matched across species. These results show that AAC analysis is not yet a reliable protein

Fig. 5. Two-dimensional gel electrophoresis separation of total *Rhizobium leguminosarum* bv. *trifolii* (ANU843) protein after 7,4'-dihydroxyflavone (DHF) exposure. The total proteins of ANU843 produced in cells exposed to DHF is shown. Isoelectric focusing in the first dimension was carried out on nonlinear pH 3 to 10, 18-cm immobilized pH gradient (IPG) strips loaded with 100 μ g of total cellular protein. Comparisons of pH 3 to 10 DHF-treated and nontreated protein maps did not reveal differentially displayed proteins in addition to those identified in Figure 2. Extended pH range was tested because several known Nod proteins have predicted isoelectric points outside the pH 4 to 7 range. M.m. = molecular mass.

identification tool for poorly characterized species, especially if protein homology is not strongly conserved. Similar conclusions have been reached with AAC analysis and cross-species matching for *Mycoplasma genitalium* (Wasinger et al. 1994).

This study introduces the use of proteome analysis as a tool for monitoring global changes in *Rhizobium* gene expression in response to defined stimuli. In other studies, the effect of mutation has been examined. Some mutants show little change in protein patterns, whereas others show multiple pleiotrophies in protein expression (S. Worland, N. Guerreiro, M. Djordjevic, and B. Rolfe, *unpublished data*) that would be difficult to demonstrate or examine via other methodologies. We have also applied proteome analysis to subterranean clover and identified several proteins induced or up-regulated by the action of chewing mites (J. J. Weinman, P.M.I. lab, *unpublished*). The high resolving power, reproducibility of protein patterns, and high purity of the excised proteins, together with the ability to sequence at <100 fmol levels, permit a rapid study of the gene products of interest, especially when looking for DDPs. Moreover, proteins that are regulated as a result of imposed conditions can be identified without resort to use mutagenesis. Thus, this technique should have wide utility in examining plant-microbe interactions.

MATERIALS AND METHODS

Assays of *nod* gene induction and expression.

Bacterial strains used for assay of *nod* gene induction and expression were derivatives of the wild-type strain ANU843 containing either a translational fusion of Mud II 1734 insertion in *nodA* (ANU894) (McIver et al. 1989) or a transcriptional fusion of Mud I 1734 to *nodE* (*nodE1027*; Djordjevic et al. 1987). The β -galactosidase activity generated by bacterial cultures was assayed as previously described (Djordjevic et al. 1987). Cultures were inoculated to an OD₆₀₀ of 0.01 in BIII medium (Dazzo 1982) with and without 2 μ M DHF. Samples were taken subsequently every 2 to 4 h throughout the growth

cycle and *nod* gene expression assessed by assaying for β -galactosidase activity (Djordjevic et al. 1987).

Growth conditions and sample preparation.

All chemicals, unless stated, were purchased from Sigma (St. Louis, MO) and were of analytical or electrophoresis grade. Strain ANU843 was grown at 28°C, in BIII medium to late lag phase before DHF was added to 2 μ M and cells incubated for 6 h before harvesting. Cells were harvested by centrifugation at $6,000 \times g$ for 10 min at 4°C, and the pellet was washed three times for 10 min at the same speed in low-salt washing buffer containing 68 mM NaCl, 3.0 mM KCl, 1.5 mM KH₂PO₄, 9.0 mM NaH₂PO₄ (Humphery-Smith et al. 1994). The cells were lysed by sonication after a fivefold dilution in a lysis solution containing 8 M urea, 4% wt/vol 3-[(3-cholamidopropyl)-dimethyl-ammonio]-1-propanesulfonate (CHAPS), 1% wt/vol dithiothreitol (DTT), 0.8% wt/vol Bio-Lyte 3-10 (Bio-Rad, Richmond, CA), 35 mM Tris, 1 mM phenylmethylsulfonyl fluoride, and 5 mM EDTA. Insoluble material and nucleic acids were removed by ultracentrifugation at $200,000 \times g$ for 1 h at 15°C. The protein concentration of the supernatant was determined by the Bradford dye-binding assay (Bio-Rad) with bovine serum albumin as the standard.

2-DE.

2-DE was done as previously described (Görg et al. 1988; Görg et al. 1995) with minor modifications. For the first dimension, IPG gel strips (180 mm, linear pH 4 to 7 and nonlinear pH 3 to 10; Pharmacia-Biotechnology, Uppsala, Sweden) were rehydrated overnight in a solution containing 8 M urea, 0.5% wt/vol CHAPS, 0.2% wt/vol DTT, 0.52% wt/vol Bio-Lyte 3-10 and a trace of bromophenol blue. For analytical gels, an aliquot containing 100 μ g of protein was diluted with sample buffer (8 M urea, 4% wt/vol CHAPS, 1% wt/vol DTT, 0.8% wt/vol Bio-Lyte 3-10, 35 mM Tris, and 0.02% wt/vol bromophenol blue) to a volume of 50 μ l for application to each IPG gel strip. For preparative gels, an aliquot containing

Fig. 6. Enlargement of the framed areas illustrated in the protein maps. The area on the two-dimensional gel showing the 7,4'-dihydroxyflavone-induced protein changes is enlarged. Spot differences are marked with arrows and numbered corresponding to those in Table 2.

0.5 to 1 mg of protein was diluted with sample buffer to a loading volume of 100 μ l. Sample application was at the anodic end and isoelectric focusing (IEF) was run with the Immobililine DryStrip kit in a Multiphore II electrophoresis unit (Pharmacia-Biotechnology) for 200 kVh at 20°C. Following IEF, IPG gel strips were equilibrated for 10 min in a solution of Tris-HCl buffer (0.05 M, pH 6.8) containing 6 M urea, 30% wt/vol glycerol (ICN, Costa Mesa, CA), 2% wt/vol sodium dodecyl sulfate (SDS), and 2% wt/vol DTT followed by a further 10-min equilibration in the same buffer containing 4% iodoacetamide (2.5% for preparative gels) instead of DTT and 0.02% bromophenol blue. The equilibrated IPG gel strips

were then placed onto Pharmacia ExcelGel SDS gels (T = 12 to 14% acrylamide) for second dimension SDS-polyacrylamide gel electrophoresis on the Multiphore II unit.

Image processing.

Proteins on analytical 2-DE gels were visualized by silver staining (Rabilloud 1992) and digitized at 600 dpi with a UMAX PS-2400X lamp scanner run under Photoshop 3.0 (Adobe, Mountain View, CA). The resulting images were transferred into the UNIX environment of a SPARC station 2 (Sun Microsystems, Mountain View, CA) for analysis. Spot detection and gel-to-gel protein spot matching were performed

Fig. 7. Chromatograms of residues 1 to 10 obtained by N-terminal sequencing of spot 70. Peak height of residue 1 corresponds to 200 fmol of initial sequencing material. Repetitive yield for each cycle was 97%. Bold single letter amino acid abbreviations above the peaks indicate residue obtained for each cycle during sequencing. Confident calls were obtained with 100 fmol of material.

with MELANIE II software (Bio-Rad, Hercules, CA) run under OpenWindows 3.0. The apparent molecular masses of the proteins were determined by co-electrophoresis with the following standard proteins (Pharmacia-Biotechnology): phosphorylase b (94 kDa), albumin (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa), trypsin inhibitor (20.1 kDa), and lactalbumin (14.4 kDa).

Electroblotting, N-terminal sequencing, and amino acid analysis.

Electroblotting onto polyvinylidene difluoride membranes (Bio-Rad) was done as previously described (Matsudaira 1987), with 10 mM 3-[cyclohexylamino]-l-propanesulfonic

acid, pH 11.0, in 10% vol/vol methanol and 0.01% SDS as transfer buffer. Electrotransfer was performed at 0.8 mA/cm² for 90 min. Membranes were stained with 0.1% wt/vol Coomassie brilliant blue R-250 (Bio-Rad) in a 40% methanol solution for 5 min and then destained in a 50% methanol, 5% acetic acid solution for 10 min. The membranes were then washed in water and allowed to air dry before storage at -20°C. Selected spots were excised from dried membranes and N-terminal amino acid sequences were done on a PROCISE 494-01 sequencer system (Perkin-Elmer Applied Biosystems, Foster City, CA). Amino acid sequence homology comparisons were made with the FASTA and TFASTA programs for the screening of protein and nucleic acid data bases, respec-

Table 1. N-terminal sequencing of constitutive *Rhizobium leguminosarum* bv. *trifolii* proteins^a

Rhizobium polypeptide characteristics				Characteristics of matching polypeptide from data base				
Spot no.	M.m. (Da)	pI	N-terminal sequence	Protein (matching residues)	Organism	M.m. (Da) ^b	pI ^b	SWISS-PROT data base ref.
13	81,257	5.40	AREYKIEDYR	Unknown				
31	67,276	5.03	MQVIETLAFG	Unknown				
36	63,280	5.23	ASKEIKFGRT	GroEL (2-11)	<i>R. leguminosarum</i> bv. <i>viciae</i>	57,883	5.11	ch60_rhiv
57	54,000	5.12	AEAATPKIXXPQVMV ^c	Unknown				
62	60,660	5.79	APRVLVPTSE	Unknown				
70	48,100	6.35	AFLADALSQV	Aspartate aminotransferase (2-11)	<i>R. meliloti</i>	43,582	5.81	aata_rhime
72	58,809	6.45	MDIWAAEIQA	Unknown				
76	37,500	4.52	TSINTNNAAM	Flagellin (2-11)	<i>R. meliloti</i>	40,590	4.35	flaa_rhime
78	36,830	4.68	HDKGTVGIAM	ChvE (26-35)	<i>Agrobacterium tumefaciens</i>	24,138	6.51	chve_agrtu
80	40,311	4.90	DVLIIVAGPL	Unknown				
84	42,810	5.35	ETLTATVNN	Unknown				
85	44,355	5.46	Blocked	Unknown				
90	39,116	5.67	GFKVAIAGAT	Malate dehydrogenase (42-51)	<i>Glycine max</i> (soybean)	33,725	7.74	mdhg_soybn
91	37,165	5.85	ARNKIALIGSGMIGG	Malate dehydrogenase (1-15)	<i>P. vesicularis</i> / <i>diminuta</i>	NA ^d	NA ^d	mdh_breve, mdh_bredi
133	20,498	6.64	SISYDLKVQ	Unknown				
140	20,304	6.28	AEIKDPENTILET	Unknown				
144	12,045	4.39	TATFDKVADIIAETSEIDRA	Hypothetical protein in 5' ORF-240 (15-34) ^e	<i>R. leguminosarum</i> bv. <i>viciae</i>	NA ^d	NA ^d	yfx1_rhiv
159	13,477	5.63	ASTNFRPLHD	GroES A (2-11)	<i>R. meliloti</i>	10,587	5.43	ch11_rhime
160	15,890	4.80	ADLAKIVDDL	50S ribosomal protein (2-11)	<i>Brucella abortus</i>	12,546	4.79	r171_bruab
161	74,148	4.98	AKVIGIDLGT	DnaK (1-10)	<i>T. ferrooxidans</i>	NA ^d	NA ^d	dnak_thife

^a Constitutive *R. leguminosarum* bv. *trifolii* proteins were characterized with N-terminal sequencing and assigned to the two-dimensional gel electrophoresis reference protein map. Spot numbers were assigned arbitrarily. M.m. (molecular mass) and pI (isoelectric point) were estimated from comparisons to standards. There was generally good agreement between the experimental M.m., pI values of the tested protein, and theoretical values of the matched protein in the data base.

^b Theoretical values obtained from the amino acid sequence of the proteins.

^c Amino acid residues 9 and 10 were lost during sequencing.

^d Unable to calculate the theoretical pI and M.m. values for these proteins because only fragments of their sequence were available in the data base.

^e This protein has homology to several acyl carrier proteins.

Table 2. List of the identified flavonoid-induced differentially displayed proteins

Spot no.	M.m. (Da) ^a	pI ^b	N-terminal sequence	Effect of DHF ^c	Protein (matching residues)	Organism	M.m. (Da) ^d	pI ^d	SWISS-PROT data base ref.
162	50,810	5.58	MDRRVVIT	Induction	NodE (1-8)	<i>R. leguminosarum</i> bv. <i>trifolii</i>	42,001	4.98	node_rhilt
163	29,033	6.16	MKRRAYISEV	Induction	NodB (1-10)	<i>R. leguminosarum</i> bv. <i>trifolii</i>	23,541	5.85	nodb_rhilt
164	20,953	6.38	QR[Y/T]Q[N/K][E/T]V	Induction	Unknown				
165 ^e	13,213	6.65	ADHHHHDNTGAA	Induction	Unknown				

^a Molecular mass.

^b Isoelectric point.

^c 7,4'-dihydroxyflavone.

^d Theoretical values obtained from the amino acid sequence of the proteins.

^e Spot identified in both mid-late lag and stationary phase cells.

tively. SWISS-PROT and GenBank data bases were searched. AAC analysis and data base matching were essentially as previously described, with cross-species isoelectric point and molecular mass windows of ± 2 units and $\pm 20\%$, respectively (Wilkins et al. 1996a; Cordwell et al. 1995).

ACKNOWLEDGMENTS

We wish to thank Jill McGovern and Denis Shaw from the Biomolecular Resource Facility, Australian National University, for their expertise in performing the amino acid sequence analysis. We would also like to thank Jin Yan, Ki Li Ou, and Andrew Gooley from Macquarie University Center for Analytical Biotechnology, for amino acid composition analysis.

LITERATURE CITED

- Ardourel, M., Demont, N., Debelle, F. D., Maillet, F., Debilly, F., Prome, J. C., Denarie, J., and Truchet, G. 1994. *Rhizobium meliloti* lipooligosaccharide nodulation factors- Different structural requirements for bacterial entry into target root hairs and induction of plant symbiotic developmental responses. *Plant Cell* 6:1357-1374.
- Bjellqvist, B., Ek, K., Righetti, P. G., Gianazza, E., Görg, A., Westermeier, R., and Postel, W. 1982. Isoelectric focusing in immobilized pH gradients: Principle, methodology and some applications. *J. Biochem. Biophys. Methods* 6:317-339.
- Bloembergen, G. V., Kamst, E., Harteveld, M., van der Drift, K. M., Haverkamp, J., Thomas-Oates, J. E., Lugtenberg, B. J., and Spaink, H. P. 1995. A central domain of *Rhizobium* NodE protein mediates host specificity by determining the hydrophobicity of fatty acyl moieties of nodulation factors. *Mol. Microbiol.* 16:1123-1136.
- Broome-Smith, J. K., Gnaneshan, S., Hunt, L. A., and Mehraein-Ghomi, F. 1994. Cleavable signal peptides are rarely found in bacterial cytoplasmic membrane proteins. *Mol. Membr. Biol.* 11:3-8.
- Caetano-Anolles, G., Crist-Estes, D. K., and Bauer, W. D. 1988. Chemotaxis of *Rhizobium meliloti* to the plant flavone luteolin requires functional nodulation genes. *J. Bacteriol.* 170:3164-3169.
- Celis, J. E., Rasmussen, H. H., Olsen, E., and Madsen, P. 1994. The human keratinocyte two dimensional protein database: Towards an integrated approach to the study of cell proliferation, differentiation and skin disease. *Electrophoresis* 15:1349-1458.
- Colonna-Romano, S., Arnold, W., Schluter, A., Boistard, P., Puhler, A., and Priefer, U. B. 1990. An FNR-life protein encoded in *Rhizobium leguminosarum* biovar *viciae* shows structural and functional homology to *Rhizobium meliloti* FixK. *Mol. Gen. Genet.* 223:138-147.
- Cordwell, S. J., Wilkins, M. R., Cerpa-Poljak, A., and Gooley, A. A. 1995. Cross-species identification of proteins separated by two-dimensional gel electrophoresis using matrix-assisted laser desorption ionisation/time-of-flight mass spectrometry and amino acid composition. *Electrophoresis* 16:438-443.
- Dazzo, F. B. 1982. Leguminous root nodules. Pages 431-446 in: *Experimental Microbial Ecology*. R. G. Burns and J. L. Slater, eds. Blackwell Scientific, Oxford, UK.
- Djordjevic, M. A., Redmond, J. W., Batley, M., and Rolfe, B. G. 1987. Clovers secrete specific phenolic compounds which either stimulate or repress *nod* gene expression in *Rhizobium trifolii*. *EMBO J.* 6:1173-1179.
- Djordjevic, M. A., and Weinman, J. J. 1991. Factors determining host recognition in the clover-*Rhizobium* symbiosis. *Aust. J. Plant Physiol.* 18:543-557.
- Dusha, I., Schroder, J., Putnoky, P., Banfalvi, Z., and Kondorosi, A. 1986. A cell-free system from *Rhizobium meliloti* to study the specific expression of nodulation genes. *Eur. J. Biochem.* 160:69-75.
- Fischer, H. M. 1994. Genetic regulation of nitrogen fixation in rhizobia. *Microbiol. Rev.* 58:352-386.
- Galat, A., Bouet, F., and Riviere, S. 1995. Amino acid compositions of proteins and their identities. *Electrophoresis* 16:1095-1103.
- Gaworzewska, E. T., and Carlisle, M. J. 1982. Positive chemotaxis of *Rhizobium leguminosarum* and other bacteria towards root exudates from legumes and other plants. *J. Gen. Microbiol.* 128:789-798.
- Görg, A., Boguth, G., Obermaier, C., and Posch, A. 1995. Two-dimensional polyacrylamide gel electrophoresis with immobilized pH gradients in the first dimension (IPG-Dalt): The state of the art and the controversy of vertical versus horizontal systems. *Electrophoresis* 16:1079-1086.
- Görg, A., Postel, W., and Gunther, S. 1988. The current state of two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 9:531-546.
- Graumann, P., Schroder, K., Schmid, R., and Marahiel, M. A. 1996. Cold shock stress-induced proteins in *Bacillus subtilis*. *J. Bacteriol.* 178:4611-4619.
- Gray, J. X., and Rolfe, B. G. 1990. Exopolysaccharide production in *Rhizobium* and its role in invasion. *Mol. Microbiol.* 4:1425-1431.
- Huang, M. L., Cangelosi, G. A., Halperin, W., and Nestor, E. W. 1990. A chromosomal *Agrobacterium tumefaciens* gene required for effective plant signal transduction. *J. Bacteriol.* 172:1814-1822.
- Humphery-Smith, I., Guyonnet, F., and Chastel, C. 1994. Polypeptide cartography of *Spiroplasma taiwanese*. *Electrophoresis* 15:1212-1217.
- Innes, R. W., Kuempel, P. L., Plazinski, J., Canter-Cremers, H., Rolfe, B. G., and Djordjevic, M. A. 1985. Plant factors induce expression of nodulation and host-range genes in *Rhizobium trifolii*. *Mol. Gen. Genet.* 201:426-432.
- Jungblut, P., Otto, A., Zeindl-Eberhart, E., Plessner, K. P., and Knecht, M. 1994. Protein composition of the human heart: The construction of a myocardial two-dimensional electrophoresis database. *Electrophoresis* 15:685-707.
- Jungblut, P., Thiede, B., Jimmy-Arndt, U., Muller, E. C., and Scheler, C. 1996. Resolution power of two-dimensional electrophoresis and identification of proteins from gels. *Electrophoresis* 17:839-847.
- Jungblut, P., and Wittmann-Liebold, B. 1995. Protein analysis on a genomic scale. *J. Biotechnol.* 41:111-120.
- Kamo, M., Kawakami, T., Miyatake, N., and Tsugita, A. 1995. Separation and characterization of *Arabidopsis thaliana* proteins by two-dimensional gel electrophoresis. *Electrophoresis* 16:423-430.
- Knecht, M., Regitz-Zagrosek, V., Pleissner, K. P., Emig, S., and Jungblut, P. 1994. Dilated cardiomyopathy: Computer assisted analysis of endomyocardial biopsy protein patterns by two-dimensional gel electrophoresis. *Eur. J. Clin. Chem. Clin. Biochem.* 32:615-624.
- Lewis-Henderson, W. R., and Djordjevic, M. A. 1991. A cultivar-specific interaction between *Rhizobium leguminosarum* bv. *trifolii* and subterranean clover is controlled by nodM, other bacterial cultivar specificity genes, and a single recessive host gene. *J. Bacteriol.* 173:2791-2799.
- Matsudaira, P. 1987. Sequence from picomole quantities of proteins electroblotted onto PVDF membranes. *J. Biol. Chem.* 262:10035-10038.
- McIver, J., Djordjevic, M. A., Weinman, J. J., Bender, G. L., and Rolfe, B. G. 1989. Extension of host range of *Rhizobium leguminosarum* bv. *trifolii* caused by point mutations in *nodD* that result in alterations in regulatory function and recognition of inducer molecules. *Mol. Plant-Microbe Interact.* 2:97-106.
- McKay, I. A., and Djordjevic, M. A. 1993. Production and excretion of Nod metabolites by *Rhizobium leguminosarum* bv. *trifolii* are disrupted by the same environmental factors that reduce nodulation in the field. *Appl. Environ. Microbiol.* 59:3385-3392.
- Pasquali, C., Frutiger, S., Wilkins, M. R., Hughes, G. J., Appel, R. D., and Bairoch, A. 1996. Two-dimensional gel electrophoresis of *Escherichia coli* homogenates: The *Escherichia coli* SWISS-2DPAGE database. *Electrophoresis* 17:547-555.
- Peters, N. K., Frost, J. W., and Long, S. R. 1986. A plant flavone, luteolin, induces expression of *Rhizobium meliloti* nodulation genes. *Science* 233:977-980.
- Philip-Hollingsworth, S., Orgambide, G. G., Bradford, J. J., Smith, D. K., Hollingsworth, R. I., and Dazzo, F. B. 1995. Mutation or increased copy number of *nodE* has no effect on the spectrum of chitolipooligosaccharide Nod factors made by *Rhizobium leguminosarum* bv. *trifolii*. *J. Biol. Chem.* 270:20968-20977.
- Rabilloud, T. 1992. A comparison between low background silver diamine and silver nitrate protein stains. *Electrophoresis* 13:429-439.
- Redmond, J. W., Batley, M., Djordjevic, M. A., Innes, R. W., Kuempel, P. L., and Rolfe, B. G. 1986. Flavones induce expression of nodulation genes in *Rhizobium*. *Nature* 323:632-635.
- Rivilla, R., Sutton, J. M., and Downie, J. A. 1995. *Rhizobium legumi-*

- nosarum* NodT is related to a family of outer-membrane transport proteins that includes TolC, PrtF, CyaE and AprF. *Gene* 161:27-31.
- Schlaman, H. R. M., Spaink, H. P., Okker, R. J. H., and Lugtenberg, B. J. J. 1989. Subcellular localization of the nodD gene product in *Rhizobium leguminosarum*. *J. Bacteriol.* 171:4686-4693.
- Schmidt, J., John, M., Wieneke, U., Krüssmann, H.-D., and Schell, J. 1986. Expression of the nodulation gene *nodA* in *Rhizobium meliloti* and localization of the gene product in the cytosol. *Proc. Natl. Acad. Sci. USA* 83:9581-9585.
- Schmidt, J., Wingender, R., John, M., Wieneke, U., and Schell, J. 1988. *Rhizobium meliloti* *nodA* and *nodB* genes are involved in generating compounds that stimulate mitosis of plant cells. *Proc. Natl. Acad. Sci. USA* 85:8578-8582.
- Schofield, P. R., and Watson, J. M. 1986. DNA sequence of *Rhizobium trifolii* nodulation genes reveals a reiterated and potentially regulatory sequence preceding *nodABC* and *nodFE*. *Nucleic Acids Res.* 14:2891-2903.
- Sobral, B. W., Honeycutt, R. J., Atherly, A. G., and McClelland, M. 1991. Electrophoretic separation of the three *Rhizobium meliloti* replicons. *J. Bacteriol.* 173:5173-5180.
- Spaink, H. P., Weinman, J. J., Djordjevic, M. A., Wijffelman, C. A., Okker, R. J., and Lugtenberg, B. J. 1989. Genetic analysis and cellular localization of the *Rhizobium* host specificity-determining NodE protein. *EMBO J.* 8:2811-2818.
- Spaink, H. P., Wijffjes, A. H. M., and Lugtenberg, B. J. J. 1995. *Rhizobium* NodI and NodJ proteins play a role in the efficiency of secretion of lipochitin oligosaccharides. *J. Bacteriol.* 177:6276-6281.
- van Bogelen, R., Olsen, E., Wanner, B., and Neidhardt, F. 1996. Global analysis of proteins synthesized during phosphorus restriction in *Escherichia coli*. *J. Bacteriol.* 178:4344-4366.
- van Rhijn, P., and Vanderleyden, J. 1995. The *Rhizobium*-plant symbiosis. *Microbiol. Rev.* 59:124-142.
- Vasquez, M., Santana, O., and Quinto, C. 1993. The NodI and NodJ proteins from *Rhizobium* and *Bradyrhizobium* strains are similar to capsular polysaccharide secretion proteins from Gram-negative bacteria. *Mol. Microbiol.* 8:369-377.
- Wasinger, V. C., Cordwell, S. J., Cerpa-Poljak, A., Yan, J. X., Gooley, A. A., Wilkins, M. R., Duncan, M. W., Harris, R., Williams, K. L., and Humphery-Smith, I. 1995. Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. *Electrophoresis* 16:1090-1094.
- Wheeler, C. H., Berry, S. L., Wilkins, M. R., Corbett, J. M., Ou, K., Gooley, A. A., Humphery-Smith, I., Williams, K. L., and Dunn, M. J. 1996. Characterisation of proteins from two-dimensional electrophoresis gels by matrix-assisted laser desorption mass spectrometry and amino acid compositional analysis. *Electrophoresis* 17:580-587.
- Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., and Golaz, O. 1996a. From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Bio/Technology* 14:61-65.
- Wilkins, M. R., Sanchez, J. C., Williams, K. L., and Hochstrasser, D. F. 1996b. Current challenges and future applications for protein maps and post-translational vector maps in proteome projects. *Electrophoresis* 17:830-838.
- Yan, J., Wilkins, M. R., Ou, K., Gooley, A. A., Williams, K. L., Sanchez, J. C., and Golaz, O. 1996. Large-scale amino acid analysis for proteome studies. *J. Chromatogr.* 736:291-302.