

Metabolic Discrimination of *Catharanthus roseus* Leaves Infected by Phytoplasma Using $^1\text{H-NMR}$ Spectroscopy and Multivariate Data Analysis¹

Young Hae Choi, Elisabet Casas Tapias, Hye Kyong Kim, Alfons W.M. Lefeber, Cornelis Erkelens, Jacobus Th.J. Verhoeven, Jernej Brzin, Jana Zel, and Robert Verpoorte*

Division of Pharmacognosy, Section Metabolomics, Institute of Biology, Leiden University, 2300 RA Leiden, The Netherlands (Y.H.C., E.C.T., H.K.K., R.V.); Division of NMR, Institute of Chemistry, Gorlaeus Laboratories, 2300 RA Leiden, The Netherlands (A.W.M.L., C.E.); Plant Protection Service, 6700 HC Wageningen, The Netherlands (J.Th.J.V.); and National Institute of Biology, 1000 Ljubljana, Slovenija (J.B., J.Z.)

A comprehensive metabolomic profiling of *Catharanthus roseus* L. G. Don infected by 10 types of phytoplasmas was carried out using one-dimensional and two-dimensional NMR spectroscopy followed by principal component analysis (PCA), an unsupervised clustering method requiring no knowledge of the data set and used to reduce the dimensionality of multivariate data while preserving most of the variance within it. With a combination of these techniques, we were able to identify those metabolites that were present in different levels in phytoplasma-infected *C. roseus* leaves than in healthy ones. The infection by phytoplasma in *C. roseus* leaves causes an increase of metabolites related to the biosynthetic pathways of phenylpropanoids or terpenoid indole alkaloids: chlorogenic acid, loganic acid, secologanin, and vindoline. Furthermore, higher abundance of Glc, Glu, polyphenols, succinic acid, and Suc were detected in the phytoplasma-infected leaves. The PCA of the $^1\text{H-NMR}$ signals of healthy and phytoplasma-infected *C. roseus* leaves shows that these metabolites are major discriminating factors to characterize the phytoplasma-infected *C. roseus* leaves from healthy ones. Based on the NMR and PCA analysis, it might be suggested that the biosynthetic pathway of terpenoid indole alkaloids, together with that of phenylpropanoids, is stimulated by the infection of phytoplasma.

In 1967 it was found that phytoplasmas, previously termed mycoplasma-like organisms (MLO), were the cause of some plant yellowing diseases (Doi et al., 1967). Phytoplasma are minute bacteria (200–800 μm) that have no cell wall and inhabit phloem sieve elements in infected plants. These noncultivable plant pathogens belong to the class of Mollicutes (McCoy et al., 1989; Lee and Davis, 1992; Garnier et al., 2001). Comparison of 16S rDNA sequences showed the MLOs to be phylogenetically close to the achleplasma/anaeroplasm group of the Mollicutes, and the trivial name phytoplasma was adopted in 1994 to replace MLO (Garnier et al., 2001). 16S rDNA sequences were determined and used in the late 1990s to classify the phytoplasmas into 20 phylogenetic clusters (Seemüller et al., 1998). They have been associated with diseases in more than 300 plant species belonging to 98 families. They are introduced directly inside the sieve tubes of plants via homopterous insect vectors, primarily belonging to the family Cicadellidae (leafhoppers; Kummert and Ruffart, 1997).

Plants infected by phytoplasmas exhibit an array of symptoms suggesting profound disturbances in the normal balance of plant metabolism, including yellowing, chlorosis, or bronzing of foliage, stunting (reduction of internodes and leaf size), virescence (the development of green flowers and the loss of normal pigments), phyllody (the development of floral parts into leafy structures), sterility of flowers, proliferation of secondary auxiliary buds often resulting in a witch-broom effect, proliferation of secondary roots, abnormal fruits and seeds, and abnormal elongation of internodes leading to slender shoots (Lee et al., 2000). The symptoms induced in diseased plants vary with the species of phytoplasma and host plants and with the stage of infection. Internally, phytoplasmal infections can cause extensive swollen veins in phloem tissues. In general, all symptoms have clearly detrimental effects on plants. However, some plant species are tolerant or resistant to phytoplasmal infections, being thus asymptomatic or exhibiting mild symptoms only (Lee et al., 2000). The resistance of plants to infection with phytoplasma and the specificity of the vector-phytoplasma-plant interaction is a cause for the fact that most plants do not harbor more than one type in natural conditions. This specificity, together with their noncultivable characteristic, has constituted a problem for the investigation of the plant-phytoplasma interaction. *Catharanthus roseus*, however, is known as

¹ This work was supported by the van Leersumfonds (KNAW).

* Corresponding author; e-mail verpoorte@chem.leidenuniv.nl; fax 31-71-527-4511.

Article, publication date, and citation information can be found at www.plantphysiol.org/cgi/doi/10.1104/pp.104.041012.

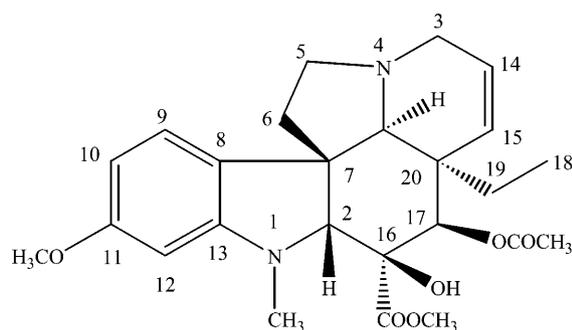
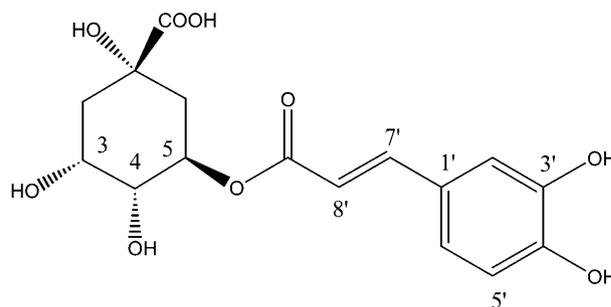
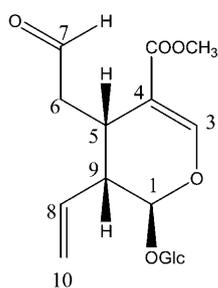
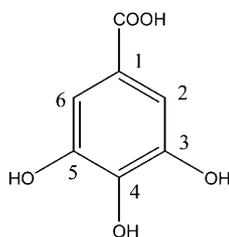
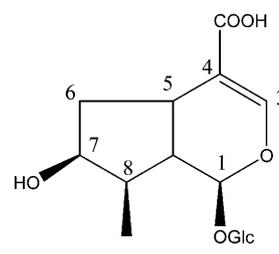
Table 1. Original host plant, origin, and grafting date of phytoplasma evaluated in the study

Phytoplasma Group or Subgroup	Disease	Original Host Plant	Origin	Grafting Date
Apple proliferation group 16SrX-A	Apple proliferation (AP)	<i>Malus domestica</i>	Italy	June 2, 2002
Clover proliferation group 16SrVI-A	Bringal little leaf (BLL)	<i>Solanum melongena</i>	India	November 27, 2002
Stolbur group 16srII-A	Stolbur (DYON)	<i>Lycopersicum esculentum</i>	France	November 14, 2002
	Molière diseases (MOL)	<i>Prunus avium</i>	France	September 27, 2002
	Stolbur (STOF)	<i>L. esculentum</i>	France	September 27, 2002
	Stolbur (STOL)	<i>L. esculentum</i>	Croatia	November 27, 2002
Faba bean phyllody group 16SrII-E	Stolbur (UDINESE)	<i>L. esculentum</i>	Italy	September 27, 2002
	Australian tomato big bud (TBB)	<i>L. esculentum</i>	Australia	November 6, 2002
Unclassified	Potato purple top (PPT)	<i>Solanum tuberosum</i>	France	October 24, 2002
Unclassified	<i>Solanum marginatum</i> big bud (SMBB)	<i>S. marginatum</i>	Ecuador	November 6, 2002

a source plant that can harbor many phytoplasmas and is thus used for the maintenance of phytoplasma cultures.

To determine possible alterations, some general approaches have been carried out. Cell wall degradation and tissue maceration by enzymatic hydrolysis are well known to be associated with pathological processes. Starch content in the roots of diseased trees was found to be only about one-half to one-third of that of healthy trees in pear decline-affected pear trees (Batjer

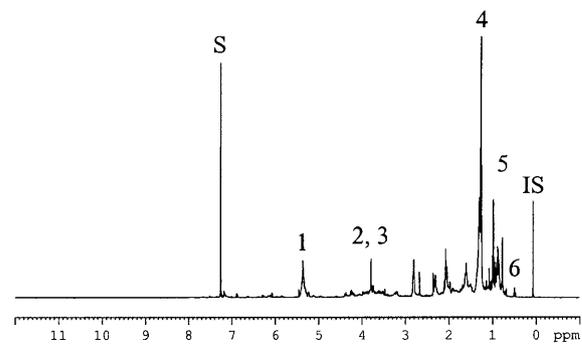
and Schneider, 1960) and proliferation-diseased apple trees (Kartte and Seemüller, 1991). Catlin et al. (1975) also reported that there was a considerable accumulation of carbohydrates and starch in leaves of pear decline-diseased pear trees and a greatly reduced transport of photosynthetically fixed ^{14}C from the leaves of affected trees. In *C. roseus* roots, soluble carbohydrates were not markedly altered, but starch was considerably reduced when infected with the grapevine yellowing and apple proliferation phytoplasmas

**vindoline****chlorogenic acid****secologanin****gallic acid****loganic acid****Figure 1.** Chemical structures of vindoline, chlorogenic acid, secologanin, loganic acid, and gallic acid.

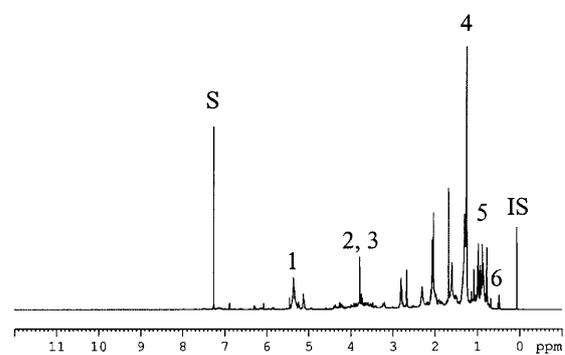
(Lepka et al., 1999). In tobacco roots, there was a general decrease in soluble carbohydrates and starch following infection by the apple proliferation phytoplasma (Lepka et al., 1999). Higher amino acid content was found in the source and sink leaves of ash yellows phytoplasma-infected *C. roseus* and source leaves of apple proliferation infected tobacco (Lepka et al., 1999). The carotenoids content in aster yellows phytoplasma-infected leaves began to diminish 6 weeks after infection. The anthocyanin content in flowers infected with aster yellows phytoplasma significantly decreased (Yu, 1997). Despite these reports, the metabolomic alterations in phytoplasma-infected plants are still unclear, especially in the case of secondary metabolites.

Quantitative and qualitative measurements of large numbers of plant metabolites can provide a broad view of the biochemical status of an organism (Fiehn et al., 2000), and it would be of great interest for the detection of plants infected by phytoplasmas evaluated in this article. Whereas genomics and proteomics can provide insights into the potential of a biological system to interact with external perturbations, it is the resulting changes in the metabolic profile of the system that are potentially more useful for the understanding of the biochemical reaction to stress (Bailey et al., 2003).

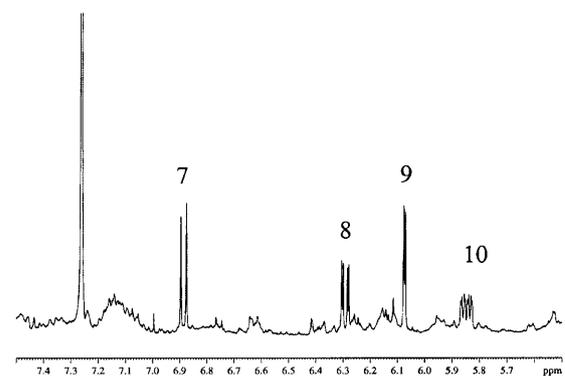
It is generally accepted that a single analytical technique will not provide sufficient visualization of the metabolome and, therefore, multiple technologies are needed for a comprehensive view (Summer et al., 2003). However, the lack of the target biological material or the instability of the metabolome forces us to choose an optimum analytical tool for the metabolomic profiling. Therefore, it is preferable to use a wide spectrum chemical analysis technique, which is rapid, reproducible, and stable in time while needing only a very basic sample preparation. NMR is one of the techniques that could meet those requirements. In the last few decades, a number of techniques have been devised to develop NMR spectroscopy as a fingerprinting tool for the interpretation and quality assessment of industrial and natural products. At the same time, multivariate or pattern recognition techniques such as the well-described principal component analysis (PCA) and hierarchical cluster analysis have been specifically designed to analyze complex data sets (Summer et al., 2003). Ward et al. (2003) reported that the various ecotypes of *Arabidopsis* could be distinguished using ^1H -NMR and multivariate data analysis. The analysis of the consequences of the genetic manipulation and strain differentiation in strains of yeast was also carried out with ^1H -NMR (Raamsdonk et al., 2001). Roessner et al. (2000) reported the metabolite profiling of transgenic potato tubers overproducing invertase using gas chromatography/mass spectrometry, and they also determined the major biochemical phenotypes of transgenic potato lines using statistical data analysis (Roessner et al., 2001). Gavaghan et al. (2000) reported that the NMR-based metabolomic approach could be applied



A



B



C

Figure 2. ^1H -NMR spectra of CHCl_3 extract of healthy *C. roseus* leaves (A), phytoplasma (BLL)-infected *C. roseus* leaves (B), and expansion part [phytoplasma (BLL)-infected leaves] in the range of δ 5.5 to δ 7.5 (C). 1, olefinic signals of fatty components or terpenoids; 2, OCH_3 of C-11 of vindoline; 3, OCH_3 of C-22 of vindoline; 4, long chain CH_2 of fatty component; 5, steroidal or triterpenoidal CH_3 ; 6, H-18 of vindoline; 7, H-9 of vindoline; 8, H-10 of vindoline; 9, H-12 of vindoline; 10, H-14 of vindoline; S, residual CHCl_3 signal; IS, internal standard (HMDS).

to differentiate between genetic strains in mouse lines.

Here, we report a $^1\text{H-NMR}$ spectroscopy method, coupled with multivariate analysis for the metabolic analysis of 2 healthy and 10 types of phytoplasma-infected *C. roseus* leaves. The phytoplasmas evaluated in this study are listed in Table I. This approach may lead to the identification of metabolic pathways connected with the defense response to phytoplasma.

RESULTS

Visual Inspection of $^1\text{H-NMR}$ Spectra and Assignments of CHCl_3 Extract of Healthy and Infected *C. roseus* Leaves by Phytoplasma

For the identification of indole alkaloids, steroids or triterpenoids, and fatty components, CHCl_3 extracts were investigated (Fig. 1). Similar metabolomic patterns were observed by visual inspection of $^1\text{H-NMR}$

spectra of the CHCl_3 extracts of the various *C. roseus* leaves infected by phytoplasmas and those of healthy plants (Fig. 2, A and B). The signals of vindoline (Fig. 1) are well distinguishable in the $^1\text{H-NMR}$ spectrum of CHCl_3 extracts. H-9 at δ 6.89 (*d*, $J = 8.2$ Hz), H-10 at δ 6.29 (*dd*, $J = 8.5$ Hz, 2.3 Hz), and H-12 at δ 6.07 (*d*, $J = 2.2$ Hz) are observed as major signals in the aromatic region of the leaves CHCl_3 extract (Fig. 2C). In addition to these aromatic signals, other characteristic signals of vindoline, such as H-14 at δ 5.85 (*ddd*, $J = 10.2$ Hz, 4.9 Hz, 1.7 Hz), OCH_3 of C-11 at δ 3.79 (*s*), OCH_3 of C-22 at δ 3.78 (*s*), and H-18 at δ 0.49 (*t*, $J = 7.4$ Hz) are clearly identified in the spectra.

The signals of catharanthine, stemmadenine, and tabersonine have been known as other main indole alkaloids of *C. roseus*. The levels of these alkaloids estimated by the $^1\text{H-NMR}$ signal intensity was relatively low compared to that of vindoline. Several methyl groups that might originate from steroids or triterpenoids showed high intensity in the range of

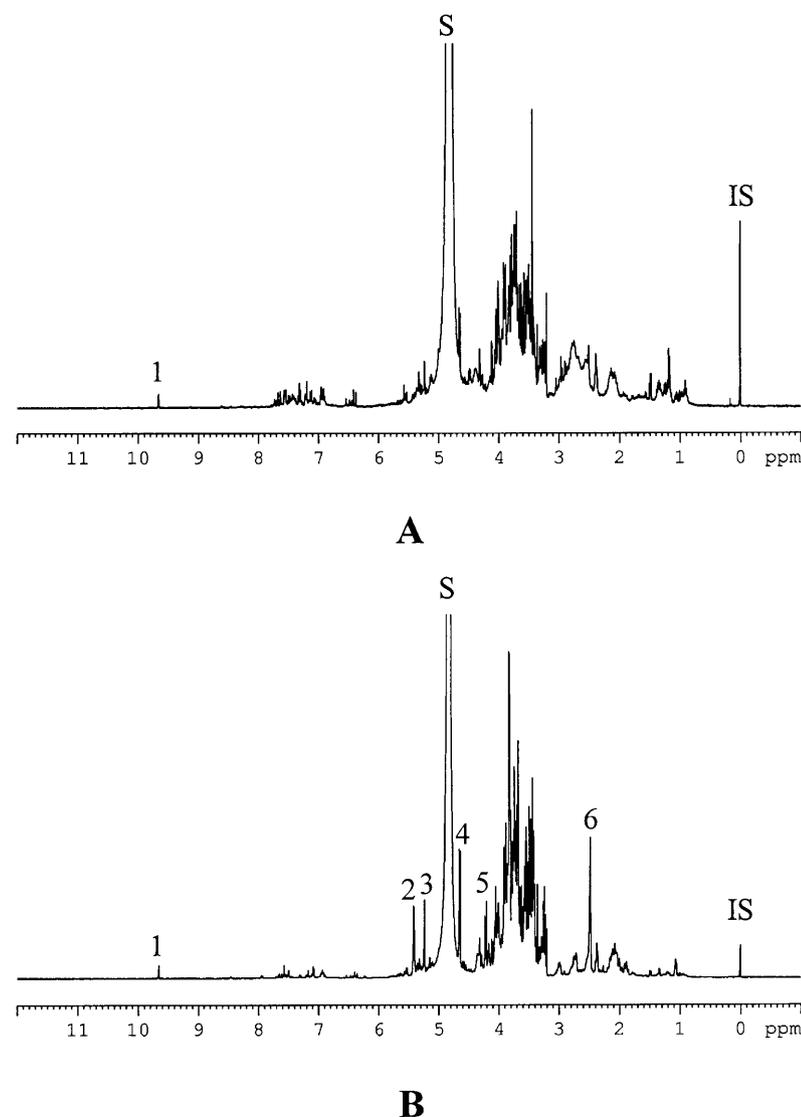


Figure 3. $^1\text{H-NMR}$ spectra of water extract of healthy *C. roseus* leaves (A) and phytoplasma (UDINESE)-infected *C. roseus* leaves (B). 1, aldehyde signal of secologanin; 2, H-1 of Suc; 3, H-1 of α -Glc; 4, H-1 of β -Glc; 5, H-1 of Fru in Suc; 6, succinic acid; S, residual water signal; IS, internal standard (TSP).

δ 0.8 to δ 1.2 of the $^1\text{H-NMR}$ spectra. However, there was no big difference in the signal patterns of these methyl groups between healthy and infected leaves. In addition to these signals, the methyl signal of fatty components at δ 1.2 to δ 1.4 and olefinic signals of fatty components, steroids, or triterpenoids at δ 5.0 to δ 5.5 were also detected as major signals in the $^1\text{H-NMR}$ spectra of CHCl_3 extract.

Identification of Chlorogenic Acid, Glc, Glu, Loganic Acid, Polyphenols, Secologanin, Suc, and Succinic Acid in $^1\text{H-NMR}$ Spectra of Water Extract of Healthy and Phytoplasma-Infected *C. roseus* Leaves

The $^1\text{H-NMR}$ spectra of water extracts for the healthy and phytoplasma-infected leaves are shown in Figure 3. As an example, the healthy leaves were compared with the plant infected by UDINESE phytoplasma in this figure. The differences between the healthy and infected plants were found to be larger than those detected in CHCl_3 extracts. The major

differences are observed in the anomeric signals of carbohydrates such as δ 5.42 (*d*, $J = 3.8$ Hz), δ 5.24 (*d*, $J = 3.7$ Hz), and δ 4.64 (*d*, $J = 9.5$ Hz). These were assigned to be the anomeric protons of Suc, α -Glc, and β -Glc, respectively (Agrawal, 1992). Another anomeric signal obtained from the Fru moiety of Suc is also well distinguishable at δ 4.22 (*d*, $J = 8.8$ Hz). The residual proton signals of the sugars shown in the crowded region (δ 3.0– δ 4.0) were assigned by the comparison of $^1\text{H-NMR}$ spectra of the reference compounds, $^1\text{H-}^1\text{H-COSY}$ (correlated spectroscopy) and TOCSY (total correlation spectroscopy) spectra. Besides the chemical shift data of $^1\text{H-NMR}$, HMBC (heteronuclear multiple bond correlation) spectra can give evidence for the identification of amino acids. H-2 or H-3 can correlate with the carbonyl group of the amino acid. Ala at δ 1.48 (H-3*d*, $J = 4.8$ Hz) correlated with δ 178.6, Glu at δ 2.14 (*m*) and δ 2.38 (*m*) correlated with δ 179.2, and Gly at δ 3.56 (*s*) correlated with δ 174.7. These could thus be identified as the most abundant amino acids in the $^1\text{H-NMR}$ spectra of the water extract.

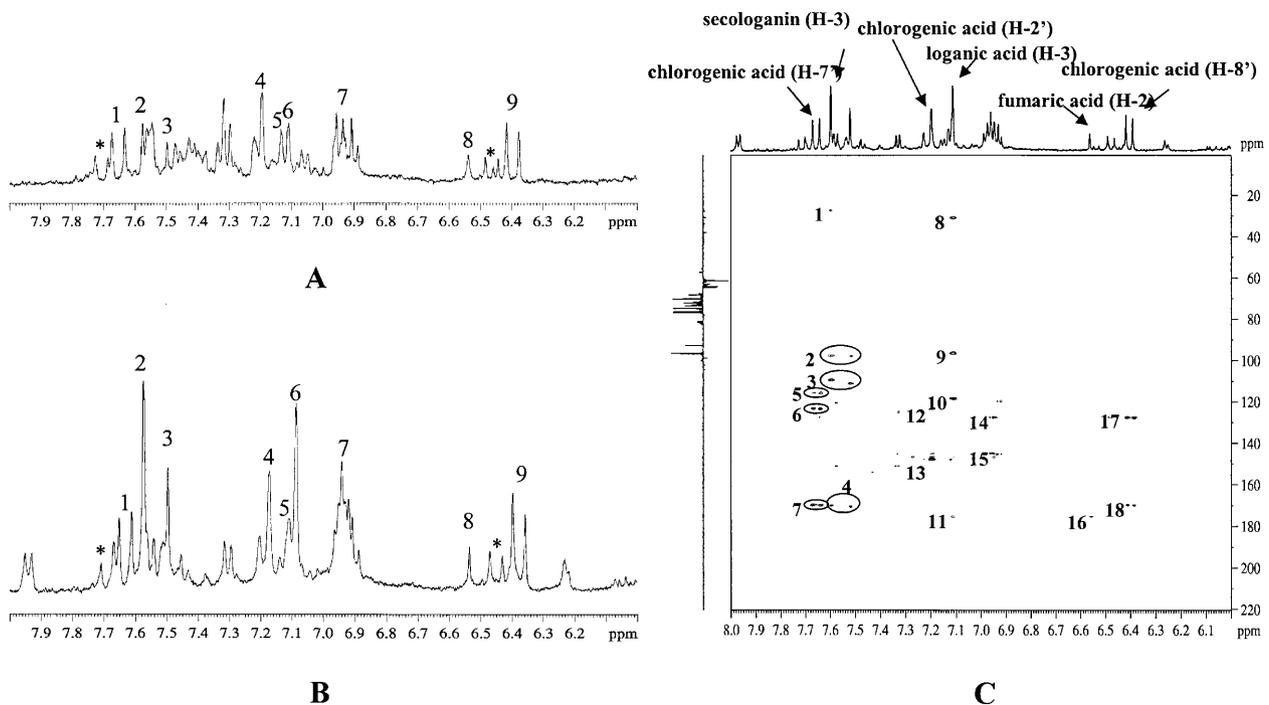
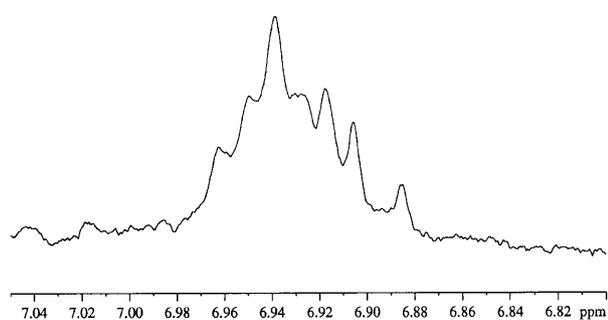
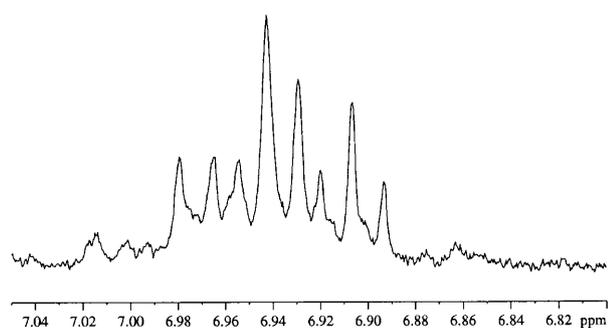


Figure 4. $^1\text{H-NMR}$ spectra of water extract of healthy *C. roseus* leaves (A) and phytoplasma (UDINESE)-infected *C. roseus* leaves (B) in the range of δ 6.0 to δ 8.0, and HMBC spectra of water fraction of phytoplasma (UDINESE)-infected (C) *C. roseus* leaves. In $^1\text{H-NMR}$ spectra (A and B): 1, H-7' of chlorogenic acid; 2 and 3, H-3 of secologanin; 4, H-2' of chlorogenic acid; 5, H-6' of chlorogenic acid; 6, H-3 of loganic acid; 7, H-5' of chlorogenic acid and aromatic signals of polyphenols; 8, fumaric acid; 9, H-8' of chlorogenic acid; *, possible signals of chlorogenic acid derivatives. In HMBC spectra (C): 1, correlation of H-3 and C-5 of secologanin; 2, correlation of H-3 and C-1 of secologanin; 3, correlation of H-3 and C-4 of secologanin; 4, correlation of H-3 and carbonyl group of secologanin; 5, correlation of H-7' and C-2' of chlorogenic acid; 6, correlation of H-7' and C-6' of chlorogenic acid; 7, correlation of H-7' and carbonyl group of chlorogenic acid; 8, correlation of H-3 and C-5 of loganic acid; 9, correlation of H-3 and C-1 of loganic acid; 10, correlation of H-3 and C-4 of loganic acid; 11, correlation of correlation of H-3 and carbonyl group of loganic acid; 12, correlation of H-2' and C-1' of chlorogenic acid; 13, correlation of H-2' and C-3'; 14, correlation of H-2 and C-1 of gallic acid derivatives; 15, correlation of H-2 and C-3 of gallic acid derivatives; 16, correlation of H-2 and carbonyl group of fumaric acid; 17, correlation of H-8' and C-1' of chlorogenic acid; 18, correlation of H-8' and carbonyl group of chlorogenic acid.



A

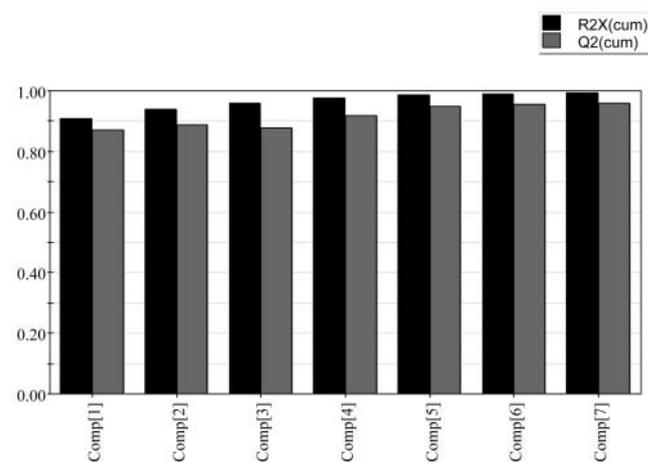


B

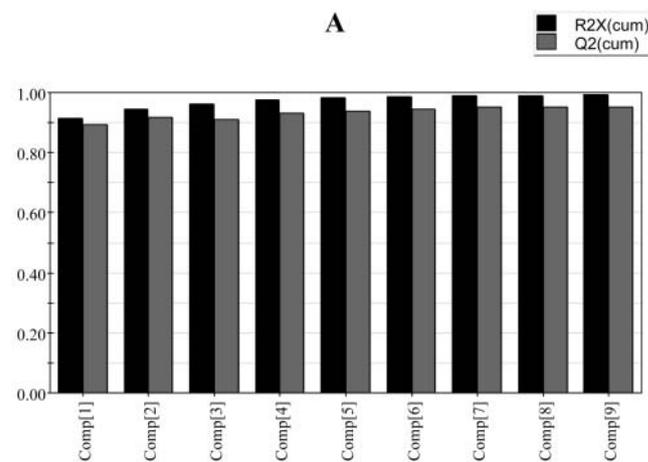
Figure 5. ^1H -NMR spectra of water extract of *C. roseus* leaves infected by phytoplasma (UDINESE) with 400 MHz (A) and with 600 MHz (B) in the range of δ 7.05 to δ 6.80.

The signals of the main aromatic compound in the aqueous extract were assigned to chlorogenic acid (Fig. 4, A and B). The ^1H -NMR spectrum is in accordance with a phenylpropanoid, showing the characteristic signals due to two trans-olefinic protons (1H each, d , $J = 15.9$ Hz at δ 7.64, H-7' and δ 6.39, d , $J = 15.9$ Hz, H-8'). In addition, three aromatic protons at δ 7.18 (1H, s), δ 7.11 (1H, d , $J = 8.5$ Hz), and δ 6.93 (1H, d , $J = 8.5$ Hz) correspond to H-2', H-6', and H-5' of the aromatic ring of chlorogenic acid (Fig. 1), respectively. Other signals were detected close to those of chlorogenic acid. They are shifted approximately 0.05 ppm downfield from the chlorogenic signals and assumed to be those of other chlorogenic acid isomers such as 4-*O*-caffeoylquinic acid or 5-*O*-caffeoylquinic acid because of the same coupling constants and correlation patterns in the ^1H - ^1H -COSY spectrum. These assignments were confirmed with the HMBC spectrum. The two olefinic protons correlate with each other in the ^1H - ^1H -COSY spectrum and with a carbonyl group at δ 174.8 in the HMBC spectrum. When compared to the ^1H -NMR spectrum of the water extract of the healthy *C. roseus* leaves, the spectrum of the phytoplasma-

infected leaves shows three more additional signals at δ 7.57, δ 7.49, and δ 7.09. These signals have the same HMBC pattern, correlate with the carbonyl groups at δ 167.4 (δ 7.57 and δ 7.49) and δ 176.3 (δ 7.09), with olefinic carbons at δ 109.4 (δ 7.57), δ 110.6 (δ 7.49), and δ 118.1 (δ 7.09), with oxygenated carbons at δ 98.2 (δ 7.57, δ 7.49, and δ 7.09). This correlation pattern suggests that these signals could be due to the H-3 of iridoids or secoiridoids (Inouye, 1991; Fig. 4C). When compared with reference compounds, those signals were assigned to be H-3 of secologanin (δ 7.57 and δ 7.49; Fig. 1) and loganic acid (δ 7.09). In the case of secologanin, H-3 is changeable, especially at higher pH values because the aldehyde proton can form the dimethylacetal, and this results in the effect on the chemical shift of H-3 of secologanin (Kim et al., 2004; Tomassini et al., 1995). So, the signal at δ 7.49 might be due to the artifact of secologanin. The presence of secologanin could be confirmed by aldehyde signal at δ 9.65.



A



B

Figure 6. Principal components explaining variances used in PCA of ^1H -NMR data set of *C. roseus* leaves. A, CHCl_3 extract; B, water extract.

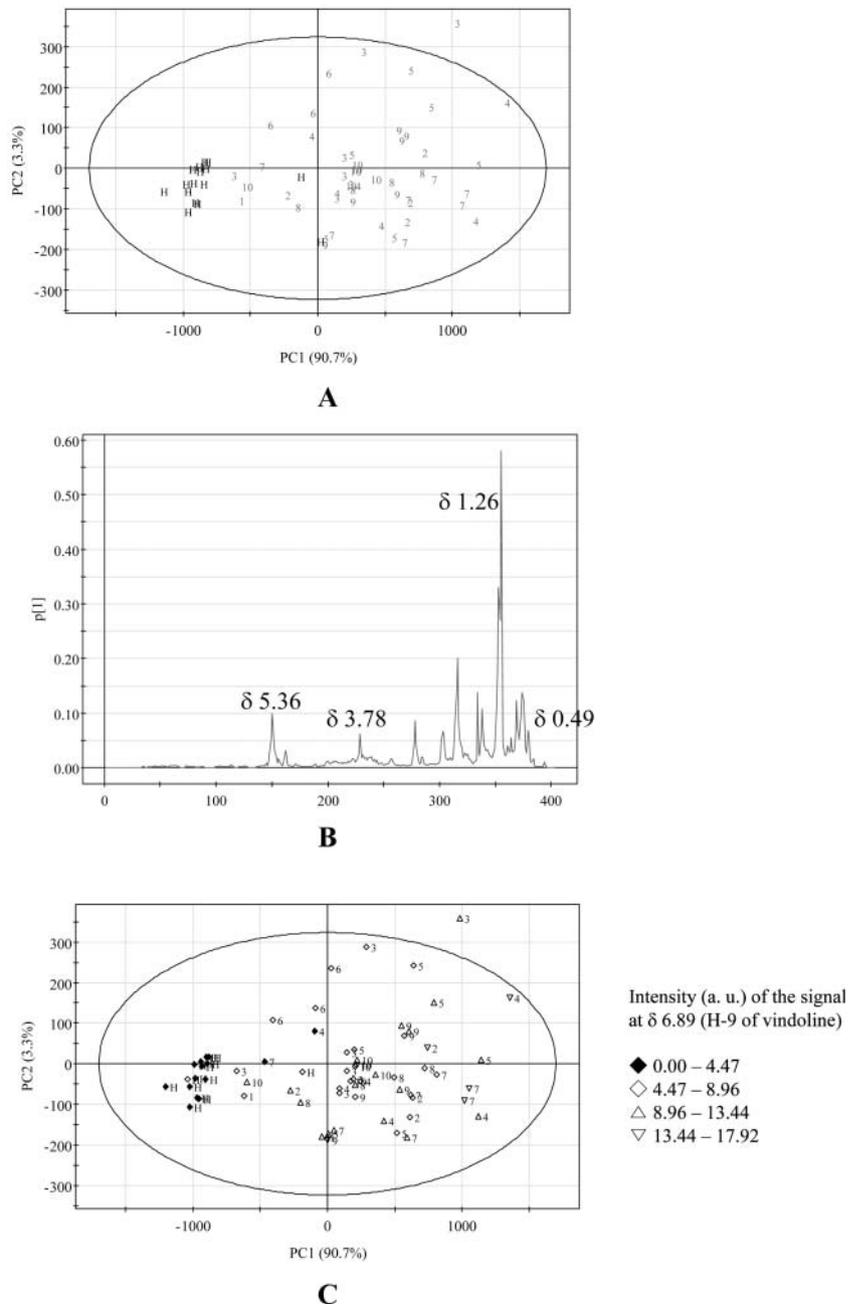
In the $^1\text{H-NMR}$ spectra of the water extract, a cluster of singlets was detected in the range of δ 6.8 to δ 7.0. These characteristic phenol signals were assumed to be those of polyphenols, such as gallic acid derivatives (Fig. 1). Unfortunately, 400 MHz $^1\text{H-NMR}$ spectra of the extract showed only a broad cluster of the signals that were not clearly recognizable. Higher resolution 600 MHz $^1\text{H-NMR}$, however, produced clear separated resonances (Fig. 5) that being characteristic phenol signals were assumed to be those of polyphenols. Fumaric acid at δ 6.54 (s), Glu at δ 2.38 (m), and succinic acid at δ 2.49 (s) are also shown as differentiating components in the *C. roseus* leaves infected by

phytoplasma. Both of them were confirmed with reference compound and HMBC spectra.

PCA of CHCl_3 Extract: Vindoline Is a Discriminating Metabolite to Different Phytoplasma-Infected *C. roseus* Leaves

PCA is an unsupervised clustering method requiring no knowledge of the data set and acts to reduce the dimensionality of multivariate data while preserving most of the variance within it (Goodacre et al., 2000). The principal components can be displayed graphically as a scores plot. This plot is useful for

Figure 7. Score and loading plot of PCA of the CHCl_3 extracts of healthy and phytoplasma-infected *C. roseus* leaves. A, score plot; B, loading plot of PC1; C, score plot with the intensity of H-9 of vindoline in $^1\text{H-NMR}$ spectra. 1, AP; 2, BLL; 3, DYON; 4, MOL; 5, PPT; 6, SMBB; 7, STOF; 8, STOL; 9, TBB; 10, UDINESE; H, healthy. The ellipse represents the Hotelling T2 with 95% confidence in score plots.



observing any groupings in the data set. PCA models are constructed using all the samples in the study. Coefficients by which the original variables must be multiplied to obtain the PC are called loadings. The numerical value of a loading of a given variable on a PC shows how much the variable has in common with that component (Massart et al., 1988). Thus, for NMR data, loading plots can be used to detect the metabolites responsible for the separation in the data. Generally, this separation took place in the first two principal components (PC1 and PC2). For the data set obtained from the analysis of the CHCl_3 extracts, a six-component model explained 99% of the variance, with the first two components explaining 94% (Fig. 6A). Examination of the scores and loading plots for PC1 versus PC2 showed that healthy *C. roseus* leaves are clearly separated from the phytoplasma-infected leaves (Fig. 7A). The separation is due mainly to PC1. Investigation of the loading plot of PC1 indicated that the first component explained the variance in fatty components due to the signals at δ 1.2 to δ 1.4 (CH_2) and at δ 5.0 to δ 5.5 (olefinic CH_2) and in indole alkaloids such as vindoline because of the signals at δ 3.79 (OCH_3 of C-11), at δ 3.78 (OCH_3 of C-22), and at δ 0.49 (H-18; Fig. 7B). It is evident that *C. roseus* leaves infected by phytoplasmas contain less fatty components and higher vindoline compared to healthy leaves. This observation was also confirmed by the investigation of the intensity of target $^1\text{H-NMR}$ signals. Figure 7C shows that infected

leaves have 2 to 4 times increased level of vindoline relative to healthy plants by the comparison of the intensity of H-9 of vindoline at δ 6.89. The difference obtained by PC2 has no effect on the separation of healthy and infected leaves because the same samples show a less reproducible value when compared to PC1 value.

PCA of Water Extract: Chlorogenic Acid, Glc, Loganin, Polyphenols (Gallic Acid Derivatives), Secologanin, Succinic Acid, and Suc Are Discriminating Metabolites to Different Phytoplasma-Infected *C. roseus* Leaves

A variety of metabolites such as Ala, chlorogenic acid, fumaric acid, Glu, Glc, loganin, polyphenols (gallic acid derivatives), secologanin, succinic acid, and Suc were detected in $^1\text{H-NMR}$ spectra of water extracts. The detailed analysis of the difference in the content of these diverse metabolites could make it possible to differentiate the infected leaves from healthy ones. For the water extract, a nine-component model explained 99% of the variance, with the first two components explaining 95% (Fig. 6B). Score plot of PC1 versus PC2 shows that healthy leaves are well separated from infected plants by both PC1 and PC2 (Fig. 8A). The healthy leaves have lower PC1 and higher PC2 relative to infected ones. Examination of the loading plot of PC1 shows that the first component explains the variance in the amount of carbohydrates because high values were detected in sugar

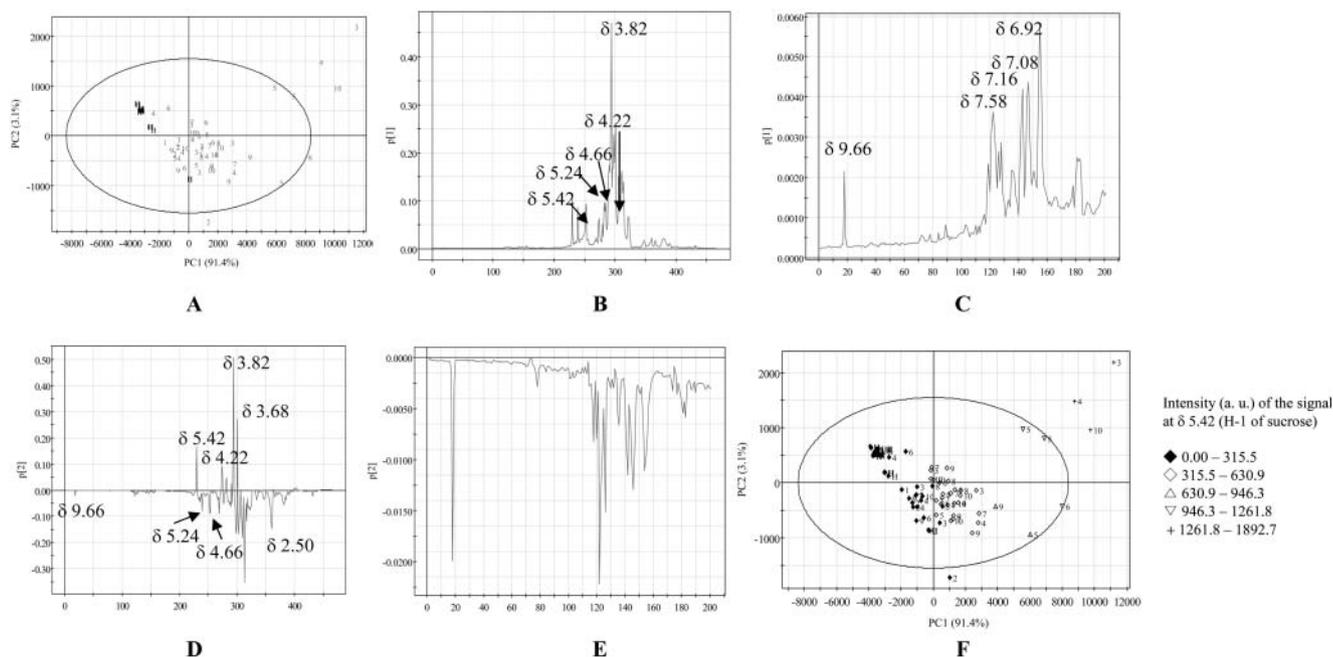
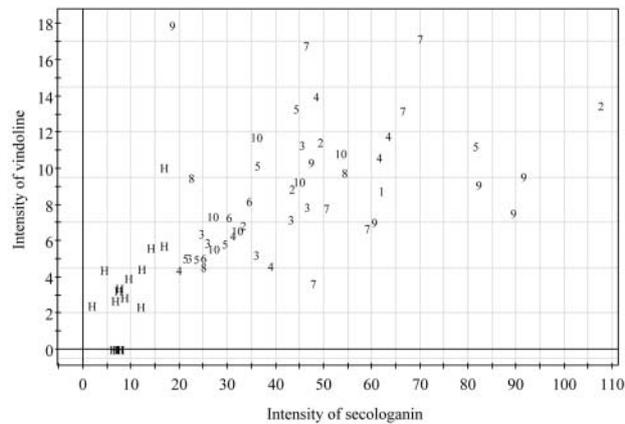
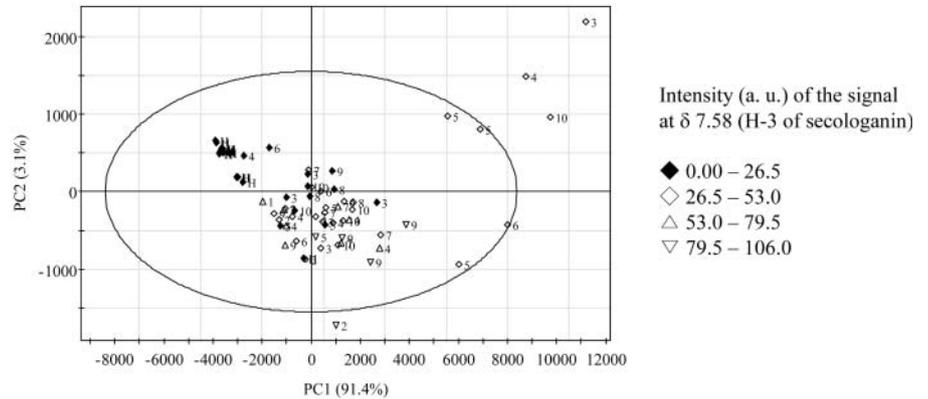


Figure 8. Score and loading plot of PCA of water extracts of healthy and infected *C. roseus* leaves by phytoplasmas. A, score plot; B, loading plot of PC1 in the range of δ 0.3 to δ 10.0; C, loading plot of PC1 in the range of δ 6.0 to δ 10.0; D, loading plot of PC2 in the range of δ 0.3 to δ 10.0; E, loading plot of PC1 in the range of δ 6.0 to δ 10.0; F, score plot with the intensity of H-1 signal at δ 5.42 of Suc in $^1\text{H-NMR}$ spectra. 1, AP; 2, BLL; 3, DYON; 4, MOL; 5, PPT; 6, SMBB; 7, STOF; 8, STOL; 9, TBB; 10, UDINESE; H, healthy. The ellipse represents the Hotelling T2 with 95% confidence in score plots.

Figure 9. Score plot with the intensity of H-3 signal at δ 7.58 of Suc in $^1\text{H-NMR}$ spectra (A), and correlation plot between secologanin and vindoline (B). 1, AP; 2, BLL; 3, DYON; 4, MOL; 5, PPT; 6, SMBB; 7, STOF; 8, STOL; 9, TBB; 10, UDINESE; H, healthy. The ellipse represents the Hotelling T2 with 95% confidence in score plot. The $^1\text{H-NMR}$ signals at δ 5.24 (H-1), δ 7.58 (H-3), δ 5.42 (H-1), δ 6.89 (H-10) were used for the analysis of the intensity of Glc, secologanin, Suc, and vindoline, respectively. The ellipse represents the Hotelling T2 with 95% confidence in score plots.

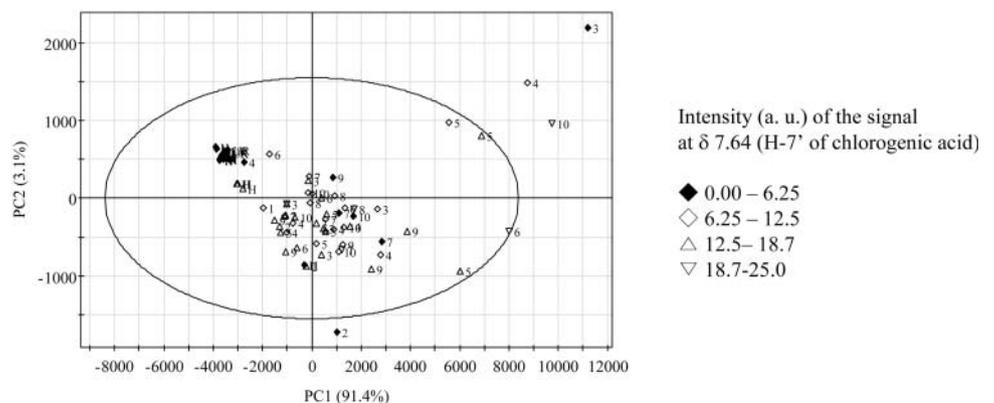


B

region (δ 5.5– δ 3.0; Fig. 8B). The anomeric protons of Suc at δ 5.42, α -Glc at δ 5.24, β -Glc at δ 4.66, and Fru at δ 4.22 show higher PC1 values in the loading plot. Additionally, the loading plot of PC1 illustrated some small positive values in the range of δ 6.0 to δ 10.0 (Fig. 8C). This region contains many peaks attributed to chlorogenic acid, fumaric acid, loganic acid, poly-

phenols, and secologanin. Most of *C. roseus* leaves infected by phytoplasma are well separated from healthy plants because of lower PC2 values, together with higher PC1 value. The lower PC2 values of infected leaves are due to the signals of α -Glc at δ 5.24, β -Glc at δ 4.66, and succinic acid at δ 2.50 (Fig. 8D). Small negative PC2 values also detected in the

Figure 10. Score plot with the intensity of H-7' signal at δ 7.64 of Suc in $^1\text{H-NMR}$ spectra. 1, AP; 2, BLL; 3, DYON; 4, MOL; 5, PPT; 6, SMBB; 7, STOF; 8, STOL; 9, TBB; 10, UDINESE; H, healthy. The ellipse represents the Hotelling T2 with 95% confidence in score plot.



range of δ 6.0 to δ 10.0 (Fig. 8E). It means that chlorogenic acid, loganic acid, secologanin, and polyphenols are shown at a relatively higher abundance in infected leaves. The Suc signals in $^1\text{H-NMR}$ spectra such as δ 5.42, δ 4.22, δ 3.82, and δ 3.68 show positive PC2 value. It thus seems to be more abundant in healthy plants compared to the infected ones. However, this contribution of Suc is quite opposite to the result of PC1 by which infected leaves are shown to contain more Suc than healthy ones. For the confirmation of the effect of Suc, the intensity of the $^1\text{H-NMR}$ signal at δ 5.42 (H-1 of Suc) is plotted in Figure 8F. In most infected *C. roseus* leaves, this signal is higher, indicating an increase in the amount of Suc. The positive PC2 value might be due to the fact of an unusually higher signal detected in three plant samples (marked by + in Fig. 8F) located outside of Hotelling T2 with 95% confidence.

DISCUSSION

$^1\text{H-NMR}$ spectroscopy has proved to be a valuable tool for unbiased metabolite fingerprinting of healthy and phytoplasma-infected *C. roseus* leaves. This work demonstrates that the combination of $^1\text{H-NMR}$ spectroscopy with multivariate data analysis is readily amenable to the rapid screening profile, which at its most basic level can allow metabolic fingerprints to be generated. Further, the implementation of chemometric approaches to interpret the resulting complex data

allows significant biochemical changes to be readily extracted from the data. By virtue of the NMR spectra already obtained, it is then possible to elucidate the nature of the metabolites that are the key to the separation between sample groups. The data for PCA can be scaled in different ways. If the data are mean centered then a covariance matrix is produced, but if the mean-centered data are scaled to unit variance, a correlation matrix is obtained. An advantage of the covariance matrix is that the loadings retain the scale of the original data. For the correlation method, however, a weaker signal possessing discriminatory power can be considered at the same level to stronger signals. In this study, both methods were evaluated but the covariance method showed a better separation.

Concerning the changes in carbohydrate metabolism caused by phytoplasma infection, it is noted that Glc and Suc were considerably increased in infected *C. roseus* leaves. Some infected plants showed a carbohydrate level 4 times higher than that of healthy plants. This increase of carbohydrates is confirmed by previous results. Lepka et al. (1999) reported that source leaves of *C. roseus* infected by grape vine yellows, apple proliferation, or ash yellows phytoplasma had a marked increased level of Glc, Fru, and Suc, and starch. In the case of young leaves, the amount of carbohydrates was found to be largely decreased (sink leaves were not analyzed in this study because of the limited amount available for analysis). Mature leaves generally export photosynthates to young leaves in healthy plants. However, the translocation was found to be

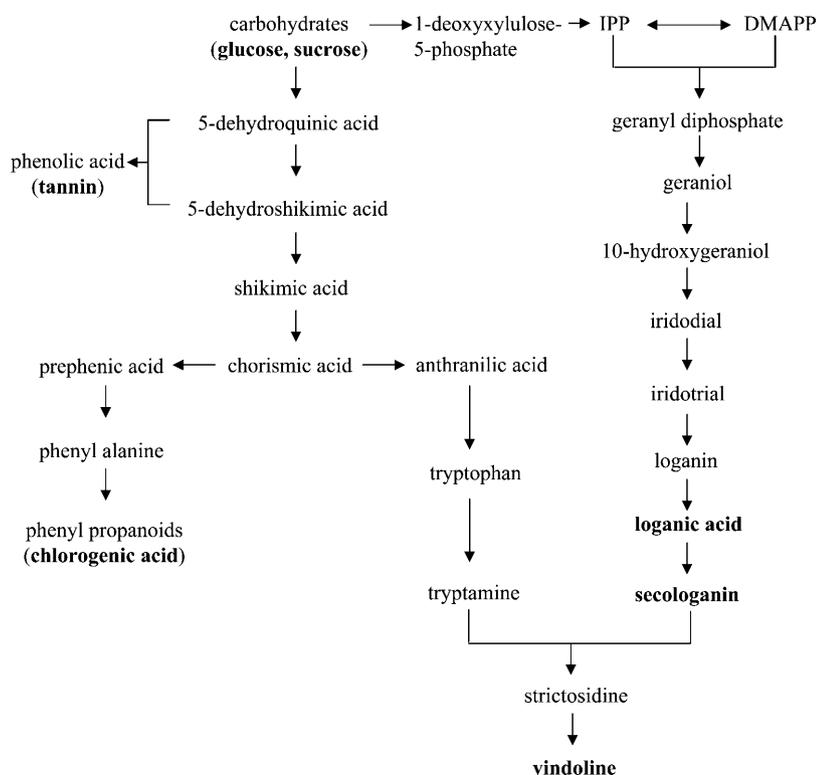


Figure 11. Schematic pathway for the biosynthesis of terpenoid indole alkaloid, phenylpropanoid, and phenolic acid. The increased metabolites in *C. roseus* leaves infected by phytoplasma are in boldface.

severely impaired in phytoplasma-infected leaves of *C. roseus* leaves (Lepka et al., 1999). The levels of carbohydrates induce the infected plants to find a way to consume the Suc or Glc that is accumulating in the source leaves. This could be achieved by using them in secondary metabolic pathways, such as those leading to terpenoid indole alkaloids or phenylpropanoids. Indeed, the increased $^1\text{H-NMR}$ intensities of loganic acid, secologanin, chlorogenic acid, and polyphenols in phytoplasma infection support this hypothesis. After the ubiquitous biosynthetic pathway leading to geranyl diphosphate, a series of specific steps restricted to a few plant species lead to secologanin through loganic acid. Secologanin and tryptamine are stereospecifically condensed to strictosidine by strictosidine synthase and further converted to other indole alkaloids such as ajmalicine, catharanthine, or vindoline (Smith, 1968; Scott et al., 1977; Stöckigt and Zenk, 1977a, 1977b). The level of loganic acid and secologanin in phytoplasma-infected *C. roseus* leaves was highly increased, and this increased level is well correlated with those of Glc and Suc. Most of the phytoplasma-infected *C. roseus* leaves showed higher levels of secologanin (Fig. 9A). Phytoplasma-infected plants showed a marked bleaching of older leaves. This bleaching of older (source) leaves might be a consequence of accumulated carbohydrates in the source leaves (Lepka et al., 1999). The loss of chlorophyll is usually accompanied by a general sugar-mediated repression of genes involved in photosynthesis (Krapp et al., 1993; Lerchl et al., 1996). Vindoline that is derived from secologanin in the terpenoid indole alkaloid pathway showed higher level in the infected leaves when compared to healthy plants. The amounts are well correlated with that of secologanin (Fig. 9B).

The analysis of phenolic metabolites showed chlorogenic acid to be clearly increased in the phytoplasma-infected leaves (Fig. 10). Phenylpropanoids including chlorogenic acid serve as inducible preformed phytoanticipins in many plant species (Dixon, 2001). Tobacco plants overexpressing L-Phe ammonia-lyase produce high levels of chlorogenic acid and exhibit markedly reduced susceptibility to infection with the fungal pathogen *Cerospora nicotianae* (Shadle et al., 2003). Wounding during the preparation of fresh-cut lettuce induced the synthesis and accumulation of chlorogenic acid (Kang and Saltveit, 2003). In case of the cell suspension cultures of *C. roseus* elicited with *Pythium aphanidermatum*, an increased amount of phenolic compounds was found in the culture medium (Moreno et al., 1996). Chlorogenic acid was also proposed as one of the allelochemicals increasing the resistance of cotton to larvae (Kranthi et al., 2003). As in these previous reports, chlorogenic acid is produced at higher levels in the infected plants. The $^1\text{H-NMR}$ spectra of *C. roseus* leaves showed that phytoplasma infection increased the accumulation of chlorogenic acid by 2 to 4 times (Fig. 10).

Musetti et al. (2000) reported that polyphenols, lignin, and suberin were highly increased in plum

and apple tree infected by apple proliferation and plum leptoncrosis. They were proposed to be defense-related metabolites in the systemic disease caused by these phytoplasmas. In this study, the characteristic signals of gallotannin in $^1\text{H-NMR}$ spectra (cluster of singlets in the range of δ 6.8– δ 7.0) were found to be increased in most of phytoplasma-infected *C. roseus* leaves. As mentioned above, one of the symptoms of phytoplasma-infected plants is decoloration of leaves, especially of source leaves. This lack of chlorophyll in the infected leaves might be due to the blockage of the biosynthesis of chlorophyll by Glu or succinic acid. This may explain why $^1\text{H-NMR}$ spectra of the infected leaves showed higher amount of these compounds.

This study shows the great potential of NMR for metabolic profiling. Although minor compounds are not covered by this approach, one single analysis allows a number of quite different secondary metabolite pathways to be covered as well as the level of a series of important primary metabolites. The results gave clear leads for further studies of the effect of the phytoplasma infections. In case of *C. roseus*, the $^1\text{H-NMR}$ spectra showed that the metabolites related to the biosynthesis of terpenoid indole alkaloid (loganic acid, secologanin, and vindoline) and phenylpropanoids (chlorogenic acid and polyphenols) are present in higher amounts in the leaves infected by phytoplasma, as well as Glc and Suc (Fig. 11). These metabolites may relate to the defense mechanism to phytoplasma in *C. roseus*.

MATERIALS AND METHODS

Plant Materials

Twelve *Catharanthus roseus* L. G. Don Peppermint White Cooler (2 types of healthy plants and 10 types of phytoplasma-infected plants) were analyzed to study their metabolic profile. The plants were graft inoculated with one of the following phytoplasma strains: apple proliferation (AP), Bringal little leaf (BLL), Molière disease (MOL), potato purple top (PPT), *Solanum marginatum* big bud (SMBB), Stolbur (DYON), Stolbur (STOF), Stolbur (STOL), Stolbur (UDINESE), and Australian tomato big bud (TBB). The strains had previously been provided by Dr. E. Seemüller, BBA, Dossenheim, Germany, and Dr. N. Petrovic, National Institute of Biology, Ljubljana, Slovenia. The strains had previously been transmitted from their original host (Table I) to *C. roseus*. AP-infected *C. roseus* plants were maintained at Leiden University at 20°C to 24°C with a photoperiod of 12 h a day. The other phytoplasma were maintained in *C. roseus* at the Plant Protection Service in Wageningen under the following conditions: during the day temperatures were 20°C to 24°C and during the night 18°C, with a photoperiod of 14 to 16 h a day. Samples were collected from symptomatic plants at three different times, i.e. January, March, and May 2003.

Solvents and Chemicals

First grade chloroform and methanol were purchased from Merck Biosolve (Valkenswaard, The Netherlands). CDCl_3 (99.96%) and D_2O (99.00%) were obtained from Cambridge Isotope Laboratories (Miami) and NaOD from Cortec (Paris). Potassium dihydrogen phosphate, hexamethyl disilane (HMDS), and trimethyl silane propionic acid sodium salt (TSP) were purchased from Merck (Darmstadt, Germany).

Extraction for Plant Materials

Three hundred milligrams of ground material were transferred to a centrifuge tube. Five milliliters of 50% water-methanol mixture and 5 mL of

chloroform were added to the tube, followed by vortexing for 30 s and sonication for 1 min. The sample was then centrifuged at 3,000 rpm for 20 min. This procedure was performed twice, and the aqueous and organic fractions were collected separately. Each fraction was placed in a 10-mL round-bottom evaporation flask and dried in a rotary vacuum evaporator. The dried fractions were dissolved in 1 mL of deuterium solvent (CDCl_3 or KH_2PO_4 buffer in D_2O).

NMR Measurements

KH_2PO_4 was added to D_2O as a buffering agent. The pH of the D_2O for NMR measurements was adjusted to 6.0 using a 1 M NaOD solution. All spectra were recorded on a Bruker (Billerica, MA) AV-400 NMR and DMX 600 spectrometer operating at a proton NMR frequency of 400.13 MHz and 600.13 MHz, respectively. For each sample, 128 scans were recorded with the following parameters: 0.126 Hz/point, pulse width (PW) = 30° (4.0 μs), and relaxation delay (RD) = 1.0 s. FIDs were Fourier transformed with line broadening factor = 0.3 Hz. The window functions have been optimized for the analysis. For quantitative analysis, peak height was used. The spectra were referenced to residual solvent signal of CDCl_3 (7.26 ppm) for CHCl_3 extract and TSP at 0.00 ppm for water extract. Hexamethyl disilane (HMDS, 0.01%, v/v) for CDCl_3 and trimethyl silane propionic acid sodium salt (TSP, 0.01%, w/v) were used for internal standard.

Data Analysis

The ^1H -NMR spectra were automatically reduced to ASCII files using AMIX (version 3.7, Bruker Biospin). Spectral intensities were scaled to HMDS for CHCl_3 extract and TSP for water extract and reduced to integrated regions of equal width (0.02 ppm) corresponding to the region of δ -0.30 to δ 10.00. The region of δ 4.7 to δ 5.0 was excluded from the analysis because of residual signal of water. PCA b were performed with the SIMCA-P software (version 10.0; Umetrics, Umeå, Sweden).

Received February 13, 2004; returned for revision May 24, 2004; accepted May 25, 2004.

LITERATURE CITED

- Agrawal PK (1992) NMR spectroscopy in the structural elucidation of oligosaccharides and glycosides. *Phytochemistry* **31**: 3307–3330
- Bailey NJ, Oven M, Holmes E, Nicholson JK, Zenk MH (2003) Metabolomic analysis of the consequences of cadmium exposure in *Silene cucubalus* cell cultures via ^1H -NMR spectroscopy and chemometrics. *Phytochemistry* **62**: 851–858
- Batjer LP, Schneider H (1960) Relation of pear decline to rootstocks and sieve-tube necrosis. *Proc Am Soc Horticult Sci* **76**: 85–97
- Catlin PB, Olsson EA, Beutel JA (1975) Reduced translocation of carbon and nitrogen from leaves with symptoms of pear curl. *J Am Soc Horticult Sci* **100**: 184–187
- Dixon RA (2001) Natural products and disease resistance. *Nature* **411**: 843–847
- Doi YM, Teranaka M, Yora K, Asuyama H (1967) Mycoplasma or PLT-group-like microorganisms found in the phloem elements of plants infected with mulberry dwarf, potato wishes' broom, aster yellows, or paulownia wishes' broom. *Ann Phytopathol Soc Jpn* **33**: 259–266
- Fiehn O, Kopka J, Dormann P, Altmann T, Trethewey RN, Willmitzer L (2000) Metabolite profiling for plant functional genomics. *Nat Biotechnol* **18**: 1157–1161
- Garnier M, Foissac X, Gaurivaud P, Laigret F, Renaudin J, Saillard C, Bové JM (2001) Mycoplasmas, plants, insects vectors: a matrimonial triangle. *Life Sci* **32**: 923–928
- Gavaghan CL, Holmes E, Lenz E, Wilson ID, Nicholson J (2000) A NMR based metabolomic approach to investigate the biochemical consequences of genetic strain differences: application to the C57BL10J and Alpk:ApfCD mouse. *FEBS Lett* **484**: 169–174
- Goodacre R, Shann B, Gilbert RJ, Timmins EM, McGovern AC, Kell DB, Logan NA (2000) Detection of the dipicolic acid biomarker in *Bacillus* spores using Curie-point pyrolysis mass spectrometry and Fourier transform infrared spectroscopy. *Anal Chem* **72**: 119–127
- Inouye H (1991) Iridoids. In PM Dey, JB Harbone, eds, *Methods in Plant Biochemistry*, Vol 7. Academic Press, London, pp 99–143
- Kang H-M, Saltveit ME (2003) Wound-induced increase in phenolic content of fresh-cut lettuce is reduced by a short immersion in aqueous hypertonic solutions. *Postharvest Biol Technol* **29**: 271–277
- Karte S, Seemüller E (1991) Histopathology of apple proliferation in *Malus taxa* and hybrids of different susceptibility. *J Phytopathol* **131**: 149–160
- Kim HK, Choi YH, Luijendijk TJC, Vera Rocha RA, Verpoorte R (2004) Comparison of secologanin extraction methods and quantitative analysis of secologanin from *Symphoricarpos albus* by using ^1H -NMR. *Phytochem Anal* (in press)
- Kranthi S, Kranthi KR, Wanjari RR (2003) Influence of semiooper damage on cotton host-plant resistance to *Helicoverpa armigera* (Hub). *Plant Sci* **164**: 157–163
- Krapp A, Hofmann B, Schafer C, Stitt M (1993) Regulation of the expression of rbcS and other photosynthetic genes by carbohydrates: a mechanism for the "sink-regulation" of photosynthesis. *Plant J* **3**: 817–828
- Kummert J, Rufflart G (1997) A preliminary report on the detection of phytoplasma by PCR. *Biochemica* **1**: 19–22
- Lee I-M, Davis RE (1992) Mycoplasmas which infect plants and insects. In J Maniloff, RN McElhaney, LR Finch, JB Baseman, eds, *Mycoplasmas: Molecular Biology and Pathogenesis*. Am Soc Microbiol, Washington, DC, pp 379–390
- Lee I-M, Davis RE, Gundersen-Rindal DE (2000) Phytoplasma: phytopathogenic mollicutes. *Annu Rev Microbiol* **54**: 221–255
- Lepka P, Stitt M, Moll E, Seemüller E (1999) Effect of phytoplasmal infection on concentration and translocation of carbohydrates and amino acids in periwinkle and tobacco. *Physiol Mol Plant Pathol* **55**: 59–68
- Lerchl J, Geigenberger P, Stitt M, Sonnewald U (1996) Inhibition of long distance sucrose transport by inorganic pyrophosphatase can be complemented by phloem specific expression of cytosolic yeast-derived invertase in transgenic plants. *Plant Cell* **7**: 259–270
- Massart DL, Vandeginste BGM, Deming SN, Michotte Y, Kauffman L (1988) *Chemometrics: A Textbook*. Elsevier, New York
- McCoy REA, Caudwell CJ, Chang TA, Chen LN, Chiykowski MT, Cousin JL, Dale GTN, De Leeuw DA, Golino KJ, Hackett BC, et al (1989) Plant diseases associated with mycoplasma-like organisms. In RF Whitcomb, JG Tully, eds, *The Mycoplasmas*, Vol 5. Academic Press, San Diego, pp 545–560
- Moreno PRH, Poulsen C, van der Heijden R, Verpoorte R (1996) Effects of elicitation of different metabolic pathways in *Catharanthus roseus* (L.) Don cell suspension cultures. *Enzyme Microb Technol* **18**: 99–107
- Musetti R, Favali MA, Pressacco L (2000) Histopathology and polyphenol content in plants infected by phytoplasmas. *Cytobios* **102**: 133–147
- Raamsdonk LM, Teusink B, Broadhurst D, Zhang N, Hayes A, Walsh MC, Berden JA, Brindle KM, Kell DB, Rowland JJ, et al (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat Biotechnol* **19**: 45–50
- Roessner U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie AR (2001) Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* **13**: 11–29
- Roessner U, Wagner C, Kopka J, Trethewey RN, Willmitzer L (2000) Simultaneous analysis of metabolites in potato tuber by gas chromatography-mass spectrometry. *Plant J* **23**: 131–142
- Scott AJ, Lee SL, de Capite P, Culver MG (1977) The role of isovincoside (strictosidine) in the biosynthesis of the indole alkaloids. *Heterocycles* (Tokyo) **7**: 979–984
- Seemüller E, Marccone C, Lauer U, Ragozzino A, Göschl M (1998) Current status of molecular classification of the phytoplasmas. *J Plant Pathol* **80**: 3–26
- Shadle GL, Wesley SV, Korth KL, Chen F, Lamb C, Dixon RA (2003) Phenylpropanoid compounds and disease resistance in transgenic tobacco with altered expression of L-phenylalanine ammonia-lyase. *Phytochemistry* **64**: 153–161
- Smith GN (1968) Strictosidine: a key intermediate in the biogenesis of indole alkaloids. *J Chem Soc Chem Comm*: 912–914

- Stöckigt J, Zenk MH (1977a)** Isovincoside (strictosidine), the key intermediate in the enzymatic formation of indole alkaloids. *FEBS Lett* **79**: 233–237
- Stöckigt J, Zenk MH (1977b)** Strictosidine (isovincoside): the key intermediate in the biosynthesis of monoterpenoid indole alkaloids. *J Chem Soc Chem Comm*: 646–648
- Summer LW, Mendes P, Dixon A (2003)** Plant metabolomics: large-scale phytochemistry in the functional genomics era. *Phytochemistry* **62**: 817–836
- Tomassini L, Cometa MF, Serafini M, Nicoletti M (1995)** Isolation of secoiridoid artefacts from *Lonicera japonica*. *J Nat Prod* **58**: 1756–1758
- Ward JL, Harris C, Lewis J, Beale MH (2003)** Assessment of ¹H-NMR spectroscopy and multivariate analysis as a technique for metabolite fingerprinting of *Arabidopsis thaliana*. *Phytochemistry* **62**: 949–958
- Yu F (1997)** Pigment content in vitro culture of periwinkle infected with aster yellows phytoplasma or *Spiroplasma citri*. MS thesis. University of Georgia, Griffin