

Fractionation Protocol for the Isolation of Polypeptides from Plant Biomass

Per Claeson,* Ulf Göransson, Senia Johansson, Teus Luijendijk, and Lars Bohlin

Division of Pharmacognosy, Department of Pharmacy, Biomedical Center, Uppsala University, Box 579, S-751 23 Uppsala, Sweden

Received July 18, 1997

A fractionation protocol for the isolation of a highly purified polypeptide fraction from plant biomass is described. The procedure dereplicates ubiquitous substance classes known to interfere with bioassays often used in natural product-based drug discovery programs. The protocol involves pre-extraction with dichloromethane, extraction with ethanol (50%), removal of tannins with polyamide, removal of low-molecular-weight components with size-exclusion chromatography over Sephadex G-10, and final removal of salts and polysaccharides with solid-phase extraction using reversed-phase cartridges. The method has been applied to the aerial parts of *Viola arvensis*, resulting in the isolation of a peptide fraction that on further separation yielded a novel 29-residue macrocyclic polypeptide named varv peptide A, cyclo(-TCVGGTCNT-PGCSWVCTRNLPLVCGE-).

Strategies and methodology for the isolation of polypeptides (defined as peptides containing between 10 and 50 amino acid residues) from plant biomass have recently received attention for three main reasons. First, plants containing unique pharmacologically active polypeptides have been found within natural product-based drug discovery programs.^{1,2} Second, plants, like animals, are now known to make use of peptides as signal substances.^{3,4} Finally, genetically transformed plants ("transgenic plants") are now considered an attractive and cost-efficient alternative to fermentation-based systems for production of, for example, high value, recombinant polypeptides.^{5,6}

An extensive number of biologically important polypeptides, such as hormones, neurotransmitters and snake toxins, have been isolated from human and animal sources. Until now, however, only a limited number of polypeptides have been reported from plants.^{e.g., 1,2,7–9} Consequently, numerous methods for the isolation of polypeptides from animal materials are described in the literature, whereas a discussion of procedures for the isolation of polypeptides from plant biomass is virtually lacking. The plant biomass constitutes a highly complex matrix containing many components, e.g., photosynthetic pigments, polysaccharides, tannins, and secondary metabolites, that are not present in animal materials. Isolation procedures described for animal polypeptides therefore are generally not directly applicable to the isolation of polypeptides from plant materials.

As part of a research program aimed at finding novel biologically active plant polypeptides in biomass of intact plants and plant cell cultures, we have developed a fractionation protocol that efficiently dereplicates most ubiquitous plant constituents and enables isolation of a highly purified polypeptide fraction from plant biomass. This paper describes the fractionation protocol that presently has been applied to a large number of plant materials in our laboratory. To illustrate the utility of the procedure, we also report the isolation of

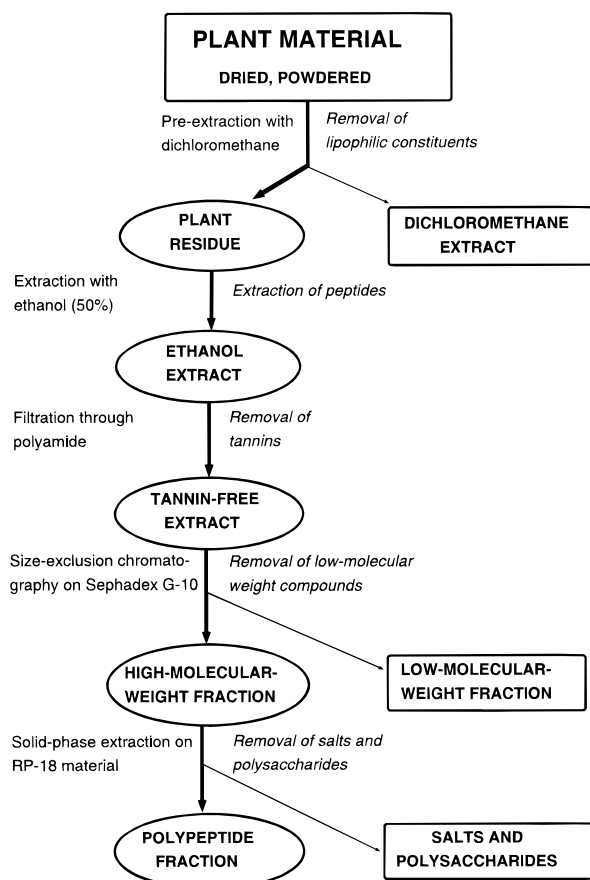


Figure 1. Fractionation protocol for isolation of polypeptides from plant biomass. For details see the Experimental Section.

a new 29-residue macrocyclic peptide from the polypeptide fraction of *Viola arvensis* Murray (Violaceae).

Results and Discussion

A flowchart of the procedure is outlined in Figure 1. The fresh plant material was dried at 55–60 °C to preserve it and prevent enzymatic degradation of peptides. Previously reported plant peptides contain sev-

* To whom correspondence should be addressed. Phone: +46 18 471 44 91. FAX: +46 18 50 91 01. E-mail: per.claeson@pharmacog.uu.se.

eral disulfide bonds and appear largely heat stable. Several of them are reportedly stable to aqueous boiling.^{8,9} The dried plant material was ground to a fine powder and preextracted in a flask with dichloromethane on a shaking table. Four consecutive extraction periods (1 h each) with fresh solvent were found to give an extractive yield (gravimetrically determined) amounting to more than 90% of the yield obtained by continuous Soxhlet extraction for 8 h (exhaustive extraction). Polypeptides are not soluble in dichloromethane, but ubiquitous lipophilic substances such as chlorophyll, lipids, and other low-molecular-weight substances (e.g., terpenoids, phenylpropanoids, etc.) are extracted and thus removed.

The air-dried plant residue was then extracted with 50% aqueous ethanol. This solvent has several advantages. The solubility of known polypeptides is generally better in 50% ethanol than in pure water or alcohol. It needs no preservation from microbial growth. Furthermore, it does not extract most polysaccharides^{10,11} or enzymes.

The 50% ethanol extract contained ubiquitous polyphenols (tannins), and to remove them, the extract was passed through a column of polyamide. Strong hydrogen bonding occurs between polyphenolics and polyamide, and the tannins are thus practically irreversibly bound to the column. Several studies have shown that this is a highly efficient method to remove undesired tannins.^{12–17} Peptides are not retained on this column support.^{11,14} The binding of tannins to polyamide is highly pH-dependent. By lowering the pH from 7.5 to 3, an approximately 10-fold increase in binding capacity of polyamide for hydrolyzable tannins has been reported.¹³ Our experiments confirmed this observation, and therefore, after evaporation of the ethanol, the extracts were acidified (pH ca. 2.7) by addition of acetic acid to a final concentration of 2% before application onto the column. Peptides were then eluted using acetic acid (2%) as eluent. The column was washed afterward with ethanol (50%)/acetic acid (2%) to elute peptides insoluble in 2% acetic acid.

The tannin-free extract was subjected to size-exclusion chromatography on a calibrated column packed with Sephadex G-10. Peptides (and other compounds) with a molecular weight above approximately 700 Da were eluted at the void volume of the column (high-molecular-weight fraction). Compounds with molecular weights lower than ca. 700 Da (e.g., amino acids, glycosides, alkaloids, etc.) were retarded on the column and eluted as the low-molecular-weight fraction.

Several experiments were performed to establish a suitable mobile phase for this gel filtration step. The reference substances bacitracin (mol wt ca. 1300 Da), bradykinin (mol wt 1060 Da), G-strophanthin (mol wt 585 Da), and different amino acids were used to calibrate the column. Pure water or conventional buffers (phosphate, ammonium acetate) were found unsuitable as the extract samples could only be dissolved partially. Ethanol (50%) was found to provide incomplete separation (overlapping of high- and low-molecular-weight markers). Plant pigments in the extracts also bound strongly to the gel. Inclusion of 2% acetic acid in the mobile phase was found to suppress most of this interaction between pigments and the gel.

In 1967, Eaker and Porath¹⁸ showed that Sephadex G-10, as supplied by the manufacturer, behaved as a weak cation exchanger and that this capacity could be practically eliminated by washing with 1 M aqueous pyridine. Furthermore, they showed that the basic amino acids (arginine, lysine, and histidine) eluted anomalously at the void volume when 0.2 M acetic acid was used as the mobile phase on such a pyridine-washed column. After increasing the ionic strength of the mobile phase by including 0.5 M NaCl, the basic amino acids eluted, as expected, in the low-molecular-weight fraction after the void volume. We repeated some of these earlier experiments with positive results. Making use of the information obtained, we found that a mobile phase consisting of ethanol (50%)/acetic acid (2%)/NaCl (0.2 M) exhibited suitable properties for the size-exclusion chromatography step. Bacitracin and bradykinin appeared at the void volume, whereas G-strophanthin and the amino acids eluted in the low-molecular-weight fraction. The plant extracts readily dissolved, and nonspecific interactions between the plant pigments and the gel were largely suppressed.

The resulting high-molecular-weight fraction contained, besides the polypeptides, a large amount of NaCl and, as shown by NMR and TLC, polysaccharides. The salt and the polysaccharides were removed by solid-phase extraction (SPE) using a C₁₈ reversed-phase column. The lyophilized sample was dissolved in a large volume of 50 mM NH₄HCO₃, applied to the column, and then eluted with the same buffer. Experiments with commercially available polypeptides (oxytocin, gramicidin, bacitracin, and insulin) showed that they were not eluted in this step. Salts and polysaccharides were, however, easily eluted. The polypeptides were then eluted from the column in three steps by adding 20%, 50%, and 80% ethanol, respectively. None of the reference peptides required higher concentrations of ethanol in order to be eluted.

The highly purified fraction thus obtained (fraction P) can be submitted directly to analytical (chromatographic and spectroscopic) procedures to check whether the material contains polypeptides. Fraction P can also directly be subjected to bioassay. The main advantages of performing this rather laborious fractionation procedure prior to bioassay are that higher test concentrations of the peptides can be obtained and that a number of known bioactive, but presently undesired, nonpeptide plant constituents are dereplicated. The procedure removes several compound classes that interfere with bioassays ("false positive hits"). Thus, tannins, for example, which are known to nonspecifically interfere with various enzyme inhibition assays,^{12,15,19} and anti-HIV-active polysaccharides¹⁰ were routinely excluded prior to bioassay. Chlorophylls and other lipophilic plant pigments that might conceivably interfere with photometrically evaluated bioassays and bioactive low-molecular-weight substances, such as alkaloids and glycosides, were also removed.

To validate the fractionation protocol, we applied the described procedure to plant material [the aerial parts of *V. arvensis* (Violaceae)] that has been reported previously to contain a cyclic 29-residue polypeptide, viola peptide-I.²⁰ Fraction P of *V. arvensis* yielded free amino acids on hydrolysis, and a RP-HPLC chromato-

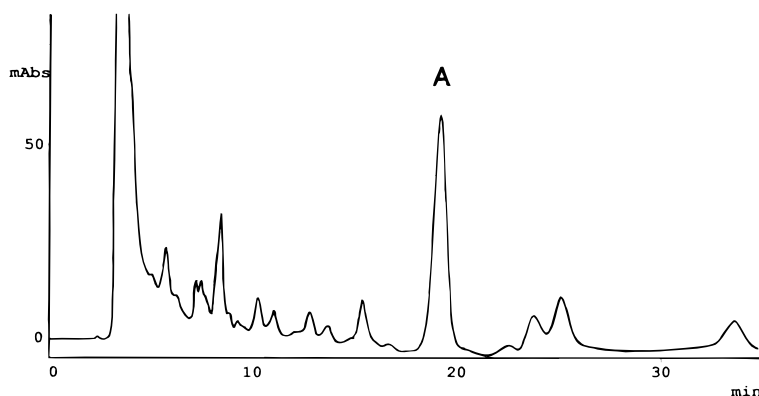


Figure 2. Semipreparative RP-HPLC of the polypeptide fraction obtained from *Viola arvensis*. The peak labeled A in the chromatogram corresponds to varv peptide A. Mobile phase: 40% CH₃CN-*i*-PrOH (6:4)/0.1% TFA (adjusted to pH = 2.25 by addition of NH₄OH). UV detection at 215 nm. Flow rate 4 mL/min. (For further details see the Experimental Section).

Table 1. Amino Acid Analysis of Varv Peptide A Isolated from *Viola arvensis*

amino acid	residues from amino acid analysis	residues from sequencing ^a
Asx	2.0	2 (2 Asn)
Thr	3.7	4
Ser	2.0	2
Glx	1.0	1 (Glu)
Pro	2.9	3
Gly	4.9	5
Cys	5.2 ^b	6 ^c
Val	3.0	3
Leu	1.0	1
Arg	1.0	1
Trp	1.0 ^d	1

^a A total of 29 residues were determined by sequencing. ^b Half-cysteine was determined as cysteic acid with a separate sample following oxidation with performic acid. ^c Cys was determined as (pyridylethyl)cysteine following alkylation with vinylpyridine. ^d Trp was determined photometrically.

gram of the fraction showed several peaks with UV spectra consistent with tryptophan-containing peptides.

To obtain sufficient material for characterization of the components in fraction P, the same isolation strategy was applied on a larger amount of plant material (for details, see the Experimental Section). The polypeptide fraction thus obtained was subjected to reversed-phase chromatography (Figure 2), and the substance corresponding to the major peak in the chromatogram was collected and designated varv peptide A. Quantitative amino acid analysis of the hydrolysate of varv peptide A indicated the amino acid composition shown in Table 1. The calculated average mass of a linear peptide with this composition would be between 2901.3 Da (2 Asn; 1 Gln) and 2904.3 Da (2 Asp; 1 Glu). Mass spectrometry of varv peptide A, however, provided a lower molecular weight of 2879.4 Da, suggesting the peptide to be macrocyclic, like the previously reported *viola* peptide-I.²⁰ Varv peptide A was then reduced with mercaptoethanol and subsequently alkylated with 4-vinylpyridine. Cleavage of the alkylated peptide with endoproteinase Glu-C resulted in a single linear product, consistent with the opening of a macrocyclic ring. The linear product was subjected to automated Edman degradation, defining the 2 Asx and 1 Glx from the quantitative amino acid analysis as 2 Asn and 1 Glu (cf., Table 1). Consistent with the amino acid sequence obtained for the linear product, the structure for varv peptide A follows below:

cyclo-
(-Thr-Cys-Val-Gly-Gly₅-Thr-Cys-Asn-Thr-Pro₁₀-
Gly-Cys-Ser-Cys-Ser₁₅-Trp-Pro-Val-Cys-Thr₂₀-
Arg-Asn-Gly-Leu-Pro₂₅-Val-Cys-Gly-Glu-)

The calculated average mass of a linear peptide with this amino acid composition would be 2902.3 Da. Assuming the peptide to be macrocyclic (-18 Da) and the six cysteine residues to be engaged in intramolecular disulfide bonds (-6 Da) would give a molecular weight of 2878.3 Da, which was in agreement with the molecular weight of 2879.4 Da experimentally determined by MALDI-TOF MS for varv peptide A.

Varv peptide A and the previously reported *viola* peptide-I share a very high degree of sequence homology. They differ only at two amino acid positions. The Trp and the Arg residues in varv peptide A are substituted for an Arg residue and an X (unidentified amino acid) residue, respectively, in the published sequence of *viola* peptide-I.²⁰

The isolation and identification of a 29-residue polypeptide from fraction P of *V. arvensis* validates the above-described fractionation protocol.

Experimental Section

Plant Material. *V. arvensis* Murray (Violaceae) was collected in July 1996 near the Ångström Laboratory, Uppsala. A voucher specimen (labeled VM-107) was identified by Dr. Ö. Nilsson, the Botanical Garden, Uppsala, and deposited at the herbarium of Uppsala University.

Extraction. The dried, powdered plant material (4.0 g) was placed in a Soxhlet thimble and positioned in an E-flask (100 mL) with 75 mL of CH₂Cl₂. The E-flask was mounted on a shaking table for 1 h. The procedure was repeated three times with fresh solvent. The CH₂-Cl₂-soluble extractives were discarded. The plant residue was dried at room temperature, and the main extraction was then carried out three times in an analogous manner with EtOH (50%).

Removal of Tannins by Filtration through Polyamide. The 50% EtOH extract was evaporated in vacuo to a volume of ca. 50 mL and acidified by addition of HOAc to a final concentration of 2%. A 4.0 g quantity of polyamide 6S (Riedel-de Haen, Seelze, Germany) was preswollen in 2% HOAc overnight and packed in a 50 mL glass column. The gel was rinsed with 200 mL of

EtOH (50%)/HOAc (2%) and then with 200 mL of HOAc (2%). The sample was applied and the column eluted by gravity flow with 100 mL of HOAc (2%) followed by 100 mL of EtOH (50%)/HOAc (2%). The eluates were combined, evaporated in vacuo, and lyophilized after addition of 50 mL of demineralized water.

Size-Exclusion Chromatography on Sephadex G-10. All experiments were performed on a fast protein liquid chromatograph (FPLC; Pharmacia Biotech, Uppsala, Sweden). Sephadex G-10 (Pharmacia Biotech) (95 g) was preswollen in 210 mL of the mobile phase EtOH (50%)/HOAc (2%)/NaCl (0.2 M) and packed in a glass column with length-variable end-pieces (Baeckström SEPARO AB, Stockholm, Sweden) to give a column with an effective size of 3 × 30 cm. The column was first washed with 1 M pyridine in EtOH (50%) and then equilibrated with several column volumes of the mobile phase. The void volume of the column was determined to be 87 mL in separate experiments using the markers bacitracin and bradykinin (both from Sigma). The tannin-free plant extract (550 mg) was dissolved in 5.5 mL of the mobile phase and centrifuged, and a 5-mL aliquot was injected into the FPLC, operated at a flow rate of 0.6 mL/min. The high-molecular-weight fraction eluting at the void volume was collected and lyophilized.

Solid-Phase Extraction on C₁₈ Material. The high-molecular-weight fraction described above was dissolved in NH₄HCO₃ buffer (50 mM) in a 40:1 (v/w) ratio of buffer:sample and applied to a C₁₈ Isolute SPE column (5 g/20 mL; Sorbent AB, Sollentuna, Sweden) previously conditioned in EtOH (95%) and preequilibrated in the same buffer. The column was washed with 50 mL of the buffer and the eluate discarded. The column was then eluted sequentially with 50 mL each of 20%, 50%, and 80% EtOH in NH₄HCO₃ buffer (50 mM). The combined eluates were evaporated in vacuo and lyophilized to yield fraction P. Typical yields of fraction P amounted to ca. 0.5–1% (w/w) of the dried plant materials.

Isolation of Varv Peptide A. The dried plant powder (130 g) was extracted with 500 mL of dichloromethane in a beaker on a gyratory shaker for 1 h. This was repeated five times with fresh solvent. The CH₂Cl₂ extract was discarded. The dried plant residue was extracted in a similar manner with 50% ethanol. This extract was concentrated to a volume of ca. 500 mL and acetic acid was added to a final concentration of 2%. It was then loaded onto a column of 24.1 g of Polyamide 6S. Elution took place using 450 mL of 2% (v/v) acetic acid, followed by 625 mL of 2% (v/v) acetic acid/50% (v/v) ethanol. The combined eluates were concentrated and subsequently freeze-dried (yield 52.5 g). Approximately 4.6 g of this lyophilisate was then loaded onto a 28 × 3 (i.d.) cm column packed with Sephadex G-10 and eluted with the mobile phase described above (flow rate 0.6 mL/min). The high-molecular-weight fraction was collected (yield: 410 mg). The sample was redissolved in 50 mM NH₄HCO₃, and the solution was then subjected to solid-phase extraction on a 5 g C₁₈ cartridge using increasing concentrations of ethanol (20 to 80%) in 50 mM NH₄HCO₃ for elution of the peptides. The combined eluates were concentrated and freeze-dried to yield 156 mg of the polypeptide fraction.

Semipreparative HPLC was performed with a Shimadzu system, equipped with an SPD-M10Avp photodiode array detector, and a 250 × 10 (i.d.) mm Dynamax column (C₁₈, 5 μm, pore size 300 Å). Crude varv peptide A was obtained after repeated injections of the polypeptide fraction (42 mg total) using the conditions shown in Figure 2. Final purification of the peptide was achieved by gel filtration on a Superdex Peptide HR 10/30 column (Pharmacia Biotech, Uppsala, Sweden) using a mobile phase of 40% CH₃CN/0.1% TFA followed by RP-HPLC on an analytical column (Dynamax, C₁₈, 5 μm, pore size 300 Å, 250 × 4.6 mm i.d., mobile phase: 36% CH₃CN/0.1% TFA). Final yield of varv peptide A: 150 μg. To estimate the concentration of varv peptide A in fraction P, a standard curve for analytical HPLC was constructed using isolated varv peptide A (data not shown). By using this standard curve the concentration of varv peptide A in fraction P was estimated to ca. 3.5% (w/w).

Structure Determination of Varv Peptide A. The quantitative determination of the amino acid content of the peptide was performed at the Amino Acid Analysis Centre, Department of Biochemistry, Uppsala University. The peptide was hydrolyzed for 24 h at 110 °C with 6 N HCl containing 2 mg/mL phenol, and the hydrolysates were analyzed with an LKB model 4151 Alpha Plus amino acid analyzer using ninhydrin detection.

For the amino acid sequence analysis, the peptide was reduced with mercaptoethanol in 0.25 M Tris-HCl containing 1 mM EDTA and 6 M guanidine-HCl (pH 8.5, 24 °C, 2 h) and subsequently *S*-pyridylethylated by addition of 4-vinylpyridine to the same solution (24 °C, 2 h). The reduced and alkylated peptide was cleaved with endoproteinase Glu-C in 0.1 M ammonium bicarbonate buffer (pH 8.1, 37 °C, 2.5 h). The sequence analysis was performed via automated Edman degradation using an Applied Biosystems model 477A protein/peptide sequencer coupled on-line to an Applied Biosystems 120A PTH analyzer.

Mass spectra were obtained using a Kratos Kompact IV MALDI-TOF mass spectrometer. The spectrometer was externally calibrated and operated in the linear mode (experimental accuracy: ±0.1%).

Acknowledgment. We thank Dr. Å. Engström, Department of Medical and Physiological Chemistry, Uppsala University, for assistance with MS and amino acid sequence analysis. We also thank Pharmacia Biotech, Uppsala, for support with separation equipment.

References and Notes

- Gustafson, K. R.; Sowder, R. C., II; Henderson, L. E.; Parsons, I. C.; Kashman, Y.; Cardellina, J. H., II; McMahon, J. B.; Buckheit, R. W.; Pannell, L. K.; Boyd, M. R. *J. Am. Chem. Soc.* **1994**, *116*, 9337–9338.
- Witherup, K. M.; Bogusky, M. J.; Anderson, P. S.; Ramjit, H.; Ransom, R. W.; Wood, T.; Sardana, M. *J. Nat. Prod.* **1994**, *57*, 1619–1625.
- Marx, J. *Science* **1996**, *273*, 1338–1339.
- Bergey, D. R.; Howe, G. A.; Ryan, C. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 12053–12058.
- Goddijn, O. J. M.; Pen, J. *Tibtech* **1995**, *13*, 379–387.
- Whitelam, G. *J. Sci. Food Agric.* **1995**, *68*, 1–9.
- Bohlman, H.; Apel, K. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **1991**, *42*, 227–240.
- Samuelsson, G. *System. Zool.* **1973**, *22*, 566–569.
- Gran, L. *Lloydia* **1973**, *36*, 174–178.

- (10) Beutler, J. A.; McKee, T. C.; Fuller, R. W.; Tischler, M.; Cardellina, J. H., II; Snader, K. M.; McCloud, T. G.; Boyd, M. R. *Antivir. Chem. Chemother.* **1993**, *4*, 167–172.
- (11) Thunberg, E.; Samuelsson, G. *Acta Pharm. Suec.* **1982**, *19*, 285–292.
- (12) Cardellina, J. H., II; Munro, M. H. G.; Fuller, R. W.; Manfredi, K. P.; McKee, T. C.; Tischler, M.; Bokesch, H. R.; Gustafson, K. R.; Beutler, J. A.; Boyd, M. R. *J. Nat. Prod.* **1993**, *56*, 1123–1129.
- (13) Loomis, W. D.; Battaile, J. *Phytochemistry* **1966**, *5*, 423–438.
- (14) Samuelsson, G.; Ekblad, M. *Acta Chem. Scand.* **1967**, *21*, 849–856.
- (15) Tan, G. H.; Pezzuto, J. M.; Kinghorn, A. D.; Hughes, S. H. *J. Nat. Prod.* **1991**, *54*, 143–154.
- (16) Wall, M. E.; Taylor, H.; Ambrosio, L.; Davis, K. *J. Pharm. Sci.* **1969**, *58*, 839–841.
- (17) Wall, M. E.; Wani, M. C.; Brown, D. M.; Fullas, F.; Olwald, J. B.; Josephson, F. F.; Thornton, N. M.; Pezzuto, J. M.; Beecher, C. W. W.; Farnsworth, N. R.; Cordell, G. A.; Kinghorn, A. D. *Phytomedicine* **1996**, *3*, 281–285.
- (18) Eaker, D.; Porath, J. *Separation Sci.* **1967**, *2*, 507–550.
- (19) Haslam, E. *J. Nat. Prod.* **1996**, *59*, 205–215.
- (20) Schöpke, Th.; Agha, M. I. H.; Kraft, R.; Otto, A.; Hiller, K. *Sci. Pharm.* **1993**, *61*, 145–153.

NP970342R