Isolation and characterization of anti-HIV peptides from *Dorstenia contrajerva* and *Treculia obovoidea*

Heidi R. Bokesch[a,b], Romila D. Charan[b], Karina M. Meragelman[b], John A. Beutler[b], Roberta Gardella[a,b], Barry R. O'Keefe[b], Tawnya C. McKee[b,*], James B. McMahon[b]

[a] *Basic Research Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD 21702-1201, USA*
[b] *Molecular Targets Development Program, Center for Cancer Research, NCI-Frederick, Frederick, MD 21702-1201, USA*

Received 4 December 2003; revised 9 April 2004; accepted 26 April 2004

Available online 12 May 2004

Edited by Hans-Dieter Klenk

Abstract Using a high throughput screen based on the interaction of the HIV-1 gp41 ectodomain with the virucidal protein cyanovirin-N (CV-N), we isolated two new peptides which inhibited the binding of CV-N to gp41 and which subsequently showed anti-HIV activity in a whole cell assay. A 5-kDa (contrajervin) and 10 kDa (treculavirin) peptide were isolated from *Dorstenia contrajerva* and *Treculia obovoidea*, respectively. Treculavirin was composed of two subunits, each containing 50 amino acid residues, which are covalently linked by at least one disulfide bond between the subunits. Both peptides were shown to bind to gp41 and gp120 and to inhibit the cytopathic effects of HIV-1_{RF} infection in a human T-lymphoblastoid cell line (CEM-SS).

Keywords: Peptide; gp41; Moraceae; *Dorstenia contrajerva*; *Treculia obovoidea*

1. Introduction

Infection of cells by HIV-1 requires virus-cell fusion, a process which is mediated in part by the viral envelope glycoprotein gp41. Compounds which inhibit this interaction by binding to gp41 are capable of inhibiting viral entry into host cells [1]. Cyanovirin-N (CV-N) is an anti-HIV protein originally isolated from the cyanobacterium *Nostoc ellipsosporum* [2], whose activity is due to high-affinity binding to gp120 and gp41 envelope glycoproteins [3]. Based on this interaction, an assay was developed to find compounds that could bind competitively to gp41 in a manner similar to CV-N, therefore providing new antiviral leads [4]. The assay involves competition for CV-N binding to passively immobilized gp41, using time-resolved fluorescence detection of europium-labeled CV-N. Extracts that showed activity in the initial screen were further prioritized by eliminating those whose gp41 binding could be blocked by the sugars mannose and *N*-acetylglucosamine (glcNAc), indicating monosaccharide-specific lectin activity. Two plant extracts that met the criteria were selected for bioassay-guided fractionation and subsequently yielded 5- and 10-kDa peptides.

Dorstenia contrajerva and *Treculia obovoidea* are plants in the Moraceae family, which are found in tropical and subtropical areas. Various plant parts of *D. contrajerva* have been used in traditional medicines to reduce fever and to treat colds, diarrhea, and snakebites, and the reported activity was attributed to coumarins, chalcones, flavones and flavanones [5]. The genus *Treculia* contains three species, the best known being *Treculia africana* (African breadfruit). Reported components of this staple food include polyphenols and polysaccharides as well as a high content of fatty acids and essential amino acids. There is no reported literature on *T. obovoidea*. To the best of our knowledge, this is the first report of peptides or anti-HIV activity from either genus.

This paper describes two peptides isolated from the leaves of *D. contrajerva* and the bark of *T. obovoidea*. The *Dorstenia* peptide, contrajervin, is a 5-kDa, 46 amino acid monomer containing eight disulfide bridge-linked cysteine residues. The *Treculia* peptide, treculavirin, is a 10-kDa, 100 amino acid dimer containing 20 cysteine residues, all of which are involved in disulfide bonds, and at least one of these disulfide bonds is between the two 50 amino acid monomers. The two peptides show high primary sequence similarity to each other as well as to chitinase and hevein-like proteins.

2. Materials and methods

2.1. General experimental procedures

All solvents were high-performance liquid chromatography (HPLC) grade and were purchased from EM Science. Endoproteinases were obtained from Roche Molecular Biochemicals (Indianapolis, IN). The gp41 (recombinant, glycosylated, HIV-1_{HxB2} gp41 and ecto domain) was purchased from Viral Therapeutics, Inc. (Ithaca, NY) and the gp120 (native, glycosylated and HIV-1_{IIIB} gp120) from Advanced Biotechnologies Inc. (Columbia, MD). The gp120 (recombinant, glycosylated and HIV-1_{BaL} gp120) was obtained through the NIH AIDS Research and Reference Reagent Program, Division of AIDS, NIAID, NIH. Electrospray ionization mass spectra (ESI-MS) were acquired on a Hewlett–Packard HP1100 integrated liquid chromatography–mass spectrometry (LC–MS) system. Amino acid sequences were determined by Edman degradation using an Applied Biosystems Model 494 sequencer. All HPLC mobile phases included 0.05% (v/v) trifluoroacetic acid (TFA).

* Corresponding author. Fax: +1-301-846-6177.
E-mail address: manuscripts@ncifcrf.gov (T.C. McKee).

Abbreviations: CV-N, cyanovirin-N; LC–MS, liquid chromatography–mass spectrometry; HPLC, high-performance liquid chromatography; TFA, trifluoroacetic acid; HTS, high throughput screening; UDA, *Urtica diocia* agglutinin; PEC, *S*-(β-4-pyridylethyl)cysteine; ESI-MS, electrospray ionization mass spectra; BSA, bovine serum albumin; PBS, phosphate buffered saline; glcNAc, *N*-acetylglucosamine

0014-5793/$22.00
doi:10.1016/j.febslet.2004.04.085

2.2. Plant material

Plant material was obtained under NCI contract and voucher specimens are maintained at the Smithsonian Institute. Leaves of *D. contrajerva* L. were collected at the Altun-Ha Ruins in Belize, British Honduras in March 1990. The dried plant material (449 g) was ground and sequentially extracted with methanol–methylene chloride 1:1 v/v, followed by 100% methanol. The plant material was then dried, extracted with water, and lyophilized to yield 23.6 g of aqueous extract.

Stembark of *T. obovoidea* N. E. Br. was collected in Woleu Ntem province, Gabon, Africa in February 1992. The dried plant material (446 g) was treated as described above to yield 8.4 g of aqueous extract.

2.3. Isolation and purification of peptides

A 5.0-g portion of the aqueous extract of *D. contrajerva* was subjected to vacuum-liquid chromatography on Bakerbond wide pore C_4 media (8 × 5 cm), eluting with a total of 2 L of a stepwise gradient of 0–100% methanol. The bioactive water–methanol 2:1 v/v fraction was lyophilized to yield 430 mg of powder which was precipitated with 50 ml of water–ethanol 1:1 v/v at −20 °C overnight. The resulting precipitate (190 mg) was resuspended in 15 ml of water and filtered through an Amicon Centriprep 3000 molecular weight cutoff filter. The >3000 Da retentate (93 mg) was purified using reversed-phase HPLC (Dynamax C_4, 1 × 25 cm, 300 Å) eluting with a linear gradient of 0–60% acetonitrile in water over 40 min at a flow rate of 3 ml/min to yield 7.5 mg of contrajervin.

A 2.0-g portion of the aqueous extract of *T. obovoidea* was subjected to vacuum-liquid chromatography as above. The bioactive water–methanol 2:1 v/v fraction was lyophilized to yield 417 mg of powder. A 100-mg portion of this was chromatographed on Sephadex LH-20 (95 × 2.5 cm), eluting with water–methanol 3:1 v/v. An early eluting fraction (13.3 mg) was further purified by reversed-phase HPLC (Hamilton PRP-3, 1 × 25 cm, 300 Å) eluting with a gradient of 5–30% acetonitrile in water over 45 min at a flow rate of 3 ml/min to yield 0.6 mg of treculavirin.

2.4. Sequence determination

For contrajervin, disulfide bonds were reduced and alkylated as previously described [6]. The derivatized peptide was purified by reversed-phase HPLC (Dynamax C_4, 300 Å) using a gradient elution of 0.05% aqueous TFA for 40 min, then increasing to 50% acetonitrile in water over 60 min. The *S*-(β-4-pyridylethyl)cysteine (PEC) derivative (250 μg) was subjected to endoproteinase Asp-N and trypsin (both from Boehringer–Mannheim GmbH) digestion as per the manufacturer's instructions at an enzyme/substrate ratio of 1:40 and 1:10, respectively. The cleaved peptide products were purified by reversed-phase HPLC (Dynamax C_4, 300 Å) using a gradient of water for 20 min, then increasing to 50% acetonitrile over 90 min. The PEC derivative and digested fractions were sequenced by Edman degradation and analyzed by LC–MS to confirm the complete sequence.

For treculavirin, disulfide bonds were reduced and alkylated as previously described [7] with the following change. The reaction mixture was maintained at 37 °C for 17 h for reduction and 25 h for alkylation. The reduced and alkylated products were each analyzed by ESI-MS. The PEC derivative was purified by reversed-phase HPLC (Hamilton PRP-3, 300 Å) using a gradient elution of water for 30 min, then increasing to 50% acetonitrile in water over 30 min. The PEC derivative (250 μg) was subjected to endoproteinase Asp-N and trypsin digestion as per the manufacturer's instructions at an enzyme/substrate ratio of 1:50 and 1:20, respectively. The cleaved peptide products were purified by reversed-phase HPLC (Hamilton PRP-3, 300 Å) using a gradient of water for 30 min, then increasing to 40% acetonitrile over 30 min and to 100% acetonitrile over an additional 10 min. The digested fractions were sequenced by Edman degradation and analyzed by LC–MS to confirm the complete sequence.

2.5. Biological assays

Fractions and pure peptides were screened in a high throughput screening (HTS) assay that measured inhibition of the binding of labeled CV-N to gp41 as previously described [4]. This is a solid-phase competition-binding assay using time-resolved fluorescence detection of europium Eu^{+3}-labeled CV-N. The recombinant gp41 utilized for this assay was fully glycosylated and was in the form of a fusion-active six-helix bundle representing the ectodomain of gp41 [8]. In brief, samples in 1% bovine serum albumin (BSA) in phosphate buffered saline (PBS) were preincubated with plate-bound gp41 in a 96-well plate format at room temperature. Eu^{+3}-labeled CV-N was then added and the plates were incubated again at room temperature before being washed with PBS containing 0.05% Tween 20 (TBPS). DELFIA® enhancer solution (Perkin–Elmer Life Sciences) was added to solubilize the bound europium and fluorescence was measured in time-resolved mode using a Perkin–Elmer Life Sciences Victor^{2TM} fluorometer.

An XTT-tetrazolium based assay was used to confirm the anti-HIV activity of the pure peptides. This assay, as previously described [9], measures the viability of virus-infected CEM-SS cells by the metabolic reduction of the tetrazolium salt, XTT, to a soluble, colored formazan product which can be quantitated by spectrophotometry.

Further studies on biological activity were done with contrajervin only, due to the limited mass of treculavirin. To determine whether contrajervin also inhibited the binding of labeled CV-N to gp120, ELISA plates were treated as in the gp41 assay with the following modifications. 25 ng of $gp120_{BaL}$ and $gp120_{IBaL}$ was bound to separate 96-well plates, which were then rinsed with TBPS and blocked with PBS containing BSA. The plate was rinsed with TPBS, then incubated with eight serial dilutions of contrajervin starting at a high concentration of 50 μg/ml for 30 min before adding 100 ng/well of Eu^{+3}-labeled CV-N. The plate was again rinsed with TPBS before the addition of the DELFIA® enhancer solution.

To determine whether the peptide may be an inhibitor of HIV binding and/or fusion, the effect of delayed addition of contrajervin to HIV-1_{RF} infected cells was studied using a previously described method [10]. An XTT assay was carried out with contrajervin added to quadruplicate wells at 0, 2, 4, 6, 8, 12, and 24 h after addition of the virus. After seven days of incubation, the cultures were observed microscopically for the presence of cytopathic effects (syncytia) and then stained with XTT.

Contrajervin was subjected to a chitin-binding assay as previously described [11]. The peptide and a wheat germ agglutinin (WGA) control were applied to chitin microcolumns and eluted sequentially with PBS and 0.1 M acetic acid. Samples were desalted and analyzed by reversed-phase HPLC (Dynamax C_4, 300 Å) using a gradient of 0–60% acetonitrile in water over 45 min at a flow rate of 3 ml/min.

3. Results and discussion

3.1. Peptide characterization and structure

The aqueous extract of *D. contrajerva* was subjected to reversed-phase vacuum-liquid chromatography, ethanol precipitation, size exclusion filtration and reversed-phase HPLC to yield 7.5 mg of contrajervin. ESI-MS analysis of the native protein provided a molecular weight of 5103.8 Da, which is in agreement with the theoretical calculated molecular weight of 5104.5. The native protein was reduced and alkylated, yielding an ESI-MS weight of 5952.7 Da, which was consistent with the presence of eight disulfide-linked cysteines. Direct sequencing of the PEC derivative provided a primary sequence of 33 amino acids. The PEC derivative was subjected to endoproteinase Asp-N digestion to yield three fragments, two of which provided overlapping confirmation of the sequence, while the third fragment extended the primary sequence to 43 amino acids. Trypsin digestion of the PEC derivative also resulted in three fragments, two providing overlapping confirmations, with the third fragment extending the sequence to a final count of 46 amino acids (Fig. 1(A)).

The aqueous extract of *T. obovoidea* was subjected to vacuum-liquid chromatography, size exclusion column chromatography and reversed-phase HPLC to yield 0.6 mg of treculavirin. ESI-MS analysis of pure treculavirin gave two molecular ions of 10 369 and 10 427 Da, indicating that it exhibited heterogeneity. Reduction of the native protein yielded two fragments with molecular weights of 5194.6 and 5252.0, suggesting the presence of two monomers covalently linked via one or more disulfide bonds. ESI-MS of the reduced and

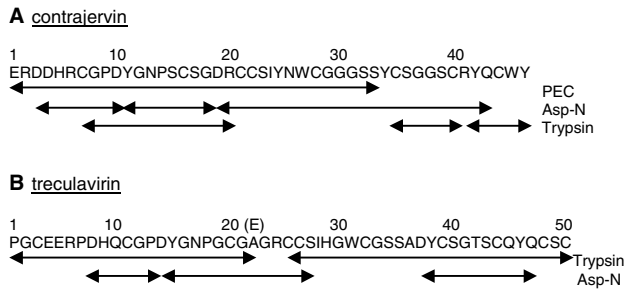

Fig. 1. (A) Primary amino acid sequence of native contrajervin obtained after digestion with endoproteinases Asp-N and trypsin. Arrows indicate sequenced peptide fragments. (B): Primary amino acid sequence of one of the monomers of treculavirin obtained after digestion of the PEC derivative with endoproteinases, trypsin and Asp-N. Parentheses above residue 22 indicates interchanging alanine and glutamic acid residues.

alkylated peptide gave molecular weights of 6243.7 and 6302.3, indicating that each monomer contained 10 cysteines.

One of the PEC monomers was subjected to trypsin digestion which yielded two fragments. Edman degradation of the first fragment provided a sequence of 22 amino acids and defined the microheterogeneity, with the monomer existing with interchanging alanine and glutamic acids at residue 22. The second fragment provided a sequence of 23 amino acids. Digestion of the second PEC monomer gave similar results. The PEC monomers were then digested with endoproteinase Asp-N to yield three fragments, all of which overlapped the trypsin fragments and provided the final monomer sequence of 50 amino acids (Fig. 1(B)). The calculated theoretical molecular weights of the native protein, 10371.3 and 10429.3 Da, supported the deduced 100 amino acid primary sequence.

3.2. Biological activity

Contrajervin and treculavirin showed 50% displacement of labeled CV-N from gp41 at a concentration of 0.59 and 0.20 μM, respectively. Contrajervin also showed 50% displacement of CV-N from native gp120 obtained from a T-tropic virus ($gp120_{IIIB}$) and recombinant gp120 derived from an M-tropic virus ($gp120_{BaL}$) at a concentration of 0.20 μM. In a whole cell assay, contrajervin inhibited the cytopathic effects of HIV-1_{RF} infection in a human T-lymphoblastoid cell line (CEM-SS), with an EC_{50} value of 1.0 μM and an IC_{50} value of >4.9 μM. Treculavirin showed anti-HIV activity at an EC_{50} value of <0.02 μM and an IC_{50} value of > 2.5 μM. For both compounds, an in vitro "therapeutic index" was not obtained as no toxicity was observed at the highest dose tested. Attempts to identify an IC_{50} were not pursued due to the limited amounts of material available.

A time course experiment was performed in which aliquots of contrajervin were added at various times following virus addition to target cells. A concentration of contrajervin previously shown in dose–response studies to be 100% protective was utilized. Complete protection was confirmed when the compound was added immediately following virus addition. Surprisingly, delayed addition of contrajervin, even up to eleven hours after virus addition, resulted in nearly complete protection of the cells. These data suggest that the compound is acting by a second mechanism/target which contributes to its cytoprotective activity.

Microcolumns packed with chitin were loaded with contrajervin or WGA as a positive control, and eluted at neutral and low pH. Reversed-phase HPLC of the resulting fractions showed that contrajervin was present in the initial fraction that was recycled through the column, while WGA was present only in the fraction eluted with low pH buffer (data not shown). These results indicated that contrajervin does not exhibit strong binding affinity toward chitin.

3.3. Sequence similarity

The Moraceae family is known to contain lectins, which have anti-HIV activity of micromolar concentrations [7]. As part of the original HTS screen to prioritize extracts, active hits were tested to see if their activity was due to monosaccharide-specific lectin-like properties. The original binding assay was run with the extracts in the presence of mannose and glcNAc (the only monosaccharides present on the recombinant gp41 used in this assay), to see if the sugars decreased the ability of the extracts to compete with CV-N binding to gp41. Neither of the extracts displayed reduced binding to gp41 in the presence of the sugars, therefore they did not appear to contain monosaccharide-specific lectins (Fig. 2).

The monomers of treculavirin share a high sequence similarity (68%) with contrajervin (Fig. 3). A search of the GenBank non-redundant database, BLASTP [12], for identification of protein sequence homology indicated strong

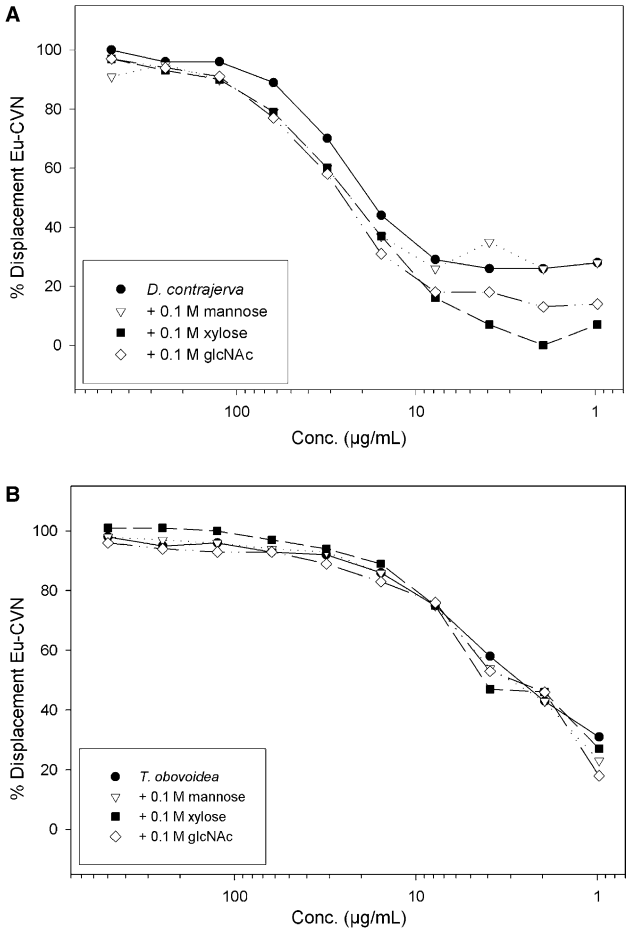

Fig. 2. Effect of addition of simple sugars on the ability of *D. contrajerva* and *T. obovoidea* to compete with CV-N binding to gp41.

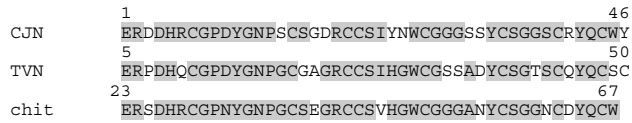

Fig. 3. Aligned amino acid sequences of contrajervin (CJN) and treculavirin (TVN) and the similar region of a deduced sequence of a class I chitinase (CHI) from *Ficus carica*. Numbers to the top of the sequences indicate amino acid residue numbers. Sequence identities are indicated by gray shading.

homology for both peptides (contrajervin, 73%; treculavirin, 72%) to a class I chitinase from *Ficus carica* (Moraceae) whose sequence was deduced from isolated cDNA [13] (Fig. 3). Contrajervin and treculavirin also showed homology (66% and 65%, respectively) to a hevein-like protein from *Sambucus nigra* (Caprifoliaceae) [14] as well as a precursor of *Urtica diocia* (Moraceae) agglutinin (UDA) isolectin (64% and 66%, respectively) [15]. The amino acid sequences of both peptides are cysteine/glycine rich which is characteristic of the chitin binding domain of chitin binding proteins. Because of the primary sequence similarity of the peptides to chitin binding proteins, we investigated the ability of contrajervin to bind to a chitin substrate. Assay results indicated that although contrajervin contains a primary structural motif similar to chitin-binding proteins, it does not exhibit strong binding affinity toward chitin.

The gene encoding the class I chitinase from *Ficus carica* has been characterized as a stress-related gene [13] and UDA has been suggested to play a role in plant defense [16]. Although the ecological function of the secondary metabolites from *Dorstenia* and *Treculia* is unknown, they may also play a role in plant defense.

Acknowledgements: We thank Gordon Cragg (Natural Products Branch) for coordinating plant collections, Tom McCloud (Natural Products Extraction Laboratory) for extractions, and Jennifer Wilson for HIV testing. This project has been funded in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract no. NO1-CO-12400. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

References

[1] deRosny, E., Vassell, R., Wingfield, P.T., Wild, C.T. and Weiss, C.D. (2001) J. Virol. 75, 8859–8863.
[2] Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Mori, T., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M., Currens, M.J., Cardellina 2nd, J.H., Buckheit Jr., R.W., Nara, P.L., Pannell, L.K., Sowder 2nd, R.C. and Henderson, L.E. (1997) Antimicrob. Agents Chemother. 41, 1521–1530.
[3] Shenoy, S.R., O'Keefe, B.R., Bolmstedt, A.J., Cartner, L.K. and Boyd, M.R. (2001) J. Pharmacol. Exp. Ther. 297, 704–710.
[4] Beutler, J.A., McMahon, J.B., Johnson, T.R., O'Keefe, B.R., Buzzell, R.A., Robbins, D., Gardella, R., Wilson, J. and Boyd, M.R. (2002) J. Biomol. Screen 7, 105–110.
[5] Abegaz, B.M., Ngadjui, B.T., Dongo, E. and Bezabih, M.-T. (2000) Curr. Org. Chem. 4, 1079–1090.
[6] Bokesch, H.R., Pannell, L.K., Cochran, P.K., Sowder 2nd, R.C., McKee, T.C. and Boyd, M.R. (2001) J. Nat. Prod. 64, 249–250.
[7] Charan, R.D., Munro, M.H., O'Keefe, B.R., Sowder, R., McKee, T.C., Currens, M.J., Pannell, L.K. and Boyd, M.R. (2000) J. Nat. Prod. 63, 1170–1174.
[8] Weissenhorn, W., Wharton, S.A., Calder, L.J., Earl, P.L., Moss, B., Aliprandis, E., Skehel, J.J. and Wiley, D.C. (1996) EMBO J. 15, 1507–1514.
[9] Gulakowski, R.J., McMahon, J.B., Staley, P.G., Moran, R.A. and Boyd, M.R. (1991) J. Virol. Methods 33, 87–100.
[10] O'Keefe, B.R., Beutler, J.A., Cardellina 2nd, J.H., Gulakowski, R.J., Krepps, B.L., McMahon, J.B., Sowder 2nd, R.C., Henderson, L.E., Pannell, L.K., Pomponi, S.A. and Boyd, M.R. (1997) Eur. J. Biochem. 245, 47–53.
[11] Bokesch, H.R., O'Keefe, B.R., McKee, T.C., Pannell, L.K., Patterson, G.M., Gardella, R.S., Sowder 2nd, R.C., Turpin, J., Watson, K., Buckheit Jr., R.W. and Boyd, M.R. (2003) Biochemistry 42, 2578–2584.
[12] Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Nucleic Acids Res. 25, 3389–3402.
[13] Kim, J.S., Kim, Y.O., Ryu, H.J., Kwak, Y.S., Lee, J.Y. and Kang, H. (2003) Plant Cell Physiol. 44, 412–414.
[14] Van Damme, E.J.M., Barre, A., Rouge, P., Van Leuven, F. and Peumans, W.J. (1997) J. Biol. Chem. 272, 8353–8360.
[15] Lerner, D.R. and Raikhel, N.V. (1992) J. Biol. Chem. 267, 11085–11091.
[16] Peumans, W.J. and Van Damme, E.J. (1995) Plant Physiol. 109, 347–352.