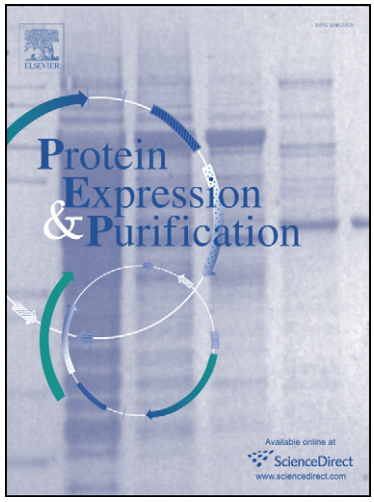

High-yield expression and purification of isotopically labeled Norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies

Hanna Berkner, Julia Engelhorn, David K. Liscombe, Kristian Schweimer, Birgitta M. Wöhrl, Peter J. Facchini, Paul Rösch, Irena Matečko

PII: S1046-5928(07)00185-4
DOI: 10.1016/j.pep.2007.07.010
Reference: YPREP 3123

To appear in: *Protein Expression and Purification*

Received Date: 12 June 2007
Revised Date: 16 July 2007
Accepted Date: 23 July 2007

Please cite this article as: H. Berkner, J. Engelhorn, D.K. Liscombe, K. Schweimer, B.M. Wöhrl, P.J. Facchini, P. Rösch, I. Matečko, High-yield expression and purification of isotopically labeled Norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies, *Protein Expression and Purification* (2007), doi: 10.1016/j.pep.2007.07.010

High-yield expression and purification of isotopically labeled Norcoclaurine synthase, a Bet v 1-homologous enzyme, from *Thalictrum flavum* for NMR studies

Hanna Berkner[a], Julia Engelhorn[a], David K. Liscombe[b], Kristian Schweimer[a], Birgitta M. Wöhrl[a], Peter J. Facchini[b], Paul Rösch[a], and Irena Matečko[a*]

[a] Department of Biopolymers and Research Center for Bio-Macromolecules, Universität Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany

[b] Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 1N4, Canada

[*]Corresponding author:

Irena Matečko, Universität Bayreuth, Universitätsstraße 30, 95440 Bayreuth, Germany

E-mail: irena.matecko@uni-bayreuth.de

Phone: +49-921-553868

Fax: +49-921-553544

ABSTRACT

The enzyme Norcoclaurine synthase (NCS) found in the common meadow rue, *Thalictrum flavum*, and other plants shows sequence homology to members of the class 10 of pathogenesis related (PR 10) proteins that contains allergens such as the major birch pollen allergen Bet v 1, the major cherry allergen Pru av 1, and the major apple allergen Mal d 1. The enzyme is involved in the plant's secondary metabolism and is required for the production of bioactive secondary metabolites like morphine. Whereas the physiological function of PR 10 class allergens is still unknown, NCS activity has been studied in detail. Investigation of the structural properties of NCS by NMR spectroscopy can thus not only provide new information concerning the reaction mechanism of the enzyme, but is also expected to help clarify the long standing and heavily debated question on the physiological function as well as the reasons for the allergenic potential of members of this protein family. As the first important step towards the three-dimensional solution structure, we optimized expression of recombinant NCS in *E. coli* and established an efficient purification protocol yielding high amounts of pure isotopically labeled active enzyme. The identity of NCS was confirmed by electrospray ionization mass spectrometry, and activity of the purified enzyme was determined by an assay detecting the radiolabeled reaction product. Spectroscopic analysis by NMR spectroscopy showed that the protein was properly folded with well defined tertiary structure.

Abbreviations: NCS, Norcoclaurine synthase; NMR, nuclear magnetic resonance; PR 10, class 10 of pathogenesis related proteins; 4-HPAA, 4-hydroxyphenylacetaldehyde; PCR, polymerase chain reaction; His_6-tag, hexahistidine-tag; LB medium, Luria-Bertani medium; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; OD_{600}, optical density at 600 nm; IPTG, isopropyl-β-D-thiogalactoside; IMAC, immobilized metal ion affinity chromatography; FPLC, fast performance liquid chromatography; ESI-MS, electrospray ionization mass spectrometry; TLC, thin-layer chromatography; DTT, dithiothreitol; HSQC, heteronuclear single quantum coherence; MALDI-TOF-MS, matrix assisted laser desorption ionization time-of-flight mass spectrometry.

Key words: Norcoclaurine synthase, PR-10 protein family, protein purification, isotope labeling, NMR

INTRODUCTION

The enzyme Norcoclaurine synthase (NCS) found among other plants in opium poppy (*Papaver somniferum*) and in the common meadow rue (*Thalictrum flavum*) shows up to 38 % sequence identity with members of class 10 of pathogenesis related (PR 10) proteins which includes the family of Bet v 1 homologous allergens [1-3]. PR proteins represent a group of plant proteins divided into 14 classes [4]. Their main common characteristic is the induction of their expression under stress conditions [5].

So far, little is known about the physiological function of members of the Bet v 1 family in plants, but a growing number of people, at present about 25 % of the total population of the industrialized countries [6-8], experiences the effect these proteins can have on the human immune system: They cause type I allergies involving weaker symptoms like allergic rhinitis as well as asthma and sometimes even severe anaphylactic reactions. The molecular mechanism leading to these symptoms has been investigated in detail. The dimerization of allergen-specific type E immunoglobulins provoked by binding of two allergen epitopes represents the crucial step in a signal cascade that finally leads to the release of mediators like histamine and serotonin [9-12]. A promising approach to develop a therapy fighting this type of allergy is the hyposensibilisation of patients using so-called hypoallergens. These artificially designed proteins have the ability to provoke an immune response involving type G instead of type E immunoglobulins and thus to prevent the subsequent signal cascade and the resulting symptoms [13]. An essential prerequisite for the construction of such a hypoallergen is detailed information about the structure of the naturally occurring allergen.

Three-dimensional structures of a number of allergens are already known, among them the structure of the major birch pollen allergen Bet v 1 [14, 15], the cherry allergen Pru av 1 [16], the minor birch pollen allergen Bet v 4 [17], and others [18]. Based on their amino acid

sequence homology to Bet v 1, several other pollen allergens and also different food allergens like Pru av 1 or Mal d 1 are grouped into the family of Bet v 1-homologous allergens. The members of this family share a number of common structural features. Consisting of 153-160 amino acids, their sequence contains a highly conserved P-loop motive with a G-X-G-G-X-G-T consensus sequence [19, 20]. Whereas the P-loop is also present in NCS, this enzyme comprising 210 amino acids contains extensions at both the amino- and the carboxy-terminus as compared to the members of the Bet v 1 family. The determination of the three-dimensional structure of NCS which is unique in the Bet v 1 family as its function in the plant is known can thus shed some light on the evolutionary connections between enzymatically active members of the PR 10 class and proteins such as Bet v 1 and Pru av 1 for which an enzymatic activity has not been established yet. Such studies are also intended to help answer the therapeutically relevant question about the structural basis of allergic reactions.

Additionally, structural data can also provide new information about the reaction mechanism of the enzyme. NCS catalyzes the first step in benzylisoquinoline alkaloid synthesis which consists of the condensation of dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) to norcoclaurine [21]. Benzylisoquinoline alkaloids form an interesting group of plant secondary metabolites because many of them show pharmacological activity, for example the analgesics morphine and codeine, the muscle relexant papaverine, and the antibiotic sanguinarine [21]. The common meadow rue, *Thalictrum flavum*, produces the antimicrobial berberine and magnoflorine which has been suggested to possess anti-HIV activity [22, 23]. In 2004, the gene coding for NCS was isolated from a *Thalictrum flavum* cDNA library, the gene was cloned and the protein was expressed in *E. coli*. Enzyme activity assays were carried out using the crude extract obtained after cell lysis [1]. With the aim to determine the NMR solution structure of this in many respects interesting protein, we optimized the expression of NCS in

E. coli and established a purification protocol yielding sufficient amounts of properly folded and highly pure isotopically labeled enzyme.

MATERIALS AND METHODS

Plasmid construction

The plasmid pET29b-*Δ19NCS* [1] containing the cDNA of *Thalictrum flavum* Norcoclaurine synthase coding for amino acids 20-210 cloned with *Bam*HI/*Xho*I was used as a template for construction of pET29b-*Δ29NCS* coding for amino acids 30-210. The corresponding gene fragment was amplified by PCR using a forward primer including an *Nde*I site (underlined; 5'-AGGTAGGCATATGCTGCACCACCAGGGCATAATAAACC-3') and a reverse primer containing an *Xho*I site (underlined; 3'- GTGGTGCTCGAGGACTGTTATTATTGC-5'). The temperature cycling for 35 cycles was set up as follows: 94 °C for 30 s, 54 °C for 30 s and 72 °C for 1 min. The amplified fragment was inserted into expression vector pET29b (Novagen, Madison, USA) using *Nde*I/*Xho*I restriction sites, and successful cloning was confirmed by DNA sequencing. In contrast to plasmid pET29b-*Δ19NCS* coding for an amino-terminal S-tag fusion with a carboxy-terminal hexahistidine-tag (His_6-tag), expression of pET29b-*Δ29NCS* yields only C-terminally His_6-tagged Δ29NCS.

Expression in LB medium

Plasmid pET29b-Δ*19NCS* was either expressed in *E. coli* ER2566pLysS [1] or transformed into *E. coli* BL21(*DE3*) and *E. coli* Rosetta(*DE3*) host cells. For expression of Δ29NCS, plasmid pET29b-Δ*29NCS* was transformed into *E. coli* Rosetta(*DE3*) host cells. After

transformation host cells were grown on LB agar containing the appropriate antibiotics (ER2566pLysS/pET29b-Δ*19NCS* and BL21(*DE3*)/pET29b-Δ*19NCS*: 20 µg/ml kanamycin; Rosetta(*DE3*)/pET29b-Δ*29NCS*: 20 µg/ml kanamycin, 34 µg/ml chloramphenicol) for selection at 37 °C overnight. For induction tests, 10-20 mL LB medium containing the respective antibiotics was inoculated with the expression strain and grown overnight at 37 °C. This starter culture was used to inoculate 100 mL of the same medium to an OD_{600} of 0.2. Cells were then incubated at 37 °C under continuous shaking. At an OD_{600} of 0.6, expression was induced by addition of 1 mM IPTG followed by continuous incubation at 37 °C for 4-6 hours. To monitor the expression, aliquots containing equal amounts of cells were taken from the culture every hour and applied to a 15 % SDS-polyacrylamide gel. For preparative expression, the same procedure was carried out using final culture volumes up to 2.5 L. Cells were harvested 4-6 h after induction by centrifugation at 5000 x g for 15 min at 4 °C and the cell pellet was stored at -20 °C.

Preparative expression in M9 minimal medium

For expression of Δ19NCS and Δ29NCS in minimal medium, 100 mL of M9 [24] containing 2 mM $MgSO_4$, 10 µM Fe(III)citrate, 0.1 mM $CaCl_2$, 4 g/L glucose, trace element solution TS2 [25], MEM vitamin solution (Gibco, Invitrogen, Karlsruhe), 20 µg/mL kanamycin, 34 µg/mL chloramphenicol and 1.5 g/L $(NH_4)_2SO_4$ as the only source of nitrogen were inoculated with *E. coli* Rosetta(*DE3*)/pET29b-Δ29NCS and grown overnight at 37 °C. 2 L of the same medium were inoculated with the overnight culture to an OD_{600} of 0.2 and incubated under shaking at 37 °C. Induction was carried out at an OD_{600} of 0.6 by adding 1 mM IPTG and cells were harvested after 5-6 h of expression and stored as described above. Expression of ^{15}N-labeled protein was achieved replacing $(NH_4)_2SO_4$ by $(^{15}NH_4)_2SO_4$.

Purification of Δ19NCS and Δ29NCS by affinity chromatography

Frozen cells were resuspended in binding buffer (20 mM sodium phosphate buffer pH 7.8, 500 mM NaCl, 20 mM imidazole; Δ19NCS: 5 mL/g pellet; Δ29NCS: 7 mL/g pellet) containing lysozyme, DNase, and one protease inhibitor cocktail tablet (Complete, EDTA-free, Roche) and stirred on ice for 30 min followed by sonication on ice. The lysate was centrifuged at 19100 x g for 45 min at 4 °C. The pellet was resuspended in binding buffer, stirred on ice, sonicated, and centrifuged again as described above, and supernatants were pooled. Purification of the soluble his-tagged protein was performed by immobilized metal ion affinity chromatography (IMAC) on an ÄKTA *purifier* 10-FPLC system (Amersham Biosciences, Freiburg, Germany). The supernatant was loaded on a 5 mL HisTrap column (Amersham Biosciences, Freiburg, Germany) pre-equilibrated with binding buffer. After a washing step with binding buffer, the bound protein was eluted using a stepwise gradient including 5, 10, 15, 20, 25, 30, 50, and 100 % elution buffer (20 mM sodium phosphate buffer pH 7.8, 500 mM NaCl, 300 mM imidazole). The collected 5 mL fractions were analyzed by SDS-PAGE. In case of Δ29NCS the fractions containing the protein were dialyzed as first step against 5 L 20 mM sodium phosphate buffer pH 7.8 containing 100 mM NaCl and then twice against 5 L millipore water, every step at least for 3 h at 4 °C. After shock-freezing in liquid nitrogen, the protein was lyophilized and stored at 4 °C. In case of Δ19NCS, the S-tag-fusion was subjected to thrombin cleavage.

Thrombin cleavage of Δ19NCS

After affinity chromatography, the fractions containing Δ19NCS were pooled and dialyzed against 5 L of binding buffer without imidazole for at least 3 h at 4 °C. Thrombin cleavage removing the amino-terminal S-tag from the Δ19NCS-fusion was carried out at room

temperature overnight during dialysis against 5 L fresh binding buffer without imidazole by adding 0.2 U thrombin (Novagen, Madison, USA) per mg protein inside the dialysis membrane (Spectra/Por, MWCO 3500, ROTH, Karlsruhe, Germany). Successful cleavage was confirmed by SDS-PAGE. To remove the cleaved S-tag, the sample was again subjected to nickel ion affinity chromatography under the conditions described above. Subsequent dialysis and lyophilization was also carried out as described above.

Gel electrophoresis

Denaturing SDS-PAGE (15 % w/v polyacrylamide) was performed as described by Lämmli [26]. Samples were denatured by dilution with Roti-Load 1 (ROTH, Karlsruhe, Germany) and incubation at 95 °C for 5-10 min. Proteins were visualized by Coomassie blue staining.

Protein Mass Fingerprint and N-terminal sequencing

In order to identify the fragments of Δ19NCS visible on the SDS-polyacrylamide gel before and after thrombin cleavage, gel pieces with 1 mm diameter were cut from the corresponding band and subjected to protein mass fingerprinting including tryptic digestion, determination of the resulting peptide masses by MALDI-TOF-MS and analysis of the covered protein sequence using the MASCOT program. N-terminal sequencing of the Δ19NCS degradation product was carried out by Edman degradation on an ABI Procise 491-Sequenzer. Both methods were performed by ZMMK Koeln, Zentrale Bioanalytik.

Electrospray Ionization Mass Spectrometry

Mass spectra were recorded on a LCT Micromass time-of-flight electrospray mass spectrometer from Waters with an electrospray interface, in positive ion-mode. The spray tip potential was 2000 V, sample cone potential was 50 V, and desolvation temperature 150 °C. The software used for recording and processing the data was Masslynx v3.4 (Waters). For data analysis MagTran 1.02 programme was used [27].

Stock solution was prepared by dissolving lyophilized protein in 1 mM ammonium acetate buffer pH 7.0. Concentration of the final stock solution was 1 mM. For the ESI-MS experiments 100 µM protein solution in 1 mM ammonium acetate, 5% ethanol and 0.1% acetic acid was used.

Enzyme activity assay

Enzymatic activity of purified Δ29NCS was tested as described [1]. Lyophilized protein was resuspended in assay buffer (100 mM Tris-HCl, pH 7.0; 12 mM β-mercaptoethanol). Protein concentration was determined according to the Bradford method [28]. Each assay contained 15 µg of pure Δ29NCS, 311 pmol (187 Bq) [8-^{14}C] dopamine and 25 nmol 4-HPAA in a total volume of 30 µl assay buffer. The negative control contained protein solution boiled for 10 min prior to assay setup. The assays were incubated for 1.5 h at 37 °C. Subsequently, the whole assay mixture was subjected to thin-layer chromatography (TLC) and radiolabeled compounds were visualized by autoradiography.

NMR spectroscopy

NMR samples were prepared by dissolving the required amount of lyophilized ^{15}N-labeled Δ29NCS in 20 mM sodium phospate buffer pH 7.0, 2.5 mM DTT, 0.04 % NaN_3 and 10 % D_2O. Final sample concentration was determined by measuring absorption of the protein solution at 280 nm. Two-dimensional ^{15}N-^{1}H HSQC NMR spectra were recorded according to the FHSQC-scheme [29] on a Bruker Avance 800 MHz NMR spectrometer equipped with a triple-resonance cryoprobe with pulsed field-gradient capabilities.

RESULTS AND DISCUSSION

Optimization of Δ19NCS expression in E. coli

Three different *E. coli* host strains were tested for expression of Δ19NCS (25.8 kDa) comprising an amino-terminal S-tag followed by a thrombin cleavage site and a carboxy-terminal His_6-tag. *E. coli* ER2566pLysS was already used as host strain for Δ19NCS expression [1] in order to investigate the enzymatic activity of recombinant NCS present in the cell lysate. Whereas only small amounts of enzyme are necessary to test its activity, highly concentrated and pure protein is needed for structural studies by NMR spectroscopy. For this reason, expression of Δ19NCS using plasmid pET29b-*Δ19NCS* had to be optimized. In addition to *E. coli* ER2566pLysS, *E. coli* BL21(*DE3*) and *E. coli* Rosetta(*DE3*) were tested as alternative expression hosts (Fig. 1). Only low expression rates and, as consequence, no accumulation of significant amounts of protein over expression time could be observed for *E. coli* ER 2566pLysS (Fig. 1A) and *E. coli* BL21(*DE3*) (Fig. 1B). However, expression in *E. coli* Rosetta(*DE3*) yielded high amounts of Δ19NCS augmenting over expression time (Fig. 1C). As the sequence coding for Δ19NCS is identical to that of the NCS gene from

Thalictrum flavum, differences in codon usage between the originating organism and the expression host might lead to problems during translation of the recombinant protein due to the lack of certain tRNAs [30]. Analyzing the sequence coding for Δ19NCS including both tags, a number of rare codons can be identified, especially coding for leucine and isoleucine. The special feature of *E. coli* Rosetta(*DE3*) is its ability to provide the lacking tRNAs corresponding to these *E. coli* rare codons and thus to facilitate expression of heterologous genes [31]. In case of Δ19NCS, the strategy to use this strain as expression host significantly improved the expression yield compared to the other strains tested. Concerning the solubility of Δ19NCS expressed in *E. coli* Rosetta(*DE3*), more than 50 % of the recombinant protein was present in the supernatant after cell lysis (Fig. 2, lane 2 and 4).

Purification of Δ19NCS

Purification of soluble Δ19NCS with C-terminal His_6-tag was carried out using nickel ion affinity chromatography. The protein eluted at relatively low imidazole concentrations of 45-60 mM, probably due to bad accessibility of the affinity tag in the folded protein. The corresponding fractions and the flow-through were analyzed by SDS-PAGE (Fig. 2, lane 6-9). Lanes 7-9 show two intense bands corresponding to molecular masses of about 30 and 24 kDa, respectively. The corresponding proteins are not present in the flow-through (lane 6). The upper band runs at the same height as the over-expressed protein detected in the cells after induction (lane 1), whereas the lower band is almost invisible at this state, but appears in lanes 2 and 4 analyzing the protein content of the supernatant after cell lysis. As for pure Δ19NCS a single band corresponding to a molecular mass of 25.8 kDa was expected, the two bands were subjected to protein mass fingerprinting in order to identify Δ19NCS. The analysis showed that the upper band contained full-length Δ19NCS, while the lower band

corresponded to a specific degradation product ΔΔ19NCS. The exact N-terminal sequence of the degradation product could be determined by Edman degradation as LHHQG, revealing that the S-tag and 10 additional N-terminal amino acids of Δ19NCS had been cleaved off. In consequence, ΔΔ19NCS was termed Δ29NCS. Though all lysis and purification steps were carried out thoroughly on ice or with ice-cold buffers and in presence of protease inhibitors, specific degradation of Δ19NCS could not be avoided.

In order to separate the S-tag from full-length Δ19NCS, thrombin cleavage was performed during dialysis of the dilute protein solution against binding buffer overnight, which turned out to be as efficient as using concentrated protein solution in the special thrombin cleavage buffer containing $CaCl_2$ recommended by Novagen. Obviously, no protease inhibitor could be used during this step. SDS-PAGE analysis indicates nearly 100 % efficiency of thrombin cleavage because no S-tagged Δ19NCS is detectable after cleavage, but an additional band corresponding to the correct 23 kDa thrombin cleavage product as confirmed by protein mass fingerprinting appears (Fig. 3). As the intensity of the lower band corresponding to Δ29NCS increases slightly after cleavage, degradation as described above still takes place during cleavage, but slowly. During the second nickel ion affinity chromatography step performed in order to remove the cleaved S-tag, Δ19NCS and Δ29NCS showed slightly different elution behavior, that is Δ29NCS started to elute at lower imidazole concentrations compared to Δ19NCS. In consequence, the two NCS species could be at least partially separated.

During sample preparation for NMR spectroscopic analysis of the folding state of NCS, samples consisting in major parts of Δ29NCS turned out to be more stable against macroscopic aggregation than samples containing an excess of Δ19NCS. NMR spectra of both samples containing mainly Δ19NCS and samples with an excess of Δ29NCS indicated existence of a well defined tertiary structure. Thus, we decided to construct an expression vector coding for Δ29NCS lacking the N-terminal S-tag but including a C-terminal His_6-tag

with two central aims. First, we wanted to shorten the purification procedure by avoiding thrombin cleavage of the S-tag not needed for our purposes and, secondly, we wanted to obtain only one defined NCS species suitable for structural NMR spectroscopic studies as well as for quantitative substrate binding studies.

Expression and purification of Δ29NCS

The plasmid pET29b-*Δ29NCS* coding for amino acids 30-210 of NCS linked with a C-terminal His_6-tag was constructed as described above, and the expected sequence was confirmed by DNA sequencing. Prior to preparative expression of labeled protein in M9 medium, expression tests in *E. coli* Rosetta(*DE3*) in LB medium were performed. The expression rate of Δ29NCS was as high as for Δ19NCS in the same host strain. A solubility test comparing levels of soluble NCS after 4, 6, and 20 h induction time showed that an induction period of 6 h is suitable for maximal yields of soluble protein. Preparative expression of Δ29NCS in minimal medium is documented by SDS-PAGE in figure 4A. The expression level in M9 medium was as high as in LB medium. Around 3 g cells per liter M9 medium (wet weight) could be harvested after 6 h expression. More than 50 % of the overexpressed Δ29NCS was found in the soluble fraction after cell lysis (Fig. 4B, lane 1). IMAC was used as single purification step. Highly pure Δ29NCS eluted at imidazole concentrations from 45 to 60 mM (Fig. 4B, lane 6 and 7). Although the running height of a minor impurity corresponding to a molecular weight of about 45 kDa could be interpreted as a hint on dimerization of Δ29NCS, this impurity has also been observed after purification of other proteins using a HisTrap column, and thus it is more likely an abundant *E. coli* protein with weak affinity to the nickel column. As lyophilized protein is easy to store and resistant against degradation, Δ29NCS was dialyzed against H_2O, frozen in liquid nitrogen,

lyophilized, and stored at 4 °C until further use. Purification yields ranged from 40 to 50 mg Δ29NCS per liter M9 medium, corresponding to about 15 mg protein per gram cells.

Characterization of Δ29NCS by ESI-MS

Mass spectrometry was used to check identity and purity of recombinant Δ29NCS. ESI-MS spectra of recombinant Δ29NCS (Figure 5) indicate the presence of highly pure species in the sample. Molecular mass of Δ29NCS was determined to be 21179.5 Da which is in agreement with that calculated from the amino acid sequence (theoretical molecular mass: 21181.3 Da).

Enzymatic activity of Δ29NCS

To investigate whether the purified recombinant protein shows enzymatic activity, non-labeled 4-HPAA and radiolabeled dopamine was incubated in presence of native Δ29NCS, and, as a negative control, in presence of thermally denatured Δ29NCS. Components of the reaction mixture were separated by thin-layer chromatography and radiolabeled reaction products were detected by autoradiography [1] (Fig. 6). In addition to the radiolabeled dopamine visible in the presence of native (lane 1) and denatured (lane 2) Δ29NCS , a second radiolabeled reaction product is detectable only in the reaction mixture containing native Δ29NCS. The corresponding spot on lane 1 marked with an arrow shows an R_f of 0.60 which is identical to the expected migration distance of the reaction product (S)-norcoclaurine [21], showing that the removal of 10 aminoterminal amino acids from the enzymatically active Δ19NCS did not affect enzyme activity. The activity of the recombinant enzyme is further indication for a largely intact native conformation after purification.

NMR spectroscopy

Proper folding of the purified recombinant Δ29NCS was examined by measuring an ^{1}H-^{15}N HSQC spectrum of uniformly ^{15}N-labeled protein. The protein consists of 190 amino acids (Fig. 7A) including nine prolines that do not contribute signals to the ^{1}H-^{15}N HSQC spectrum. In consequence, together with the side chain amide protons, about 200 cross peaks would be expected to show up in the spectrum (Fig. 7B), a number that compares well to the actual number of 194 ^{1}H-^{15}N cross correlation peaks detected. Overlapping signals in the spectral region around 8 ppm in the proton dimension and around 121 ppm in the nitrogen dimension indicate the presence of unstructured areas in the protein, but the high dispersion of signals in both dimensions indicates a well defined tertiary structure of Δ29NCS. Concluding, the expression and purification protocol described for the newly designed NCS construct Δ29NCS is suitable to yield enough properly folded and enzymatically active protein to allow determination of the three-dimensional solution structure by NMR spectroscopy which is currently in progress in our department.

ACKNOWLEDGEMENTS

We thank Katrin Weiss for professional technical assistance. This work was supported by grant DFG RO 617/11 – 4 from Deutsche Forschungsgemeinschaft.

REFERENCES

[1] N. Samanani, D.K. Liscombe, P.J. Facchini, *Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis*, Plant J. 40 (2004) 302-13.

[2] D.K. Liscombe, B.P. MacLeod, N. Loukanina, O.I. Nandi, P.J. Facchini, *Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms*, Phytochemistry 66 (2005) 2501-20.

[3] K. Hoffmann-Sommergruber, *Plant allergens and pathogenesis-related proteins. What do they have in common?*, Int. Arch. Allergy Immunol. 122 (2000) 155-66.

[4] H. Breiteneder, C. Ebner, *Molecular and biochemical classification of plant-derived food allergens*, J. Allergy. Clin. Immunol. 106 (2000) 27-36.

[5] C. Ebner, K. Hoffmann-Sommergruber, H. Breiteneder, *Plant food allergens homologous to pathogenesis-related proteins*, Allergy. 56 Suppl 67 (2001) 43-4.

[6] A.B. Kay, *Allergy and allergic diseases – First of two parts*, N. Engl. J. Med. 344 (2001) 30-37.

[7] A.B. Kay, *Allergy and allergic diseases – Second of two parts*, N. Engl. J. Med. 344 (2001) 109-113.

[8] M. Wills-Karp, J. Santeliz, C.L. Karp, *The germless theory of allergic disease: revisiting the hygiene hypothesis*, Nat. Immunol. 1 (2001) 69-75.

[9] D. Venarske, R.D. deShazo, *Molecular mechanisms of allergic disease*, South Med. J. 96 (2003) 1049-54.

[10] H.J. Gould, B.J. Sutton, A.J. Beavil, R.L. Beavil, N. McCloskey, H.A. Coker, D. Fear, L. Smurthwaite, *The biology of IGE and the basis of allergic disease*, Annu. Rev. Immunol. 21 (2003) 579-628.

[11] S.M. Miescher, M. Vogel, *Molecular aspects of allergy*, Mol. Aspects Med. 23 (2002) 413-62.

[12] I.K. Budde, R.C. Aalberse, *Histamine-releasing factors, a heterogeneous group of different activities.* Clin. Exp. Allergy. 33 (2003) 1175-82.

[13] R. Valenta, S. Vrtala, *Recombinant allergens for specific immunotherapy*, Allergy 54 Suppl 56 (1999) 43-4.

[14] M. Gajhede, P. Osmark, F.M. Poulsen, H. Ipsen, J.N. Larsen, R.J. Joost van Neerven, C. Schou, H. Lowenstein, M.D. Spangfort, *X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy*, Nat. Struct. Biol. 3 (1996) 1040-5.

[15] C. Faber, A. Lindemann, H. Sticht, A. Ejchart, A. Kungl, M.Susani, R.W. Frank, D. Kraft, M. Breitenbach, P. Rosch, *Secondary structure and tertiary fold of the birch pollen allergen Bet v 1 in solution*, J. Biol. Chem. 271 (1996) 19243-50.

[16] P. Neudecker, K. Schweimer, J. Nerkamp, S. Scheurer, S. Vieths, H. Sticht, P. Rosch, *Allergic cross-reactivity made visible: solution structure of the major cherry allergen Pru av 1*, J. Biol. Chem. 276 (2001) 22756-63.

[17] P. Neudecker, J. Nerkamp, A. Eisenmann, A. Nourse, T. Lauber, K. Schweimer, K. Lehmann, S. Schwarzinger, F. Ferreira, P. Rosch, *Solution structure, dynamics, and hydrodynamics of the calcium-bound cross-reactive birch pollen allergen Bet v 4*

reveal a canonical monomeric two EF-hand assembly with a regulatory function, J. Mol. Biol. 336 (2004) 1141-57.

[18] P. Rosch, *Nuclear magnetic resonance studies of allergens*, J. Chromatogr. B Biomed. Sci. Appl. 756 (2001) 165-77.

[19] M. Saraste, P.R. Sibbald, A. Wittinghofer, *The P-loop--a common motif in ATP- and GTP-binding proteins*, Trends Biochem. Sci. 15 (1990) 430-4.

[20] P. Neudecker, K. Lehmann, J. Nerkamp, T. Haase, A. Wangorsch, K. Fotisch, S. Hoffmann, P. Rosch, S. Vieths, S. Scheurer, *Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry (Prunus avium) and celery (Apium graveolens): correlating IgE reactivity with three-dimensional structure*, Biochem. J. 376 (2003) 97-107.

[21] N. Samanani, P.J. Facchini, *Isolation and partial characterization of norcoclaurine synthase, the first committed step in benzylisoquinoline alkaloid biosynthesis, from opium poppy*, Planta 213 (2001) 898-906.

[22] T.C. Birdsall, G.S. Kelly, *Berberine: Therapeutic Potential of an Alkaloid Found in Several Medicinal Plants*, Alt. Med. Rev. 2 (1997) 94-103.

[23] M.A. Rashid, K.R. Gustafson, Y. Kashman, J.H. Cardellina, J.B. McMahon, M.R. Boyd, *Anti-HIV alkaloids from Toddalia asiatica*, Nat. Prod. Lett. 6 (1995) 153-156.

[24] J. Sambrook, E.F. Fritsch, M. Maniatis, *Molecular cloning: a laboratory manual*, Cold spring harbor laboratory press, New York, 1989.

[25] O. Meyer, H.G. Schlegel, *Biology of aerobic carbon monoxide-oxidizing bacteria*, Annu. Rev. Microbiol. 37 (1983) 277-310.

[26] U.K. Laemmli, *Cleavage of structural proteins during the assembly of the head of bacteriophage T4*, Nature 227 (1970) 680-5.

[27] Z. Zhang, A.G. Marshall, *A universal algorithm for fast and automated charge state deconvolution of electrospray mass-to-charge ratio spectra*, J Am Soc Mass Spectrom. 9 (1998) 225-33.

[28] M.M. Bradford, *A rapid and sensitive method for quantitation of microgram quantities of protein using the principle of protein dye binding*, Anal. Biochem. 72 (1976) 248-254.

[29] S. Mori, C. Abeygunawardana, M.O. Johnson, P.C. van Zijl, *Improved sensitivity of HSQC spectra of exchanging protons at short interscan delays using a new fast HSQC (FHSQC) detection scheme that avoids water saturation.* J. Magn. Reson. B. 108 (1995) 94-8.

[30] E. Goldman, A.H. Rosenberg, G. Zubay, F.W. Studier, *Consecutive low-usage leucine codons block translation only when near the 5' end of a message in Escherichia coli*, J. Mol. Biol. 245 (1995) 467-73.

[31] J.F. Kane, *Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli*, Curr. Opin. Biotechnol. 6 (1995) 494-500.

Figure legends:

Fig.1. Comparison of Δ19NCS expression levels in different *E. coli* host strains analyzed by 15 % SDS-PAGE. (A) *E. coli* ER2566pLysS/pET29b-*Δ19NCS*; (B) *E. coli* BL21(*DE3*)/ pET29b-*Δ19NCS*; (C) *E. coli* Rosetta(*DE3*)/pET29b-*Δ19NCS*; M: Molecular weight standard; lanes 1-5: before induction, 1 h, 2 h, 3 h, 4 h after induction; arrow = Δ19NCS.

Fig. 2. Purification of Δ19NCS expressed in *E. coli* Rosetta(*DE3*)/pET29b-*Δ19NCS* grown on M9 minimal medium and analyzed by 15 % SDS-PAGE. M: Molecular weight standard; lane 1: yield of Δ19NCS after 4 h expression; lanes 2, 4: soluble fraction after lysis; lanes 3, 5: insoluble fraction after lysis; lane 6-9: IMAC purification; lane 6: flow-through; lane 7-9: Δ19NCS eluted at 45-60 mM imidazole; arrow = Δ19NCS full length; asterisk = ΔΔ19NCS.

Fig. 3. Thrombin cleavage of S-tag from Δ19NCS analyzed by 15 % SDS-PAGE. M: molecular weight standard; lane 1: before thrombin cleavage; lane 2: after thrombin cleavage; arrow: Δ19NCS; asterisk: ΔΔ19NCS; double-asterisk: Δ19NCS without S-tag.

Fig. 4. Expression and purification of Δ29NCS expressed in *E. coli* Rosetta(*DE3*)/pET29b-*Δ29NCS* grown on M9 minimal medium and analyzed by 15 % SDS-PAGE. (A) Expression yield. M: molecular weight standard; lane 1: before induction; lane 2: 6 h after induction; (B) lysis and purification. M: molecular weight standard; lane 1: soluble fraction after lysis; lane 2: insoluble fraction after lysis; lane 3-7: IMAC purification. Lane 3: flow-through; lane 4-5: proteins eluted at 20-30 mM imidazole; lane 6-7: Δ29NCS eluted at 45-60 mM imidazole; arrow: Δ29NCS.

Fig. 5. Positive ion ESI mass spectrum (A) and transformed spectrum (B) of Δ29NCS (100 µM) dissolved in 1 mM ammonium acetate, 5% ethanol and 0.1% acetic acid. Experimental molecular mass of Δ29NCS is 21179.5 Da.

Fig. 6. Enzyme activity assay: autoradiograph of TLC-plate used to separate reaction compounds after 1.5 h incubation at 37 °C. Lane 1: assay containing native Δ29NCS; lane 2: negative control containing thermally denatured Δ29NCS; arrow: ^{14}C-(S)-Norcoclaurine; asterisk: [8-^{14}C] dopamine; O: sample origin; F: solvent front.

Fig. 7. Analysis of Δ29NCS by NMR spectroscopy. (A) Amino acid sequence of Δ29NCS. (B) ^{1}H-^{15}N HSQC spectrum of Δ29NCS (730 µM) at 308 K; 800 MHz; 16 accumulations; 256 x 1024 data points.

Fig.1

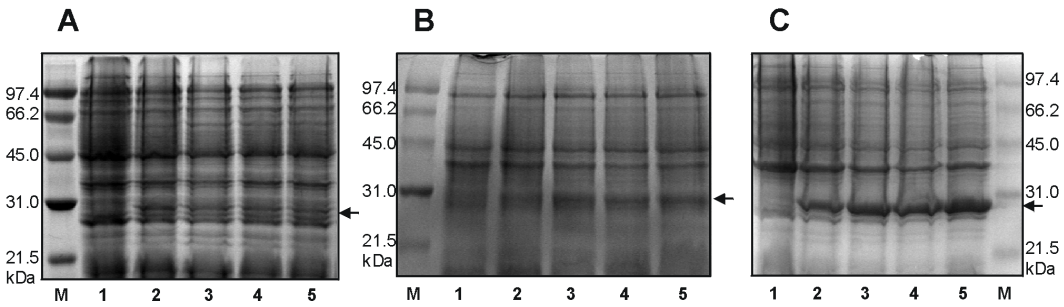

Fig.2

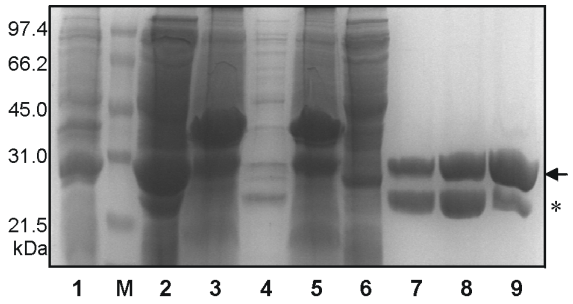

Fig.3

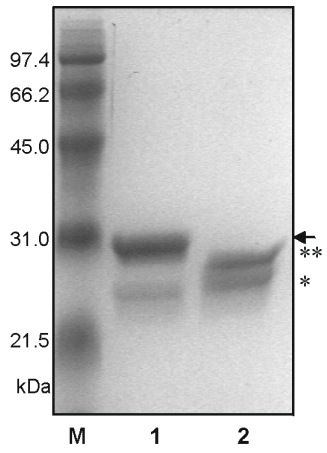

Fig.4

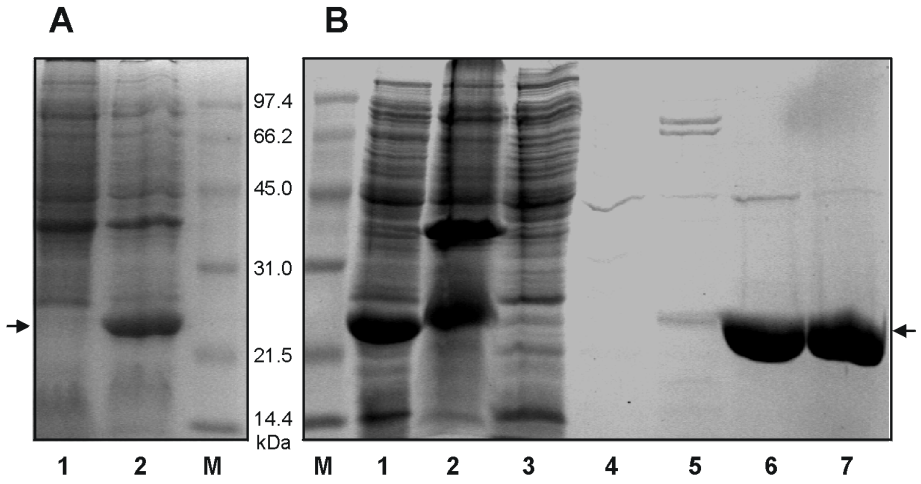

Fig.5

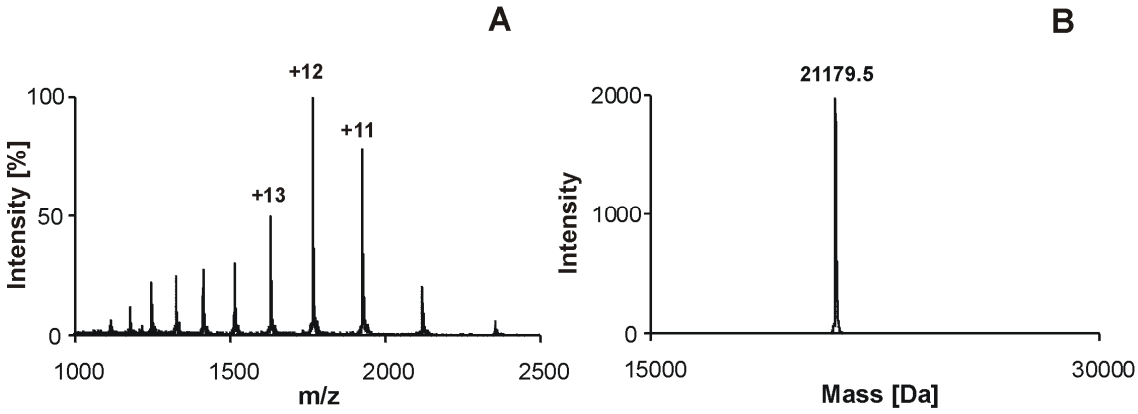

Fig.6

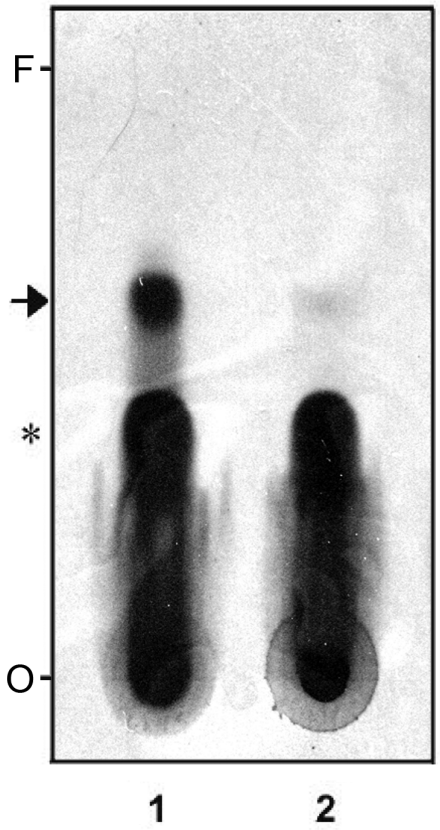

Fig.7

A

```
          10         20         30         40         50
MLHHQGIINQ VSTVTKVIHH ELEVAASADD IWTVYSWPGL AKHLPDLLPG
          60         70         80         90        100
AFEKLEIIGD GGVGTILDMT FVPGEFPHEY KEKFILVDNE HRLKKVQMIE
         110        120        130        140        150
GGYLDLGVTY YMDTIHVVPT GKDSCVIKSS TEYHVKPEFV KIVEPLITTG
         160        170        180        190
PLAAMADAIS KLVLEHKSKS NSDEIEAAII TVLEHHHHHH
```

B