Conformation, catalytic site, and enzymatic mechanism of the PR10 allergen related enzyme norcoclaurine synthase

Hanna Berkner, Kristian Schweimer, Irena Matecko, and Paul Rösch*

Lehrstuhl für Struktur und Chemie der Biopolymere &
Research Center for Bio-Macromolecules, Universität Bayreuth
Universitätsstrasse 30, 95447 Bayreuth, Germany

Running title: NCS substrate binding

*Corresponding author; Universität Bayreuth, Lehrstuhl Biopolymere, Universitätsstr. 30, 95447 Bayreuth, Germany; Phone +49 921 55-3540; Fax: +49 921 16490459; E-mail: roesch@unibt.de

Page heading title: Conformation and mechanism of norcoclaurine synthase

Abbreviations used: NCS, (S)-norcoclaurine synthase; BIAs, benzylisoquinoline alkaloids; 4-HPAA, 4-hydroxyphenylacetaldehyde; PR10 proteins, class 10 of pathogenesis related proteins; SEC, size exclusion chromatography; DTT, dithiothreitol; TROSY, transverse relaxation optimized spectroscopy; CSI, chemical shift index; HSQC, heteronuclear single quantum coherence.

SYNOPSIS

The enzyme (S)-Norcoclaurine synthase (NCS; EC 4.2.1.78) found in the common meadow rue, *Thalictrum flavum*, and other plant species is involved in the biosynthesis of benzylisoquinoline alkaloids (BIAs). This group of plant secondary metabolites comprises pharmacologically active compounds like morphine and codeine. NCS catalyzes the condensation of 4-hydroxyphenylacetaldehyde (4-HPAA) and dopamine to (S)-norcoclaurine, the common precursor of all plant BIAs. While enzymatic properties of NCS and mechanistic aspects of the reaction have been studied in detail, no structural information on NCS was available so far. The enzyme shows significant sequence homology to members of class 10 of pathogenesis related proteins (PR10 proteins) such as the major birch pollen allergen Bet v 1. Our CD and NMR spectroscopic data indicated high similarity of the NCS and the Bet v 1 fold and allowed us to model NCS using Bet v 1 as a template. Virtually complete backbone assignment of the NCS sequence was used to study substrate binding by NMR titration experiments. While binding of 4-HPAA seems to induce side chain rearrangements in an extensive part of the protein, the putative distinct interaction site for dopamine could be clearly identified. The oligomerization state of NCS that reportedly plays an important role in enzyme functionality was determined to be concentration-dependent by size exclusion chromatography as well as NMR relaxation measurements, and the enzyme is predominantly monomer at the low micromolar concentrations used for activity assays.

Keywords: Norcoclaurine synthase, substrate binding, Bet v 1 allergen, NMR spectroscopy, homology modeling

INTRODUCTION

Benzylisoquinoline alkaloids form an important group of plant secondary metabolites including about 2500 different compounds. Many of these small organic molecules show pharmacological activity and have been used for medical purposes 3000 years ago as well as today [1]. BIAs are mainly produced by representatives of the plant families *Papaveraceae, Fumariaceae, Ranunculaceae, Berberidaceae or Menispermaceae*. Opium poppy (*Papaver somniferum*) for example produces the analgesics morphine and codeine, the muscle relaxant papaverine, the antibiotic sanguinarine, and noscapine which is supposed to be anti-tumorigenic [2]. Other less well-known plant species like the common meadow rue (*Thalictrum flavum*) also produce BIAs with pharmacological activity, in this case the antimicrobial berberine and the potential anti-HIV agent magnoflorine [3, 4]. While humans use BIAs as a pharmaceutical drug, plants probably developed those substances for defense against microbial and herbivore attacks. Thus, it is not surprising that the activity of several enzymes involved in BIA biosynthesis is upregulated under environmental stress conditions that can be simulated e.g. by elicitor treatment of cell cultures. This behavior was demonstrated for native (S)-norcoclaurine synthase from *Papaver somniferum* [2]. NCS catalyzes the condensation of 4-hydroxyphenylacetaldehyde and dopamine to (S)-norcoclaurine (Figure 1) which is the central precursor to all BIAs of plant origin [5, 6].

In 2004, the gene coding for NCS in *Thalictrum flavum* could be identified [7]. The deduced amino acid sequence showed 50-60% homology to members of class 10 of pathogenesis related proteins that includes the Bet v 1 protein family of pollen and food allergens. PR10 proteins also seem to play a role in the plant defense system against pathogens [8], but, despite numerous efforts, their precise function could not be established so far. However, the enzymatic properties of native NCS isolated from a number of different BIA producing plant species as well as of recombinant NCS from *Thalictrum flavum*, *Papaver somniferum* and *Coptis japonica* have been investigated [2, 7, 9-11]. Recently, the reaction mechanism of the catalyzed Pictet-Spengler reaction could be elucidated for *Thalictrum flavum* NCS [12], but still no information was available concerning the substrate binding sites and the active center of the enzyme.

The structural characterization of NCS is a prerequisite for the determination of substrate binding sites and further functional studies on this protein. Recently, we developed a purification protocol yielding high amounts of pure isotopically labeled *Thalictrum flavum* NCS suitable for structural studies by means of NMR spectroscopy [13]. Combining

structural information from CD spectroscopy and NMR experiments, including the assignment of Δ29NCS backbone resonances, and the amino acid sequence similarity of NCS to the major birch pollen allergen Bet v 1 (Figure 2), we could built a semi-experimental homology model based on the Bet v 1 crystal structure [14]. Moreover, we could use this model to map the regions of the protein affected by binding of the substrate dopamine and analogs of the second substrate 4-HPAA determined by NMR titration experiments. Additionally, we determined the oligomerization state of recombinant NCS applying size exclusion chromatography and NMR relaxation measurements. We discuss our results in the context of the enzymatic properties of NCS reported so far and in comparison with the only enzyme of known structure catalyzing a similar Pictet-Spengler-type condensation reaction, strictosidine synthase [15].

EXPERIMENTAL

Materials

All chemicals used in this study were at least 98 % pure. Dopamine hydrochloride was purchased from Sigma, while L-phenylalanine, methyl(4-hydroxyphenyl)acetate and 2-(4-hydroxyphenyl)ethanol were from Fluka. Isotopically labeled compounds were purchased from Campro Scientific (Berlin, Germany; $(^{15}NH_4)_2SO_4$), from Spectra Stable Isotopes (Columbia, MD, USA; [U-^{13}C]glucose) or from Euriso-top (Gif-sur-Yvette, France; D_2O).

Methods

Expression and purification of Δ29NCS

Unlabeled and ^{15}N-labeled Δ29NCS was expressed using the expression system *E. coli* Rosetta(*DE3*)/pET29b-*Δ29NCS* as described before [13]. In case of $^{13}C^{15}N$-labeled Δ29NCS, unlabeled glucose was replaced by 2 g/L of [U-13C]glucose. For expression of deuterated $^{13}C^{15}N$-labeled Δ29NCS, five precultures with increasing D_2O content were grown prior to inoculation of the main culture. The first culture was grown overnight at 37 °C in 10 mL LB-medium containing 20 µg/mL kanamycin and 34 µg/mL chloramphenicol. Subsequent cultures were grown at 37 °C in M9 minimal medium [16] including 2 mM $MgSO_4$, 10 µM Fe(III)citrate, 0.1 mM $CaCl_2$, 2 g/L [U-13C]glucose, 1.5 g/L $(^{15}NH_4)_2SO_4$, trace element solution TS2 [17], 1x MEM vitamin solution (Gibco, Invitrogen, Karlsruhe), 20 µg/mL kanamycin, and 34 µg/mL chloramphenicol and contained 25, 50, 75 and 100 % D_2O, respectively. Each preculture was inoculated to an OD_{600} of 0.2 with the respective amount of the preceding preculture. The same medium in 100 % D_2O including 2 % of *E. coli* CN Standard medium ^{13}C, ^{15}N labeled (Silantes, Munich, Germany) was used for the final culture with a volume of 1.5 L. This culture was inoculated to an OD_{600} of 0.07 using the whole preceding preculture and incubated at 37°C until it reached an OD_{600} of 0.6. After induction by addition of 1 mM IPTG, incubation was continued at 25 °C and cells were harvested after 9 h of expression. Purification of Δ29NCS was achieved as described before [13]. Pure protein was dialyzed against millipore water, shock-frozen in liquid nitrogen, lyophilized and stored at 4°C.

Analytical size exclusion chromatography

In order to determine the oligomerization state of recombinant Δ29NCS, size exclusion chromatography was performed on two sequentially connected Superdex 75 HR 10/30 columns (total bed volume app. 48 ml; GE Healthcare, Munich, Germany). 50 mM potassium phosphate buffer pH 7.0 containing 200 mM NaCl and 1 mM dithiothreitol (DTT) was used as sample buffer as well as for column equilibration and for isocratic protein elution carried

out at a flow rate of 0.6 ml/min. Samples with concentrations in a range from 10 µM to 1 mM were prepared either by solubilizing lyophilized Δ29NCS or by dialyzing Δ29NCS protein solution against sample buffer. Molecular mass calibration was performed using the Low Molecular Weight Gel Filtration Calibration Kit (GE Healthcare, Munich, Germany) comprising the following proteins: albumine (67.0 kDa), ovalbumine (43.0 kDa), chymotrypsinogen (25.0 kDa), ribonuclease A (13.7 kDa).

Circular dichroism spectroscopy

Far UV circular dichroism measurements were performed on a J-810 S spectropolarimeter with a CDF-426S temperature control unit (JASCO International, Tokyo, Japan). Samples were prepared dissolving lyophilized protein in 10 mM sodium phosphate buffer, pH 7.0, containing 0.5 mM DTT. Spectra were recorded at 20 °C and 95 °C in a wavelength range of 190-260 nm with 50 nm/min scanning speed in a 1 mm path length quartz cuvette (Hellma, Mühlheim, Germany) at a protein concentration of 10 µM. Buffer spectra were subtracted and 10 spectra were accumulated. Thermal stability was analyzed by monitoring the CD signal at 220 nm during heating from 20 °C to 95 °C and cooling back to 20 °C with a heating/cooling rate of 1 °C/min. Quartz cuvettes with 1 cm path length equipped with a stirrer were used at a protein concentration of 2.5 µM. In order to normalize the measured ellipticity the following formula was used to calculate the mean residue molar ellipticity: $[\Theta]_{MRW} = \Theta/(c \cdot d \cdot N)$

with Θ: measured ellipticity; c: protein concentration; d: path length; N: number of amino acids.

NMR spectroscopy and data evaluation

All NMR experiments were performed at 305.5 K on Bruker Avance 700 MHz and Avance 800 MHz spectrometers with either room temperature or cryogenic-cooled triple-resonance probes equipped with pulsed field-gradient capabilities.

For determination of longitudinal and transversal ^{15}N relaxation rates (R_1 and R_2) at 700 MHz proton resonance frequency, pulse sequences published by Dayie and Wagner [18] were used. Measurements were performed twice with relaxation delays of 10.6 ms, 531.1 ms, 1062.2 ms, 1593.3 ms and 2124.4 ms for R_1 and 8.8 ms, 26.4 ms, 44.0 ms, 61.6 ms and 79.2 ms for R_2. Samples with concentrations of 200 µM and 1 mM ^{15}N-labeled Δ29NCS were prepared in 20 mM sodium phosphate buffer pH 7.8, 5 mM DTT, 10 % D_2O and 0.04 % sodium azide. For relaxation measurements in presence of the substrate dopamine and the 4-HPAA-analog 2-(4-hydroxyphenyl)ethanol, a 100-fold molar excess of each of the two substances was added to 200 µM ^{15}N-labeled Δ29NCS. Relaxation rates were calculated by least-squares fitting of

mono-exponential decays to the time-dependent peak intensities with the program Curvefit (A.G. Palmer III, unpublished), and errors were estimated using the jackknife procedure. For estimation of the overall rotational correlation time τ_c with the program Tensor 2.0 [19], residues with relaxation rates differing from the mean value by more than one-fold the standard deviation were excluded. As for globular proteins, τ_c is directly proportional to their dimension, the molecular mass of a protein with given τ_c can be estimated using a calibration curve calculated from a set of globular proteins of known molecular mass with determined τ_c. We used BMRB entries 5841 (14.9 kDa), 4267 (20.7 kDa) and 5720 (23.6 kDa), and the major birch pollen allergen Bet v 1 (17.4 kDa) for calculation of the calibration curve. All τ_c values were normalized to 298 K.

For backbone resonance assignment, the following TROSY-type NMR experiments [20, 21] were performed with a 400 µM deuterated ^{15}N,^{13}C-labeled Δ29NCS sample in 20 mM sodium phosphate buffer pH 7.8, 1 mM DTT, 0.04 % sodium azide, 10 % D_2O: [^{15}N,^{1}H]-TROSY, tr-HNCO, tr-HN(CA)CO, tr-HNCA, tr-HN(CO)CA, tr-HNCACB, tr-HN(CO)CACB. Additionally, NNH-NOESY and ^{15}N-NOESY-HSQC spectra were recorded with a 670 µM ^{15}N,^{13}C-labeled Δ29NCS sample in the same buffer. These spectra were principally used to determine the links between the β-strands. NMR data was processed and analyzed using in-house written software and the program NMRView 5.2.2 [22]. In order to localize secondary structure elements on the Δ29NCS sequence, the chemical shift index was calculated from the chemical shifts of backbone carbonyl and C_α atoms according to Wishart et al.[23]. As perdeuteration of the protein has an influence on chemical shift values [24], the C_α chemical shift was systematically corrected by + 0.5 ppm prior to CSI calculation.

All NMR titration experiments were performed with ^{15}N-labeled Δ29NCS in 50 mM sodium phosphate buffer pH 7.0 containing 2 mM DTT, 0.04 % sodium azide and 10 % D_2O with two exceptions: In case of the titration with the 4-HPAA-analog methyl(4-hydroxyphenyl)acetate, for solubility reasons, 5 % DMSO was added to the sample buffer that was also used to prepare the stock solutions for the respective titrations. Concerning the titration with the 4-HPAA-analog 2-(4-hydroxyphenyl)ethanol, 20 mM sodium phosphate buffer pH 7 containing 100 mM NaCl, 0.04 % sodium azide and 10 % D_2O was used. Substrate binding was investigated recording a series of [^{15}N,^{1}H]-HSQC spectra upon gradual addition of the respective substrate or substrate analog to a final 100-fold molar excess. The dissociation

constant K_D was determined by fitting the ligand concentration-dependent chemical shift changes of residues in the fast-exchange limit to the following equation which is valid for a two-state model:

$$\delta_{obs} = \delta_P + (\delta_{PL} - \delta_P) [\{K_D + (1 + r) P_0\}/(2 P_0) - \{(K_D + (1 + r) P_0)^2 - 4 P_0^2 r\}^{1/2}/(2 P_0)]$$

with δ_{obs}, δ_P, δ_{PL} as the respective chemical shifts observed in the actual titration step, for the free Δ29NCS, and for the completely bound ligand. P_0 represents the total concentration of Δ29NCS, and r stands for the molar ratio between the respective ligand and Δ29NCS. The chemical shift changes between free and bound state in the proton and nitrogen dimension, $\Delta\delta(^1H)$ and $\Delta\delta(^{15}N)$, were normalized to $\Delta\delta_{norm} = [\Delta\delta(^1H)^2 + (0.1\ \Delta\delta(^{15}N))^2]^{1/2}$. Normalized chemical shift changes larger than 0.04 ppm were considered as significant [25].

Sequence alignment and homology modeling

The sequence alignment of Δ29NCS and Bet v 1 was performed using ClustalW [26]. For the three-dimensional structure model, the alignment was modified according to the secondary structure information obtained from the CSI. The Δ29NCS structure model was built with the program Modeller 9v2 [27-29] using Bet v 1 (PDB code 1BV1) as template structure. Stereochemical quality of the resulting structure was checked with PROCHECK V3.5 [30, 31]. Figures were generated using the program PyMOL v0.99 (2006, DeLano Scientific LLC, South San Francisco, California, USA). Correct orientation of the β-strands was verified according to amide-amide NOEs observed between the respective backbone amide protons.

RESULTS

Oligomerization state of Δ29NCS

Two independent methods, size exclusion chromatography and NMR relaxation measurements, were used to determine the oligomerization state of native recombinant Δ29NCS. The apparent molecular mass of the enzyme calculated from both kinds of experiment turned out to be concentration-dependent. Most of the measurements were performed in duplicate or three times, with the respective standard deviation not exceeding 0.4 kDa (Table 1). Although the absolute values for the apparent molecular mass obtained from NMR relaxation experiments are systematically lower than those determined by SEC, they show the same tendency. At concentrations ranging from 10 µM to 1 mM, monomeric and oligomeric forms of Δ29NCS are in fast exchange on the respective timescale, as only the intermediate state can be observed by NMR as well as by SEC. At a concentration of 10 µM, the apparent molecular mass is closer to the molecular mass of the 21.2 kDa monomer, while at a concentration of 1 mM it is intermediate between the masses expected for monomeric and dimeric protein. Thus, Δ29NCS appears to be predominantly monomeric at concentrations around 10 µM. In order to study if dimerization of the enzyme occurs upon substrate binding, measurements in presence of a 20 or 100-fold excess of the substrate dopamine and the 4-HPAA analog 2-(4-hydroxyphenyl)ethanol were performed, resulting in absolutely no observable increase of the apparent molecular mass as determined by SEC and only a slight increase of 5 % as determined by NMR relaxation measurements compared to the experiments without substrate. Thus, there is no indication that substrate binding induces dimerization of Δ29NCS. The difference between the absolute values for the apparent molecular mass calculated from the two methods might result from the fact that the retention volume as well as the relaxation rate is not only dependent on the molecular mass, but also on the hydrodynamic radius of the protein. Even though we report a relatively globular fold for Δ29NCS and we only chose globular proteins for the calibration curves, the different sets of calibration proteins might at least partially explain this observation.

Secondary structure content and thermal stability of Δ29NCS versus Bet v 1

CD spectroscopy was used to estimate secondary structure content and thermal stability of Δ29NCS. The CD spectrum of native Δ29NCS shows a broad minimum at 215 nm considered characteristic for β-sheet structures (Figure 3A). At 95 °C this minimum is shifting to lower wavelengths typical for random coil conformation, indicating loss of secondary structure. Subsequent cooling to 20 °C results in a CD spectrum virtually identical to that of the native protein, showing the repairability of thermal unfolding. The shape of the CD spectra of native

Δ29NCS and native Bet v 1 is highly similar, especially in the minimum region, implying that the two proteins contain similar fractions of the same secondary structure elements (Figure 3B). Secondary structure content of the two proteins was analyzed using the CDSSTR algorithm combined with a reference data set consisting of 43 proteins [32] provided by the DICHROWEB server [33, 34]. Similar to the secondary structure content calculated for Bet v 1, namely 28 % α-helix and 29 % β-sheet, 27 % of α-helical segments and 25 % β-sheet was calculated for Δ29NCS.

Temperature dependent changes of ellipticity at 220 nm (Figure 3C) were used to calculate the melting point of the Δ29NCS structure as T_m = 66.4 °C, virtually identical to the melting point of Bet v 1 determined to be 67.7 °C. The cooling curve shows thermal unfolding to be only partially reversible, as the ellipticity decreases on cooling, but does not reach the starting point due to partial protein aggregation.

Structural characterization of Δ29NCS by NMR spectroscopy

As the concentration-dependent oligomerization of the 21.2 kDa protein leads to increasing transversal relaxation rates which in turn decrease the signal intensity in 2D and especially 3D NMR experiments used for backbone assignment, perdeuteration of Δ29NCS was necessary to yield spectra in a suitable quality for sequential resonance assignment. Nearly all signals visible in the [^{15}N,^{1}H]-TROSY spectrum could be assigned (Figure 4A). Most of the unassigned resonances correspond to residual signals of suppressed side chain amides. Thus 86 % of the Δ29NCS amino acid sequence could be assigned (Figure 4B). The major part of the signals for the unassigned aminoterminal and carboxyterminal amino acids is missing due to conformational or solvent exchange typical for residues in flexible regions and accelerated by the high pH necessarily employed for these measurements. In order to gain information on the distribution of secondary structure elements on the sequence of Δ29NCS, we calculated the chemical shift index (CSI) by combining C' and C_α chemical shifts. A downfield shift relative to the random coil value symbolized by positive values in the CSI indicates the presence of α-helical conformation, while an upfield shift symbolized by negative values in the CSI is an indication for β-strands (Figure 5A). The CSI confirms that not only secondary structure content, but also distribution of the secondary structure elements is highly similar to Bet v 1, although not entirely identical. As in Bet v 1, the secondary structure aligned to the sequence beginning at the aminoterminus starts with a β-strand followed by two α-helices which, in turn, are followed by six β-strands. Major differences between Δ29NCS are found in the carboxyterminal part: While the major birch pollen allergen contains one single long

carboxyterminal helix, the corresponding helix in Δ29NCS is surrounded by two additional short α-helices. The carboxyterminal helix is found in the region of the Δ29NCS sequence that cannot be aligned to the shorter Bet v 1. To check if the identified β-strands are connected to a seven-stranded β-sheet in the same manner as in Bet v 1, we determined their relative orientation observing amide-amide NOEs between the backbone amide protons. The NOEs prove that also the relative orientation of the Δ29NCS β-strands is the same as in Bet v 1. This means that in the antiparallel β-sheet, all strands are spatially connected to the respective sequentially succeeding strand with exception of the aminoterminal β-strand that is oriented antiparallel to the β-strand located most closely to the carboxyterminus.

Homology model of Δ29NCS

The homology model of Δ29NCS was built by combining the experimental results from NMR measurements with homology modeling using the crystal structure of the homologous major birch pollen allergen Bet v 1 [14] as template structure. On the one hand, Bet v 1 was chosen as template because of sequence similarity, and especially conservation of sequence motives typical for the Bet v 1 protein family like the P-loop (Figure 2) On the other hand, it was chosen because the CD data suggested highly similar secondary structure content and thermal stability (Figures. 3B & 3C). The chemical shift index (Figure 5A) was used to modify the alignment of Δ29NCS and Bet v 1 obtained by Clustal W according to the expected Δ29NCS secondary structure. As described before, the CSI suggested that Δ29NCS shows the same secondary structure element distribution as Bet v 1, except for two small α-helices surrounding the long C-terminal α-helix also found in Bet v 1. The locations of these two helices were therefore defined as additional restraints in the modeling procedure. The correct relative orientation of the β-strands in the model structure obtained was verified by comparison to our experimental NOE data. The stereochemical quality of the final model structure (Figure 5B) was assessed using the PROCHECK program. Concerning Phi/Psi torsion angles, 90 % of the amino acid residues were found in the most favored regions of the Ramachandran plot, while 9 % and 1 % were localized in the additional and generously allowed regions, respectively. No residues were found in the disallowed regions. This corresponds to a class 1 stereochemical quality according to Morris et al. [30]. The Δ29NCS model structure consists of an antiparallel seven-stranded β-sheet and two short α-helices α1 and α2, which pack together against a carboxyterminal 20 residue α-helix α4, in turn surrounded by two short α-helices α3 and α5. The latter represent the only secondary

structure elements that do not find a counterpart in the Bet v 1 structure. The Δ29NCS β-sheet consist of amino acids 14 → 24 (β1), 135 ← 123 (β7), 109 → 120 (β6), 100 ← 92 (β5), 79 → 88 (β4), 70 ← 64 (β3) and 53 → 58 (β2). The position of the secondary structure elements is in good agreement with the CSI.

Δ29NCS substrate binding

In order to study substrate binding, NMR titration experiments were performed observing the chemical shift changes of Δ29NCS backbone amide resonances upon gradual addition of either one of the substrates. In case of 4-HPAA, the substrate analog methyl(4-hydroxyphenyl)acetate carrying a methylester group instead of the aldehyde group was used because 4-HPAA was not available in sufficient purity for quantitative titration experiments. Purity of the oxygen-sensitive compound which was always handled under oxygen-free conditions was checked by 1D proton NMR. The addition of methyl(4-hydroxyphenyl)acetate to Δ29NCS provokes significant chemical shift changes for a number of mostly consecutive amino acids (Figure 6A) and, apart from two residues exhibiting chemical shift changes around the significance threshold, the changes of chemical shifts upon addition of dopamine could be related to two sequence areas (Figure 6B).

The amino acids affected in case of the methyl-(4-hydroxyphenyl)acetate titration comprise residues 37 to 46 corresponding to helix α2, residues 67 to 72 located in the aminoterminal part of strand β3 and the adjacent loop, residues 81 to 87 representing the major part of strand β4, residues 94 to 97 located in the middle of strand β5, residues 100 to 103 corresponding to the first part of the adjacent loop, residues 112 to 115 located in the middle of strand β6, residues 150 to 158 representing the aminoterminal part of the long carboxyterminal helix α4 and residues 180 to 183 corresponding to part of helix α5. For helix α2, it is not really obvious from the chemical shift perturbation that a cluster of amino acids is affected, because the two amino acids subsequent to W37 are not assigned, and the chemical shift changes for L40 and A41 could not be determined as the corresponding signals disappear during titration, indicating that these residues show intermediate exchange in contrast to all other residues whose exchange behavior is fast on the NMR time scale. The only residue showing a significant chemical shift change that is not integrated in a cluster of shifting residues is Y133 which is located on strand β7 and is spatially close to the affected parts of helix α4 and strand β6 (Figure 6C). A direct interaction with the relatively small substrate molecule cannot explain the chemical shift changes in such large parts of the protein. Thus, it is likely that

binding of methyl(4-hydroxyphenyl)acetate induces a rearrangement of the side chains in the affected parts of the protein that are all in spatial proximity to each other.

We also used the 4-HPAA-analog 2-(4-hydroxyphenyl)ethanol for titration, albeit under slightly different buffer conditions. In most cases, the chemical shift changes observed are assigned to the same residues shifting upon addition of methyl(4-hydroxyphenyl)acetate, indicating that the methlester- or hydroxy group replacing the aldehyde group in the respective analog does not alter substrate binding.

In case of dopamine binding, the chemical shift changes can be mainly found in a distinct part of the protein comprising amino acids 69 to 72 that represent part of strand β3 and the adjacent loop. Furthermore, residues 150 to 155 corresponding to the aminoterminal part of helix α4 are involved (Figure 6D). In this case, the chemical shift changes are probably induced by direct binding to the respective residues. Closer inspection of the type of amino acids involved and the spatial orientation of their side chains reveals a putative binding site between F71 which might be responsible for hydrophobic interactions with the aromatic ring of dopamine and M155 which might associate with the aliphatic part. A K_D value of 5 mM could be estimated by averaging the values determined from the concentration dependent chemical shift changes for residues F71 and M155, indicating rather weak binding.

To rule out the possibility that we observed non-specific substrate binding, we performed a negative control titration with phenylalanine yielding chemical shift changes of 0.04 and 0.06 ppm, that is in the range of the significance level, for only two residues, namely V72 and L152. These residues are also affected by dopamine binding, but the chemical shift changes observed in this case are more than twice as high, and in contrast to phenylalanine binding, the saturation level is almost reached for these residues. Thus, despite the rather high K_D-values, it is unlikely that we observe non-specific substrate binding.

Samanani and Facchini [9] conclude from product inhibition experiments that 4-HPAA binds to the enzyme before dopamine. Thus, we also performed titration experiments with dopamine in the presence of one of the 4-HPAA analogs at saturation level. Significant chemical shift changes could be observed except for for amino acids A153 and M155, and determination of K_Ds was hindered by the fact that we could not reach saturation level in our experiments. This is surprising at first glance, because it could be interpreted as an indication that dopamine does not bind to Δ29NCS subsequent to 4-HPAA analog binding. However, as the putative dopamine binding site is located in the part of the protein also affected by 4-HPAA analog binding, it is likely that the 4-HPAA binding site is located close to the

dopamine binding site. Thus, only small changes might occur because the chemical environment is already changed by 4-HPAA analog binding.

DISCUSSION

The plant enzyme (S)-norcoclaurine synthase catalyzes the condensation of 4-HPAA and dopamine to the central benzylisoquinoline alkaloid precursor (S)-norcoclaurine [5, 6]. So far, enzymatic properties of NCS from a number of different plant species have been investigated [2, 9, 35, 36], and in a few cases, the production of recombinant protein could be established [7, 10, 11]. The oligomerization state that plays an important role for enzyme functionality was investigated for NCS isolated from *Thalictrum flavum*. The results from size exclusion chromatography indicated that the enzyme is a dimer since the molecular mass of NCS derived from SDS-PAGE was determined to be 15 kDa and that determined by SEC was around 28 kDa [9]. However, at that time the protein sequence of NCS was unkown. The identification of the gene coding for NCS revealed that the protein consists of 210 amino acids including a potential aminoterminal signal peptide of 19 amino acids that will probably be cleaved during the subcellular localization process in the plant [7]. Consequently, one would expect a ~ 21 kDa protein to be present in the plant. Therefore, we studied oligomerization of recombinant *Thalictrum flavum* Δ29NCS by two independent methods, namely SEC and NMR relaxation measurements. The results of both methods demonstrated concentration-dependent oligomerization of Δ29NCS with an equilibrium of monomeric and oligomeric protein over a broad concentration range (Table 1). Although the total values for the apparent molecular mass of Δ29NCS differed between the two methods, it can be concluded that at low protein concentrations in the range of 10 µM, the enzyme is predominantly monomeric. Since NCS activity assays were performed at concentrations of around 80 nM [7, 12], one would expect that the prevailing monomeric form of the enzyme is in large parts responsible for the observed activity. Substrate saturation kinetics measured for natural as well as for recombinant *Thalictrum flavum* NCS indicated cooperativity in case of dopamine binding [7, 9, 12]. This observation was explained by the assumption that NCS exists as a homodimer with one dopamine binding site on each subunit. It was proposed that binding to one site could influence binding to the site on the other subunit. Our results on the Δ29NCS oligomerization state do not support this explanation. Furthermore, our experimental titration curves for dopamine binding fit well to a two-state model suggesting non-cooperative binding. Moreover, we did not find evidence for a second dopamine binding site in the enzyme. Schuhmacher et al. and Rueffer et al. [35, 36] reported non-cooperative substrate

saturation kinetics with regard to dopamine for NCS from a number of different plant species, among them the closely related *Thalictrum tuberosum*.

General enzymatic properties like pH and temperature optimum were also determined for natural as well as for recombinant *Thalictrum flavum* NCS [7, 9]. The relatively high temperature optimum at 45-55 °C for the natural and at 40 °C for the recombinant enzyme, and the rapid decrease in activity at temperatures exceeding 55 °C is in agreement with our results on the thermal stability of Δ29NCS. The melting point of 66.4 °C corresponds to cooperative unfolding of the protein in the temperature range between 60 °C and 70 °C (Figure 3C).

While the reaction mechanism of the catalyzed Pictet-Spengler reaction could be established for NCS from *Thalictrum flavum* [12], no structural information on NCS has been available. However, this is a prerequisite for the localization of substrate binding sites. Combining experimental NMR spectroscopic data with homology modeling using the homologous major birch pollen allergen Bet v 1 as template, we have built a reliable model structure of Δ29NCS (Figure 5B). Since we achieved almost complete backbone assignment of the protein (Figure 4), we could use this model to investigate substrate binding by NMR titration experiments. Using different analogs of 4-HPAA, we observed chemical shift changes for a considerable part of the protein, suggesting that a rearrangement of the respective side chains is induced by 4-HPAA binding (Figure 6A,C). This observation is in agreement with results from product analog inhibition experiments reported by Samanani and Facchini [9] who suggested that 4-HPAA binds to a different conformation of the enzyme compared to the product (S)-norcoclaurine. They proposed the changes to occur upon product release, while our data suggests that they are induced by 4-HPAA binding. From our data, the exact binding site of 4-HPAA is difficult to determine because the extent of chemical shift changes does not allow to distinguish between indirectly caused rearrangement and direct interaction. The chemical shift changes observed upon dopamine binding can be assigned to a distinct area in the protein structure, and spatial orientation of the side chains as well as lower K_D values relatively to the surrounding amino acids suggest direct interactions of dopamine with amino acid residues F71 and M155 (Figure 6B,D). The involvement of F71 is supported by the fact that binding pockets of other proteins binding dopamine also comprise phenylalanine residues [37, 38].

The K_D values estimated for dopamine binding in the millimolar range suggest only weak substrate binding. Since it was suggested that dopamine binds after 4-HPAA [9], the side chain rearrangement induced by 4-HPAA binding might increase the affinity of the enzyme for dopamine. In addition, the low binding affinities of NCS could be due to its subcellular

location. As NCS from *Thalictrum flavum* carries a putative signal peptide [7], it is probably targeted to the endoplasmatic reticulum. From there, it might be translocated to different cellular compartments like special vacuoles. Interestingly, it has been shown for *Papaver bracteatum*, which is a BIA producing plant, that dopamine accumulates in vacuolar compartments [39]. In consequence, the high local substrate concentration could compensate the low binding affinity. The efficiency of the enzyme might thus be due to its stereospecifity with respect to product formation.

The only enzyme of known structure that catalyzes a similar Pictet-Spengler-type reaction is strictosidine synthase which, among other plants, was isolated from *Rauvolfia serpentina* [15]. This enzyme catalyzes the condensation of the aldehyde secologanine and the amine tryptamine to strictosidine, a reaction analogous to that catalyzed by NCS, but in this case yielding the precursor for all plant indole alkaloids. No sequence homology is found between the two enzymes, and the protein fold is also quite different, but but there are similarities in catalytic mechanism. In the active site of strictosidine synthase, a glutamate is found that is suggested to deprotonate the amine moiety of tryptamine. This facilitates the formation of a Schiff base with the aldehyde moiety of secologanine. In the Δ29NCS structure, E75 is found in the immediate vicinity of the putative dopamine binding site, with its side chain pointing towards the binding site (Figure 6E). Consequently, this residue could play the same role in norcoclaurine formation, facilitating the Schiff base formation between the deprotonated dopamine and 4-HPAA. Considering the reaction mechanism proposed by Luk et al. [12], it is even more likely that E75 plays a role in deprotonation of the σ-intermediate formed after the electrophilic cyclization between the imminium ion and the aromatic ring of dopamine.

ACKNOWLEDGEMENTS

We thank Ramona Heissmann for expert technical assistance. We would also like to give our thanks to Peter J. Facchini (Department of Biological Sciences, University of Calgary, Calgary, Alberta T2N 1N4, Canada) for providing the original TfNCSΔ19/pET29b plasmid. This project was supported by a grant from the Deutsche Forschungsgemeinschaft (RO 617/11-4).

REFERENCES

1 Buchanan, B., Gruissem, W. and Jones, R. E. (2000) Biochemistry & Molecular Biology of Plants, American Society of Plant Physiologists, Rockville, Md.

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

3 Birdsall, T. C. and Kelly, G. S. (1997) Berberine: therapeutic potential of an alkaloid found in several medicinal plants. Alt Med Rev **2**, 94-103

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

5 Stadler, R., Kutchan, T. M., Löffler, S., Nagakura, N., Cassels, B. K. and Zenk, M. H. (1987) Revision of the early steps of reticuline biosynthesis. Tetrahedron Lett **28**, 1251-1254

6 Stadler, R., Kutchan, T. M. and Zenk, M. H. (1989) Norcoclaurine is the central intermediate in benzylisoquinoline alkaloid biosynthesis. Phytochemistry **28**, 1083-1086

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

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

9 Samanani, N. and Facchini, P. J. (2002) Purification and characterization of norcoclaurine synthase. The first committed enzyme in benzylisoquinoline alkaloid biosynthesis in plants. J Biol Chem **277**, 33878-83

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

11 Minami, H., Dubouzet, E., Iwasa, K. and Sato, F. (2007) Functional analysis of norcoclaurine synthase in Coptis japonica. J Biol Chem **282**, 6274-82

12 Luk, L. Y., Bunn, S., Liscombe, D. K., Facchini, P. J. and Tanner, M. E. (2007) Mechanistic studies on norcoclaurine synthase of benzylisoquinoline alkaloid biosynthesis: an enzymatic Pictet-Spengler reaction. Biochemistry **46**, 10153-61

13 Berkner, H., Engelhorn, J., Liscombe, D. K., Schweimer, K., Wohrl, B. M., Facchini, P. J., Rosch, P. and Matecko, I. (2007) High-yield expression and purification of isotopically labeled norcoclaurine synthase, a Bet v 1-homologous enzyme, from Thalictrum flavum for NMR studies. Protein Expr Purif **56**, 197-204

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

15 Ma, X., Panjikar, S., Koepke, J., Loris, E. and Stockigt, J. (2006) The structure of Rauvolfia serpentina strictosidine synthase is a novel six-bladed beta-propeller fold in plant proteins. Plant Cell **18**, 907-20

16 Sambrook, J., Fritsch, E. F. and Maniatis, M. (1989) Molecular cloning: a laboratory manual, Cold spring harbor laboratory press, New York

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

18 Dayie, K. T. and Wagner, G. (1994) Relaxation-rate measurements for 15N-1H groups with pulsed-field gradients and preservation of coherence pathways. J Magn Reson **111A**, 121-126

19 Dosset, P., Hus, J. C., Blackledge, M. and Marion, D. (2000) Efficient analysis of macromolecular rotational diffusion from heteronuclear relaxation data. J Biomol NMR **16**, 23-8

20 Salzmann, M., Pervushin, K., Wider, G., Senn, H. and Wuthrich, K. (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins. Proc Natl Acad Sci U S A **95**, 13585-90

21 Pervushin, K., Riek, R., Wider, G. and Wuthrich, K. (1997) Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. Proc Natl Acad Sci U S A **94**, 12366-71

22 Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. Methods Mol Biol **278**, 313-52

23 Wishart, D. S., Sykes, B. D. and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure. J Mol Biol **222**, 311-33

24 Gardner, K. H. and Kay, L. E. (1998) The use of 2H, 13C, 15N multidimensional NMR to study the structure and dynamics of proteins. Annu Rev Biophys Biomol Struct **27**, 357-406

25 Hajduk, P. J., Dinges, J., Miknis, G. F., Merlock, M., Middleton, T., Kempf, D. J., Egan, D. A., Walter, K. A., Robins, T. S., Shuker, S. B., Holzman, T. F. and Fesik, S. W. (1997) NMR-based discovery of lead inhibitors that block DNA binding of the human papillomavirus E2 protein. J Med Chem **40**, 3144-50

26 Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D. G. and Thompson, J. D. (2003) Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res **31**, 3497-500

27 Marti-Renom, M. A., Stuart, A. C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. Annu Rev Biophys Biomol Struct **29**, 291-325

28 Sali, A. and Blundell, T. L. (1993) Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol **234**, 779-815

29 Fiser, A., Do, R. K. and Sali, A. (2000) Modeling of loops in protein structures. Protein Sci **9**, 1753-73

30 Morris, A. L., MacArthur, M. W., Hutchinson, E. G. and Thornton, J. M. (1992) Stereochemical quality of protein structure coordinates. Proteins **12**, 345-64

31 Laskowski, R. A., Moss, D. S. and Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst **26**, 283-291

32 Sreerama, N. and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal Biochem **287**, 252-60

33 Lobley, A., Whitmore, L. and Wallace, B. A. (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra. Bioinformatics **18**, 211-2

34 Whitmore, L. and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res **32**, W668-73

35 Rueffer, M., El-Shagi, H., Nagakura, N. and Zenk, M. H. (1981) (S)-Norlaudanosoline Synthase: The First Enzyme in the Benzylisoquinoline Biosynthetic Pathway. FEBS Lett **129**, 5-9

36 Schumacher, H. M., Ruffer, M., Nagakura, N. and Zenk, M. H. (1983) Partial Purification and Properties of (S)-Norlaudanosoline Synthase from Eschscholtzia tenuifolia Cell Cultures. Planta Med **48**, 212-20

37 Lu, J. H., Li, H. T., Liu, M. C., Zhang, J. P., Li, M., An, X. M. and Chang, W. R. (2005) Crystal structure of human sulfotransferase SULT1A3 in complex with dopamine and 3'-phosphoadenosine 5'-phosphate. Biochem Biophys Res Commun **335**, 417-23

38 Erlandsen, H., Flatmark, T., Stevens, R. C. and Hough, E. (1998) Crystallographic analysis of the human phenylalanine hydroxylase catalytic domain with bound catechol inhibitors at 2.0 A resolution. Biochemistry **37**, 15638-46

39 Kutchan, T. M., Rush, M. and Coscia, C. J. (1986) Subcellular Localization of Alkaloids and Dopamine in Different Vacuolar Compartments of Papaver bracteatum. Plant Physiol **81**, 161-166

Figure Legends

Figure 1 **Condensation reaction of 4-hydroxyphenylacetaldehyde and dopamine catalyzed by the enzyme (S)-norcoclaurine synthase**

Figure 2 **Sequence alignment of Δ29NCS and the major birch pollen allergen Bet v 1**

Δ29NCS and Bet v 1 amino acid sequences were aligned with Clustal W; Black background = Identical amino acid; Grey background = Conservative exchange; Box = P-loop consensus sequence; The two homologous proteins show 25 % sequence identity.

Figure 3 **Comparison of Δ29NCS and Bet v 1 secondary structure content and thermal stability by means of far UV CD spectroscopy**

(**A**) Δ29NCS in native state at 20°C (——), in denatured state at 95°C (— —) and renatured at 20°C (------), calibrated to soluble protein concentration. (**B**) Native Δ29NCS (——) and native Bet v 1 (------) at 20°C. (**C**) Thermal denaturation of Δ29NCS; ♦ = heating curve; Δ = cooling curve.

Figure 4 **Backbone assignment of Δ29NCS**

(**A**) [^{15}N,^{1}H]-TROSY NMR spectrum of deuterated ^{15}N,^{13}C-labeled Δ29NCS (400 µM) measured at 305.5 K and 800 MHz; 4 accumulations; assigned backbone amide resonances are labeled with amino acid type in one letter code and residue number; sc = side chain amide resonance; W sc = tryptophane side chain resonance. (**B**) Amino acid sequence of Δ29NCS with assigned residues underlined.

Figure 5 **Chemical shift index and homology model of Δ29NCS**

(**A**) Combined chemical shift index calculated from C' and C_α chemical shifts derived from HNCO, HN(CA)CO, HN(CO)CA and HNCA spectra of Δ29NCS; +1 was assigned to either C' or C_α residues showing a downfield shift related to the random coil value, while –1 was assigned to C' or C_α residues showing an upfield shift related to the random coil value; 0 was assigned to chemical shifts in the range of values typical for random coil; obtained values for C' and C_α were added. (**B**) Structure of Δ29NCS built by combining homology modeling using the Bet v 1 crystal structure (PDB code 1BV1) as template with experimental information on secondary and tertiary structure as the CSI and backbone amide-amide NOEs in β-strand regions; secondary structure elements are labeled according to their sequence position.

Figure 6 **Δ29NCS substrate binding**

(**A**) Chemical shift perturbation expressed in normalized shift derived from the titration of Δ29NCS with up to a 100-fold excess of the 4-HPAA-analog methyl(4-hydroxyphenyl)-acetate; the dotted line represents the significance level of 0.04 ppm; asterisk = residue not assigned. (**B**) Chemical shift perturbation expressed in normalized shift derived from the titration of Δ29NCS with up to a 100-fold excess of dopamine; the dotted line represents the significance level of 0.04 ppm; asterisk = residue not assigned. (**C**) Δ29NCS model structure with residues affected by 4-HPAA analog binding labeled red. (**D**) Δ29NCS model structure with residues involved in dopamine binding labeled in red and schematic representation of potentially involved side chains of F71 and M155. (**E**) Detail of the Δ29NCS model structure showing the orientation of the potential catalytic residue E75 relative to the suggested binding site of dopamine; distances are in Å.

Table 1 Concentration-dependence of the apparent molecular mass of Δ29NCS

Protein concentration [μM]	Apparent molecular mass (Size exclusion chromatography) [kDa]	Apparent molecular mass (NMR relaxation measurements) [kDa]
10	27.2 ± 0.1*	n.d.
200	28.8 ± 0.4*	23.9 ± 0.2*
400	29.9 ± 0.2*	n.d.
600	31.6	n.d.
1000	32.5	29.3 ± 0.2*

*Measurements were repeated once or twice; given value = mean value with standard deviation

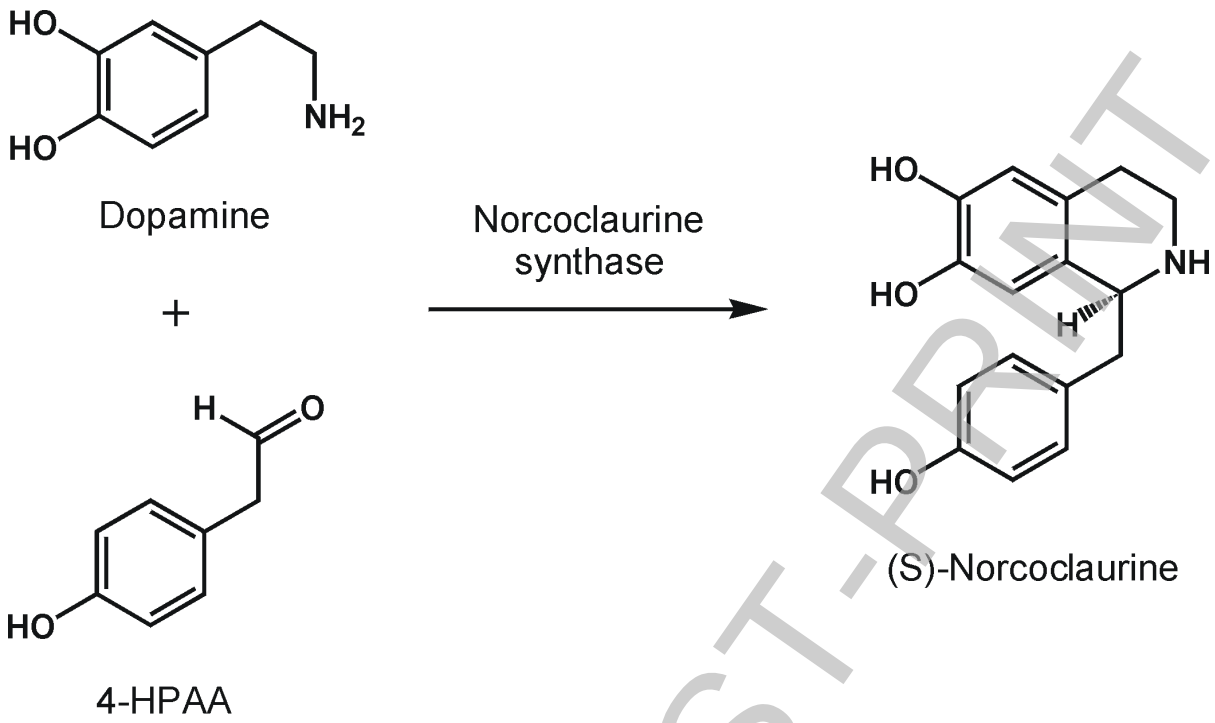


Figure 1

Figure 2

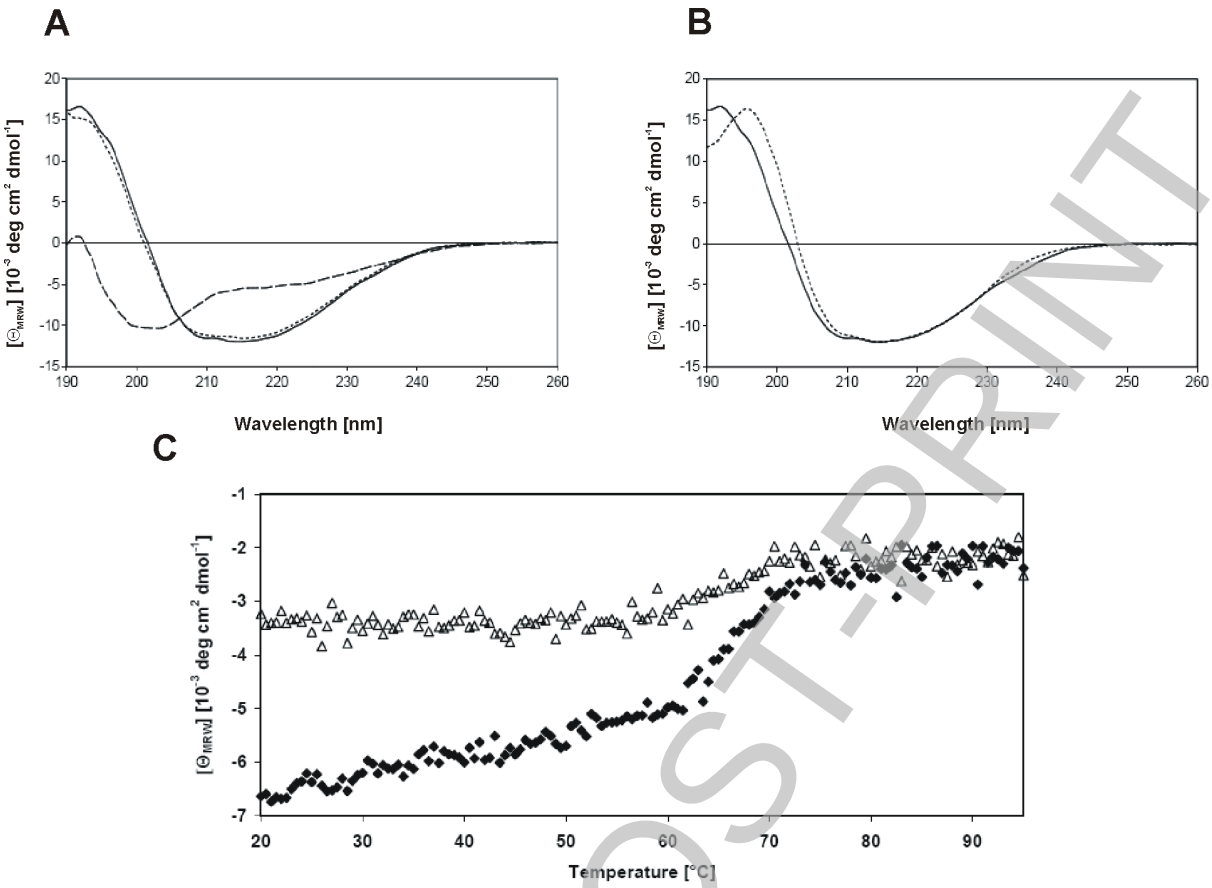
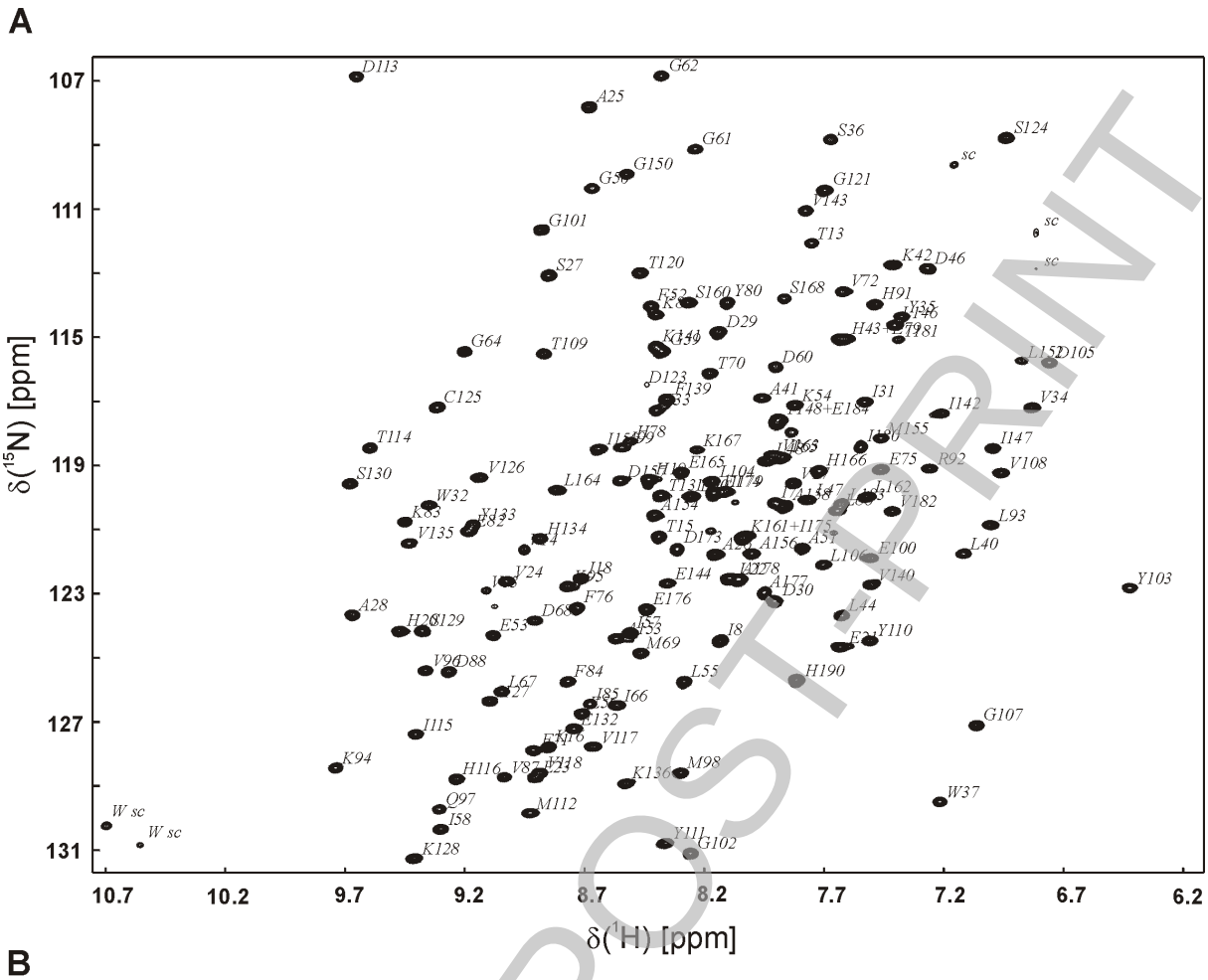


Figure 3



B

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

Figure 4

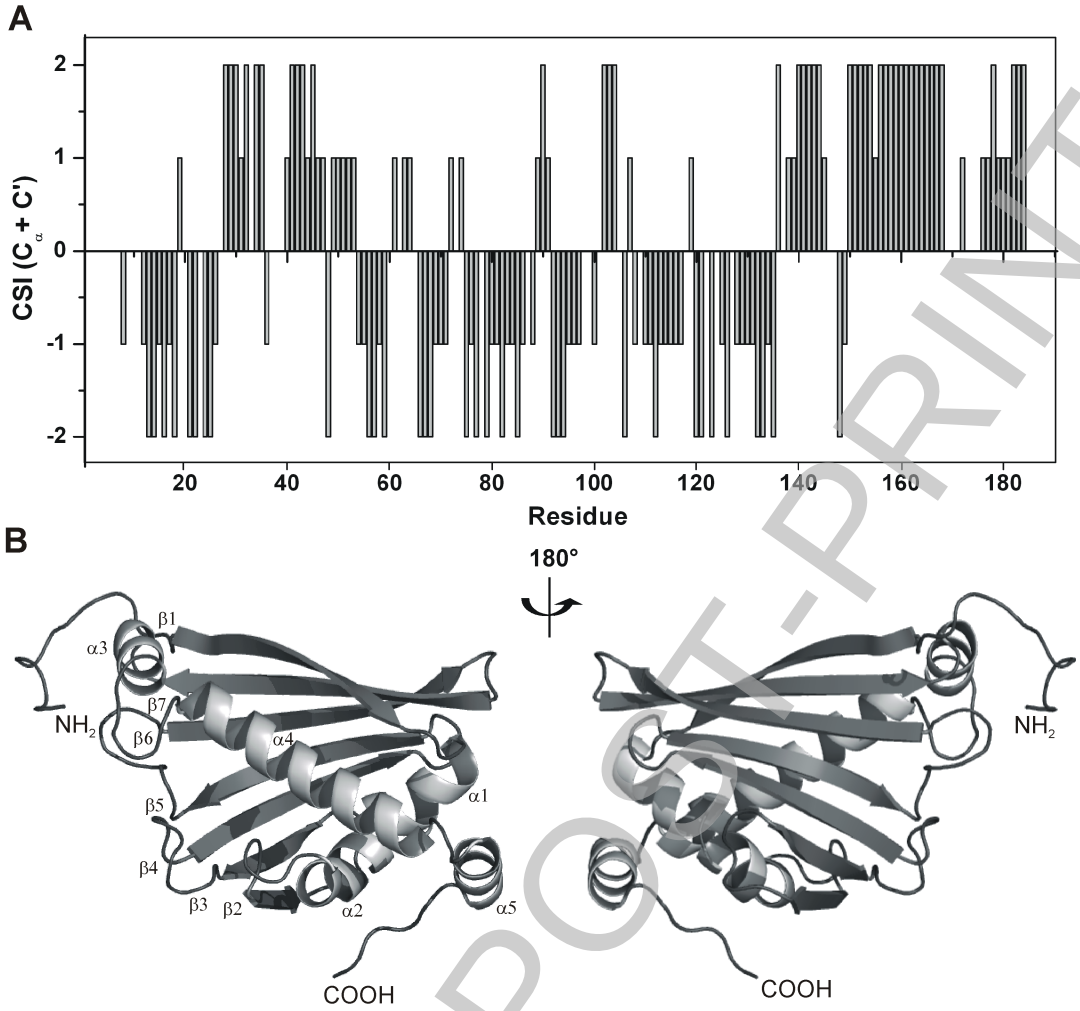


Figure 5

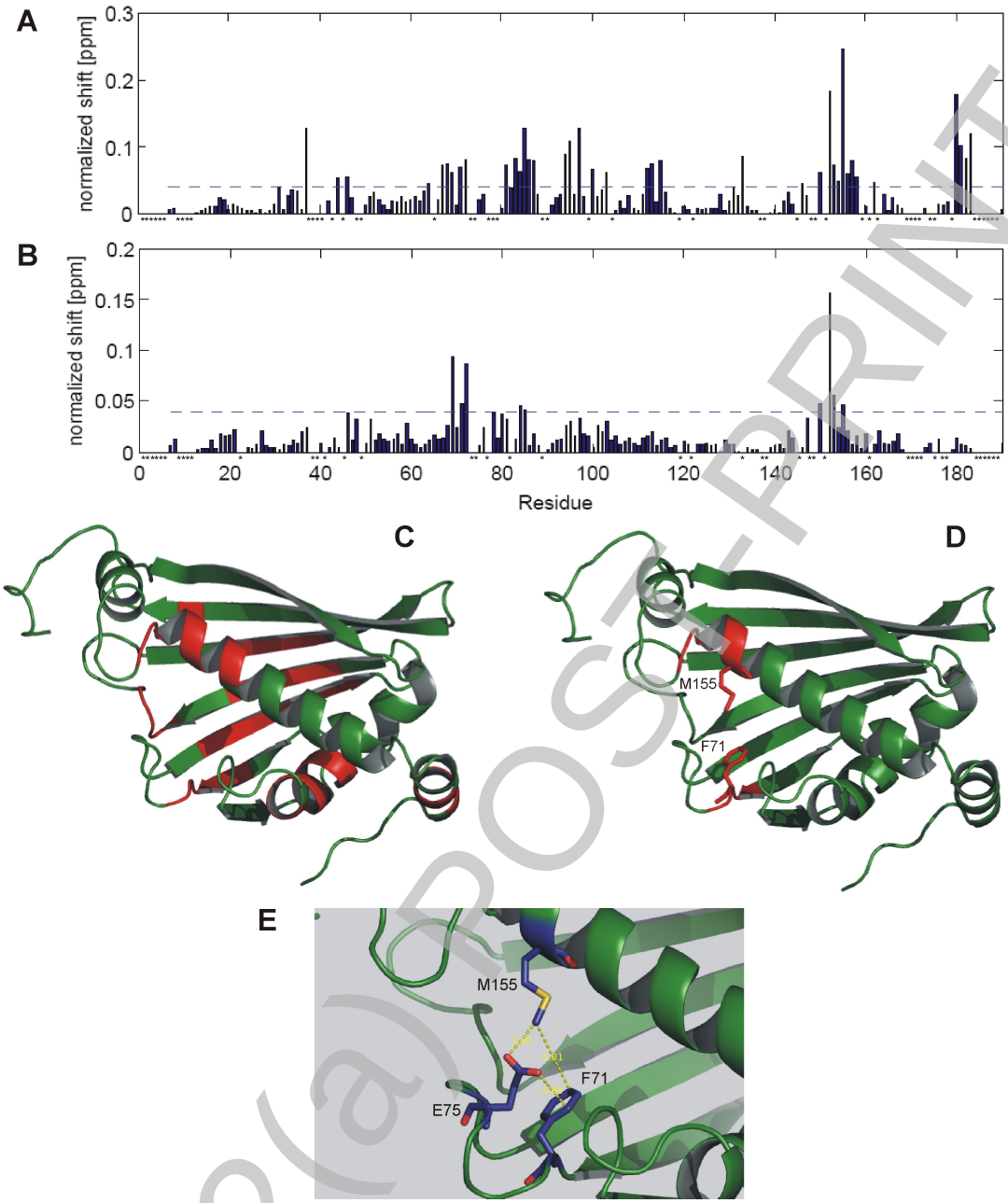


Figure 6