Structural Basis of Enzymatic (S)-Norcoclaurine Biosynthesis**

Andrea Ilari†, Stefano Franceschini†, Alessandra Bonamore‡, Fabio Arenghi‡, Bruno Bottà‡, Alberto Macone‡, Alessandra Pasquò, Luca Bellucci**, and Alberto Boffi†

From the †Dipartimento di Scienze Biochimiche and Istituto di Biologia e Patologia Moleculari, CNR, Sapienza University, Piazzale Aldo Moro 5, 00185 Roma, Italy, ‡CPC Biotech s.r.l., via dei Mille 111, 00121, Napoli, Italy, the †Dipartimento di Chimica e Tecnologia del Farmaco Sapienza University, Piazzale Aldo Moro 5, 00185, Roma, Italy, ‡ENEA Casaccia Research Centre, Dipartimento BIOTEC, Sezione Genetica e Genomica Vegetale via Anguillarese 301, I-00100 Roma, Italy, and the **Istituto Tecnico Chimico Farmacologico, University of Siena, via Aldo Moro 2, 53110, Siena, Italy

The enzyme norcoclaurine synthase (NCS) catalyzes the stereospecific Pictet-Spengler cyclization between dopamine and 4-hydroxyphenylacetaldehyde, the key step in the benzylisoquinoline alkaloid biosynthetic pathway. The crystallographic structure of norcoclaurine synthase from Thalictrum flavum in its complex with dopamine substrate and the nonreactive substrate analogue 4-hydroxybenzaldehyde has been solved at 2.1 Å resolution. NCS shares no common features with the functionally correlated “Pictet-Spenglerases” that catalyze the first step of the indole alkaloids pathways and conforms to the overall fold of the Bet v1-like protein. The active site of NCS is located within a 20-Å-long catalytic tunnel and is shaped by the side chains of a tyrosine, a lysine, an aspartic, and a glutamic acid. The geometry of the amino acid side chains with respect to the substrates reveals the structural determinants that govern the mechanism of the stereospecific Pictet-Spengler cyclization, thus establishing an excellent foundation for the understanding of the finer details of the catalytic process. Site-directed mutations of the relevant residues confirm the assignment based on crystallographic findings.

Alkaloids are among the most important plant secondary metabolites, comprising approximately 12,000 compounds grouped into several families (1). Most alkaloids are derived from amines produced by the decarboxylation of amino acids such as histidine, lysine, ornithine, tryptophan, and tyrosine. The coupling of the amines to other metabolites represents the first step in the biosynthesis of alkaloids belonging to diverse families. So far, however, this important entry reaction has been fully characterized only for a few biosynthetic pathways.

Benzylisoquinoline alkaloids are tyrosine-derived compounds and include a number of biologically active substances that are widely employed as pharmaceuticals such as morphine, codeine, berberine, papaverine, etc. The enzymatic pathways leading to the amazing diversity of benzylisoquinoline derivatives have been shown to originate from a common route in which the first committed step consists of the Pictet-Spengler condensation of dopamine with 4-hydroxyphenylacetaldehyde (4-HPAA) to yield the benzylisoquinoline central precursor, (S)-norcoclaurine (Fig. 1). The reaction is highly stereospecific, and the chirality of the (S)-norcoclaurine product is essential to drive the intricate pathway of substrate stereoselective enzymatic reactions toward the terminal metabolites (2, 3).

The Pictet-Spengler reaction entails the acid-catalyzed electrophilic addition of an iminium ion to a substituted benzyl (benzylisoquinoline alkaloids) or indole (indole alkaloids) species. The reaction mechanism consists of a two-step process in which the iminium ion is generated first from the condensation between the aldehyde carbonyl and the phenylethyl amine (or tryptamine) substrate followed by a Mannich type cyclization to yield tetrahydro-benzylisoquinolines (or tetrahydro-β-carboline) (4). Enzymes that catalyze the synthesis of (S)-strictosidine (the most important natural β-carboline), a central intermediate for indole alkaloids, have been recently characterized from the structural and mechanistic point of view. Strictosidine synthase, which catalyzes the formation of (S)-strictosidine from tryptamine and secologalin, has been cloned from Catharanthus roseus and Rauwolfia serpentina (5). Co-crystallization of strictosidine synthase from R. serpentina in the presence of tryptamine or secologalin substrates yielded key structural information into substrates binding and orientation within the active site of the enzyme (6). These studies provided the first example of an enzyme-catalyzed Pictet-Spengler reaction within the indole alkaloids biosynthesis.

In contrast, little structural information is presently available for norcoclaurine synthases (NCS), the key enzymes in the pathway that leads to the biosynthesis of benzylisoquinoline alkaloids. NCS from Thalictrum flavum (NCS; EC 4.2.1.78) has been identified recently (7) and has been fully characterized in

* This work was supported by Project FIRB 2003 from Ministero dell’Istruzione dell’Università e della Ricerca Scientifica (to A. B., B. B., and F. A.), local university grants (to A. B. and B. B.), and the European Community Research Infrastructure Action under the F6P Structuring the European Research Area Programme (through the Integrated Infrastructure Initiative* Integrating Activity on Synchrotron and Free Electron Laser Science, Contract R II 3-CT-2004-50608). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

** The on-line version of this article (available at http://www.jbc.org) contains supplemental materials (Text and Figures).

The atomic coordinates and structure factors (codes 2VNE and 2VQ5) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

† To whom correspondence should be addressed: Dipartimento di Scienze Biochimiche, Sapienza University, P. le Aldo Moro 5, 00185 Roma, Italy. Fax: 39-06-4991-0990; E-mail: alberto.boffi@uniroma1.it.

‡ The abbreviations used are: 4-HPAA, 4-hydroxyphenylacetaldehyde; NCS, norcoclaurine synthase; SeMet, selenomethionine; PHB, 4-hydroxybenzaldehyde.
The expression of the SeMet NCS protein was designed, synthesized, and optimized for Escherichia coli codon usage by the Gene Optimizer Assisted Sequence Analysis (Genart-Arg). The NCS protein was truncated at the first 19 amino acids (12), with a N-terminal MTGS sequence and a His tag at the C terminus was subcloned into the NdeI and Xhol restriction site of the vector pET22-b. Transformation of chemically competent E. coli strain BL21DE3 and protein expression have been performed as described by Pasquo et al. (12). The protein has been purified using a 5-ml HisTrap Fast Flow column (GE) equilibrated with a 50 mM Tris HCl buffer at pH 7.5. A linear gradient of imidazole concentration from 0 to 0.5 M (buffered at pH 7.5) was applied. The protein eluted at an imidazole concentration comprised between 0.25 and 0.30 M (12).

Crystalization—SeMet NCS crystals grew under conditions similar to those identified for the His-tagged wild type protein (12). Protein concentration was —11 mg/ml, and crystallization buffer was 0.1 M acetate buffer at pH 4 containing 1.4 M ammonium sulfate and 0.2 M sodium chloride. Crystallization temperature was 298 K. The SeMet crystals were cryo-protected in a solution containing 75% v/v of the reservoir solution and 25% v/v of glycerol and mounted on nylon loops. Then the crystals were flash-frozen by quick submersion into liquid nitrogen.

Data Collection and Data Analysis of the SeMet Derivative—A three-wavelength multiple wavelength anomalous diffraction data set was collected from SeMet-NCS on the ID14-2 beamline at the synchrotron radiation source Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung (Berlin, Germany), using a CCD detector. Complete data sets were collected at 120° of ϕ rotation each were collected at the peak (λ = 0.97966 Å), inflection (λ = 0.97984 Å), and remote (λ = 0.97800 Å) wavelengths, at a temperature of 100 K. Each frame was collected with an exposure time of 2 s and a 1° oscillation range. The three data sets were processed with DENZO (13) and scaled with SCALEPACK (13). The three-wavelengths multiple wavelength anomalous diffraction data set of SeMet-NCS was further scaled and analyzed with SCALEIT from the CCP4 suite (14). The autoindexing procedure indicated that the crystals belonged to the trigonal space group P3,21 with unit cell dimensions a = b = 86.31 Å, c = 118.36 Å, α = β = 90°, γ = 120°. All three of the data sets were more than 96% complete with Rmerge values below 10%. All of the statistics of the scaling procedure are reported in Table 1. A value of V_M = 2.89 Å³ Da⁻¹ has been calculated according to Matthews (15), assuming 6 asymmetric units within the unit cell. Each asymmetric unit contained two monomers with molecular masses of 24 kDa each.

Structure Solution and Refinement of Norcoclaurine Synthase—A heavy atom site search and phase determination were performed by the program SOLVE (16) using data collected at the three wavelengths (indicated under “Data Collection and Data Analysis of the SeMet Derivative”) in the 20-2.7 Å resolution range. A single solution was found with six heavy atom sites with a Z score of 14.3 for a signal-to-noise ratio of 0.3, leading to a mean figure of merit of 0.46. The program ARP-WARP (17)
was used to improve the initial phase and build automatically part of the protein skeleton. The program disclosed the presence of two molecules in the asymmetric unit as indicated by the $V_m$ calculation. The phases thus obtained were improved using the solvent-flattening and 2-fold averaging density modification (18). Refinement was performed using the program REFMACS (19), which applies the maximum likelihood method. The refinement statistics are presented in Table 1. Model building was performed using the program COOT (20). The final model (a dimer) includes 327 residues (19–27 and 31–194 residues in monomer A and 40–194 in monomer B), each monomer contains four SeMet and 18 water molecules. The final $R_{	ext{free}}$ for all resolution shells (74–2.72 Å), calculated using the working set reflections (12882), is 23.9%, and the free $R$ value calculated using the test set reflections (676) is 28.5%. The final $R_{	ext{cryst}}$ calculated for the highest resolution shell (2.72–2.79 Å) using the working set reflections (898) is 33.7%, and the free $R$ value calculated using the test set reflections (30) is 49%. The quality of the model was assessed by the program PROCHECK (21). The most favored regions of the Ramachandran plot contained 88.4% of nonglycine residues. The atomic coordinates and the structure factors have been deposited in the Protein Data Bank (accession number 2VNE).

**Structure Solution and Refinement of Norcoclaurine Synthase in Complex with Dopamine and 4-Hydroxybenzaldehyde**—The crystals of SeMet norcoclaurine synthase, obtained as reported above, were soaked in the mother liquor solution containing dopamine (2 mM) and 4-hydroxybenzaldehyde (2 mM) at 294 K for 24 h. Attempts to soak the crystals using the natural substrate, 4-HPAA, were unfruitful because of the intrinsic instability of the substrate itself (half-life of ~1 h) under the soaking conditions described above. The data were collected as 0.65° oscillation frames using the CCD detector on the x-ray beamline ID-23-1 at the European Synchrotron Radiation Facility (Grenoble, France), at a wavelength of 0.972 Å at 100 K using 25% v/v of glycerol as cryoprotectant. Data analysis, performed with DENZO (13), indicated that the crystal belonged to the trigonal space group P3$_2$1 and had the following unit cell dimensions: $a = b = 86.474$, $c = 117.977$ Å. The data were scaled using SCALEPACK (13) and had an $R_{	ext{merge}} = 7.5\%$ and an $\chi^2 = 1.012$.

The structure was solved by molecular replacement using the monomer B of SeMet norcoclaurine synthase without ligands as search probe (Protein Data Bank entry 2VNE). The rotational and translational searches, performed with the program MOLREP (22) in the resolution range of 10–3.0 Å, produced a clear solution corresponding to a dimer in the asymmetric unit. Refinement was performed using the program REFMACS (19), and model building was performed using the program COOT (20) (Table 1). The final model includes 327 residues (19–27 and 31–194 residues in monomer A, 40–194 in monomer B), a dopamine molecule, two 4-hydroxybenzaldehyde molecules, 176 water molecules, two acetate anions, and five chloride anions. The final $R_{	ext{cryst}}$ calculated for the highest resolution shell (2.144–2.089 Å) using the working set reflections (2045) is 23.9%, and the free $R$ value calculated using the test set reflections (1508), is 26.67%. The final $R_{	ext{cryst}}$ calculated for the highest resolution shell (2.144–2.089 Å) using the working set reflections (2045) is 26.5%, and the free $R$ value calculated using the test set reflections (99) is 35.6%. The most favored regions of the Ramachandran plot contain 92.8% of nonglycine residues. The atomic coordinates and the structure factors have been deposited in the Protein Data Bank (2VQ5).

**Catalytic Activity and Site-directed Mutagenesis**—K122A, E110A, Y108F, and Y108A mutants were obtained on the same PET 22-b vector, expressed and purified as described under “Protein Expression and Purification of SeMet Derivative.” The Y108A mutant, however, resulted in the expression of insoluble, possibly unfolded protein and could not be recovered in soluble form by guanidine denaturation/refolding experiments. The three soluble mutants and the wild type protein were thus analyzed for enzymatic activity using the circular dichroism-based assay proposed by Luk et al. (9). A Jasco J-715 spectropolarimeter was used for kinetic measurements (see supplemental

<table>
<thead>
<tr>
<th>Complex</th>
<th>Peak</th>
<th>Inflection</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P3$_2$1</td>
<td>P3$_2$1</td>
<td></td>
</tr>
<tr>
<td>Unit-cell parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a$ (Å)</td>
<td>86.41</td>
<td>86.31</td>
<td></td>
</tr>
<tr>
<td>$b$ (Å)</td>
<td>86.47</td>
<td>86.38</td>
<td></td>
</tr>
<tr>
<td>$c$ (Å)</td>
<td>118.00</td>
<td>118.36</td>
<td></td>
</tr>
<tr>
<td>No. of molecules in ASU</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>$&lt;R&gt;$ for atomic model (Å$^2$)</td>
<td>37.4</td>
<td>58.7</td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>2.1</td>
<td>2.70</td>
<td>2.85</td>
</tr>
<tr>
<td>Total observations</td>
<td>94324</td>
<td>77375</td>
<td>69000</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>30,118 (2978)</td>
<td>14,034 (1344)</td>
<td>12,015 (1169)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.9 (98.3)</td>
<td>98.9 (97.7)</td>
<td>99 (98.4)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.3 (3.1)</td>
<td>5.5 (4.8)</td>
<td>5.8 (5.3)</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>0.075 (0.51)</td>
<td>0.114 (0.51)</td>
<td>0.088 (0.55)</td>
</tr>
<tr>
<td>$C_{	ext{Rfree}}$</td>
<td>1.01 (0.84)</td>
<td>1.3 (0.9)</td>
<td>1.0 (0.8)</td>
</tr>
<tr>
<td>$&lt;I/i(s)&gt;$</td>
<td>13.7 (2.0)</td>
<td>16.4 (2.5)</td>
<td>17.0 (3.0)</td>
</tr>
<tr>
<td>$R_{	ext{cryst}}$</td>
<td>22.07</td>
<td>23.9</td>
<td></td>
</tr>
<tr>
<td>$R_{	ext{free}}$</td>
<td>26.67</td>
<td>28.5</td>
<td></td>
</tr>
<tr>
<td>Residues in core region of Ramachandran plot (%)</td>
<td>92.8</td>
<td>88.4</td>
<td></td>
</tr>
<tr>
<td>Residues in most allowed region (%)</td>
<td>7.2</td>
<td>11.6</td>
<td></td>
</tr>
<tr>
<td>Residues in generously allowed region (%)</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

* $R_{merge} = \sum_{i} \frac{\sum_{hkl} I(hkl) - \langle I(hkl) \rangle}{\sum_{hkl} I(hkl)}$ where $I(hkl)$ is the $i$th observation of the reflection (hkl), and $\langle I(hkl) \rangle$ is the mean intensity of the (hkl) reflection.
Structure of Norcoclaurine Synthase

Dopamine hydrochloride was obtained from Sigma-Aldrich, 4-hydroxyphenylacetaldehyde was prepared from tyrosine according to the method of Hazen et al. (Ref. 23; see also supplemental materials paragraph 1). Kinetic data were obtained by following CD signal changes at 285 nm and 37 °C. Data analysis and fitting procedures were carried out with the Matlab program (The Math Works inc. South Natick, MA) according to a simple Michaelis and Menten scheme. Substrate consumption and (S)-norcoclaurine formation were also measured by gas chromatography-mass spectrometry using an Agilent 6850A gas chromatograph coupled to a 5973N quadrupole mass selective detector (Agilent Technologies, Palo Alto, CA). Ethyl acetate extracts at the end of the reaction were also analyzed by electrospray ionization mass spectrometry apparatus (Thermo Finnigan LXQ) equipped with a linear ion trap analyzer.

RESULTS

Analysis of the x-ray structures 2VNE and 2VQ5 revealed a dimeric assembly in the asymmetric unit. A crystallographic tetramer composed of two asymmetric units, each containing a single dimer (A1-B1 or A2-B2), is represented in Fig. 2. Each NCS monomer revealed that NCS conforms to the overall fold adopted by proteins belonging to the Bet v1-like superfamily, which includes plant phytohormone carriers, pathogen-related proteins (PR10), MLN64-START domains, and the recently characterized tetracenomycin aromatase/cyclase (24, 25). Accordingly, the ensemble of secondary elements consists of a seven-stranded antiparallel β-sheets wrapped around a long C-terminal helix (α3) and two smaller α-helical segments (α1 and α2) (26). However, the C-terminal helix is longer in NCS than in Bet-v1 homologous proteins and is formed by two helical segments joined by an extended stretch (residues 173–177). Further, NCS has an additional N-terminal domain that forms a short α-helix and a long flexible segment that yields interpretable diffraction patterns only in one monomer within a single dimeric asymmetric unit. This segment folds in a β-strand secondary element that stabilizes the interface of the crystallographic 2-fold symmetry related dimer within the tetramer (Fig. 2).

Each monomer shows an accessible cleft, located between the seven-stranded antiparallel β-sheets and the three α-helices, that extends through the protein matrix forming a 23.4-Å-long tunnel (Fig. 3a). The wider opening (4.2 Å diameter), is formed by an array of hydrophobic residues and a polar patch composed by Tyr108, Tyr131, Tyr139, and Glu103 side chains located at the entrance of the cavity. Deeper in the cavity, the side chain of Lys122 protrudes toward the interior of the tunnel forming a “hook” capable of intercepting the carbonyl group of the aldehyde substrate (Fig. 3b). In correspondence with Lys122, the tunnel is thus restricted to a diameter of 1.2 Å. The smaller opening of the catalytic tunnel (3.2 Å) lies besides the Lys122 side chain and Asn117 and is solvent-accessible. X-ray data obtained on crystals soaked with the dopamine substrate and the nonreactive substrate analogue 4-hydroxybenzaldehyde (PHB) indicate that the two molecules adopt a stacked configuration with the respective aromatic rings lying on almost parallel planes (Figs. 4 and 5). PHB carbonyl oxygen is hydrogen-bonded to Lys122 amino group, whereas the phenolic oxygen is in contact (2.44 Å) with the carboxyl moiety of Asp441. Dopamine is hold in place by the stacking interaction with PHB and by hydrogen bonding of the C-1 phenol hydroxyl with the Tyr108 phenol hydroxyl. Most significantly, dopamine C-5 carbon atom lies at 2.7 Å from the carboxyl group of Glu140, suggesting a key role for this residue in the catalytic mechanism (Fig. 6). Comparison between the structures of the unliganded (2VNE) and substrate-bound (2VQ5) derivatives brings about that only very small adjustments of the residues just described occurred upon substrates binding in the crystal state.

Enzyme kinetic measurements on the wild type protein and on the three mutants K122A, E110A, and
Y108F are summarized in Table 2 and described in detail within the supplemental materials (paragraph 3). No signal relative to the formation of (S)-norcoclaurine was observed in the K122A mutant, whereas lower but significant activity was detected in both Y108F and E110A mutants. $k_{cat}$ values obtained on the wild type enzyme were lower than those obtained by Luk et al. (9), possibly because of variability in the determination of the absolute protein concentration.

**DISCUSSION**

The enantioselective synthesis of plant indole or benzylisoquinoline alkaloids has always been a challenging task for organic chemists and has led, within the past three decades, to the development of innovative processes in asymmetric synthesis (27, 28). In general, the cyclization step of the Pictet-Spengler reaction is a typical acid-base-catalyzed mechanism in which an iminium ion (Schiff base) attacks an electron rich aromatic carbon with subsequent release of the aromatic proton. The propensity of the iminium carbon to form a $\Lambda$ bond with the appropriate aromatic carbon is enhanced in the case of indole species because of the acidity of the 1 and 2 C-H bonds. Thus, the first step in indole alkaloids synthesis is thermodynamically and kinetically favored under physiological conditions, and catalysis is required essentially to drive the correct stereochemistry in the ring closure step. In contrast, harsher reaction conditions are required in the case of substituted phenylethyl derivatives because of the lower acidity of the benzenoid proton (27). In this case, the positional contribution of ring substituents and their electron donating or withdrawing effects play a key role in determining the reaction rate of the ring closure step. In this framework, it is of extreme interest to understand the strategy employed by nature to evolve enzymes capable of catalyzing the latter type of Pictet-Spengler reaction. The present data offer a fairly complete view of the structural determinants that govern the reaction in the NCS enzyme.

The geometry of the NCS active site is dominated by the presence of three strong proton exchanger, Lys$^{122}$, Asp$^{141}$, and Glu$^{110}$, and of a hydrogen bonding donor, Tyr$^{108}$. These residues shape the binding site of the two aromatic substrates
and dictate the mechanism proposed in Fig. 6. The reciprocal orientations of dopamine and PHB (Fig. 3) and their relationships with neighboring amino acid side chains immediately suggest a general acid-base reaction mechanism that matches closely the classical two-step Pictet-Spengler scheme and eventually leads to the stereoscopic ring closure to yield (S)-norcoclaurine. The presence of a strong interaction (2.6 Å) between the amino group of Lys122 and the carbonyl oxygen of the aldehylaurine. The presence of a strong interaction (2.6 Å) between the amino group of Lys122 and the carbonyl oxygen of the aldehyde, coupled to the off-plane position of the carbonyl with respect to the phenyl ring, is suggestive of a proton transfer from the ammonium ion to the carbonyl oxygen and consequent stabilization of a partial positive charge on the carbon atom. Such a configuration supports the idea that Lys122 is also involved in the water molecule release from the carbionolamine moiety subsequent to the nucleophilic attack of the dopamine amino group to the aldehyde carbonyl (Fig. 6, a and b). The electrophilicity of the imine double bond thus formed is the driving force of the subsequent Mannich type cyclization. Ring closure entails a rotameric rearrangement of the iminium ion adduct (steps c and d of Fig. 6) followed by electrophilic substitution at the C-5 position (steps e and f of Fig. 6). The rotameric arrangement (clockwise rotation of the bond connecting the iminium nitrogen and the adjacent dopamine carbon atom) occurring within steps c and d is a prerequisite for the effective ring closure in that it brings the iminium carbon atom in proximity of the ring C-5 atom. On the basis of the starting configuration of the two substrates, it can be inferred that the protein exerts a steric constraint on the adduct (step c) by allowing free rotation in the clockwise direction only. After deprotonation, assisted by the carboxyl moiety of Glu110, the S-stereospecific product is formed (Fig. 6f). The scheme indicated in Fig. 6 is in fair agreement with the mechanism proposed by Luk et al. (9), based on kinetic isotope effects. The authors suggested correctly that step d is driven by the transient formation of a phenolate ion (C-2 phenolate) that attacks the iminium ion in the first step of the aromatic substitution process. Thus, phenolate formation is judged essential to favor the release of the C-5 proton. However, no proton acceptor is found in the vicinity of the C-2 dopamine hydroxyl that could serve to stabilize the phenolate species. It may be envisaged that solvent water molecules from the wider opening of the catalytic tunnel eventually scavenge the phenolic C-2 proton. In turn, the Tyr108 phenol hydroxyl is hydrogen-bonded to the C-1 hydroxyl group of the dopamine substrate, thus increasing the partial positive charge on the C-1 oxygen. This interaction may favorably contribute to the transient loss of aromaticity necessary to the ring closure step (Fig. 6d). Thus, the cyclization step may be envisaged as a concerted process in which Glu110 acts as a base on the catecholate moiety transiently stabilized by the Tyr108 hydrogen bonding on the C-1 hydroxyl. The mechanism thus proposed is exquisitely stereospecific in that, given the position of Glu110 with respect to the dopamine ring orientation, C-5 proton abstraction may only occur from a single possible configuration of the intermediate (Fig. 6e).

The proposed scheme is supported by the kinetic data on selected site-specific mutants given in Table 2. In particular, alanine substitution of residue Lys122 completely abolishes the stereoselective synthesis of (S)-norcoclaurine. The residual activity observed in Y108F and E110A mutants may be interpreted as impaired dopamine binding, as demonstrated by the significant increase in the $k_{cat}$ value for dopamine in both mutants. Alternatively, the lack of hydrogen bonding to the C-1 hydroxyl of dopamine and the consequent destabilization of the phenolate intermediate may be taken into account for the reduced activity of the Y108F mutant. The kinetic data, however, were obtained by monitoring selectively the (S)-norcoclaurine formation by CD spectroscopy and do not allow quantitative discrimination on the stereochemical control of the reaction in the mutants with respect to the wild type protein. A careful assessment of the enantiomeric excess of (S)-norcoclaurine chiral product versus the R product (still not reported even for the native protein) will be necessary to distinguish between background reaction contributions and possible nonstereoselective catalysis in some of the mutants. In this framework, the finer details of NCS catalytic mechanisms, such as the nature of the rotameric arrangement (steps c and d of Fig. 6) and the role of Tyr108 on the electronic configuration of the dopamine substrate, still need to be fully clarified. Moreover, further studies will be necessary to understand the observed cooperativity in enzyme kinetics (7–9). At present, only minor ligand-linked tertiary structural changes have been detected at the active site (different rotameric arrangement of Phe122 in subunit A versus subunit B), whereas the quaternary structures of the substrate-free and substrate-bound proteins appear to be essentially superimposable.

On the basis of the present understanding of S-norcoclaurine biosynthesis, it is of interest to compare the proposed reaction scheme and the structural determinants that govern the catalytic mechanism of NCS to those established for the functionally related strictosidine synthase STR1. As shown in Fig. 7, the basic tenet that envisage an acid-base catalysis as a common conduit for Pictet-Spengler cyclizations is respected in both NCS and STR1. However, the strategy for achieving the asymmetric condensation is reversed in the two enzymes. NCS employs the positive charge of Lys122 as a strong polarizing agent for the carbonyl group of the aldehyde substrate upon which the amine substrate acts as a nucleophilic agent. Conversely, STR1 employs the negatively charged carboxyl moiety of Glu309 to hold in place and eventually deprotonate the nitrogen atom of the amine substrate that subsequently reacts with the incoming aldehyde substrate. It might be inferred, however, that also NCS displays a glutamic acid residue (Glu110) within the catalytic site, and hence alternative reaction schemes could be envisaged that entail the direct participation of the carboxyl...
Structure of Norcoclaurine Synthase

...molecule (a cytokinin with adenine scaffold) within the tunnel of zeatin binding protein from Vigna radiata provides a particularly striking example of structural analogy with substrate-bound NCS (26). In fact, the zeatin-binding site is formed within the large cavity inside the protein molecule, between the β-sheet and the C-terminal helix α3. The zeatin molecule is accommodated within the binding site by hydrogen bonding interactions with a threonine (Thr139) and a glutamate (Glu69) and is stabilized by extensive van der Waals’ contacts with aromatic side chains lining the bottom of the binding cavity (Phe26, Phe56, Phe102, Tyr98, and Tyr142). Comparison of the amino acid sequences among members of the cytokinin-binding protein family also reveals that the residues forming hydrogen bonds and hydrophobic residues engaged in van der Waals’ interactions with the inner ligand are conserved. It follows that the topological positions important for cytokinins binding are at least partially superimposable to those relevant for catalysis in NCS. This is not true, in general, for other members of the PR10 superfamily, because structural data show high variability in shape, volume, and chemical properties of the ligand binding tunnel. More structural data on members of this protein family will be necessary to unveil whether there are common features in the topology of residues involved in ligand binding or catalysis.

Acknowledgments—We thank the European Synchrotron Radiation Facility and the Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung (Berlin, Germany) where the data were collected. In particular, we thank Dr. Georg Zocher for help in multiple wavelength anomalous diffraction data collection.

REFERENCES

Structure of Norcoclaurine Synthase

demic Publishers, Dordrecht, The Netherlands
logr. 52, 43–4814
logr. 60, 2126–2132
1022–1025
23. Hazen, S. L., Gaut, J. P., Hsu, F. F., Crowley, J. R., d’Avignon, A., and
Nat. Struct. Biol. 3, 1040–1045
25. Ames, B. D., Korman, T. P., Zhang, W., Smith, P., Vu, T., Tang, Y., and
26. Pasternak, O., Bujacz, G. D., Fujimoto, Y., Jelen, F., Otlewski, J., Sikorski,
1086–1087
25, 1989) U. S. Patent 4851537
Carlos, CA
30. Petřek, M., Otýpka, M., Banaš, P., Košinová, P., Koča, J., and Damborský,
J. (2006) BMC Bioinformatics 7, 316