Yeast β-Alanine Synthase Shares a Structural Scaffold and Origin with Dizinc-dependent Exopeptidases*

Stina Lundgren‡, Zoran Gojković§, Jure Piškur§, and Doreen Dobritzsch‡¶

From the §Division of Molecular Structural Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-17177 Stockholm, Sweden and the §BioCentrum-DTU, Technical University of Denmark, DK-2900 Lyngby, Denmark

Received for publication, August 6, 2003, and in revised form, October 6, 2003
Published, JBC Papers in Press, October 8, 2003, DOI 10.1074/jbc.M308674200

β-Alanine synthase (βAS) is the final enzyme of the reductive pyrimidine catabolic pathway, which is responsible for the breakdown of pyrimidine bases, including several anticancer drugs. In eukaryotes, βASs belong to two subfamilies, which exhibit a low degree of sequence similarity. We determined the structure of βAS from Saccharomyces kluyveri to a resolution of 2.7 Å. The subunit of the homodimeric enzyme consists of two domains: a larger catalytic domain with a dizinc metal center, which represents the active site of βAS, and a smaller domain mediating the majority of the intersubunit contacts. Both domains exhibit a mixed α/β-topology. Surprisingly, the observed high structural homology to a family of dizinc-dependent exopeptidases suggests that these two enzyme groups have a common origin. Alterations in the ligand composition of the metal-binding site can be explained as adjustments to the catalysis of a different reaction, the hydrolysis of an N-carbamyl bond by βAS compared with the hydrolysis of a peptide bond by exopeptidases. In contrast, there is no resemblance to the three-dimensional structure of the functionally closely related N-carbamyl-D-amino acid amidohydrolases. Based on comparative structural analysis and observed deviations in the backbone conformations of the eight copies of the subunit in the asymmetric unit, we suggest that conformational changes occur during each catalytic cycle.

In most living organisms, the degradation of uracil and thymine is achieved by the three-step reaction sequence of the reductive pyrimidine catabolic pathway. The first and rate-limiting enzyme is dihydropyrimidine dehydrogenase (dihydrouracil dehydrogenase (NADP+), EC 1.3.1.2), which catalyzes the NADPH-dependent reduction of uracil and thymine to the corresponding dihydropyrimidines (Scheme 1). In the second step, dihydropteridinase (EC 3.5.2.2) generates N-carbamyl-β-alanine and N-carbamyl-β-aminooxyacid, respectively, via reversible hydrolytic cleavage of the dihydropyrimidine ring. β-Alanine synthase (βAS; N-carbamyl-β-alanine amidohydrolase and β-ureidopropionase, EC 3.5.1.6), catalyzes the third and final step: the hydrolysis of the N-carbamylated β-amino acids to β-alanine or aminoisobutyrate under the release of carbon dioxide and ammonia (1).

The key role of the pyrimidine catabolic pathway is the regulation of the pyrimidine pool in the cell and thus the maintenance of a balanced supply of precursors for nucleic acid synthesis. This pathway is also the main clearance route for cytotoxic pyrimidine analogs such as 5-fluorouracil, which are widely used for the treatment of a variety of common tumors such as colorectal, head/neck, and breast cancer (2). The enzymes of the pyrimidine catabolic pathway have therefore a considerable influence on the efficacy and pharmacokinetics of these anti-tumor agents (3). In addition, the pathway provides the cell with β-alanine, which is an essential component of pantothentic acid (vitamin B₃) and hence of coenzyme A and acyl carrier protein (4). Biosynthesis of pantothentic acid occurs in bacteria, yeast, and plants, whereas animals obtain it through their diet. Nevertheless, β-alanine biosynthesis is also of critical importance in animal tissue. Because it is widely distributed in the central nervous system of vertebrates and is a structural analog of γ-amino-n-butyric acid and glycine, major inhibitory neurotransmitters (5), β-alanine itself has been suggested to be involved in synaptic transmissions. Recent studies showed that β-alanine activates the same system of receptors on the membranes of rat sacral dorsal commissural nuclei as glycine (6). In humans, abnormal β-alanine production is associated with neural dysfunction, seizures, and death (7). The reductive degradation of uracil and thymine represents the major pathway providing β-alanine in plants and filamentous fungi and the sole source of β-alanine in mammalian tissues (8), whereas microorganisms produce it mostly via direct α-decarboxylation of L-aspartate (9).

βAS is an amidohydrolase that acts on an N-carbamylated β-amino acid substrate. Based on the comparison of the primary structures and sequence phylogenetic analysis of βASs and related amidohydrolases, these enzymes can be assigned to one of three subfamilies (10). The majority of eukaryotic βASs belong to one subfamily, with pairwise sequence identities as low as 55%. βAS from the yeast Saccharomyces kluveri (SkβAS) shares only a limited sequence similarity with these enzymes, but is instead closely related to bacterial N-carbamyl-

---

The abbreviations used are: βAS, β-alanine synthase; SkβAS, S. kluveri β-alanine synthase; DCases, N-carbamyl-D-amino acid amidohydrolase; SmMet, selenomethionine; MES, 4-morpholineethanesulfonic acid; SGAP, S. griseus aminopeptidase; APAP, A. proteolytica aminopeptidase; CPG2, carboxypeptidase G2 (Pseudomonas sp.); PepT, peptidase T (S. typhimurium); PepV, peptidase V (L. delbrueckii).

*This work was supported by grants from the Swedish Research Council, the Wallenberg Konsortium North, the Royal Swedish Academy of Sciences, the Foundation Lars Härtas Minne, the Karolinska Institutes Research Foundation, and the Danish Research Council. 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 atomic coordinates and structure factors (code 1R3N and 1R43) 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: Dept. of Medical Biochemistry and Biophysics, Karolinska Institutet, Tomtebodavägen 6, S-17177 Stockholm, Sweden. Tel.: 46-8-728-7278; Fax: 46-8-32-7628; E-mail: Doreen.Dobritzsch@mbb.ki.se.
t-aminoc acid amidohydrolases (30–40% sequence identity), which belong to the second subfamily. Although the involvement of the bacterial enzymes in pyrimidine catabolism is questionable because many of these enzymes cannot use N-carbamyl-β-alanine as substrate, genetic studies have clearly shown that SkβAS represents the final enzyme in the pyrimidine degradation pathway. A yeast βAS mutant was no longer able to grow on N-carbamyl-β-alanine as the sole nitrogen source (10, 11). However, the presence of a functional pyrimidine degradation pathway makes S. kluiveri exceptional among the Saccharomyces species. The third subfamily of amidohydrolases consists of bacterial and archaeabacterial N-carbamyl-α-aminoc acid amidohydrolases (DCases), with a pairwise sequence identity as low as 20%. With the recent determination of the crystal structures of the DCases from two Agrobacterium species, this subfamily has thus far been the only one for which structural models are available (12, 13).

Hydrolytic reactions similar to that catalyzed by βAS frequently involve a divalent metal ion. Metal analyses performed for the rat and maize enzymes suggested that both contain Zn²⁺ ions (14, 15). Furthermore, kinetic studies demonstrated that rat βAS is an allosteric enzyme with positive cooperativity toward the substrate N-carbamyl-β-alanine, which triggers a change in the oligomeric state from homohexamer to homodecamer, whereas the presence of the reaction product β-alanine induces the dissociation into inactive trimers (16). The two characterized plant βASs exist as homodecamers at pH 7, with subunit molecular masses comparable with those of the mammalian enzymes (43–45 kDa) (15). In contrast, SkβAS is a homodimer, and kinetic studies do not support an allosteric regulation of the enzyme. Analysis of its metal-binding capabilities remain inconclusive.²

Here, we report the crystal structure of recombinant SkβAS determined to 2.7-Å resolution by the multiwavelength anomalous dispersion method. The structure reveals a binuclear zinc center representing the active site of the enzyme. The overall fold is not related to that of bacterial DCases of the third subfamily of amidohydrolases, but to a family of dizinc-dependent exopeptidases.

EXPERIMENTAL PROCEDURES

Purification and Crystallization—The overexpression and purification of native and selenomethionine (SeMet)-substituted SkβAS were carried out as described (10, 17). Crystals of both proteins were obtained at 20 °C by the hanging drop vapor diffusion method. Briefly, a 6-μl drop of an equal volume of the protein solution (4.0–4.5 mg/ml in 50 mM Tris (pH 7.5), 1 mM dithiothreitol, and 100 mM NaCl) and of the reservoir solution was equilibrated against 1 ml of either 0.88 M trisodium citrate, 0.1 M sodium citrate (pH 6.0), and 5% (w/v) dioxane or 1.0 M trisodium citrate, 0.1 M MES (pH 6.5), and 5% (w/v) dioxane. The crystals appeared after 5–7 days and belonged to space group P2₁.

Crystals of native SkβAS showed cell parameters of a = 117.2 Å, b = 77.1 Å, c = 225.5 Å, and β = 95.0°, with eight molecules in the asymmetric unit. SeMet-substituted SkβAS crystals exhibited a 4-fold smaller unit cell (a = 61.0 Å, b = 77.9 Å, c = 110.1 Å, and β = 97.2°), with only two molecules in the asymmetric unit.

Data Collection and Phasing—Prior to data collection, all crystals of SkβAS were flash-frozen in a nitrogen gas stream at 100 K. A cryoprotectant was not required. Native data were collected to 2.7-Å resolution on beamline ID14-1 at the European Synchrotron Radiation Facility (Grenoble, France) using an ADSC Q4R CCD detector. A multwavelength anomalous dispersion experiment was performed on a single crystal of SeMet-substituted SkβAS. A fluorescence spectrum was recorded, and data for multiwavelength anomalous dispersion phasing were collected to 2.8-Å resolution at beamline ID14-4 at the European Synchrotron Radiation Facility at three different wavelengths: λ₁ = 0.9792 Å and λ₂ = 0.9794 Å near the Se K edge to maximize the anomalous f' and dispersive f' signals, respectively, and λ₃ = 0.9393 Å as a remote data set. All data sets were processed with MOSFLM and scaled with SCALA of the CCP4 suite of programs (18). Structure factors were derived from the reflection intensities with the program TRUNCATE (18). Details for the data collection and processing are given in Table I. Using only the peak data, 15 of the possible 16 selenium positions were automatically determined with the program SOLVE (19), and an initial electron density map and a partial Cα trace model were obtained after statistical density modification implemented in the program RESOLVE (20). All three data sets were combined and scaled using SCALEIT (18) prior to the refinement of the heavy atom positions, occupancies, and temperature factors with SHARP (21). The phasing statistics are given in Table I. The anomalous residual map calculated after the first refinement cycle revealed the position of the sixteenth selenium in the asymmetric unit, which was included in the refinement. The resulting electron density map was further improved by density modification and non-crystallographic symmetry averaging using DM (22). The protein model for SeMet-substituted SkβAS was built with the graphics program O (23), and the first cycle of rigid body refinement, simulated annealing, and positional and temperature factor refinement was performed with the program CNS (24) using all reflections in the resolution range of 30.0 to 2.5 Å, with the exception of 5% randomly selected reflections, which were used for monitoring R cryst. Cycles of manual corrections of the model and TLS and restrained maximum likelihood refinement as implemented in REFMAC Version 5.0 (25) were repeated, with tight main chain and medium side chain non-crystallographic symmetry averaging restraints for the dimerization domain and medium main chain and loose side chain non-crystallographic symmetry averaging restraints for the catalytic domain. A total of 68 water molecules were added, and one molecule of each Tris and β-aminoisobutyrate were fitted into appropriately shaped remaining density peaks of an |Fo| - |Fc| map. Furthermore, one dithiothreitol molecule was added to the model, covalently attached via a disulfide bridge to Cys²⁶⁸ in chain A, as the modification became apparent in the electron density map. The values for R cryst and R free after refinement were 23.3 and 28.1%, respectively.

An early model for SeMet-substituted SkβAS was used to solve the structure of the native enzyme by molecular replacement with the program EPMR (26). Based on the observation that the two protomers in the asymmetric unit of SeMet-substituted SkβAS crystals are not structurally equivalent due to different angles between the domains, the following search protocol was used. First, we positioned four dimers of the dimerization domain in the asymmetric unit of the native SkβAS crystals and thereafter the eight catalytic domains. The structure of native SkβAS was refined to 2.7-Å resolution using one cycle of rigid

² S. Lundgren, Z. Gjojkovic, J. Piaskur, and D. Dobritsch, unpublished data.
body refinement in REFMAC Version 5.0 and then alternating cycles of model building with the program O and TLS and restrained refinement (REFMAC Version 5.0), resulting in final values of 20.8% for Rcryst and 26.6% for Rfree. The same sets of non-crystallographic symmetry averaging restraints as for the SeMet-substituted protein were applied. Both models have good stereochemistry, as determined by the program PROCHECK (27). There are no non-glycine residues with forbidden combinations of dihedral angles of the peptide backbone. The refinement statistics for both structures are given in Table I.

Comparisons between SkβAS and structurally related aminopeptidases from Streptomyces griseus and Aeromonas proteolytica (Protein Data Bank codes 1FZO and 1CP6, respectively), carboxypeptidase G2 (code 1CG2), peptidases T and V (codes 1FNO and 1LFW, respectively), and aminopeptidase A (code 1GYT) were accomplished using the program TOPO (28) with default parameters. Figs. 1 – 3 were generated using BOBSCRIPT (29, 30) and RASTER3D (31). The crystallographic data have been deposited in the Protein Data Bank, with code 1R3N for the native structure and code 1R43 for SeMet-substituted SkβAS.

RESULTS AND DISCUSSION

Structure Determination

The structure of SkβAS has been determined in two different crystal forms, both belonging to the monoclinic space group P21. Comparison of the final models revealed that the backbone architectures of the subunits are essentially the same, and the observed differences in domain arrangement (discussed below) between the subunits of SeMet-substituted SkβAS do not exceed those observed in the crystals of native SkβAS. For this reason, the following analysis of the three-dimensional structure is based on the model for the larger unit cell, which has been refined to a slightly higher resolution.

The SkβAS gene encodes a polypeptide chain of 455 amino acids. The electron density for the large cell model is continuous from residue 19 in chain D, residue 23 in chains B and E, residue 26 in chain G, and residue 18 in the remaining chains. The lack of electron density for the N-terminal stretches of amino acids can be attributed to their mobility because N-terminal sequencing of SkβAS from dissolved crystals revealed that these residues (except Met1) are present in the protein. The C-terminal tag of eight histidine residues does not adopt a defined structure and is not visible in the map. The final model also includes two Zn2+ ions/subunit and a total of 519 water molecules. The model for SeMet-substituted SkβAS also contains one molecule of Tris bound close to Arg51 on the surface of the structure and is not visible in the map. The model for SeMet-substituted SkβAS also contains one molecule of Tris bound close to Arg51 on the surface of the chain A and one molecule of dithiothreitol covalently attached to Cys168 near the active site of the same chain.

Overall Structure

The three-dimensional structure of the SkβAS subunit is illustrated schematically in Fig. 1a. The subunit consists of two domains: a larger catalytic domain that contains all zinc- and most putative substrate-binding residues and a smaller dimerization domain.

Catalytic Domain—The catalytic domain comprises residues 2–246 and 365–455 and has a mixed three-layer α/βα-sandwich architecture. The central mixed β-sheet consists of eight strands (strand order: s1, s2, s4, s3, s5, s6, s7; strands s2, s6, and s7 are antiparallel to the others) and can be divided into three parts. The first half of the sheet forms a closely packed...
Fig. 1. Overall structure of SkβAS. a, stereo view of the subunit of SkβAS. Secondary structure elements are shown in pink for α- and 3₁₀-helices and in green for β-strands and are labeled as follows: h and g for α- and 3₁₀-helices, respectively (with successive numbering according to the sequence for each domain), and s for β-strands (successive numbering for those forming the central β-sheets in both domains and Roman numerals for additional β-strands). D indicates secondary structure elements of the dimerization domain. Purple spheres represent the zinc ions.
unit with only a slight twist between the strands. The positioning of the N terminus of helix h3 between the C termini of strands s3 and s5 separates this unit from the second one, comprising strands s5 and s8, which are more twisted than the first four strands. The curvature of the β-sheet at strands s6 and s7 is even more pronounced. Only the N-terminal part of strand s6 interacts with strand s8, whereas the C-terminal part is involved in hydrogen bonds to an additional strand, strand sIII, which is oriented almost perpendicular to the second unit. The central sheet is flanked on one side by four long α-helices, whereas two shorter α-helices are packed against its opposite face. Besides four additional α-helices, the catalytic domain also contains two 310-helical segments and a β-hairpin located on the surface of the protein.

**Dimerization Domain**—The dimerization domain (residues 247–364) is inserted between strands s6 and sIII of the catalytic domain and folds into a four-stranded antiparallel sheet flanked on one side by two long α-helices. A two-amino acid-long loop segment splits one of the strands into two (strands Ds2-1 and Ds2-2). A short 3_10-helix and two β-strands are the only additional secondary structure elements. There is one cis-peptide bond in the structure, between Lys304 and Pro305, which creates a sharp turn in the loop following strand Ds2-2.

**Dimer Interactions and Crystal Packing**

The interactions between the subunits in the SkβAS dimer involve van der Waals contacts and hydrogen bonding and lead to 2280 Å² (−12%) of buried surface of the monomer. The contacts are mediated mostly via structural elements of the dimerization domain (Fig. 1b). Strands Ds2-2 of both subunits are hydrogen-bonded to each other, but because the bonds do not involve backbone nitrogen and oxygen atoms, no continuous eight-stranded β-sheet is formed across the dimer interface. Helices Dh1 and Dg1 and strand DsI also participate in extensive monomer-monomer contacts, as well as strand DsII and its proceeding loop. Enabled by the cis-peptide bond between Lys304 and Pro305, the loop turns sharply and places itself and the following strand on top of the β-sheet in the dimerization domain of the partner subunit. Arg232, which is most likely involved in substrate binding (discussed below), forms hydrogen bonds with two residues from the other subunit in close proximity to the active site. In addition, several amino acids from the catalytic domain (positions 228–232, 235–236, 394–396, and 417–421) contribute to the interface.

The pairwise superposition of the dimer-forming subunits in the asymmetric unit of native SkβAS crystals reveals that their backbone structure is not entirely equivalent. The differences are mainly due to a rigid body rotation of the catalytic domain with respect to the dimerization domain. The two domain-separating amino acids Val324 and Val365 serve as hinge residues. For example, superposition of molecule B onto molecule A results in an overall root mean square deviation of 0.81 Å, whereas separate superpositions of their catalytic or dimerization domains give rise to much lower root mean square deviations of 0.35 and 0.07 Å, respectively. After superposition of all eight crystallographically independent subunits (Fig. 1c), the maximum distance between equivalent C-α atoms was measured for Thr365 and corresponds to 6 Å.

The differences in domain arrangement in the protomers eventually lead to asymmetry in the dimer. The proper 2-fold symmetry relating the dimerization domains does not apply to the catalytic domains, although the deviation is rather small.
(α = 178.9°, with a screw along the rotation axis of 2.2 Å as determined for the dimer formed by chains A and B).

For both crystal forms, we observed that, in each dimer, the electron density for the catalytic domain of one of the monomers (A/C/F/H) is better defined than for the same domain of the other monomer, indicating greater flexibility of the latter. Analysis of the crystal contacts gives a reasonable explanation for the greater flexibility of one of the catalytic domains in the dimer. It makes only about one-third of the number of crystal contacts compared with the other and is due to the hinge-like limited connection to the dimerization domain relatively free to adopt different positions in the crystal. In contrast, the electron density observed for the dimerization domain, which is extensively involved in intermolecular contacts, is well defined in all cases.

**Structural Homology to Dizinc-dependent Exopeptidases**

A structural similarity search using the TOP server (28) revealed a striking similarity of SkβAS to dizinc-dependent exopeptidases of the M20 and M28 families (Fig. 2) (32). The single domain structures of the aminopeptidases from Streptomyces griseus (SGAP) (33) and Pseudomonas sp. (APAP) (34) align well with the catalytic domain of SkβAS (Table II). In contrast, carboxypeptidase G2 (CPG2; Pseudomonas sp.) (35) and peptidases T (PepT; Salmonella typhimurium) (36) and V (PepV, Lactobacillus delbrueckii) (37) consist of two domains, which are both structurally related to their counterparts in SkβAS. The arrangement of the two domains relative to each other does, however, vary between the proteins. In CPG2 and PepT, the dimerization domain extends away from the active site in the catalytic domain, which, as a consequence, is considerably more solvent-accessible than that of SkβAS. For both enzymes, the deviation from the domain arrangement in SkβAS is based on a rigid body rotation of the dimerization domain in the same direction and by approximately the same angle. In contrast, the corresponding “lid” domain of PepV is rotated in the other direction with respect to SkβAS. Because the biologically active form of PepV is the monomer, the lid domain does not have a function in subunit oligomerization. However, with 202 residues, it is about twice as big as the dimerization domain of SkβAS, and its topology (comprising a central eight-stranded antiparallel β-sheet flanked on one side by four α-helices) resembles the structure of a dimer of SkβAS dimerization domains. As in SkβAS, residues of the lid domain are close enough to assist in the creation of a cavity, which harbors part of the substrate-binding site; but because of the different arrangement of this domain relative to the catalytic domain, the active-site accessibility is more restricted.

**Metal-binding Site**

Prior to the determination of the three-dimensional structure, no information was available about the location of the active site or the metal-binding capabilities of SkβAS. Two zinc sites had been proposed for rat βAS (14); however, the amino acids putatively involved in zinc binding are not conserved in SkβAS (10). The presence of a dimetal site in this enzyme became apparent after calculation of the initial electron density map, which showed two connected peaks of strong density. The high structural similarity of SkβAS to dizinc-dependent exopeptidases, the results of the metal analyses performed for other βASs (14, 15), and the type of ligands coordinating the metal ions suggested zinc as the most probable type of metal. After inclusion of two zinc ions/subunit in the model, the distances to the coordinating residues refined to those usually observed in zinc-binding sites. The presence of zinc was later confirmed by an x-ray absorption scan performed at beamline

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|}
\hline
 & Catalytic domain & & & Dimerization domain & \\
\hline
 & No. aligned & r.m.s.d. & No. aligned & r.m.s.d. & \\
\hline
 & C-o atoms & Å & & C-o atoms & Å & \\
\hline
SGAP & 186 & 1.70 & & & \\
APAP & 188 & 1.95 & & & \\
CPG2 & 161 & 1.75 & 89 & 1.53 & \\
PepT & 124 & 2.14 & 70 & 1.18 & \\
PepV & 161 & 2.34 & 51 & 1.81 & \\
\hline
\end{tabular}
\caption{Structural alignment of the two SkβAS domains with the corresponding domains of dizinc-dependent exopeptidases.}
\end{table}
The utilization of a dimetal center for activation of a water molecule is a common theme in enzyme catalysis of carbon–nitrogen bond hydrolysis. We therefore assume that the dizinc center represents the active site of SkβAS. It is situated on the surface of the cleft between the catalytic and dimerization domains and thus is freely solvent-accessible. With the exception of His226, which belongs to strand s5, all zinc-binding residues originate from loops decorating the C-terminal end of the β-sheet in the catalytic domain. The zinc coordination is illustrated in Fig. 3a. The carboxyl group of Asp125 serves as a bridging ligand between both zinc ions, which are separated by a distance of 3.4 Å. Zinc 1 is further coordinated by the imidazole groups of His114 and His397, and zinc 2 by His421 and the bridging water plus an additional zinc-coordinated solvent molecule (W1 and W2) are shown as smaller red spheres. A $|F_o| - |F_c|$ map calculated before inclusion of α-aminoisobutyrate in the model has been contoured at 3.0 $σ$. The reaction product has been fitted into the density and is shown with yellow carbon atoms. In contrast, carbon atoms of amino acids are gray. b, stereo view of the active sites of SkβAS and PepV after superposition of their catalytic domains compared with a shown in a different orientation. The carbon atoms of PepV residues are yellow, and those of the inhibitor Asp(PO2CH2)Ala-OH are magenta. Zinc-binding and putative substrate-binding residues of SkβAS are depicted with gray carbons. The zines are cyan for SkβAS and green for PepV. Zinc-bound solvent molecules in the active site of SkβAS are shown as red spheres. The coordination of the zinc ions and water molecules is indicated by dotted lines for SkβAS only. All amino acid residues of SkβAS are labeled. The corresponding residues of PepV are indicated in parentheses. The rigid body movement required for simultaneous binding of the SkβAS substrates to the dizinc center and the putative carboxylate-binding residues is indicated by arrows.
the carboxylate group of Glu\textsuperscript{160}. As a consequence, the two zinc ions are not equivalent because they do not share the same chemical environment. The moderate resolution of the collected data rendered it difficult to unambiguously determine whether or not a water molecule or hydroxyl moiety bridges the zinc ions. The refined distances and B-values for a bridging water molecule and the electron density map did, however, support a bound water/hydroxyl molecule. No peaks of negative density appear in the $|F_o| - |F_c|$ map for the bridging water after the refinement. Zinc 1 has a second water ligand (W2 in Fig. 3a), in turn forming hydrogen bonds with the carboxyl group of Glu\textsuperscript{159}, the imidazole of His\textsuperscript{107}, and another water, which mediates additional interactions with the side chains of Glu\textsuperscript{724} and Asp\textsuperscript{998} as well as with the backbone nitrogen of Asp\textsuperscript{398}. Thus, whereas the coordination of zinc 2 is distorted tetrahedral, zinc 1 has a five-member coordination sphere.

Substrate-binding Residues and Proposed Catalytic Mechanism

The presence of a dimetal center in Sk\textsuperscript{AS} and the structural similarity of the protein to enzymes using a corresponding mechanism suggest that cleavage of the substrates N-carbamyl-\beta-alanine and N-carbamyl-\beta-aminoisobutyrate is achieved by activation of a zinc-coordinated water molecule to a hydroxyl ion nucleophile, which subsequently attacks the bond between the carbon of the carbamyl moiety and the nitrogen of the \beta-amino acid moiety of the substrate. Modeling of substrate molecules into the active site of Sk\textsuperscript{AS} and comparison with inhibitor binding in PepV (37) enabled us to identify potential substrate-binding residues (Fig. 3, a and b). Both substrates (N-carbamyl-\beta-alanine and N-carbamyl-\beta-aminoisobutyrate) can be fitted equally well into the cleft between the dizinc center in the catalytic domain and the residues of the dimerization domains facing it. In the model, the carbonyl oxygen of the N-carbamyl group replaces the water molecule coordinated to zinc 1 (W2). Together with the imidazole of His\textsuperscript{107}, which is located within hydrogen bond distance, the metal will polarize the carbonyl bond and increase the electrophility of the carbonyl carbon, facilitating attack by the activated water molecule. Both ligands will also have a function in stabilization of the oxyanion created upon formation of the tetrahedral intermediate. The side chain of Gln\textsuperscript{229} is a likely candidate for interaction with the primary carbamyl amide, whereas the backbone oxygen of Gly\textsuperscript{398} could form a hydrogen bond with the secondary amide, i.e., the amino group of the reaction product. Only minimum side chain movements are required to bring the carboxylate of Glu\textsuperscript{159} within hydrogen bond distance to the same substrate moiety as well as to the bridging water molecule. Glu\textsuperscript{159} represents the carboxylate group hydrogen-bonded to the metal-activated water molecule, which is conserved in all structurally characterized zinc peptidases. It has been proposed to accept a proton from the water nucleophile and to facilitate the breakdown of the tetrahedral intermediate by subsequent shuttling of the proton to the amino group of the generated \beta-amino acid (38). Hence, the location of Glu\textsuperscript{159} is in agreement with its proton-shuttling function. There are most likely three residues involved in binding of the carbonyl group of the substrate: Arg\textsuperscript{225}, His\textsuperscript{362}, and Asn\textsuperscript{295}. These three amino acids belong to the dimerization domains of either the same subunit as the zinc-coordinating residues (Arg\textsuperscript{225}) or the other subunit in the dimer (His\textsuperscript{362} and Asn\textsuperscript{295}). The dimeric state of Sk\textsuperscript{AS} seems thus to be essential for catalytic activity.

In both crystal forms, we observed a peak of electron density in this “carboxylate-binding site,” which remains unoccupied by protein residues. Again, because of the modest resolution of the data, the density cannot be unambiguously assigned to a specific molecule. From the compounds present in the solutions used for crystallization of the protein, Tris and citrate are possible candidates. However, introduction of either one into the model and subsequent refinement resulted in peaks of negative difference density in the $|F_o| - |F_c|$ map. Although not added to the crystallization solution, the reaction product \beta-aminoisobutyrate seems to fit best into the density. The $|F_o| - |F_c|$ map calculated after refinement with \beta-aminoisobutyrate included in the active sites reveals well-defined density for this molecule. But because additional peaks of positive difference density appear connected to it, the true nature of the bound compound remains unclear.

In the modeled enzyme-substrate complex, the placement of the carboxamyl moiety of the substrate in close proximity to the two zinc ions results in a gap between its carboxyl group and the three putative carboxylate-binding residues, which is too large to allow hydrogen bond formation. However, the size of this gap varies in the eight protomers present in the asymmetric unit of the Sk\textsuperscript{AS} crystals due to the observed differences in domain arrangement. In chain A, a maximum distance of 3.7 Å is observed between the carboxyl carbons of the modeled substrate and of the \beta-aminoisobutyrate fitted into the density peak, whereas a minimum distance of only 2.5 Å is measured in chain B. We assume that, in the true enzyme-substrate complex, the dimerization domain is moved even closer toward the catalytic domain, without steric clashes or major local conformational changes achieved by a modest rigid body rotation of the catalytic domain. The resulting narrowing of the substrate-binding cleft would then allow a simultaneous interaction of the substrate molecule with the zinc ions and the three residues in the carboxylate-binding pocket.

It is tempting to speculate that the observed flexibility in domain arrangement is a component of the catalytic mechanism of Sk\textsuperscript{AS}, which allows opening and closing of the active site during each reaction cycle. Substrate binding and catalysis could thus be achieved in several steps: 1) electrostatic attraction between the negatively charged carbonyl group of the substrate and the positively charged Arg\textsuperscript{225} and formation of hydrogen bonds with the other carboxylate-binding residues; 2) closure of the active site, allowing interaction of the carboxyl group of the substrate with the dimeric center; 3) nucleophilic attack of the activated water molecule and formation of the tetrahedral intermediate; 4) hydrolysis of the carbon–nitrogen bond; 5) spontaneous decomposition of the generated carbamate to carbon dioxide and ammonia; and 6) opening of the active site, release of the \beta-amino acid, and addition of a new bridging water molecule. The proposed mechanism is illustrated in Scheme 2.

Amidohydrolase Versus Peptidase Activity

The significant similarity in three-dimensional structure of Sk\textsuperscript{AS} and dizinc-dependent exopeptidases is not mirrored by an equally high homology in amino acid sequence. In Fig. 4, the sequences of Sk\textsuperscript{AS}, SGAP, APAP, CPG2, PepT, and PepV are aligned based on the structural superposition. Basically, only the zinc-binding residues corresponding to His\textsuperscript{114}, Asp\textsuperscript{126}, and Glu\textsuperscript{160} in Sk\textsuperscript{AS} are conserved at the same spatial position in all six proteins. His\textsuperscript{121} has also counterparts in the exopeptidases; although in CPG2, the localization of its C-\alpha atom deviates beyond the default values defining structural homology in the program TOP (28), which was used for the structural alignment. Nevertheless, changes in side chain conformation place the imidazole amide interacting with zinc 2 at the same position as seen in the other five enzymes. Most interesting is the last of the zinc-coordinating residues, corresponding to His\textsuperscript{126} in Sk\textsuperscript{AS}. Although the position of the zinc-binding function is
SCHEME 2. Proposed reaction mechanism of SkβAS. Step 1, binding of N-carbamyl-β-alanine first to the carboxylate-anchoring residues in the active site; Step 2, simultaneous interaction of the substrate with the dizinc center allowed by domain movements, displacement of a zinc-bound water by the substrate, and activation of the bridging water molecule by Glu159; Step 3, nucleophilic attack of the carbonyl carbon of the substrate by the activated water molecule and formation of the tetrahedral intermediate; Step 4, cleavage of the carbon–nitrogen bond, product formation, and spontaneous decomposition of the generated carbamate; Step 5, opening of the active site and release of carbon dioxide and ammonia; Step 6, release of the reaction product β-alanine and addition of a new bridging water molecule.
FIG. 4. Structure-based alignment of the sequences of SkβAS and dizinc-dependent exopeptidases. Row 1, SkβAS; row 2, SGAP; row 3, APAP; row 4, CPG2; row 5, PepV; row 6, PepT. The secondary structure elements (labeled as described for Fig. 1a) are indicated by arrows for β-strands and boxes for α- and 3_10-helices, which are white for the catalytic domain and light blue for the dimerization domain of SkβAS. For the

Crystal Structure of Yeast β-Alanine Synthase
conserved, the chemistry is not. In all exopeptidases, the histidine is exchanged with an acidic residue, either aspartate (SGAP, APAP, PepT, and PepV) or glutamate (CPG2). Replacement of an acidic group with histidine results in a decreased negative charge from the protein ligands, which stabilizes the metal-bound hydroxide. As a consequence, the zinc-bound water in SkβAS should exhibit a lower pKₐ and higher reactivity than the metal-bound solvent in the exopeptidases (39). Indeed, the hydrolysis of N-carbamyl-β-alanine and N-carbamyl-β-aminonooctanoyl tripeptide requires an increased nucleophilicity of the attacking water compared with the hydrolysis of a peptide bond due to the better resonance stabilization of the ureido group and the lower electrophilicity of its carbonyl carbon.

Two other residues are strictly conserved at almost the same position: the proton-shuttling Glu₁⁵⁹ and Asp₁¹⁶ of SkβAS and their corresponding residues in the dizinc-dependent exopeptidases. Asp₁¹⁶ appears to be more of structural importance rather than playing a role in catalysis. Its side chain makes hydrogen bonds with the backbone oxygens of the catalytic Glu¹⁵⁹ and the zinc ligand Glu¹⁶⁰ and might thus ensure the maintenance of a proper zinc-binding site geometry.

Interestingly, the putative substrate-binding Arg₃²₂ is also strictly conserved between SkβAS and the exopeptidases with a two-domain structure: CPG2, PepV, and PepT. The complex of the dipeptidase PepV with the transition state analog Asp₅¹⁹Phe₅⁰₁⁶Lala-OH (37) revealed that the arginine interacts with the terminal carboxyl group of the substrate, together with a histidine and an asparagine (His₃⁶² and Asn₃⁰⁹ in SkβAS) (Fig. 3b). Moreover, residues forming hydrogen bonds with Arg₃²₂ in SkβAS are also found in PepV. Superposition of the zinc-binding sites of, for example, chain A of SkβAS and the PepV complex does not automatically result in superposition of the carboxylate-anchoring residues. In SkβAS, Arg₃²₂, His₃⁶², and Asn₃⁰⁹ are located 4.2, 4.4, and 5.6 A farther away from the zinc site compared with their counterparts in PepV. After taking into account that SkβAS-catalyzed hydrolysis generates a β-amino acid instead of an α-amino acid, which requires a difference in localization of about one carbon–carbon single bond length, a distance of >2.5 Å remains to be bridged to allow simultaneous interaction of the substrate with the zinc ions and the carboxylate-binding residues. This distance corresponds to that estimated by modeling of the substrate into the active site, supporting our hypothesis that an active-site opening and closure event is an integral part of the catalytic cycle of SkβAS.

No structures of complexes with transition state analogs or inhibitors are available for CPG2 and PepT. Also the spatial positions of their dimerization domains, which differ significantly from that observed in SkβAS, make a comparison of putative substrate-binding sites rather difficult.

The few other highly conserved residues in the sequences of SkβAS and dizinc exopeptidases (Fig. 4) appear to have a function in correct folding and stabilization of the three-dimensional structure of the proteins. Apart from the aspartate/glutamate replacement for one of the histidine ligands of zinc 1, the structure of SkβAS is distinct from those of the exopeptidases in another noteworthy detail: it does not contain a cis-peptide bond between the zinc-bridging Asp₁²₅ and the following residue. In CPG2, APAP, PepT, and PepV, this residue is an aspartate, which is replaced by asparagine in SGAP. Due to the formation of the cis-peptide bond, the carboxyl/carboxamide side chain adopts a position that allows the formation of hydrogen bonds with two catalytically crucial residues: with the main chain amide of the histidine corresponding to the zinc ligand His₁¹⁴ in SkβAS and with the second carboxyl oxygen of a zinc-coordinating aspartate or glutamate residue. In SkβAS, residue 126 is a glycine. Nevertheless, apparently, the absence of the carboxylate/carboxamide side chain and the cis-peptide bond does not result in a disturbance of the zinc coordination site. The hydrogen bond to His₁¹⁴ is preserved by involvement of the carbonyl oxygen of Gly¹²₆, and the formation of the second hydrogen bond is not required due to the replacement of the zinc-binding aspartate/glutamate with a second histidine ligand for zinc 1 in SkβAS.

Thus, by a few subtle changes within the environment of the zinc ions, the active site of S. kluwyeri β-alanine synthase has adopted to the necessity of hydrolyzing a carbon–nitrogen bond that is more stable than a peptide bond. However, in general, the same basic structural scaffold as in the dizinc-dependent exopeptidases is employed to perform the reaction.

**Evolution of Amidohydrolases**

On the basis of sequence phylogenetic analysis, the family of amidohydrolases can be divided into three subfamilies, with mammalian and most other eukaryotic βASs constituting the first subfamily, SkβAS and N-carbamyl-l-amino acid amidohydrolases constituting the second, and DCases constituting the third (10). Comparison of the crystal structures of SkβAS and two bacterial DCases revealed that the second and third subfamilies are not structurally related despite the fact that both enzymes catalyze the same reaction on very similar substrates. DCase from *Agrobacterium* sp. exists as an associated homotetramer in the crystal (12, 13). Each subunit of ~34 kDa folds into a single domain comprising a sandwich of parallel β-sheets surrounded by two layers of α-helices. Furthermore, DCase does not contain metal ions. A triad of residues well conserved within the third subfamily of amidohydrolases (Cys-Glu-Lys) has been proposed to be essential for the catalysis of N-carbamyl-l-amino acid breakdown. A model for the catalytic mechanism has been suggested in which the cysteine represents the catalytic nucleophile; the glutamate enhances its nucleophilicity; and the ε-amino group of the lysine stabilizes the tetradecapeptidal intermediate. Thus, the members of the second and third subfamilies of amidohydrolases also do not share a common catalytic mechanism.

This finding raises an interesting question: whether or not the βASs of the first subfamily are structurally and/or mechanistically related to members of the other subfamilies. The overall sequence identity of, for example, rat βAS to SkβAS and DCase is well below 20% in both cases. On the other hand, the same applies to SkβAS and the dizinc-dependent exopeptidases, which share basically only the zinc-coordinating residues and a few structurally important residues. Rat and maize βASs have been shown to contain zinc ions, whereas DCase does not. However, with the single exception of a glutamate corresponding to Glu¹⁶⁰, none of the zinc-coordinating residues of SkβAS is conserved in the sequences of βASs from the first subfamily. This could indicate that zinc binding in these en-

---

*Exopeptidases, only the sequences of the structurally aligned parts are shown (uppercase letters). Pink number signs indicate the zinc-coordinating residues. Amino acids labeled by boldface letters indicate sequence conservation in all structurally aligned sequence parts of exopeptidases. If the amino acid is also conserved in the sequence of SkβAS, the residue is marked additionally by pink background shading. Gray background shading indicates sequence conservation between SkβAS and any of the exopeptidases. For residues that are conserved and aligned only in some of the six proteins, the corresponding amino acid in the non-aligned proteins is given as a lowercase italic letter if its spatial position only barely does not fulfill the requirements defining structural alignment. A hypen is used when the residue is not present at the same position due to either considerable differences in structure or deletions in the sequence. Dotted lines indicate the absence of the dimerization domain in SGAP and APAP.*
enzymes is not achieved in the same fashion or at least does not involve amino acids at equivalent positions in the sequence. In contrast, residues corresponding to the catalytic triad in DCase are well conserved in the mammalian βAS sequences, with the glutamate residue again being the same as that aligning with Glu^{159} of SkβAS. The slightly higher degree of sequence conservation between the βASs of the first subfamily and DCase rather than SkβAS, including amino acids crucial for catalysis and substrate recognition, suggests that the first and third subfamilies of amidohydrolases are more closely related to each other. However, structural characterization of a mammalian or plant βAS is required to confirm this suggestion.

A few interesting phylogenetic questions also need to be answered. Why does SkβAS not belong to the same subfamily as mammalian and other eukaryotic βASs? Why does it group with bacterial N-carbamyl-l-amino acid amidohydrolases instead, which are otherwise believed not to participate directly in pyrimidine catabolism? A plausible explanation is that mammalian, plant, and other eukaryotic βASs have evolved from the same ancestral protein present in a common progenitor of all eukaryotic life forms. The gene encoding the ancestral protein has then apparently been lost in the yeast lineage, as indicated by the absence of “true” βAS-encoding genes in the fully sequenced genomes of Saccharomyces cerevisiae, Schizosaccharomyces pombe, and several other yeasts. Instead, the structural scaffold of an ancestor of the dizinc-dependent peptidases and N-carbamyl-l-amino acid amidohydrolases might have been recruited and its catalytic machinery modified to obtain a functional βAS in S. kluiveri.

Apparantly, the same structural scaffold is used by a variety of enzymes, which have recently been grouped together into the aminoacylase-1 family (40). These enzymes are human and porcine aminoacylase-1; SGAP; CPG2; PepV; N-carbamyl-l-amino acid amidohydrolases from Bacillus stearothermophilus; the precursors of carboxypeptidase S and aminopeptidase Y from S. cerevisiae; succinyllyaminopimelate desuccinylase from Escherichia coli, Corynebacterium glutamicum, and Hemophilus influenzae; acetylornithine decacylase from E. coli and Dictyostelium discoideum; and hydantoin utilization protein C from Pseudomonas species. The amino acid sequences of all members contain three signature sequence motifs, which are also found in the SkβAS: (S/G)A/HDXV, GXDX, and XEE (with H, D, and E representing His^{114}, Asp^{126}, and Glu^{159} in SkβAS, i.e. two zinc ligands and the catalytic glutamate). The presence of the signature motifs and the proven structural similarity to CPG2, PepV, and SGAP clearly group SkβAS in the same enzyme family.

The sequences of mammalian and other eukaryotic βASs do not contain these motifs, and it is therefore very likely that the presented three-dimensional structure of SkβAS does not reflect the scaffold used by the common progenitor of the majority of βAS enzymes. It does, however, add to our knowledge about evolutionary processes as an excellent example of how “novel” enzymes can originate or be reinvented by recruiting and remodeling an already available structural scaffold.

Acknowledgments—We acknowledge access to synchrotron radiation at the European Synchrotron Radiation Facility and the MAX-lab of Lund University (Lund, Sweden) and are grateful to Gordon Leonard for assistance during multiwavelength anomalous dispersion data collection. We are indebted to Birgit Andersen for excellent technical assistance with protein purification.

REFERENCES

1. Wasternack, C. (1980) Pharmacol. Ther. 8, 629–651
2. Heggie, G. D., Sommadossi, J. P., Cross, D. S., Huster, W. J., and Diasio, R. B. (1987) Cancer Res. 47, 2203–2206
3. Milano, G., and Etienne, M.-C. (1994) Anticancer Res. 14, 2295–2297
4. Cronan, J. E., Jr., Littel, K. J., and Jackowski, S. (1989) J. Bacteriol. 169, 916–922
5. Enna, S. J., and Gallagher, J. P. (1983) Int. Rev. Neurobiol. 24, 181–212
6. Wang, D.-S., Zhu, H.-L., and Li, J.-S. (2003) Int. J. Neurosci. 113, 283–305
7. van Gennip, A. H., Abeling, N. G. M., Vreken, P., and van Kuilenburg, A. B. P. (1997) J. Inherit. Metab. Dis. 20, 203–213
8. Traut, T. W., and Jones, M. E. (1996) Prog. Nucleic Acid Res. Mol. Biol. 53, 1–78
9. Williamson, J. M., and Brown, G. M. (1979) J. Biol. Chem. 254, 8074–8082
10. Gojkovic, Z., Sandrini, M. P. F., and Piskur, J. (2001) Genetics 158, 999–1011
11. Gojkovic, Z., Paracchini, S., and Piskur, J. (1998) Adv. Exp. Med. Biol. 431, 475–479
12. Nakai, T., Hasegawa, T., Yamashita, E., Yamamoto, M., Kamasaka, T., Ueki, T., Namba, H., Henaka, Y., Takahashi, S., Sato, M., and Teukhara, I. (2000) Structure 8, 729–737
13. Wang, W.-C., Hsu, W.-H., Chien, F.-T., and Chen, C.-Y. (2001) J. Biol. Chem. 276, 251–261
14. Kvalnes-Krick, K. L., and Traut, T. W. (1993) J. Biol. Chem. 268, 5686–5693
15. Walsh, T. A., Green, S. B., Larrinua, I. M., and Schmitzer, P. R. (2001) Plant Physiol. 125, 1001–1012
16. Matthews, M. M., and Traut, T. W. (1978) J. Biol. Chem. 232, 7232–7237
17. Dobritzsch, D., Gojkovic, Z., Andersen, B., and Piskur, J. (2003) Acta Crystallogr. Sect. D69, 1267–1269
18. Collaborative Computational Project Number 4 (1994) Acta Crystallogr. Sect. D 50, 760–763
19. Terwilliger, T. C., and Berendzen, J. (1999) Acta Crystallogr. Sect. D 55, 849–861
20. Terwilliger, T. C. (2000) Acta Crystallogr. Sect. D 56, 803–823
21. de la Fortelle, E., and Bricogne, G. (1997) Methods Enzymol. 276, 472–494
22. Cowtan, K., and Main, P. (1996) Acta Crystallogr. Sect. D 52, 43–48
23. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) Acta Crystallogr. Sect. A 47, 110–119
24. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D 54, 905–921
25. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Acta Crystallogr. Sect. D 53, 240–255
26. Kissinger, C. R., Gehiiaar, D. K., and Fogel, D. B. (1999) Acta Crystallogr. Sect. D 55, 484–491
27. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) J. Appl. Crystallogr. 26, 283–291
28. Lu, G. (2000) J. Appl. Crystallogr. 33, 176–183
29. Krasilov, P. J. (1991) J. Appl. Crystallogr. 24, 946–950
30. Eonuff, R. M. (1997) J. Mol. Graph. Model. 15, 132–134
31. Merritt, E. A., and Bacon, D. J. (1997) Methods Enzymol. 277, 505–524
32. Rawlings, N. D., and Barrett, A. J. (1995) Methods Enzymol. 248, 183–228
33. Gilboa, R., Spungin-Bialik, A., Wohlfahrt, G., Schomburg, D., Blumberg, S., and Shoham, G. (2001) Protein 44, 490–504
34. De Paola, C. C., Bennett, S., Holz, R. C., Ringe, D., and Petsko, G. A. (1999) Biochemistry 38, 9048–9053
35. Rowell, S., Puipit, R. A., Tucker, A. D., Melton, R. G., Blow, D. M., and Brick, P. (1997) Structure 5, 337–347
36. Skandalakis, K., and Miller, C. G. (2002) Eur. J. Biochem. 269, 443–450
37. Jozic, D., Bourenkov, G., Bartunik, H., Scholze, H., Dive, V., Henrich, B., Huber, R., Bode, W., and Maskos, K. (2002) Structure 10, 1097–1106
38. Lebherz, W. N., and Strater, N. (1996) Chem. Res. 96, 2375–2383
39. CHRISTIANSON, D. W., and FIERKE, C. A. (1996) Acta Crystallogr. Sect. A 52, 331–339
40. Biagini, A., and Puigserver, A. (2001) Comp. Biochem. Physiol. B 130, 469–481