Structure, Mechanism, and Substrate Profile for Sco3058: The Closest Bacterial Homologue to Human Renal Dipeptidase‡,*

Jennifer A. Cummings,§ Tinh T. Nguyen,§ Alexander A. Fedorov,‖ Peter Kolb,‖ Chengfu Xu,§ Elena V. Fedorov,‖ Brian K. Shoichet,‡ David P. Barondeau,§ Steven C. Almo,‖ and Frank M. Rauschel*†

Department of Chemistry, P.O. Box 30012, Texas A&M University, College Station, Texas 77843, ‖ Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, and *Department of Pharmaceutical Chemistry, University of California, 1700 4th Street, San Francisco, California 94158-2330

Received November 10, 2009; Revised Manuscript Received December 11, 2009

ABSTRACT: Human renal dipeptidase, an enzyme associated with glutathione metabolism and the hydrolysis of β-lactams, is similar in sequence to a cluster of ~400 microbial proteins currently annotated as nonspecific dipeptidases within the amidohydrolase superfamily. The closest homologue to the human renal dipeptidase from a fully sequenced microbe is Sco3058 from Streptomyces coelicolor. Dipeptide substrates of Sco3058 were identified by screening a comprehensive series of L-Xaa-L-Xaa, L-Xaa-D-Xaa, and D-Xaa-L-Xaa dipeptide libraries. The substrate specificity profile shows that Sco3058 hydrolyzes a broad range of dipeptides with a marked preference for an L-amino acid at the N-terminus and a D-amino acid at the C-terminus. The best substrate identified was L-Arg-D-Asp (κcat/Km = 7.6 × 105 M-1 s-1). The three-dimensional structure of Sco3058 was determined in the absence and presence of the inhibitors citrate and a phosphinate mimic of L-Ala-D-Asp. The enzyme folds as a (β/α)6 barrel, and two zinc ions are bound in the active site. Site-directed mutagenesis was used to probe the importance of specific residues that have direct interactions with the substrate analogues in the active site (Asp-22, His-150, Arg-223, and Asp-320). The solvent viscosity and kinetic effects of D2O indicate that substrate binding is relatively sticky and that proton transfers do not occur during the rate-limiting step. A bell-shaped pH-rate profile for κcat and κcat/Km indicated that one group needs to be deprotonated and a second group must be protonated for optimal turnover. Computational docking of high-energy intermediate forms of L/D-Ala-L/D-Ala to the three-dimensional structure of Sco3058 identified the structural determinants for stereochemical preferences for substrate binding and turnover.

The mammalian renal dipeptidase catalyzes the general reaction outlined in Scheme 1 (1–6). This enzyme is also able to catalyze the hydrolysis of leukotriene D4 (2, 3, 7) as well as the oxidized degradation product of glutathione, l-Cys-Gly (2–4). Human renal dipeptidase (hRDP)† is the only enzyme in humans that has been shown to catalyze the hydrolysis of β-lactams such as SCH 29482, imipenem, meropenem, and DA-1131 (8, 9). However, dipeptidases are consistently better substrates for hRDP than are β-lactams (2, 7, 10). Glycine is the preferred amino acid at the C-terminus, relative to an amino acid in the L-configuration, and glycyldihydrophenylalanine is a better substrate for hRDP than glycylphenylalanine (10). With rat renal dipeptidase, l-Ala-Gly is hydrolyzed faster than L-Ala-l-Ala and Gly-d-Ala is hydrolyzed faster than Gly-l-Ala (3). Mammalian renal dipeptidases and some of their microbial homologues have been assayed with various substrates, but a complete substrate specificity profile for these types of enzymes has not been determined (2–4, 6, 10–12).

The crystal structure of hRDP has been determined with the inhibitor cilastatin bound in the active site (PDB entry 1ITU) (13). Cilastatin is considered to be a dipeptide analogue, but it does not contain a free amino group at the N-terminus and is not hydrolyzed by hRDP. The metal center in the active site consists of two Zn ions that are bridged by a molecule from solvent (water or hydroxide) and the side chain carboxylate of Glu-125. The α-Zn is additionally ligated by His-20 and Asp-22, whereas the β-Zn is coordinated to His-198 and His-219. It has been suggested that the water that bridges the two metal ions attacks the carbonyl of the peptide bond to form a tetrahedral intermediate that is stabilized by an electrostatic interaction with His-152. This residue is found at the end of β-strand 4 of the (β/α)6 barrel structure and has been shown to be important for the binding of the cilastatin inhibitor (13, 14).

The structure of hRDP confirms that this enzyme is a member of the amidohydrolase superfamily (AHS). This is a rather large enzyme superfamily that has been shown to catalyze the hydrolysis of ester and amide bonds found within carbohydrate-, peptide-, and nucleic acid-containing substrates (15). The hRDP is contained within a large cluster of proteins designated as cog2355E. This cluster of proteins is typified by an HxD motif at

© 2009 American Chemical Society
the end of β-strand 1 that is used to coordinate one of the two metal ions in the active site, whereas most of the other enzymes in this superfamily have an HxH motif. In hRDP, the invariant aspartate at the end of β-strand 8 is not coordinated to the α-metal, whereas in nearly all of the other members of the AHS that have been structurally characterized, this residue is a direct ligand to the α-metal ion. The proposed chemical mechanism for the hydrolysis of dipeptides by hRDP is at variance with those mechanisms of amide bond hydrolysis proposed previously for other members of the AHS (16–18). In hRDP, the carbonyl group of the amide bond about to be hydrolyzed is proposed to be polarized by an electrostatic interaction with His-152, whereas in other members of the AHS, this function is contributed by a direct interaction with the β-metal ion (13, 16–18). The renal dipeptidase-like enzyme BCs-1 from Brevibacillus borstelensis has a tryptophan in place of the residue equivalent to His-152. BCs-1 is reported to have respectable activity (kcat/Km = 2 × 105 M–1 s–1), which supports the conclusion that either His-152 does not play a crucial role in the chemical mechanism or the mechanism for substrate hydrolysis of enzymes such as BCs-1 varies from that of hRDP (11).

The putative dipeptidase Sco3058 from Streptomyces coelicolor is the closest bacterial homologue to the human enzyme from a completely sequenced microbial organism. The level of sequence identity is ~41%. S. coelicolor is a member of a family of soil bacteria known for their versatile metabolism and production of antibiotics and pharmaceutical agents (19). In this paper, we unravel the substrate specificity of Sco3058 by determining the catalytic activity with 51-L-Xaa-L-Xaa, L-Xaa-D-Xaa, and D-Xaa-L-Xaa dipeptide libraries. The results demonstrate that this enzyme is a rather promiscuous dipeptidase that is able to hydrolyze a large fraction of the ~1000 compounds screened. Nevertheless, Sco3058 exhibited some degree of specificity, and the best substrate identified was L-Arg-D-Asp (kcat/Km = 7.6 × 105 M–1 s–1). The structure of Sco3058 has been determined to a resolution of 1.3 Å in the presence and absence of dipeptide mimics. The chemical mechanism has been addressed through the utilization of pH-rate profiles, changes in solvent viscosity, and the utilization of site-directed mutants.

MATERIALS AND METHODS

Materials. All chemicals were obtained from Sigma Aldrich, unless otherwise stated. The coupling enzymes, glutamate-oxaloacetate transaminase and malate dehydrogenase, were purchased from Sigma Aldrich and Calbiochem, respectively. Glycyl-D-Leu was purchased from TCI America. ICP standards were obtained from Inorganic Ventures Inc. Nitrocefin was purchased from Sigma Aldrich, and all other chemicals were obtained from Sigma Aldrich, unless otherwise stated. The coupling enzymes, glutamate-oxaloacetate transaminase and malate dehydrogenase, were obtained from Inorganic Ventures Inc. Nitrocefin was purchased from Sigma Aldrich, and all other chemicals were obtained from Sigma Aldrich, unless otherwise stated.

Synthesis of Dipeptide Libraries. The L-Xaa-D-Xaa and L-Xaa-L-Xaa dipeptide libraries were synthesized as described previously (21, 22). The 17 L-Xaa-L-Xaa libraries except that

Scheme 2

N-Fmoc-D-Xaa-OH amino acids were substituted for N-Fmoc-L-Xaa-OH amino acids. These compound libraries contain a single amino acid at the N-terminus and all of the common amino acids at the C-terminus. However, amino acids L-Cys, D-Cys, and D-Ile were not incorporated into any of the dipeptide libraries. A sample of these libraries, expected to contain ~2 nmol of each library component, was submitted for amino acid analysis (AAA) to quantify the amount of every amino acid in these libraries except for tryptophan. The amount of tryptophan in these libraries was determined by the absorbance at 280 nm.

Synthesis of Phosphinate Pseudodi-peptides. Three raceemic phosphinate pseudodi-peptide mimics of the putative tetra-hedral intermediates for the hydrolysis of Ala-Asp (1), Phe-Asp (2), and Tyr-Asp (3) were created as potential inhibitors of Sco3058. These compounds were synthesized according to previously described protocols (23–26). The structures of these compounds are presented in Scheme 2. Compound 1: 31P NMR (121.4 MHz, D2O) δ 36.24, 36.20; (MS (ESI negative mode) m/z 238.0129 (M–H)–; MS (ESI positive mode) m/z 240.0977 (M + H)+; C11H10O5NP requires m/z 239.0559. Compound 2: 31P NMR (121.4 MHz, D2O) δ 35.01; (MS (ESI negative mode) m/z 314.1116 (M–H)–; MS (ESI positive mode) m/z 316.0786 (M + H)+; C11H10O5NP requires m/z 315.0872. Compound 3: 31P NMR (121.4 MHz, D2O) δ 34.71; (MS (ESI negative mode) m/z 330.0984 (M–H)–; MS (ESI positive mode) m/z 332.0828 (M + H)+; C11H16O5NP requires m/z 331.0821.

Cloning and Expression of Sco3058. The gene for Sco3058 (gi:2121500) was amplified from S. coelicolor A3(2) genomic DNA with the primers 5’-GCCAGAGCCATATGACATCGCTGGAGAGGCCTGTGCCTGCGG-3’ and 5’-GCCGAGATTCCAGGCTTCCCGTCGCTGCGGCGCCGTGCGG-3’. The Pfx platinum polymerase (Invitrogen), 0.05 unit/μL PCR mixture, was used according to the manufacturer’s instructions except the annealing step was omitted. The PCR product was digested with Ndel and EcoRI and ligated into the pET30a(+) vector (Novagen) with T4 ligase (New England Biolabs). Escherichia coli BL21 Rosetta 2 (DE3) cells were electro-transformed with the pET30a(+)Sco3058 plasmid. Cells harboring the pET30a(+)Sco3058 plasmid were grown at 30 °C in Terrific Broth containing 25 and 50 μg/mL chloramphenicol and kanamycin, respectively. When the optical density at 600 nm reached ~0.6, 2.5 mM Zn(acetate)2 was added and the expression of Sco3058 was induced with 1.0 mM isopropyl β-thiogalactoside (IPTG). The cells were grown overnight at room temperature. Site-directed mutagenesis was conducted on the gene for Sco3058 at residues Asp-22, His-150, Arg-223, and Asp-320 using the QuikChange kit from Stratagene.

Purification of Sco3058. The cells containing Sco3058 were suspended in a purification buffer [20 mM Tris with 100 μg/mL phenylmethylsulfonyl fluoride (pH 7.5)] and lysed via sonication at 0 °C. The cell lysate was obtained by centrifugation and subsequently treated with protease sulfate (2% (w/w) of cell mass) and removed by centrifugation. The protein was precipitated at 60% ammonium sulfate (371 g/L at 0 °C), isolated by centrifugation, resuspended in 20 mM Tris (pH 7.5), and purified.
by size exclusion chromatography using a 3 L column of AcA34 in 20 mM Tris (pH 7.5) at a flow rate of 1.0 mL/min. The fractions containing the protein of expected size, as assessed by SDS–PAGE, were combined. Sco3058 was further purified by ion exchange chromatography with a Resource Q column (Pharmacia) using a gradient of NaCl in 50 mM Tris (pH 7.5). A final gel filtration step, with a column of Superdex 200 (Amersham Biosciences) in 50 mM HEPES (pH 7.5), was utilized to remove the salt.

**Ninhydrin-Based Enzyme Assays.** The measurement of dipeptidase activity was conducted in 50 mM HEPES (pH 7.5). For measurement of the inhibitory properties of the phosphatase pseudodipeptides, 0.40 mM L-Leu-d-Ala was used as the substrate and the inhibitor concentrations were varied from 0 to 100 μM. All assays were conducted at 30 °C. The formation of free amino acids was quantified using a ninhydrin-based assay as previously described (27). One volume of the enzymatic reaction was mixed with 2 volumes of ice-cold Cd-ninhydrin reagent [0.9% (w/v) ninhydrin dissolved in 1:10:80 (v/g/mL CdCl2/acetic acid/ETHOH mixture), heated at 80 °C for 5 min, and cooled prior to the absorbance being read at 507 nm. The correlation between the absorbance at 507 nm and the concentration of free amino acids was determined for each amino acid.

**β-Lactamase Assay.** Sco3058 was tested for its ability to hydrolyze the β-lactam nitrocefin. The hydrolysis of nitrocefin was monitored by the increase in absorbance at 486 nm (ε = 20500 M⁻¹ cm⁻¹) in the presence of 100 mM potassium phosphate (pH 7.0). The kcat/Km was estimated by fitting the data to eq 2.

**Coupled Enzyme Assays.** The hydrolysis of L-Arg-L-Asp was monitored by coupling the formation of L-aspartate to the oxidation of NADH in a system that includes glutamate-oxaloacetate transaminase and malate dehydrogenase. The change in the concentration of NADH was measured spectrophotometrically using a SPECTRAMax-340 plate reader (Molecular Devices Inc.) by following the decrease in absorbance at 340 nm. The standard assay conditions included 100 mM Tris (pH 7.5), varying concentrations of L-Arg-L-Asp, 0.36 mM NADH, 7 units of glutamate-oxaloacetate transaminase, 1.0 unit of malate dehydrogenase, 3.7 mM α-ketoglutarate, 100 mM KCl, and Sco3058 in a final volume of 250 μL at 30 °C. This assay was used to determine the kinetic parameters of L-Arg-L-Asp and L-Ala-L-Asp and the Km of citrate versus the L-Arg-L-Asp substrate.

The pH dependence of the kinetic parameters, kcat and kcat/Km, was measured over the pH range of 5.5–10.0 at 0.25 pH unit intervals for the zinc-substituted enzyme. The buffers used for the pH–rate profiles were MES, Bis-Tris, Tris, and CABS. The pH values were recorded after the completion of the assays. The effects of solvent viscosity on the activity of Sco3058 were determined at pH 7.5 using sucrose as the microviscogen at 25 °C. The concentrations of sucrose were 0, 10, 14, 20, 24, and 27% (w/w), with corresponding rates determined at pH 7.5 using sucrose as the microviscogen at 25 °C. The concentrations of enzyme used were 0–1000 nM (L-Ala-L-Xaa), 0–200 nM (L-Arg-d-Xaa), 0–250 nM (L-Arg-l-Xaa, L-Trp-l-Xaa, and L-Met-l-Xaa), or 20–400 nM (d-Leu-l-Xaa). The sample preparation and data analysis were conducted as previously described (21, 22).

**Data Analysis.** The kinetic parameters, kcat, Km, and kcat/Km, were determined by fitting the initial velocity data to eq 2.

\[
v/E_i = (k_{cat}[A])/(K_m + [A])
\]

where v is the initial velocity, Ei is the total enzyme concentration, kcat is the turnover number, [A] is the substrate concentration, and Km is the Michaelis constant. The profiles for the variation of kcat and kcat/Km with pH were fit to eq 3.

\[
\log y = \log[c/(1 + H/K_a + K_b/H)]
\]

where c is the pH-independent value of y, Klic, and Kb are the dissociation constants of the ionizable groups, and H is the proton concentration. The kinetic constants for the inhibition of Sco3058 by the pseudodipeptide mimics were obtained by a fit of the data to eq 4.

\[
v/E_i = (k_{cat}[A])/[K_m(1 + I/K_i) + A]
\]

where Ki is the competitive inhibition constant.

**Crystallization and Data Collection.** Crystals of Sco3058 were grown by hanging drop vapor diffusion at room temperature using 0.7–0.9 M citrate, 0.1 M imidazole (pH 8.0), and 12.0 mg/mL enzyme. Crystals were observed within 2 weeks and exhibited diffraction consistent with space group P3121, with one molecule of Sco3058 per asymmetric unit and 62% solvent. Prior to data collection, the citrate-grown crystals were transferred to a cryoprotectant solution composed of 80% mother liquor and 20% glycerol and flash-cooled and stored in liquid nitrogen. Data for the glycerol-citrate-Sco3058 complex were collected to a maximum resolution of 1.7 Å using the single-wavelength SSRL 7-1 beamline (Stanford Synchrotron Radiation Lightsource), and intensities were integrated and scaled with DENZO and SCALEPACK (28).

The crystals of wild-type Sco3058 complexed with Zn²⁺ were grown by vapor diffusion at room temperature using the following crystallization conditions. The protein solution contained wild-type Sco3058 (14.0 mg/mL) in 50 mM HEPES (pH 7.5); the precipitant contained 22% poly(acrylic acid), 0.1 M HEPES
(pH 7.5), and 20 mM ZnCl₂. Crystals appeared in 3 days and exhibited diffraction consistent with space group P₂₁2₁, with one molecule of Sco3058 per asymmetric unit. Prior to data collection, the crystals were transferred to cryoprotectant solutions composed of mother liquors supplemented with 20% glycerol and flash-cooled in a nitrogen stream. Data for Sco3058 with Zn were collected to 1.3 Å resolution at the NSLS X4A beamline (Brookhaven National Laboratory) on an ADSC CCD detector. Intensities were integrated and scaled with DENZO and SCALEPACK (Brookhaven National Laboratory) on an ADSC CCD detector.

Structure Determination and Model Refinement. The Sco3058-glycerol-citrate structure was determined by molecular replacement using AMoRe with the renal dipeptidase from humans (PDB entry 1ITQ) as the template (29, 30). Manual model building and solvent building were performed with Xfit, and the structure was refined to 1.7 Å with CNS (31–33). Electron density corresponding to the bound citrate and Zn ions was evident in the active site.

The structure of wild-type Sco3058 complexed with Zn was similarly determined by molecular replacement with PHASER (34), using the coordinates of human renal dipeptidase (PDB entry 1ITQ) as the search model. Iterative cycles of automatic model rebuilding with ARP (35), manual rebuilding with TOM (36), and refinement with CNS (31) were performed. The model was refined at 1.3 Å with an R cryst of 0.184 and an R free of 0.198. The final structure contained protein residues 1–391, 390 water molecules, and two well-defined Zn ions. Structure determination and refinement for the Sco3058-Zn pseudodipeptide ternary complex were analogous to the procedures used for the Sco3058-Zn complex. The model of the Sco3058-Zn pseudodipeptide complex was refined at 1.4 Å with an R cryst of 0.193 and an R free of 0.206. The final structure includes protein residues 1–391, 417 water molecules, two Zn ions, and one well-defined pseudodipeptide inhibitor bound in the active site. Final crystallographic refinement statistics for all complexes are listed in Table 1.

Docking Calculations. In parallel with the experimental assays, a comprehensive library of dipeptides and capped amino acids was generated computationally. This library was composed of all L-, L-D-, D-L-, and D-D dipeptides as well as N-acetylated, N-succinylated, N-formylated, N-carbamoylated, N-formimidoylated, and hydantoin variants of individual L- and D-amino acids. All molecules were transformed to their high-energy intermediate (HEI) forms by in silico reaction with a hydroxide anion. The HEI form of a potential substrate mimics the tetrahedral state of the molecule immediately after the attack of water or hydroxide (37, 38). This state of a substrate should show higher complementarity toward the enzyme, as it is the form that the enzyme is preconfigured to stabilize (39). Since the reaction center is tetrahedral, two enantiomers of each HEI exist, bringing the total number of diastereomers of any given dipeptide to eight. This HEI library was docked into the binding site of Sco3058 with DOCK 3.5.54 as described previously (40). We note that the score assigned to each pose of a docked dipeptide is based on its fit to the structure (the score consists of van der Waals and interactions).
electrostatic terms and is corrected for desolvation) and does not contain any terms reflecting differences in reactivity. For a library in which the reactive center is consistent, the latter term can be assumed to be approximately constant, allowing conclusions to be drawn from complementarity alone. To investigate the structural determinants for binding in more detail, the four diastereomers of the Ala-Ala dipeptide (L-L, L-D, D-L, and D-D) were analyzed separately from the rest and directly compared with the experimentally determined turnover rates.

RESULTS

Cloning, Expression, and Purification of Sco3058. The gene for Sco3058 was successfully cloned into the pET30a(+) vector. The protein was expressed at moderate levels and purified to homogeneity. The identity of Sco3058 was confirmed by sequencing the first five amino acids from the N-terminus of the purified protein. The amino acid sequence (90% MTSLE and 10% TSLEK) matched that of MTSEK predicted from the DNA sequence for this enzyme. ICP-MS confirmed the presence of 2.0 equiv of Zn per subunit.

N-Terminal Substrate Specificity of Sco3058. The substrate specificity of Sco3058 for the amino acid at the N-terminus of the various dipeptide libraries was determined by measuring the composite rate of dipeptide hydrolysis in these libraries. A typical enzyme-course plot for the hydrolysis of the L-Arg-L-Xaa library is presented in Figure 1A. An enzyme-course plot for the slower hydrolysis of the D-Xaa-L-Xaa dipeptide libraries using D-Tyr-L-Xaa as an example is presented in Figure 1B. The relative hydrolysis rates for the L-Xaa-L-Xaa, L-Xaa-D-Xaa, and D-Xaa-L-Xaa dipeptide libraries are presented in Figure 2A–C. These libraries contain a fixed amino acid at the N-terminus and 17–19 different amino acids at the C-terminus. For the L-Xaa-L-Xaa dipeptide libraries, the N-terminal preference was for a leucine residue with slightly lower turnovers for methionine, arginine, and glutamine. For the L-Xaa-D-Xaa series of dipeptides, the preference at the N-terminus was for arginine, glutamine, methionine, and leucine. With the series of D-Xaa-L-Xaa dipeptide libraries, the preferred amino acids at the N-terminus are methionine, leucine, arginine, and lysine.

C-Terminal Substrate Specificity for Sco3058. The observed rates for the liberation of specific amino acids from the C-terminus of the dipeptides contained within L-Arg-L-Xaa, L-Arg-D-Xaa, L-Ala-L-Xaa, L-Trp-L-Xaa, L-Met-L-Xaa, and D-Leu-L-Xaa are presented in Figure 3A–F, respectively. The relative rates were obtained from fits of the data to eq 1. Within each library, dipeptides that terminated with L-Asp were hydrolyzed the fastest. Proline was not detected in any of the hydrolysis products, and thus, dipeptides that terminate in a proline are not hydrolyzed. Dipeptides with L-Val, L-Thr, L-Arg, or L-Lys at the C-terminus were consistently hydrolyzed more slowly than the 14 remaining L-amino acids given the same amino acid at the N-terminus.
Kinetic Constants for Selected Substrates. The kinetic constants, $k_{cat}$, $K_m$, and $k_{cat}/K_m$ for 25 dipeptide substrates are presented in Table 2. Of the compounds tested, 1-Ala-1-Asp has the highest value of $k_{cat}$ (1400 s$^{-1}$) and 1-Arg-d-Asp has the highest value of $k_{cat}/K_m$ (7.6 × 10$^5$ M$^{-1}$ s$^{-1}$). For the four enantiomers of the Ala-Ala dipeptide, the relative rates of hydrolysis are in the following order: 1-Ala-d-Ala > 1-Ala-L-Ala > d-Ala-d-Ala > d-Ala-L-Ala. The rates of hydrolysis of N-formyl-d-Glu, N-acetyl-d-Glu, and N-propionyl-d-Leu at a concentration of 2.0 mM were less than 2 × 10$^{-3}$ s$^{-1}$. Nitrocefin was slowly hydrolyzed by Sco3058, but the value of $k_{cat}/K_m$ was less than 0.1 M$^{-1}$ s$^{-1}$. The phosphinate pseudodipeptide mimics 1, 2, and 3 were found to be competitive inhibitors of Sco3058 with $K_i$ values of 390 ± 21 nM, 2.3 ± 0.3 μM, and 1.1 ± 0.2 μM, respectively. The $K_i$ of citrate was 25 ± 15 mM.

pH—Rate Profiles. The kinetic constants for the hydrolysis of the L-Arg-L-Asp dipeptide were obtained as a function of pH. The pH—rate profiles for $k_{cat}$ and $k_{cat}/K_m$ are presented in panels A and B of Figure 4, respectively. The pH profiles are bell-shaped and are consistent with a single functional group that must be protonated for catalytic activity and another group that must be protonated for catalytic activity. The pK$a$ values for the effect of pH on $k_{cat}$ (Figure 4A) were 6.2 ± 0.1 and 8.4 ± 0.1. In Figure 4B, the two ionizations observed for Sco3058 are less than 2 pK$a$ units apart, and therefore, only an average of the two pK$a$ values of 7.6 ± 0.4 can be determined (4f).

Site-Directed Mutagenesis. Site-directed mutagenesis was utilized to identify residues important for enzymatic activity. Several conserved residues in the active site, including Asp-22, His-150, Arg-223, and Asp-320, were altered, and the kinetic constants for the purified mutants are listed in Table 3. When Asp-22 from the HxD motif at the end of β-strand 1 was changed to histidine, no activity was detected even though the purified protein contained 2 equiv of zinc. His-150 was mutated to asparagine and alanine; both of these mutants exhibited a significant decrease in $k_{cat}$ (200- and 20-fold, respectively), and the value of $K_m$ more than doubled. The conserved Arg-223 was changed to lysine and methionine. The lysine mutant was reduced in activity by 3 orders of magnitude relative to the wild-type enzyme. No hydrolytic activity was detected with the methionine

<table>
<thead>
<tr>
<th>substrate</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}/K_m$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-Arg-o-Asp</td>
<td>107 ± 3.2</td>
<td>0.14 ± 0.02</td>
<td>(7.6 ± 0.9) × 10$^4$</td>
</tr>
<tr>
<td>1-Met-o-Glu</td>
<td>340 ± 4</td>
<td>0.89 ± 0.04</td>
<td>(3.8 ± 0.2) × 10$^5$</td>
</tr>
<tr>
<td>1-Met-p-Glu</td>
<td>194 ± 4</td>
<td>0.56 ± 0.03</td>
<td>(3.5 ± 0.3) × 10$^5$</td>
</tr>
<tr>
<td>1-Leu-o-Glu</td>
<td>135 ± 2</td>
<td>0.55 ± 0.03</td>
<td>(2.5 ± 0.1) × 10$^5$</td>
</tr>
<tr>
<td>1-Leu-o-Ala</td>
<td>160 ± 3</td>
<td>0.75 ± 0.04</td>
<td>(2.1 ± 0.1) × 10$^5$</td>
</tr>
<tr>
<td>1-Leu-p-Ser</td>
<td>400 ± 10</td>
<td>2.0 ± 0.2</td>
<td>(2.0 ± 1.0) × 10$^5$</td>
</tr>
<tr>
<td>1-Ala-o-Ala</td>
<td>750 ± 25</td>
<td>4.4 ± 0.3</td>
<td>(1.7 ± 0.1) × 10$^5$</td>
</tr>
<tr>
<td>1-Tyr-o-Leu</td>
<td>24 ± 0.2</td>
<td>0.17 ± 0.01</td>
<td>(1.5 ± 0.07) × 10$^6$</td>
</tr>
<tr>
<td>1-Met-o-Leu</td>
<td>310 ± 6</td>
<td>2.3 ± 0.1</td>
<td>(1.4 ± 0.1) × 10$^5$</td>
</tr>
<tr>
<td>1-Asn-o-Glu</td>
<td>288 ± 10</td>
<td>2.4 ± 0.2</td>
<td>(1.2 ± 0.1) × 10$^5$</td>
</tr>
<tr>
<td>1-Ala-o-Asp</td>
<td>1400 ± 60</td>
<td>13.8 ± 0.9</td>
<td>(1.0 ± 1.0) × 10$^5$</td>
</tr>
<tr>
<td>1-Leu-o-Leu</td>
<td>140 ± 6</td>
<td>1.4 ± 0.1</td>
<td>(1.0 ± 1.0) × 10$^5$</td>
</tr>
<tr>
<td>1-Leu-o-Asp</td>
<td>160 ± 11</td>
<td>2.0 ± 0.3</td>
<td>(8.3 ± 1.4) × 10$^4$</td>
</tr>
<tr>
<td>1-Arg-o-Asp</td>
<td>132 ± 2</td>
<td>1.8 ± 0.1</td>
<td>(7.2 ± 0.4) × 10$^5$</td>
</tr>
<tr>
<td>1-Leu-o-Asp</td>
<td>21 ± 0.3</td>
<td>0.37 ± 0.03</td>
<td>(5.6 ± 0.4) × 10$^4$</td>
</tr>
<tr>
<td>Gly-o-Asp</td>
<td>220 ± 14</td>
<td>9.5 ± 1.0</td>
<td>(2.4 ± 0.3) × 10$^4$</td>
</tr>
<tr>
<td>1-Ala-o-Glu</td>
<td>220 ± 14</td>
<td>9.5 ± 1.0</td>
<td>(2.4 ± 0.3) × 10$^4$</td>
</tr>
<tr>
<td>1-Ala-o-Asp</td>
<td>140 ± 17</td>
<td>16 ± 3.3</td>
<td>(8.4 ± 1.9) × 10$^2$</td>
</tr>
<tr>
<td>1-Ala-o-Asp</td>
<td>140 ± 17</td>
<td>16 ± 3.3</td>
<td>(8.4 ± 1.9) × 10$^2$</td>
</tr>
</tbody>
</table>

*From fits of the data to eq 2. For those entries without values of $k_{cat}$ and $K_m$, saturation was not achieved at concentrations up to 5 mM.


The solvent isotope effects for D$_2$O for the wild-type enzyme in H$_2$O and D$_2$O with L-Arg-L-Asp effects were utilized to determine the effect of proton transfers (mutant.

pseudodipeptide Ala-Asp inhibitor (\textit{mutant}).

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|c|}
\hline
enzyme & $k_{\text{cat}}$ (s$^{-1}$) & $K_m$ (mM) & $k_{\text{cat}}/K_m$ (M$^{-1}$ s$^{-1}$) \\
\hline
wild type & 132 ± 2 & 1.8 ± 0.1 & (7.2 ± 0.4) x 10$^4$ \\
D22H & <0.03 & & \\
H150N & 0.7 ± 0.01 & 4.9 ± 2.1 & (1.4 ± 0.4) x 10$^2$ \\
H150A & 7.2 ± 0.4 & 4.3 ± 0.5 & (1.7 ± 0.2) x 10$^3$ \\
R223K & 0.4 ± 0.04 & 8.8 ± 1.4 & 46 ± 9 \\
R223M & <0.03 & & \\
D320N & <0.03 & & \\
D320A & <0.03 & & \\
\hline
\end{tabular}
\caption{Kinetic Parameters for Mutants of Sco3058$^a$}
\end{table}

$^a$These constants were obtained by a fit of the data to eq 2.

The conserved Asp-320 from \textit{beta}-strand 8 was modified to asparagine and alanine, and no activity was observed with either mutant.

\textbf{Solvent Isotope and Viscosity Effects.} Solvent isotope effects were utilized to determine the effect of proton transfers during substrate turnover. The kinetic parameters were obtained for the wild-type enzyme in H$_2$O and D$_2$O with L-Arg-L-Asp as the substrate. The solvent isotope effects for D$_2$O/$k_{\text{cat}}$ and D$_2$O/($k_{\text{cat}}/K_m$) are 1.30 ± 0.02 and 1.10 ± 0.05, respectively (data not shown). Alterations in solvent viscosity using sucrose were utilized to probe the degree of rate limitation for the binding and dissociation of products and substrates on the kinetic constants of Sco3058 (42, 43). The plot of $^6k_{\text{cat}}/^6k_{\text{cat}}$ versus the relative solvent viscosity for the wild-type enzyme exhibits a slope of 0.05 ± 0.03, and the plot for the effect of viscosity on $k_{\text{cat}}/K_m$ is 0.50 ± 0.05.

\textbf{Structure of Sco3058.} The crystal structures of Sco3058 and two additional complexes with either glycerol and citrate or the pseudodipeptide Ala-Asp inhibitor (I) in the active site were determined to 1.3, 1.7, and 1.4 Å, respectively. A ribbon representation of the dimeric (β/α)$_2$ barrel structure is presented in Figure 5. The zinc in the α-site is coordinated to His-20 and Asp-22 from the HxD motif that originates from the end of \textit{beta}-strand 1. The zinc in the β-site is ligated to the two conserved histidines (His-191 and His-212) from the ends of \textit{beta}-strands 5 and 6, respectively. The two metals are bridged by Glu-123 from \textit{beta}-strand 3 and a hydroxide/water molecule. The next water molecule nearest to either of the two metal ions is 2.4 Å from M$_β$. The conserved aspartate (Asp-320) from the end of \textit{beta}-strand 8 is 3.6 Å from M$_α$ and thus does not coordinate this metal. However, this residue interacts with the bridging solvent molecule at a hydrogen bonding distance of 3.0 Å. The binuclear metal center is shown in Figure 6A.

The active site of Sco3058 in the presence of glycerol and citrate is shown in Figures 6B and 7A. In the glycerol- and citrate-bound structure, the α-metal is tetrahedral in coordination while the β-metal exhibits a distorted trigonal bipyramidal geometry. In addition to the four ligands described for the native structure, the β-metal is additionally coordinated to the C3 carboxylate of citrate. This carboxylate is also ion-paired with Arg-223 and occupies the same position as the water molecule, which is 2.4 Å from M$_β$ in the native structure. The carboxylate of citrate from the \textit{pro-S} arm is ion-paired to His-150, while the carboxylate from the \textit{pro-R} arm is interacting with Thr-324. The hydroxyl group attached to C3 of citrate is 2.8 Å from Asp-320 and is also hydrogen bonded with the C2 hydroxyl group of glycerol. In addition, the C1 hydroxyl group of glycerol is within hydrogen bonding distance (3.0 Å) of Asp-22 and 2.4 Å from the α-metal.

The $F_o - F_c$ electron density map that illustrates the orientation of the pseudodipeptide inhibitor I in the active site of Sco3058 is presented in Figure 8. The two chiral centers that mimic the C$_α$ stereocenters of the N- and C-terminal ends of a dipeptide are both in the R-configuration. This would correspond to the dipeptide L-Ala-D-Asp. In this complex, the β-zinc is coordinated by five ligands in a distorted trigonal bipyramidal arrangement. One of these interactions originates from the phosphinate group, and another is from the carboxylate that would correspond to the C-terminal carboxylate of a dipeptide substrate. The α-metal is in a distorted square pyrimidal arrangement, and the α-amino group of the inhibitor is 2.3 Å from this metal ion. The C-terminal carboxylate of the inhibitor is ion-paired with the guanidino group of Arg-223. His-150 makes polar interactions with the phosphinate moiety at a distance of 2.7 Å. The side chain carboxylate of the inhibitor interacts with Thr-324 at a distance of 2.7 Å. The orientation
of inhibitor 1 in the active site of Sco3058 is presented in Figures 6C and 7B.

**Docking of Dipeptides.** The docking results were postprocessed to exclude all poses in which the high-energy intermediate portion of the molecule was more than 4 Å from one of the two metal ions. The top 100 of the remaining poses were inspected visually to reject molecules with intramolecular clashes and geometries incompatible with the reaction mechanism. This left 63 molecules for further consideration. Among these, there is no clear preference for a specific residue at the N-terminus; in fact, most of the 63 molecules are capped (i.e., N-acetylated, N-formylated, etc.). At the C-terminus, 41% of the amino acids are in an L-configuration, and 59% are D (no glycine residues were observed at this position). Upon examination of the residue types in more detail, 11% are L-glutamate, 29% are D-glutamate, 14% are L-aspartate, 21% are D-aspartate, and 3% each are...
L-asparagine, D-asparagine, D-serine, and L-glutamine. These computed specificities, at the C-terminus, may be compared with those determined experimentally (Figure 3).

The computationally determined poses of the eight diastereomers of the Ala-Ala dipeptide fall into two categories based on the stereochemistry at the HEI center. These would include molecules with an HEI center that would originate from an attack of the water molecule on the re face of the amide bond and those that would originate from an attack on the si face. In all cases, except for D-Ala-D-Ala, the intermediates corresponding to an attack from the si face docked in unfavorable binding modes, either forming intramolecular clashes or pointing the negatively charged oxygen (corresponding to the attacking hydroxyl) away from the catalytic zinc ions. The C-terminal carboxy moiety is always in contact with Arg-223, whereas the N-terminal amino group can form polar interactions with either the backbone carbonyl of Gly-323, the side chain of Tyr-68, or Asp-22. The proposed orientations of L-Ala-D-Ala and D-Ala-L-Ala in the active site of Sco3058 are presented in panels A and B of Figure 9, respectively.

The overlay of the docking-derived pose of L-Ala-D-Ala with the crystallographic pose of the pseudodipeptide 1 (Figure 10) shows very good agreement in the parts that are similar. The phosphinate moiety of 1 overlaps almost perfectly with the tetrahedral HEI portion of the dipeptide, and the C-terminal carboxy moieties both interact with Arg-223. The two backbones are lined up to a high degree, with the exception of the N-terminal amino groups, which are rotated by 90° with respect to each other.

**DISCUSSION**

**Substrate Specificity.** The substrate specificity of Sco3058 was determined using an array of dipeptide libraries. It is evident from the data presented in Table 2 that this enzyme has the ability to hydrolyze a rather broad range of dipeptides having amino acids of either the D- or L-configuration. However, at the N-terminus, amino acids of the l-configuration are greatly preferred over those of the D-configuration. For example, the values of $k_{cat}/K_m$ for L-Ala-D-Ala and L-Ala-L-Ala are 2–3 orders of magnitude greater than those of their D-Ala-D-Ala and D-Ala-L-Ala counterparts. Conversely, at the C-terminus, amino acids of the D-configuration are preferred over those of the l-configuration. Thus, the values of $k_{cat}/K_m$ for l-Arg-D-Asp and l-Ala-D-Ala are ~1 order of magnitude greater than those of l-Arg-L-Asp and l-Ala-L-Ala, respectively. At the N-terminus, there is a preference for arginine, methionine, glutamine, and leucine, whereas at the C-terminus, there is a clear preference for glutamate and aspartate. The very best substrates have $k_{cat}/K_m$ values that exceed $10^5$ M$^{-1}$ s$^{-1}$. N-Acyl derivatives of D-amino acids are hydrolyzed approximately 5 orders of magnitude more slowly than are the corresponding dipeptide substrates, and thus, this enzyme appears to be a true dipeptidase rather than a carboxypeptidase.

**Structural Determinants of Substrate Specificity.** The three-dimensional structure of Sco3058 was determined in the absence of ligands and also in the presence of citrate and a pseudodipeptide analogue bound in the active site. The analogue of the proposed tetrahedral intermediate was synthesized without regard for the stereochemistry at the two chiral centers. The structure of the inhibitor bound in the active site is a mimic of the l-Ala-D-Asp dipeptide. As a structural mimic of the tetrahedral addition complex, this analogue is compromised since the amide nitrogen has been substituted with a methylene group to maintain the hydrolytic stability of the inhibitor. Nevertheless, the structure of this complex indicates that the free α-amino group at the N-terminus of the pseudodipeptide interacts with $M_α$ and the side chain carboxylate of Asp-22. Conversely, the free α-carboxylate at the C-terminus interacts with $M_β$ and the side chain guanidino...
A similar rationale can be employed at the N-terminus. The most likely interactions formed by a positively charged N-terminal amino group of a dipeptide substrate are with the backbone carbonyl of Gly-323 and the side chain of Asp-22. Under these circumstances, the side chain of an Ala in the D-configuration would come too close to itself, forming an intramolecular clash (Figure 9B). However, the structural preference at the N-terminus is more subtle, and it is likely not the only factor contributing to the observed differences in the turnover rates. This is also apparent from the overlay of the L-Ala-D-Ala dipeptide with the pseudodi peptide (1) in the X-ray structure where the most similar parts overlay well but the N-terminus shows the highest divergence.

**Mechanism of Action.** One can utilize the structure of Sco3058 in the presence of the pseudodi peptide inhibitor to propose a mode of binding for typical dipeptide substrates and a working model for the chemical mechanism of substrate activation and hydrolysis. In this model, the N-terminal α-amino group of dipeptide substrates interacts with Asp-22 and $M_\beta$ while the C-terminal α-carboxylate group interacts with $M_\delta$ and Arg-223. The carbonyl group of the amide bond about to be hydrolyzed is polarized by $M_\beta$ with assistance from His-150. Since the mutation of His-150 to an alanine reduces $k_{\text{cat}}$ by a factor of only 20, the presence of this histidine in the active site is not critical for substrate turnover. His-150 from Sco3058 aligns with His-152 in the human renal dipeptidase. The hydrolytic water is the bridging water between the two metals (see Figure 6A). The binding of the α-carboxylate group to $M_\delta$ occurs with the displacement of the nonbridging water molecule to this metal ion. The phosphinate moiety of the inhibitor mimics the addition of hydroxide to the carbonyl carbon. After attack by the bridging hydroxide on the carbonyl carbon, a proton is transferred to the side chain carboxylate of Asp-320 and another proton is transferred to the leaving group amine. Asp-320 is hydrogen bonded to the bridging hydroxide and is the closest residue to the hydroxyl at C3 of citrate. This residue is the only one in the active site that is positioned to facilitate the transfer of a proton from the bridging hydroxide to the leaving group amine. This mechanism is illustrated in Scheme 3, and it is very similar to those previously proposed for other members of the amidohydrolase superfamily (16–18).

**Rate-Limiting Steps.** The reaction mechanism was also addressed with the measurement of pH-rate profiles and the effects of solvent viscosity and D$_2$O. The changes in the kinetic parameters are consistent with a mechanism where the rate-limiting step is the hydrolysis of the amide bond.

The reaction proceeds via a threestep mechanism: (1) protonation of the amide nitrogen by an asparagine residue (Lys-247) to form a tetrahedral intermediate; (2) hydrolysis of the amide bond; and (3) elimination of water. The reaction is catalyzed by a histidine residue (His-150) that serves as a general base catalyst.

In the active site of Sco3058, the L-alanine residue at the C-terminus of the substrate interacts with the phosphate group of the inhibitor, which mimics the pseudodi peptide inhibitor. The hydroxyl group of the inhibitor interacts with Thr-324, and the side chain of Asp-320.

**Docking of Dipeptides.** The docking of the full spectrum of dipeptides reproduced the observed preference for α-enantiomers of the C-terminal amino acids found experimentally. Moreover, the statistics of the top 100 poses show a dominance of acidic residues at the C-terminus, consistent with experiment. There were also discrepancies. Experimentally, aspartate is preferred toward Lys-247. Conversely, when this Ala is in the D-configuration, the methyl is positioned more favorably in a hydrophobic pocket formed by Val-245 and Phe-248 (Figure 9A). This is consistent with the observed preference for a D-configuration of the C-terminal amino acid and is also borne out in the docking scores of both the L-1,L-1 and D-1,D-1 Ala-Ala pairs. The variants with the C-terminal amino acid in the D-configuration receive a more favorable score.

The interactions in the docking poses also help explain the differences in the enzyme specificities, measured by $k_{\text{cat}}/K_m$, for the four Ala-Ala diastereomers. The docked poses show that if the C-terminal Ala is the l-enantiomer, its methyl group points toward Lys-247. Conversely, when this Ala is in the α-configuration, the methyl is positioned more favorably in a hydrophobic pocket formed by Val-245 and Phe-248 (Figure 9A). This is consistent with the observed preference for a D-configuration of the C-terminal amino acid and is also borne out in the docking scores of both the L-1,L-1 and D-1,D-1 Ala-Ala pairs. The variants with the C-terminal amino acid in the D-configuration receive a more favorable score.
Article Biochemistry, Vol. 49, No. 3, 2010 621

Position. In the case of the renal dipeptidase-like protein from sequences having a histidine and 70% a tryptophan at this Sco3058 is semiconserved with approximately 30% of the subset of cog2355E, the residue equivalent to His-150 in sequences found in cog2355E. Within the renal dipeptidase equivalent to Arg-223 is conserved in more than 95% of the ε to His-150 from Sco3058 with the nitrogen of the tryptophan side 218766909), this tryptophan residue is structurally equivalent the reactions were conducted in D2O, there was a rather modest recognition of the Xaa-carboxylate of the dipeptide, and the residue product specificity profiles and the three-dimensional structures of several microbial renal dipeptidase-like proteins are currently under investigation using the same methods applied to Sco3058.

REFERENCES


