Structural Basis for Substrate Recognition and Hydrolysis by Mouse Carnosinase CN2*  

Hideaki Unno1,2, Tetsuo Yamashita1,3, Sayuri Ujita4, Nobuaki Okumura5, Hiroto Otani, Akiko Okumura, Katsuya Nagai6, and Masami Kusunoki7  

From the Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

1-Carnosine is a bioactive dipeptide (β-alanyl-L-histidine) present in mammalian tissues, including the central nervous system, and has potential neuroprotective and neurotransmitter functions. In mammals, two types of 1-carnosine-hydrolyzing enzymes (CN1 and CN2) have been cloned thus far, and they have been classified as metalloendoproteases of the M20 family. The enzymatic activity of CN2 requires Mn2+ and CN2 is inhibited by a nonhydrolyzable substrate analog, bestatin. Here, we present the crystal structures of mouse CN2 complexed with bestatin together with Zn2+ at a resolution of 1.7 Å and that with Mn2+ at 2.3 Å. CN2 is a homodimer in a noncrystallographic asymmetric unit, and the Mn2+ and Zn2+ complexes closely resemble each other in the overall structure. Each subunit is composed of two domains: domain A, which is complexed with bestatin and two metal ions, and domain B, which provides the major interface for dimer formation. The bestatin molecule bound to domain A interacts with several residues of domain B of the other subunit, and these interactions are likely to be essential for enzyme activity. Since the bestatin molecule is not accessible to the bulk water, substrate binding would require conformational flexibility between domains A and B. The active site structure and substrate-binding model provide a structural basis for the enzymatic activity and substrate specificity of CN2 and related enzymes.

1-Carnosine (β-alanyl-L-histidine) and structurally related dipeptides, such as homocarnosine (γ-aminobutyryl-L-histidine) and anserine (β-alanyl-L-1-methylhistidine) are distributed in a wide variety of vertebrate tissues (1). 1-Carnosine is present at particularly high concentrations in mammalian skeletal muscles and the brain, and it has been implicated in neuroprotection (2), the olfactory system (1), and hypothalamic neuronal networks (3). Our recent observations suggest that central and peripheral administration of 1-carnosine at low doses attenuates 2-deoxyglucose-induced hyperglycemia (4) and suppresses peripheral sympathetic nerve activity (5, 6). These effects of 1-carnosine are suppressed by central administration of thioperamide, a histamine H3 blocker. This suggests that 1-carnosine regulates the autonomic nervous system via the hypothalamic histaminergic neurons (4 – 6). In addition, the dipeptide exhibits antioxidant and free radical scavenger properties via complexation of transition metals, such as zinc and copper, suggesting that it is also involved in neuroprotection from oxidative stress (2, 8, 9).

1-Carnosine is synthesized from β-alanine and L-histidine by carnosine synthetase and is degraded by intra- and extracellular dipeptidases known as carnosinases. Their enzymatic activities are regulated under various physiological conditions (10). Carnosinase was first isolated (11) from the porcine kidney in 1949 and was subsequently found to be widely distributed in tissues of rodents and higher mammals (12 – 15). Recently, two types of carnosinases were identified in humans and mice: human CN1 (also known as CNPD1 or CNDP dipeptidase 1), human CN2 (CNPD2 or CNDP dipeptidase 2) (16), and mouse CN2 (17). The biochemical properties of these types were investigated in detail. CN2 is present in mammalian brain and is especially abundant in the tuberomammillary nucleus of the hypothalamus, the thalamic parafascicular nucleus, neuronal fibers, and the mitral cell layer of the olfactory bulb in the nervous system (17). The cell bodies of histaminergic neurons localize in the tuberomammillary nucleus, suggesting that CN2 is involved in histamine synthesis in the histaminergic neurons, possibly by supplying L-histidine as the substrate for the histamine-synthesizing enzyme histidine decarboxylase.

Sequence-based alignments of human CN1 and human CN2 with mouse CN2 show sequence identities of 53 and 91%, respectively. Human CN1 was identified as a dipeptidase that hydrolyzes Xaa-His dipeptides, including those with first residues β-Ala (carnosine), γ-aminobutyric acid (homocarnosine), N-methyl-β-Ala, Ala, and Gly. On the other hand, CN2 has a broader specificity than CN1, but it does not hydrolyze homocarnosine and is sensitive to inhibition by bestatin (IC50 = 7 nM).
Crystal Structure of Mouse Carnosinase CN2

Unlike most other metallopeptidases, CN2 requires Mn\(^{2+}\) for complete activity, and Zn\(^{2+}\) alone cannot activate this enzyme.

Based on the similarity in primary sequences, CN1 and CN2 have been classified as metallopeptidases belonging to the M20 family of clan MH (18). Of these, the crystal structures of PepV from Lactobacillus delbrueckii (19) and CPG2 (carboxypeptidase G2) from Pseudomonas sp. (20), which share 17 and 18% sequence identities with mouse CN2, respectively, have been reported thus far. PepV and CPG2 are composed of two domains: one catalytic domain with two Zn\(^{2+}\) ions at the active center and one noncatalytic domain known as the lid domain or the dimerization domain. The dimerization domain of CPG2 provides the surface for the same interaction to form a homodimer structure, whereas PepV is present as a monomer due to the different structural features of the lid domain. Furthermore, the crystal structure of a member of the M28 family of dinuclear zinc aminopeptidases from Aeromonas proteolytica (AAP)\(^8\) (21) is similar to the catalytic domain of PepV and CPG2. However, AAP does not have a noncatalytic domain and is present as a monomer in solution.

In an attempt to determine the structure of CN2, we obtained two types of CN2 crystals: one complexed with Zn\(^{2+}\) (Zn\(^{2+}\) complex) and the other with Mn\(^{2+}\) (Mn\(^{2+}\) complex). Here, we report the structures of the two forms of CN2 crystals. We identified the residues crucial for L-carnosine-specific binding and catalysis and gained a structural basis for explaining the differences in the substrate specificity of CN2 and the related enzymes. Then we discuss metal ion selectivity and conformational flexibility of CN2 based on the structures of CN2 and other members of the M20/M28 peptidase family.

MATERIALS AND METHODS

Purification of Mouse Carnosinase CN2—Otani et al. (17) described the cloning, expression, and purification of CN2. For crystallization, the purification procedures were slightly modified from that described by Yamashita et al. (22). The cDNA encoding CN2 was subcloned into the expression vector pGEX-4T3 (Amersham Biosciences). The CN2 protein fused with glutathione S-transferase was then overexpressed in the Escherichia coli strain BL21(DE3)pLysS in the presence of 1 mM isopropyl 1-thio-D-galactopyranoside and 2 mM MnCl\(_2\) for 12–16 h at 25 °C. The E. coli cells were collected and sonicated in a buffer containing 25 mM Tris-HCl (pH 7.4), 50 mM NaCl, 0.2 mM MnCl\(_2\), and 1 mM dithiothreitol. After centrifugation, the supernatant was mixed with glutathione-Sepharose beads (Amersham Biosciences), washed, and incubated with thrombin (Amersham Biosciences) at 25 °C for 12–16 h to obtain the full-length CN2. The soluble fraction was then separated on a gel filtration column (HiLoad 26/60 Superdex 200 pg; Amersham Biosciences) and an anion exchange column (Hi-trap Q; Amersham Biosciences). Proteins were eluted from the column with a linear gradient of 50–750 mM NaCl in the same buffer.

Crystallization—Crystallization of CN2 complexed with Mn\(^{2+}\) was carried out using the hanging drop vapor diffusion method. Two microliters of a protein solution (20 mg/ml) containing 25 mM Tris-HCl (pH 7.4), 50 mM NaCl, 0.2 mM MnCl\(_2\), 1 mM dithiothreitol, and 30 mM bestatin were mixed with an equal volume of a reservoir solution and allowed to equilibrate against 0.5 ml of the reservoir solution at 293 K. Wing-shaped crystals were obtained using a reservoir solution containing 20% polyethylene glycol 3350 and 0.2 mM KF. The crystal size and quality were further improved with a combination of the macroseeding and microseeding techniques (23). A single crystal with dimensions of 0.4 × 0.2 × 0.1 mm was obtained in a hanging drop in 20% (w/v) polyethylene glycol 3350, 20% (w/v) glycerol, and 0.2 mM KF. CN2 crystals complexed with zinc ions were prepared in the same way, but Mn\(^{2+}\) ions were removed by transferring the crystallins into a harvest solution containing 20% polyethylene glycol 3550, 20% glycerol, and 0.2 mM KF for 1 day before x-ray data collection. We obtained two types of crystals, Zn\(^{2+}\) and Mn\(^{2+}\) complexes, as confirmed by x-ray absorption fine structure spectra (supplemental Fig. S1). Since Zn\(^{2+}\) ions were not added during purification, crystallization, and crystal harvesting procedures, the Zn\(^{2+}\) ions in the Zn\(^{2+}\) complex crystal are thought to be derived from the culture medium and held by CN2 during the purification, crystallization, and crystal harvesting procedures. However, since zinc ions were not completely removed from buffers for purification and soaking, we could not completely exclude the possibility that the Zn\(^{2+}\) ions were derived from these buffers.

Data Collection, Structure Solution, and Refinement—A CN2 crystal was mounted in a nylon cryoloop (Hampton Research) and placed directly into a nitrogen stream at 100 K. The metal type in a crystal was determined by x-ray absorption fine structure spectra. A crystal containing only Mn\(^{2+}\) ions was used for Mn\(^{2+}\) multiple-wavelength anomalous dispersion, and that containing only Zn\(^{2+}\) ions was for Zn\(^{2+}\) multiple-wavelength anomalous dispersion data collection. Data collection was carried out using synchrotron radiation at beamline NW-12 of the Photon Factory (KEK, Tsukuba, Japan). For the Mn\(^{2+}\) complex, two wavelengths at the Mn-K absorption edge (1.8941 Å) and peak (1.8926 Å) and one remote wavelength (1.7926 Å) were used. For the Zn\(^{2+}\) complex, wavelengths at the Zn-K absorption edge (1.2834 Å), peak (1.2827 Å), and one remote wavelength (1.2573 Å) were used. High resolution data for the Zn\(^{2+}\) complex were collected at 1.00 Å. A complete data set was collected through contiguous rotation ranges at a particular wavelength before proceeding to the next wavelength. Rotation data were recorded in frames of 1° oscillation. The data collected at different wavelengths were processed with the program HKL2000 (24). The phases were calculated using the program SHARP (25). After solvent flattening and density modification, protein models for the Mn\(^{2+}\) and Zn\(^{2+}\) complexes were constructed independently by the program ARP/wARP (26), and the models were improved with the program O (27). Each model was refined using the program Refmac5 (28). All figures were produced using MOLSCRIPT (29), POVSCRIPT + (30), PyMOL (31), and RASTER3D (32).

Enzyme Assay—cDNA encoding a CN2 mutant carrying a His\(^{228}\) to Ala substitution was synthesized using QuickChange (Stratagene), and the nucleotide sequence was verified. The protein was then expressed as described above and purified by glutathione-Sepharose. After thrombin digestion, it was further

---

\(^8\) The abbreviation used is: AAP, aminopeptidase(s) from A. proteolytica.
purified by ion exchange chromatography on a Hitrap-Q column (Amersham Biosciences). The purified protein (20 mg/ml) was incubated in a buffer containing 50 mM Tris-HCl (pH 8.8), 10 mM MnCl₂, 5 mM dithiothreitol, and 10 mM L-carnosine for 30 min at 37 °C. The reaction was terminated by the addition of bestatin to yield a final concentration of 1 mM, and a 1-µl aliquot was spotted onto a cellulose plate for thin layer chromatography. Amino acids were separated by an ascending solvent composed of 15% water, 10% formic acid, 75% isopropl alcohol and detected by spraying 10 mg/ml ninhydrin solution in ethanol. For quantitative measurement of carnosinase activity, histidine was derivatized with o-phthalaldehyde at alkaline pH after enzyme reaction (33), and absorbance at 405 nm was determined.

**RESULTS**

CN2 is a metallopeptidase that hydrolyzes a variety of dipeptides, including L-carnosine, to form the corresponding amino acids (Fig. 1A). The enzymatic activity of CN2 is inhibited by bestatin (Fig. 1B), which is considered to act as a nonhydrolyzable substrate analog for a wide variety of peptidases. At first, we tried to obtain CN2 crystals without bestatin but could not obtain a crystal suitable for structural analysis. Next, we tried to obtain crystals of CN2 complexed with bestatin. After optimizing crystallization conditions, we obtained two forms of crystals of CN2-bestatin complex. One included Mn²⁺ ions (Mn²⁺ complex), and the other included Zn²⁺ ions (Zn²⁺ complex), as determined by x-ray absorption fine structure (supplemental Fig. S1). Both types were isomorphs to each other and crystallized in space group P2₁ with one homodimer in a crystallographic asymmetric unit. The structures of these two forms of CN2 were determined by the multiple-wavelength anomalous dispersion method independently, and they were refined to resolutions of 1.7 Å (Zn²⁺ complex) and 2.3 Å (Mn²⁺ complex). The data collection and refinement statistics are listed in Table 1.

**TABLE 1**

Data collection and refinement statistics

| Parameters | Mn²⁺ complex | Zn²⁺ complex |
|------------|--------------|--------------|
| Wavelength (Å) | 1.8941 | 1.8926 |
| Resolution (Å) | 50.0-2.8 (2.91-2.81) | 50.0-2.8 (2.90-2.80) |
| Measured | 183799 | 186077 |
| Completeness (%) | 99.9 (99.1) | 99.3 (94.4) |
| Rmerge (%) | 1.7926 | 264160 |
| a = 54.49, b = 199.18, c = 55.21, β = 118.92 | 96.1 (74.4) | 46.2 (5.4) |
| Crystallographic asymmetric unit | a = 54.41, b = 199.77, c = 55.49, β = 118.52 |
| Space group | P2₁ | P2₁ |
| Crystal structure of mouse carnosinase CN2 |

**FIGURE 1. Structures of L-carnosine and bestatin.** A, structure of L-carnosine. CN2 hydrolyzes L-carnosine (β-alanyl-L-histidine) at the indicated position to yield β-alanine and L-histidine. B, structure of bestatin. Bestatin is a substrate-mimicking inhibitor of CN2; its nitrogen and oxygen atoms are as numbered. The amino group side is indicated as P1, whereas the carboxyl group side is labeled P1'.
The asymmetric unit is composed of one homodimer shaped like a curved cylinder having dimensions of $\sim 45 \times 55 \times 115 \AA$. The overall crystal structure of the Mn$^{2+}$ complex is shown in Fig. 2. The two polypeptide chains were related by noncrystallographic 2-fold symmetry in a dimer, and a total of 478 amino acid residues were observed in the structure; of the 478 residues, 473 were residues of CN2 protein, and 5 N-terminal residues (Gly-Ser-Pro-Asn-Ser) were derived from the expression vector pGEX-4T3. The two subunits are tightly associated with each other, and each has an active site structure containing two metal ions and one bestatin molecule.

**Overall Structure**—The structure of each subunit is divided into two domains, A (residues 1–203 and 416–476) and B (residues 204–415) (Fig. 3). Each domain has an $\alpha/\beta$-fold structure consisting of $\beta$-strands and $\alpha$-helices that were numbered as shown in Fig. 3. Domain A essentially comprises 12 $\alpha$-helices ($\alpha_1$–$\alpha_9$, $\alpha_17$–$\alpha_19$) and a large twisted $\beta$-sheet in which the $\beta$-strands are arranged in the order of $\beta_1$–$\beta_4$–$\beta_8$–$\beta_5$–$\beta_9$–$\beta_17$, with $\beta4$ representing the only antiparallel strand. In addition, a second smaller two-stranded $\beta$-sheet ($\beta_6$–$\beta_7$) and a third two-stranded $\beta$-sheet ($\beta_10$–$\beta_{18}$) are located above $\alpha$-helix $\alpha_6$. Domain B also has a large twisted $\beta$-sheet ($\beta_{13}$–$\beta_{14}$–$\beta_{11}$–$\beta_{15}$–$\beta_{16}$) and seven $\alpha$-helices ($\alpha_{10}$–$\alpha_{16}$) together with an additional small two-stranded $\beta$-sheet ($\beta_{12}$, $\beta_{16}$). Domain B is connected to domain A by two $\beta$-strands, namely, $\beta_{11}$ and $\beta_{17}$. Domains A and B constitute a continuous $\beta$ sheet structure. The subunit topology is illustrated in supplemental Fig. S2.

**Structural Comparison of CN2 with Other M20/M28 Family Metallopeptidases**—The structure of mouse CN2 was compared with related dinuclear metallopeptidases of the M20 and M28 families (supplemental Fig. S3, A–D). *L. delbrueckii* aminopeptidase PepV (34) and *Pseudomonas* sp. CPG2 (carboxypeptidase G2) (20) are M20 family peptidases, and both of these peptidases have two domains consisting of the catalytic domain and the lid or dimerization domain. In contrast, AAP (A. proteolytica aminopeptidase) (21), which belongs to the M28 family, is a single domain protein having only one catalytic domain and is present in monomeric form in the crystal.

Domain B of CN2 is topologically similar to the dimerization domain of CPG2. This domain provides the binding interface for homodimer formation. Structural alignment of the $\alpha$-car-
bon atoms of 212 residues in CN2 domain B against those of the CPG2 dimerization domain revealed that 76 residues of CPG2 structurally corresponded to residues of CN2 with a root mean square deviation of 1.82 Å. Recently, a crystal structure of another M20 family member, *Streptococcus pneumoniae* metallopeptidase, was deposited with the Protein Data Bank (code 2pok). This protein also has a similar domain architecture, and its 204 residues corresponded to the 212 residues of CN2 domain B with a root mean square deviation of 3.1 Å for Cα atoms. Sequence identities of CN2 with CPG2 and *Streptococcus* peptidase for these peptide segments are 28 and 13.2%, respectively, and no structural similarity was found in other structures in a search by the Dali server (available on the World Wide Web). Despite the structural similarity, the three proteins show different orientations in the dimer architectures (supplemental Fig. 3, A and B). In CPG2 and the *Streptococcus* peptidase, the dimerization domain of one subunit interacts only with the dimerization domain of the other subunit but not with any residues of the catalytic subunits. Thus, the dimerization domain appears to be an independent structural unit that is responsible for dimer formation. In contrast, domain B of CN2 interacts not only with domain B of the other subunit but also with domain A of both subunits. In addition, several residues of domain B interact with the active site of the other subunit as described below.

**Subunit Interactions**—Two subunits of CN2 are related by noncrystallographic 2-fold rotation symmetry with an interface area of 3500 Å². These interactions are mediated mainly between four peptide segments; α11, β13, loops L1 (residues 224–237), and L2 (residues 323–331) (Figs. 2 and 3). The main interaction interface is constructed between α11 and β13 of one subunit and those of the other subunit. Loop L2 consists of 9 amino acid residues between β13 and β14 in domain B of one subunit, which interacts with β13 in domain B of the other subunit. The other side of loop L2 constructs part of the bestin-binding cleft of the other subunit. Loop L1 is located between β11 and α11 and consists of a loop, a short β-strand β12, and a short α-helix (α10). Loop L1 of one subunit interacts not only with α14 and α15 in domain B of the other subunit but also with a loop (residues 415–418) between β16 and α18 and a loop (residues 436–449) between β1 and α19 of domain A. These domain B-domain B and domain A-domain B interactions enhance the interface surface area between the two subunits in CN2 (3500 Å²) compared with that of that in CPG2 (1273 Å²).

**Active Site Structure**—A CN2 homodimer contains two active sites, each of which has one bestatin and two Mn²⁺ ions that were observed to have well defined electron densities (Fig. 5, A and B). Crystal structures of peptides in complexes with bestatin for AAP (35) and leucine aminopeptidase have been reported (36), where bestatin-binding sites were located on the protein surface of one subunit. In contrast, each active site of CN2 is located in a cleft constructed by domains A and B of one subunit together with domain B of the other subunit. In this cleft, bestatin interacts with Mn²⁺ ions and amino acid residues of several different peptide segments (Fig. 5C). Of these, 6 are residues on domain A (i.e. Glu¹⁶⁶ on α8; Asp¹⁹⁵ and Tyr¹⁹⁷ on a loop between β9 and β10; and Glu⁴¹¹, Gly⁴¹³, and Ser⁴¹⁷ on a loop between β16 and α18). On the other hand, His³⁰⁰ is from β15 in domain B and contributes to connect the two domains. Moreover, it should be noted that bestatin in one subunit also interacts with several residues on the other subunit. These include His²²⁸ and Val²³¹ on loop L1 and Thr³³⁰ on loop L2 of the other subunit, where primes refer to residues on the other subunit.

To confirm that His²²⁸ of domain B is involved in substrate recognition, we synthesized a recombinant protein carrying a mutation at His²²⁸ to Ala (H228A), and its enzymatic activity was examined (Fig. 5, D and E). Incubation of l-carnosine with recombinant wild-type CN2 resulted in l-carnosine hydrolysis, yielding β-alanine and l-histidine, as detected by thin layer chromatography. On the other hand, the activity was not detected in the H228A mutant protein. CD spectra of wild type and H228A mutant proteins had almost the same shape in the range of 198–250 nm, implying that the H228A mutant protein had no aberration in the stability and folding (supplemental Fig. S4). These results indicate that His²²⁸ is really involved in the enzymatic reaction on the active site of domain A of the other subunit.

In the CN2 complex, a bestatin molecule “rides” on top of the line connecting the two metal ions in the same way as phosphate complexed with PepV (33). P1 and P1’ side chains of bestatin are accommodated in hydrophobic pockets (S1 and S1’ pockets) adjacent to the dinuclear metal ions in the active site. The S1 pocket, which is composed of Leu²¹⁰, Gly⁴¹⁶, Tyr¹⁹⁷, Glu⁴¹⁴, Val²³¹, and His²²⁸’, is the possible N-terminal side chain pocket, as judged by the structural similarity between bestatin and l-carnosine (Fig. 1). On the other hand, the S1’ pocket, which is composed of His³⁴⁰, Thr³³⁶, Ile³³³, and Tyr²¹⁵, is possibly the C-terminal side chain pocket. Whereas the S1 pocket is almost closed with a very narrow channel to the bulk solvent, the S1’ pocket is connected to the neighboring cavity, which is composed of residues that include Gin⁴⁰³, Glu⁴⁰⁷, and Pro⁴¹⁹ and opens toward the cleft between domains A and B. This cavity is also so narrow that the release of the bestatin molecule from the enzyme does not appear to be possible without a conformational change in the protein structure. Molecular surface models are given in supplemental Fig. S5.

The bestatin binding is stabilized by metal ions and hydrogen bond networks (Figs. 4 and 5C). The O2 oxygen atom of bestatin coordinates to both the metal ions and forms hydrogen bonds with the carboxylate oxygens of Glu¹⁶⁶. The O3 oxygen atom coordinates to Metal 1, and, in addition, it also forms two hydrogen bonds with an amino acid of the other subunit, namely His²²⁸’. The N1 nitrogen atom forms hydrogen bonds with the main and side chain oxygen atoms of Ser⁴¹⁷ and Glu¹⁶⁶. The carboxylate group of the P1’ side chain interacts with a cavity created by a hydrogen bond network composed of Ser⁴¹¹, Arg³⁴⁵, Ser⁴¹⁷, Ser²²⁵, and Thr³³⁰.

**Structure of Metal-binding Sites in Mn²⁺ and Zn²⁺ Complexes**—The M20 family metallopeptidases are characterized by the presence of two metal ions in their active site. The structures of this family of proteins reported thus far are in the form of a complex with Zn²⁺ as the coordinated metal (20, 34), except that the structure of an M28 family protein has been reported in the form of a complex with Mn²⁺ (37). However, CN2 requires Mn²⁺, but not Zn²⁺, for enzymatic activity (17).
suggesting that Mn$^{2+}$/H11001 also possibly interacts with the active site. We have obtained two forms of CN2 crystals, and their structures have been determined independently; one is complexed with two Mn$^{2+}$ ions in each active site, and the other is complexed with two Zn$^{2+}$ ions. The overall structures of the Zn$^{2+}$ and Mn$^{2+}$ complexes closely resemble each other with a root mean square C$\beta$ deviation of 0.19 Å (Fig. 6). Significant differences in thermal factors between the active site structures could not be found.

Although CN2 requires Mn$^{2+}$ exclusively for its enzyme reaction, the fine structures of the metal-binding sites in the Mn$^{2+}$ and Zn$^{2+}$ complexes are highly similar. The two metal ions are coordinated octahedrally by domain A residues and bestatin both in the Mn$^{2+}$ and Zn$^{2+}$ complexes (Fig. 7, A and B). Metal 1 is coordinated by an imidazole nitrogen of His445, a carboxylate oxygen of Asp132, two carboxylate oxygens of Glu167, and two backbone carbonyl oxygens of bestatin. On the other hand, Metal 2 is coordinated by a carboxylate oxygen of Asp132, an imidazole nitrogen of His99, and two carboxylate oxygens of Asp195 together with a hydroxyl oxygen and the P1-terminal nitrogen of bestatin. Each carboxylate oxygen of Asp132 coordinates to metals 1 and 2, respectively, to form a bridge-like structure between the two metals. The distance between the two metal ions is 3.70 Å in the Mn$^{2+}$/H11001 complex and 3.63 Å in the Zn$^{2+}$/H11001 complex (Fig. 7, B and C). The differences in the metal-ligand distances are also within 0.2 Å between the two complexes.

Comparison of Primary Structures of CN2 with CPG2, PepV, and AAP—To further characterize the structural features of CN2, the amino acid sequence of CN2 was aligned with those of CPG2, PepV, and AAP by a structure-based sequence alignment program,
MATRAS (38) (supplemental Fig. S6). Although these proteins have similar folding topology, homology in the amino acid sequence was found only in limited regions. The metal ligands His 99, Asp 132, Glu 167, and His 445 are conserved among all of the four proteins, except that Asp 195 is replaced by Glu in CPG2. In addition, four additional residues are conserved among the four proteins: Asp 104 and Asp 133, which lie next to the metal ligands, and Gly 129 and Gly 414, which are components of turn structures. Moreover, 16 amino acids were conserved in CN2, CPG2, and PepV but not in AAP. On the other hand, the amino acid residues interacting with the bestatin main chain (Tyr 197, Glu 166, His 228, and His 380) and those interacting with the bestatin side chains (Arg 343, Thr 330, Ser 417, Val 231, and Leu 210) were not conserved except for Glu 166 and Arg 343.

DISCUSSION

Our crystal structure analysis presented here revealed that mouse CN2 is a homodimer related by a noncrystallographic 2-fold axis, in which each subunit has one active site as identified by the location of bestatin and metal ions. This structure is related to known structures of the M20/M28 family metallopeptidases, including CPG2, PepV, and AAP, but has several different features in terms of 1) domains contributing to the formation of the active sites, 2) metal ion selectivity, 3) structure for substrate recognition, and 4) orientation of domain B.

Subunit Organization and Active Site Architecture—Domain A of CN2 is structurally similar to the catalytic domains of PepV, CPG2, and AAP. This domain provides binding sites for two metal ions and a substrate to form the main part of the active site. Although the overall sequence homology was lower than 20%, metal ion ligands (His 99, Asp 132, Glu 167, Asp 195, and His 445) and so-called “catalytic Glu” (Glu 166) of CN2 were conserved in M20/M28 family proteins, including PepV, CPG2, and AAP except that Asp 195 is substituted with Glu in AAP. Similar to PepV, CPG2, and AAP, domain A of CN2 contains one cis-peptide bond between Asp 132 and Asp 133; this bond appears to be necessary to force the metal-bridging carboxylate such that it conforms to the correct geometry. These points suggest that the four proteins share the same evolutionary origin.
However, unlike the other M20/M28 members, the active site of CN2 of one subunit also interacts with domain B of one subunit as well as the other subunit. The active site residues deriving from the other subunit (His228, Val231, and Thr330) interact with a wide portion of bestatin, including the main chain amide nitrogen, P1-side carboxylate group, and the P1-side chain pocket. The orientation of His228 closely resembles that of His269 in the lid domain of PepV; there were no residues corresponding to Thr330 or Val231. These additional residues interacting with bestatin are likely to increase the stability or specificity of the enzyme-substrate complex.

**Reaction Mechanism**—Since the orientations of the metal ligands and bestatin in CN2 are closely similar to those of AAP and PepV, the reaction mechanism of CN2 is likely to be basically the same as that described for AAP (21, 35), CPG2 (20), and PepV (34). In CPG2, a water molecule bridging two Zn$^{2+}$ ions has been proposed to act as the nucleophile in the enzymatic reaction. On the other hand, this position is replaced by the O$_2$ oxygen of bestatin in the CN2-bestatin complex, as in the case of the AAP-bestatin (35) and leucine aminopeptidase-bestatin (39) complexes. In the active center of CN2, metal ligand residues, the catalytic Glu, and two metal ions, are organized in the same way as those of the other M20/M28 family proteins. In these peptidases, the catalytic Glu acts as a general base by promoting the nucleophilic attack of the metal-bound water on the substrate carbonyl carbon and a proton shuttle during the catalytic cycle in a manner analogous to the proposed mechanism of peptide cleavage for thermolysin (40). In addition to the catalytic Glu, structural and mutational analyses indicated that His$^{228}$ of CN2 plays an essential role in the hydrolysis reaction (i.e. His$^{228}$ polarizes the carbonyl carbon of the substrate to make the nucleophilic attack by the activated water molecule more effective and to stabilize the tetrahedral intermediate) (Fig. 8B). Amino acid residues that seem to have a similar function are also present in several other M20/M28 family peptidases, including *Streptomyces griseus* aminopeptidase and glutamate carboxypeptidase G II. Mutations in Tyr$^{246}$ of *S. griseus* aminopeptidase and Tyr$^{552}$ of glutamate carboxypeptidase G II, which are located in the position of His$^{228}$ in CN2, resulted in about 100- and 10-fold reduction of activity, respectively (41, 42), indicating that the residues are essential in the hydrolysis reaction in these peptidases. In the crystal structure of PepV in complex with a phosphinic inhibitor Asp[PO$_2$CH$_2$]AlaOH, His$^{269}$ in the lid domain may act as His$^{228}$ of CN2. Although the residue corresponding to His$^{228}$ of CN2 does not exist in the crystal structure of CPG2, His$^{229}$ in the dimerization domain could be at the position of His$^{228}$ of CN2 if it undergoes a conformational change to a closed form like a CN2-bestatin complex upon substrate binding. On the other hand, several proteins of the M28 family, including AAP,

---

**FIGURE 7.** Active site structures of CN2 complexed with Zn$^{2+}$ and Mn$^{2+}$. A, superposition of the Zn$^{2+}$ and Mn$^{2+}$ complex. Molecules with the hydrogen-bonding network in the Zn$^{2+}$ complex and Mn$^{2+}$ complex are shown in cyan and pink, respectively. Bestatin and metal-binding residues are shown as stick models. Metal ions are represented as sphere models. The hydrogen-bonding network is shown as dashed lines. B and C, schematic diagrams of bestatin bound to the Mn$^{2+}$ complex (B) and Zn$^{2+}$ complex (C) in CN2. Metal-ligand distances are shown in Å.
do not have a residue corresponding to His$^{228}$ of CN2. This implies that the M20 and M28 family have enzymes that are different in terms of the requirement of the histidine residue corresponding to His$^{228}$ of CN2.

**Metal Ion Selectivity**—The metal ion selectivity of CN2 is completely different from those of M28 family metallopeptidases. CN2 requires Mn$^{2+}$ selectively for its catalytic activity and is not activated by Zn$^{2+}$, Cu$^{2+}$, or Mg$^{2+}$ (17), whereas all of the M28 family metallopeptidases are activated by Zn$^{2+}$ (43). Metal ion selectivity of AAP was reported in detail; it is activated by Zn$^{2+}$, Co$^{2+}$, Ni$^{2+}$, and Cu$^{2+}$, whereas it is not activated by Mg$^{2+}$ or Mn$^{2+}$ (7, 44). A comparison of the active center of CN2 with those of M28 metallopeptidase family proteins revealed that the ligand residues and catalytic Glu are located in the same positions. The only remarkable difference between the active centers of CN2 and M28 peptidases is the orientation of the main chain oxygen atom of Asp$^{195}$. This position of oxygen is unique to CN2 and PepV, whereas it does not exist in any structure of the M28 family reported. In the CN2-bestatin complex, it has a hydrogen bonding interaction with the N-terminal nitrogen atom of bestatin (Fig. 8B). This suggests that the N-terminal nitrogen of bound carnosine will probably also interact with the oxygen. Furthermore, the different orientation of the main chain oxygen atom of Asp$^{195}$ could result in differential selectivity of metal ions in addition to substrate specificity. Further studies will reveal the detailed reaction mechanism explaining the metal ion selectivity of M20/M28 metallopeptidase family and participation of the main chain oxygen atom of Asp$^{195}$.

**L-Carnosine-specific Recognition**—Thus far, only a few enzymes, including CN1, CN2, and PepV, have been reported to hydrolyze l-carnosine, whereas most other peptidases presumably do not. This is probably because l-carnosine is a dipeptide containing a β-amino acid. Since the orientation of l-carnosine bound to CN2 can be regarded as similar to that of bestatin, the CN2 structures complexed with bestatin give an insight into the l-carnosine specificity of mouse CN2. One of the differences between the typical dipeptides and those containing a β-amino acid is the position of the N-terminal amino group. The N-terminal nitrogen atom of l-carnosine would have a hydrogen bond with the main chain carboxyl oxygen atom of Asp$^{195}$ (Fig. 8B). In contrast, dipeptides containing only α-amino acids could not interact with Asp$^{195}$ in CN2. Consistent with this, the orientation of the main chain oxygen atom of Asp$^{177}$ in PepV is almost the same as that of Asp$^{195}$ in CN2, but there is not any atom in this position in CPG2 and AAP. This structural analysis suggests that the main chain oxygen atom of the Asp$^{195}$ in CN2 is crucial in recognizing l-carnosine.

CN1 and CN2 are related enzymes that are able to hydrolyze dipeptides, including l-carnosine (16, 17). They have similar substrate specificity but are different in that CN1 prefers small side chains at the Xaa site of the Xaa-His peptide, whereas CN2 hydrolyzes dipeptides with more bulky amino acids at this site (16). This may be due, at least in part, to the difference in the space of the S1 pocket in the active sites. In the S1 pocket, CN2 has sufficient space for accommodating side chains, such as Ser, Val, and Ile in an X-His dipeptide. In contrast, the S1 pocket of CN1 would not have sufficient space for these side chains, because Gly$^{416}$ of CN2 is replaced by Ser$^{416}$, and this would result in collision with the side chains of X-His dipeptides. These differences in the amino acids at the P1 site correspond to the differences in the substrate specificity between CN1 and CN2 (Fig. 8, A and B).

In addition to the substrate specificity, the sensitivity to bestatin is also different between CN1 and CN2. The dipeptide-hydrolyzing activity of CN2 is inhibited by bestatin, but that of CN1 is not (16, 17). The difference in bestatin sensitivity would be due to the size of the S1 pocket, which is too small to bind to bestatin in CN1 (Figs. 1 and 8A).

**Possibility of a Conformational Change during Reaction**—Although the crystal structure of CN2 complexed with bestatin provides useful information for explaining substrate specificity, the static model appears insufficient to explain the manner in which a substrate is accommodated into the active site (supplementary Fig. S5). Currently, crystal structures of the M20 family proteins have been reported only for PepV, CPG2, and Streptococcus pneumoniae metallopeptidase. The structure of PepV complexed with an inhibitor showed that it has a closed conformation similar to that of CN2 complexed with bestatin (33). On the other hand, in CPG2 and Streptococcus pneumoniae metallopeptidase, which were crystallized as a free form, the catalytic and dimerization domains are in an orientation that exposes the
active site to bulk water. These structural features indicate that M20 family proteins undergo a conformational change upon substrate binding and hydrolysis. This conformational change could be achieved by a movement of domains A and B. This hypothesis is supported in part by a structural feature of the CN2 crystal, in which a large clearance between the two domains would allow a peptide chain movement to open the active site. We tried an ultracentrifugal analysis to detect the conformational difference between free and bestatin-bound forms of CN2, but significant difference of an axis ratio could not be detected (data not shown). One of the reasons may be a slight difference in axis ratio between free and complex forms. If the free form of CN2 is like that of S. pneumoniae metallopeptidase, the difference of axis ratio would be too small to detect by ultracentrifugal analysis despite a large difference in the active site structures. The structural analysis of CN2 possessing the free active site is required to validate this hypothesis.

In the present study, the His228 of mouse CN2 was shown to be crucial for its hydrolysis. This is probably due to its interaction with the substrate molecule at the active site in the dimer in a closed conformation. In PepV complexed with the inhibitor, a histidine residue is also present in a similar position (His229), but it does not appear to interact with the substrate in the active site in the open conformation. However, it is possible that this histidine residue may interact with the substrate if CPG2 has a closed conformation upon substrate binding. Taken together, these points raise the possibility that a conformational change and the access of a histidine residue to the amide carbon atom of the substrate are crucial to the hydrolyzing reaction of the M20 family proteins.

**Conclusion**—The present structural analysis of mouse CN2 provided the basis for its substrate specificity, metal-binding properties, and action of bestatin as an inhibitor. Since L-carnosine is a bioactive peptide with a variety of biological functions, particularly in the mammalian brain, these data are advantageous in studying the biological roles and regulatory mechanisms of CN2 in aging, neurodegenerative, and metabolic diseases.

**Acknowledgments**—We thank Drs. N. Igarashi and N. Matsugaki and the technical staff at synchrotron beamline NW-12 of the Photon Factory, High Energy Accelerator Research Organization, Japan, for assistance with the data collection. We also thank Dr. T. Hatakayama and K. Hisamatsu at Nagasaki University for assistance with measurement of CD spectra.

**REFERENCES**

1. Bonfanti, L., Peretto, P., De Marchis, S., and Fasolo, A. (1999) Prog. Neurobiol. 59, 333–353
2. Tabakman, R., Lazarowici, P., and Kohen, R. (2002) J. Neurosci. Res. 68, 463–469
3. De Marchis, S., Modena, C., Peretto, P., Giffard, C., and Fasolo, A. (2000) J. Comp. Neurol. 426, 378–390
4. Yamano, T., Niijima, A., limori, S., Tsuruoka, N., Kiso, Y., and Nagai, K. (2001) Neurosci. Lett. 313, 78–82
38. Kawabata, T. (2003) *Nucleic Acids Res.* 31, 3367–3369
39. Kim, H., Burley, S. K., and Lipscomb, W. N. (1993) *J. Mol. Biol.* 230, 722–724
40. Hangauer, D. G., Monzingo, A. F., and Matthews, B. W. (1984) *Biochemistry* 23, 5730–5741
41. Fundoiano-Hershcovitz, Y., Rabinovitch, L., Langut, Y., Reiland, V., Shoham, G., and Shoham, Y. (2004) *FEBS Lett.* 571, 192–196
42. Speno, H. S., Luthi-Carter, R., Macias, W. L., Valentine, S. L., Joshi, A. R., and Coyle, J. T. (1999) *Mol. Pharmacol.* 55, 179–185
43. Barrett, A. J., and Rawlings, N. D. (1995) *Methods Enzymol.* 248, 183–228
44. Bayliss, M. E., and Prescott, J. M. (1986) *Biochemistry* 25, 8113–8117