BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

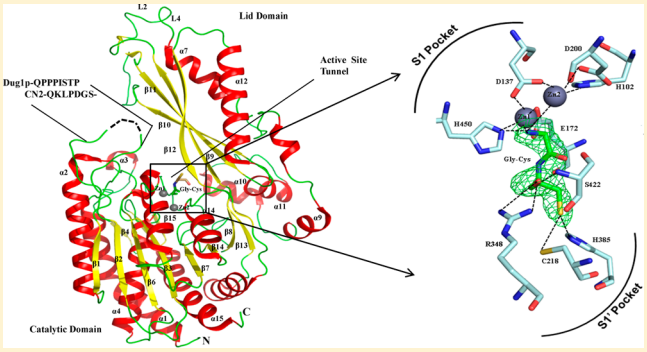
Molecular Basis of Peptide Recognition in Metallopeptidase Dug1p from *Saccharomyces cerevisiae*

Appu Kumar Singh, Mirage Singh, Vaibhav Kumar Pandya, Balasubramani G L, Vijay Singh, Mary Krishna Ekka, Monica Mittal, and S. Kumaran*

G. N. Ramachandran Protein Centre, Institute of Microbial Technology, Council of Scientific and Industrial Research (CSIR), Sector 39A, Chandigarh 160036, India

Supporting Information

ABSTRACT: Dug1p, a M20 family metallopeptidase and human orthologue of carnosinase, hydrolyzes Cys-Gly dipeptide, the last step of glutathione (GSH) degradation in *Saccharomyces cerevisiae*. Molecular bases of peptide recognition by Dug1p and other M20 family peptidases remain unclear in the absence of structural information about enzyme−peptide complexes. We report the crystal structure of Dug1p at 2.55 Å resolution in complex with a Gly-Cys dipeptide and two Zn^{2+} ions. The dipeptide is trapped in the tunnel-like active site; its C-terminus is held by residues at the S1′ binding pocket, whereas the S1 pocket coordinates Zn^{2+} ions and the N-terminus of the peptide. Superposition with the carnosinase structure shows that peptide mimics the inhibitor bestatin, but active site features are altered upon peptide binding. The space occupied by the N-terminus of bestatin is left unoccupied in the Dug1p structure, suggesting that tripeptides could bind. Modeling of tripeptides into the Dug1p active site showed tripeptides fit well. Guided by the structure and modeling, we examined the ability of Dug1p to hydrolyze tripeptides, and results show that Dug1p hydrolyzes tripeptides selectively. Point mutations of catalytic residues do not abolish the peptide binding but abolish the hydrolytic activity, suggesting a noncooperative mode in peptide recognition. In summary, results reveal that peptides are recognized primarily through their amino and carboxyl termini, but hydrolysis depends on the properties of peptide substrates, dictated by their respective sequences. Structural similarity between the Dug1p−peptide complex and the bestatin-bound complex of CN2 suggests that the Dug1p−peptide structure can be used as a template for designing natural peptide inhibitors.



Dug1p, encoded by the DUG1 gene (defective in utilization of glutathione), was shown to be essential for glutathione (GSH, L-γ-glutamyl-L-cysteinyl-glycine) metabolism in *Saccharomyces cerevisae*.[1] Degradation of the GSH into its constituent amino acids begins with the action of γ-glutamyl transpeptidase, which cleaves the γ-glutamyl moiety from the GSH and generates Cys-Gly dipeptides as intermediates. Cys-Gly dipeptide is reported to be toxic to cells, and hence, an efficient scavenging system is required for its removal. Dug1p, identified as a part of the multiprotein Dug1p−Dug2p−Dug3p GSH degradosomal assembly, catalyzes the last step in GSH degradation in *S. cerevisiae* by degrading the Cys-Gly intermediate with high specificity.[2] To improve our understanding of the peptide recognition properties of Dug1p, we studied the binding of a variety of peptides to this enzyme and found that the substrate specificity of Dug1p may be encoded at two different levels.[3] We showed that Dug1p binds a broad range of small peptides but cleaves Cys-Gly peptide with specificity, suggesting that catalytic specificity is achieved after binding. We proposed that Dug1p may employ an induced fit-like mechanism for discriminating substrates from nonsubstrate peptides. However, in the absence of structural information, molecular determinants of peptide recognition and catalysis remain unknown.

Dug1p and its two human orthologues, carnosinases CN1 and CN2, belong to the M20 metallopeptidase family.[4] Both CN1 and CN2 cleave carnosine, a bioactive dipeptide (β-alanine-L-histidine) preferably and utilize a common catalytic mechanism in which two divalent metal ions in the active site activate the catalytic water molecule for commencing nucleophilic attack on the incoming peptide bond.[5] CN2 shares a high degree of sequence identity (∼52%) with Dug1p, and the crystal structure of mouse CN2 in complex with bestatin [2-(3-amino-2-hydroxy-4-phenylbutylamino)-4-methylpentanoic acid] has been reported.[6] CN2 functions as a homodimer, and each subunit consists of one large catalytic domain and a lid domain. The catalytic domain consists of residues essential for substrate and metal binding, whereas the

Received: October 7, 2014
Revised: November 21, 2014
Published: November 26, 2014

lid domain contributes to dimerization and also to activity by donating a catalytically indispensible histidine, His228. Domain swapping between two subunits places His228 of one subunit into the active site of the other. Therefore, dimerization is essential for carnosinase activity.

Both Dug1p and carnosinase belong to the M20 family of peptidases. M20 family members include exopeptidases, aminopeptidase, carboxypeptidase, dipeptidases, and tripeptidase, which catalyze diverse peptides.[7] M20 family peptidases consist of two cocatalytic metal ions per active site and five ligand residues (His/Asp, Asp, and Glu/His) that ligate both metal ions. They control a wide variety of physiological processes in eukaryotes, and therefore, fine-tuning their activity has important biomedical applications.[8,9] They recognize diverse substrates but show specificity toward their respective substrate peptides. However, structures of CN2,[6] PepV,[10] SplA,[11] carboxypeptidase G2,[12] and aminoacyl histidine dipeptidase[13] revealed that they share a similar structural scaffold. Structures reveal a large active site cavity, formed by canonical catalytic and lid domains, and cocatalytic metal ions at the center of the active site. Sequence and structural similarities of M20 family members led to the assumption that they may recognize their respective substrate peptides in a similar manner.[14,15] Although high-resolution structures of M20 family peptidases complexed with inhibitors are available, structural determinants of peptide recognition are not known because of the lack of the structure of the enzyme−peptide complex. The presence of a large active site cleft promotes the assumption that promiscuous substrate binding may be an inherent property of M20 family enzymes. Structural analyses have led to the model in which members of the M20 family undergo conformational changes during substrate binding and hydrolysis.[6] The ability of M20 peptidases to bind a diverse group of peptides but hydrolyze selectively suggests that peptides are recognized in a sequence-independent manner. As proposed, the binding-induced conformational changes may determine the catalytic specificity.[3] However, the absence of structures of enzyme−peptide complexes and detailed biochemical analyses limit our understanding of peptide recognition by M20 family peptidases.

In this study, we present the crystal structure of Dug1p at a resolution of 2.55 Å in complex with a naturally bound dipeptide and two Zn^{2+} ions. The Gly-Cys dipeptide binds with its N-terminus pointed toward two Zn^{2+} ions. Both metal ions are fixed by interactions with active site residues in the S1 pocket, whereas the carboxy terminus of the peptide is bound at the S1′ site. Structural comparison with CN2 reveals that conformational features of active site residues and coordination of metal ions are altered. We show that Dug1p can bind di- and tripeptides in a sequence-independent manner but hydrolyzes them in a sequence-dependent manner. Mutagenesis experiments show that mutations of catalytic residues do not compromise peptide binding but abolish hydrolysis of the peptide bond. Structural and biochemical analyses reveal stereochemical features of peptide recognition and hydrolysis by Dug1p. We discuss in detail the structural features of Dug1p that confer selectivity and promiscuity.

MATERIALS AND METHODS

Protein Expression and Purification. The Dug1p gene, cloned into the pET23a(+) vector, was a kind gift from A. K. Bachhawat, and purification of Dug1p was conducted as described previously.[2] Briefly, protein expressed with C-terminal His tags was purified using Ni^{2+}-NTA affinity and gel filtration chromatography. The protein is >98% pure, as judged by 12% sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE). The protein was dialyzed against respective buffers for structural and biochemical studies.

Crystallization and Structure Determination. The initial search for the right crystallization condition for Dug1p was conducted by utilizing screening suits. Dug1p was concentrated to ∼50.0 mg/mL and crystallized using the sitting drop method. Crystals appeared in buffer containing 0.5 M ammonium sulfate, 0.1 M Tris-sodium citrate (pH 5.6), and 1.0 M lithium sulfate. Crystals took two months to grow to a size of ∼0.2 mm × 0.1 mm × 0.1 mm. The crystal was flash-frozen in liquid nitrogen using mother liquor with 20% glycerol as the cryoprotectant. The X-ray diffraction data for Dug1p were collected using in-house image plate detector (MAR345) mounted on a Rigaku (MicroMax-007HF micro focus) rotating anode X-ray generator at 100 K. A total of 190 frames were collected with 1° oscillation with a 10 min exposure for each image obtained. Data were processed, integrated, and scaled using the HKL2000 suite[16] and XDS.[17] Data were scaled using *P*622 as the point group, and we tested all the space groups under *P*622 using PHASER,[18] implemented under the CCP4 suite.[19] We used the mouse CN2 [Protein Data Bank (PDB) code 2ZOG] subunit as a search model.[6] The top solution had a translational function *Z* score of 32.2, and the solution was in the $P6_522$ space group. PHASER,[18] the CCP4 suite,[19] and COOT[20] were used for both model building and refinement, and figures were made using PyMOL.[21]

Fluorescence Titration Measurements. Titrations of enzymes with peptides were examined by monitoring the intrinsic tryptophan fluorescence of Dug1p and its mutants. All binding experiments were performed in standard binding buffer [25 mM Tris (pH 8.0), 50 mM NaCl, and 50 μM $ZnCl_2$] at 25 °C. The excitation and emission wavelengths for monitoring peptide interaction were set to 292 and 345 nm, respectively. Initial readings of both the protein, $F_{protein,0}$, and buffer, $F_{buff,0}$, were taken, with $F_0 = F_{protein,0} - F_{buff,0}$ defined as the initial fluorescence of the sample. The sample in the cuvette was then titrated with aliquots of peptides and mixed and equilibrated for 2 min before measurement. Data points from five such measurements were averaged to obtain $F_{ave,i}$. The relative fluorescence quenching upon peptide binding is defined as $Q_{obs,i} = (F_0 - F_{ave,i})/F_0$. All binding isotherms were fit to the following equation (eq 1)

$$Q_{obs}/Q_{max} = (Q_{obs}K_{obs}P)/(1 + K_{obs}P) \quad (1)$$

where Q_{obs} and Q_{max} are the fluorescence quenchings at any stage and at saturation, respectively, P is the concentration of the free peptide in solution, and K_{obs} is the association constant. Experimental data were fit using a nonlinear least-squares method, and errors were obtained from the fit.

Mass Spectrometry-Based Activity Assay. Assay conditions were the same as reported previously.[22] All reactions were stopped by the addition of 0.1% formic acid. Dug1p kinetic experiments were performed on an LC−MS/MS (Applied Biosystems) QTRAP mass spectrometer equipped with a Turbo IonSpray (TIS) electrospray ion source. Infusion of peptide into the QTRAP system was controlled by an automatic syringe pump at a speed of 10 μL min^{-1}. MS spectra of substrates were run, and initial spectra were acquired in Q1MS mode. A series of reference samples containing different concentrations of peptides were run, and peak areas were

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

calculated. All enzyme activity assays were conducted in activity buffer [25 mM Tris (pH 8.0), 50 mM NaCl, and 50 μM $ZnCl_2$], and size exclusion-purified protein (2−5 μg, 0.2−0.5 μM) was used. The standard curve is calibrated for substrate peptides using reference values after normalization of data. The area under the peak at a fixed time point at several concentrations was used to produce steady state kinetic data. The method has been described in detail previously.[22] To outline it briefly, we measured the ratio of peak intensities of substrates at different concentrations with calibration standard peak intensities. This ratio is multiplied by the concentration of the control sample to obtain the concentration of the substrate in the sample as described in eq 2.

$$C_s = I_s/I_c \times C_c \times (V_f/V_i) \quad (2)$$

where V_f and V_i are final and initial volumes, respectively, and their ratio yields the dilution factor, C_s is the substrate concentration in the sample, C_c is the concentration in the control or the concentration estimated from the calibration curve, and I_c and I_s are intensities of the control and sample, respectively. Concentrations of substrate peptides before reaction ($C_{s,0}$) and after reaction ($C_{s,t}$) were calculated using eq 2, and initial velocities (V_0) normalized to enzyme concentration were calculated using eq 3.

$$V_0 = (C_{s,0} - C_{s,t})/(t[E]_t) \quad (C_{s,0} - C_{s,t} = \Delta S) \quad (3)$$

where t and $[E]_t$ are the reaction time and total enzyme concentration, respectively.

Intact Mass Determination by a MALDI Experiment. MALDI MS analysis was performed on a MALDI-TOF mass spectrometer (Voyager DE-STR, model 4402, Applied Biosystems, Foster City, CA) equipped with a 337 nm nitrogen laser. MALDI measurements were performed in linear positive ion mode, and delayed extraction was applied. In linear mode, the acceleration voltage was 25 kV, the guide wire was 0.3% of the accelerating voltage, the grid voltage was 92%, and the delay time was 1500 ns. Positive MALDI-TOF mass spectra were acquired in linear mode over a mass range from 52000 to 55000 Da.

Cysteine-Based Enzyme Assay. The peptidase activity was assayed by measuring the amount of L-cysteine released after the reaction of the enzyme with cysteine-containing peptides as described previously.[2,23] The assay was based on the specific reaction of ninhydrin with free cysteine under strong acidic conditions. Under these conditions, the presence of other amino acids, including proline, or other cysteine-containing compounds, such as GSH and cystine, does not interfere with the estimations of cysteine.[23] All enzyme activity assays were conducted in activity buffer [25 mM Tris (pH 8.0), 50 mM NaCl, and 50 μM $ZnCl_2$] supplemented with purified protein (2−5 μg, 0.2−0.5 μM). The reaction was initiated by the addition of peptide (0.1−10 mM) in a reaction volume of 200 μL. The reaction was conducted at 30 °C for 30 min and terminated with 5% trichloroacetic acid, and the mixture was centrifuged at 11000g for 2 min. For the substrate specificity experiments, the assay used for the estimations of relative activity was a single-end point assay at saturating peptide concentrations (5−10 mM), and activity data are reported in relative percentage activity. For steady state kinetic experiments, the substrate concentration was varied between 0.5 and 10 mM and the rest of the assay conditions were the same. The steady state kinetic data were fit to the Michaelis−Menten model as described in eq 4.

$$v = (V_{max}[S])/([S] + K_M) \quad (4)$$

where v and V_{max} are the observed and maximal velocities, respectively, and $[S]$ and K_M are the free substrate concentration and the substrate concentration at which half-maximal velocity is achieved, respectively. Competitive titration activity assays were performed as described above except that Dug1p was preincubated with 0.12 mM competitor, the HGC tripeptide (HGC), for 30 min before the addition of Cys-Gly substrate (5 mM). The final activity is measured in the presence and absence of competitor.

Site-Directed Mutagenesis. The Dug1p construct was used as a template for amino acid substitutions by PCR-based site-directed mutagenesis. The PCR mixture for site-directed mutagenesis consisted of 50 ng of plasmid, 10 pmol of each primer, 1 unit of *Pfu* Turbo Hot Start DNA polymerase (Stratagene), dNTPs (250 μM each), and 1× *Pfu* pol buffer supplied along with ddH_2O in a 25 μL reaction mixture. PCR was conducted on an Eppendorf thermal cycler under the following conditions: 98 °C for 30 s, 95 °C for 30 s, the appropriate primer T_m-dependent annealing temperature (55 °C) for 1 min, and 68 °C for 7 min for 12 cycles (point mutations) or 18 cycles (multiple nucleotide changes), followed by a final extension at 68 °C for 10 min. After thermocycling, the original template DNA was eliminated by treatment with *Dpn*I at 37 °C for 1 h, and 4 μL of the PCR products was transformed into 50 μL of XL1-Blue *Escherichia coli* chemical competent cells and plated on kanamycin-containing LB plates. Five colonies were picked up and mini-prepped, and purified plasmid DNA was sequenced using an ABI Prism capillary DNA sequencer to confirm the mutation.

Determination of Secondary Structure by CD. CD measurements were taken with a JASCO-810 spectropolorimeter (Jasco, Tokyo, Japan) equipped with a Peltier type temperature controller (PTC-348W). Far- and near-UV spectra were obtained in a quartz cuvette with 1.0 cm path length, and an average of seven scans were collected for each spectrum. The calculated mean residue ellipticity (MRE) is plotted against wavelength. Average concentrations of Dug1p and mutants were ∼0.03 mg/mL and scanned between 195 and 250 nm. The signal from the buffer was subtracted for each sample.

Molecular Dynamics Simulations and Analysis. We performed MD simulation using the structure of Dug1p (PDB entry 4G1P at 2.55 Å), and the missing residue, Ile73, was modeled using swiss-modeler.[24] The structure was energy-minimized and used as a starting template for simulations. MD simulations were conducted using GROMACS[25] simulation suite 4.5.5 and the corresponding OPLS-AA/L all-atom force field parameter set.[26] The structure was solvated in a cubic box with dimensions of 10.372 nm × 10.372 nm × 10.372 nm (the minimal distance between the solute and the box boundary is kept at 1.0 nm). Simple-point charge (SPC) water molecules were added from an equilibrated cubic box containing 216 water molecules.[27] Seven Na^+ ions were added to compensate for the net negative charge of Dug1p. These ions were introduced by replacement of water molecules with the highest electrostatic potential. The entire system, protein and water, was initially energy-minimized for 50000 cycles using the steepest descent algorithm with harmonic position restraints applied on all heavy protein atoms (force constants of 1000 kJ mol^{-1} nm^{-2}) to allow relaxation of the solvent molecules. The final system consisted of 7423 protein atoms (including two zincs), seven Na^+ ions, and 36108 water molecules. We

employed a modified Berendsen thermostat method for temperature coupling.[28] Protein and solvent were coupled separately to temperature baths of reference temperature with a coupling time of 0.1 ps; a 100 ps pressure equilibration MD run (*NPT*) was performed using a Parrinello−Rahman barostat, and the pressure was kept constant at 1 bar. The linear constraint solver (LINCS) algorithm[29] was used for restraining bond lengths, and long-range electrostatic interactions were calculated using the fast particle-mesh Ewald (PME) electrostatics method with a cutoff of 1.0 nm.[30] The system was subjected to a 100 ps dynamics run with protein atoms restrained and was followed by free production for 10 ns. Except for root-mean-square deviation (RMSD) calculation, the last 8 ns run was used for analysis. Principal component analysis was used for extracting information about protein dynamics from 10 ns trajectories. We used a well-established method that uses diagonalization of the covariance matrix to yield corresponding eigenvalues and eigenvectors that describe collective modes of protein dynamics. Eigenvector analysis of backbone atoms (N, Cα, and C) allowed us to map large-amplitude motions of the protein.

RESULTS

Overall Structure. The size exclusion-purified Dug1p was crystallized, and crystals diffracted to 2.4 Å. Upon being subsequently processed, the data were truncated to 2.55 Å, and the structure was determined by molecular replacement using the structure of mouse CN2 (PDB entry 2ZOG) as a template.[6] The structure is resolved with one Dug1p subunit in the asymmetric unit in the $P6_522$ space group, and the final model included two Zn^{2+} ions, 180 water molecules, and one Gly-Cys dipeptide. The stereochemistry of the model is checked by PROCHECK, and 94% of the residues fall in the most favorable region and the rest in the favorable and allowed regions of the Ramachandran plot (Table 1). The polypeptide chain can be unambiguously traced from the N-terminus to the C-terminus, and 478 of 484 amino acid residues are defined in the electron density.

One subunit of Dug1p consists of one globular catalytic domain (residues 1−203 and 413−480) and an elongated lid domain (residues 204−412). Structural properties of catalytic and lid domains are similar to that of CN2, although conformational features of loop regions and active site residues differ. The catalytic domain core consists of a central β-sheet (β1−β2−β6−β3−β7 strand order), sandwiched between four long helices (α1, α2, α4, and α15) on one side and two short helices (α6 and α14) on the other (Figure 1A). This arrangement of the catalytic domain is similar to that of CN2,[6] PepV,[10] and CPG2[12] peptidases. The lid domain consists of twisted antiparallel β-strands (β9−β10−β11−β12 strand order) and six helices (α7−α13), five β-strands, and two long loops (L2 and L3) that provide the platform for substrate binding as well as protein dimerization. Superposition with CN2 and PepV shows that the overall structure of catalytic domains is similar but the lid domain of PepV is in the closed state as compared to that of Dug1p and CN2.

Subunit Interaction and Dimerization. Two subunits of the functional homodimer of Dug1p are related by a crystallographic 2-fold symmetry. The homodimer is highly elongated in shape with dimensions of 45.4 Å × 48.8 Å × 102.2 Å, and the centers of two active sites are separated by ∼49 Å. The subunit−subunit interface covers an area of ∼3400 Å^2, and the main interaction surface is constructed between α7 (residues 243−253) and β10 (residues 318−325) of two subunits. Domain swapping allows two loops of the lid domain, L2 (residues 226−244) and L3 (residues 326−339), of one subunit to be inserted into the catalytic site of the other subunit (Figure 1B). The swapped L2 loop of one subunit donates a catalytically important His233 to the active site of the other subunit. The His233 of the other subunit mediates a hydrogen bond with the amino terminus of the Gly-Cys peptide, suggesting that His233 may play a role in peptide recognition and dimerization is essential for the function of Dug1p (Figure 1B, inset). The subunit−subunit interface is stabilized by ∼42 hydrogen bonds and two salt bridges. The dimer interface also consists of a number of water molecules that connect two subunits through multiple hydrogen bonds. Although interaction between α8 and β10 of two subunits may contribute to

Table 1. X-ray Data Collection and Structure Refinement Statistics

Data Collection	
X-ray source	rotating anode, Rigaku Micromax-007 HF
wavelength (Å)	1.542
resolution range (Å)	50−2.55
last resolution shell (Å)	2.70−2.55
space group	$P6_522$
unit cell parameters (Å)	$a = b = 119.17$, $c = 176.33$
total no. of reflections	279022
no. of unique reflections	24507
redundancy	11.30 (11.21)
overall $I/\sigma(I)$	12.70 (2.27)
completeness (%)	98.60 (97.90)
R_{merge}[a] (%)	21.60 (87.4)
Structure Refinement	
resolution range (Å)	19.73−2.54
last resolution shell (Å)	3.06−2.55
no. of unique reflections	24491
completeness (%)	98.50
R_{cryst}[b]	19.73
R_{free}[c]	23.75
rmsd from ideality	
bond lengths (Å)	0.01
bond angles (deg)	1.19
Luzatti coordinate error (Å)	0.29
Ramachandran plot	
most favored (%)	92.8
allowed (%)	7.2
no. of outliers	0
model	
no. of modeled residues	477
no. of water molecules	121
no. of zinc ions	2
no. of L-glycine-L-cysteine peptides	1
average B factor (Å^2)	
protein	36.42
water	32.47
zinc ion	40.05
L-glycine-L-cysteine	43.06
PDB entry	4G1P

[a] $R_{merge} = \sum_{hkl}\sum_i|I_i(hkl) - \langle I(hkl)\rangle|/\sum_{hkl}\sum_i I_i(hkl)$, where $I(hkl)$ is the intensity of reflection hkl. [b] $R_{cryst} = \sum_{hkl}||F_o| - |F_c||/\sum|F_o|$. [c] R_{free} is the cross validation R factor computed for the test set of 5% of the unique reflections.

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

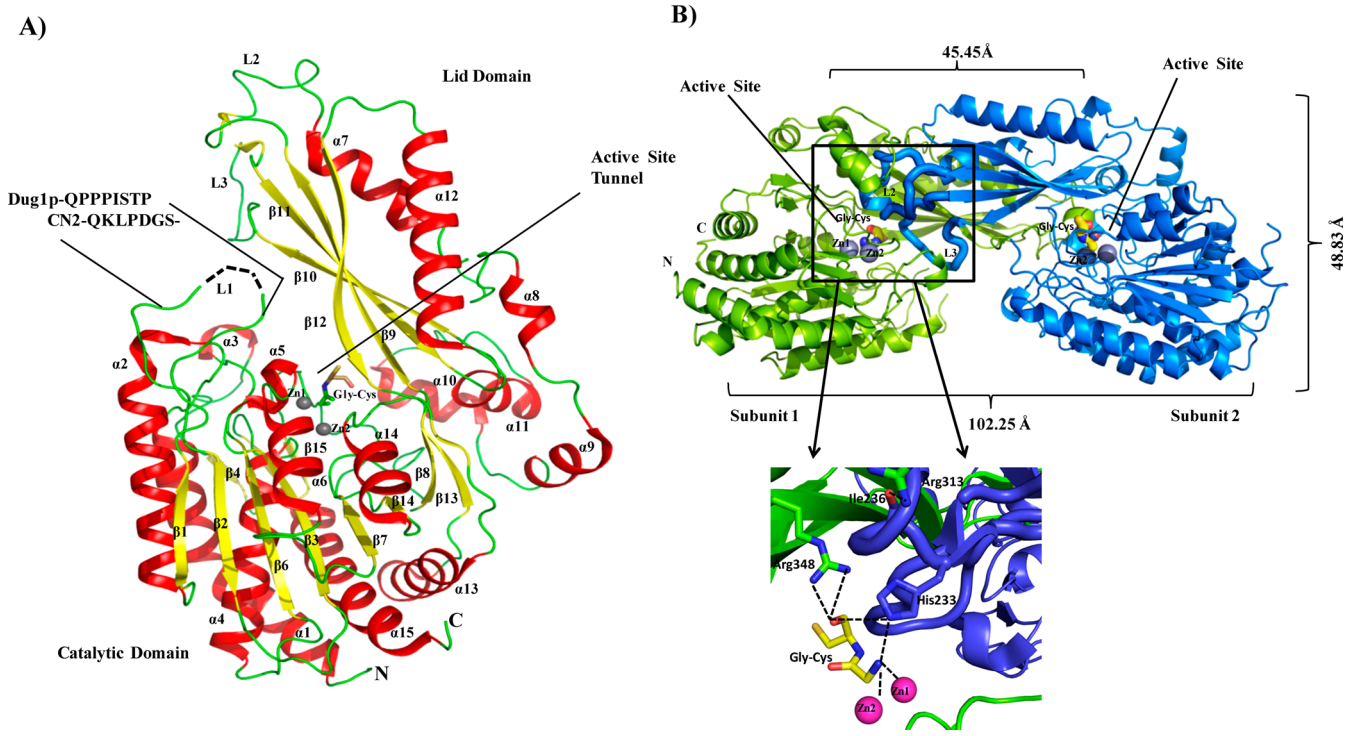


Figure 1. Structural analysis of Dug1P. (A) Cartoon representation of the overall structure of Dug1p bound with two zinc ions and a dipeptide in the active site. Active site, catalytic, and lid domains are labeled as indicated, and secondary structural elements are labeled. (B) Cartoon representation of the Dug1p dimer shown with dimensions. Two subunits, colored green and blue, interact through the lid domain. Active sites of each subunit are labeled. The structure is elongated on the *y*-axis, and the Gly-Cys dipeptide ligand is shown as sticks. The inset shows the His233 from the blue colored subunit interacts with the NH group of the peptide bound at the other subunit (green).

the stability of the dimer, domain swapping through loops L2 and L3 may dominate the energetic contribution to dimerization.

Structural Features of the Active Site and Gly-Cys Peptide. Both catalytic and lid domains enclose to form a deep active site chamber that has the metal ion binding center and S1 and S1′ binding pockets. The wide-open channel with a net volume of ∼420 Å^3 provides a sufficiently large space for accommodating di- and tripeptides. Two Zn^{2+} ions, located 3.5 Å from each other, can be modeled into electron densities found at the active site. The $F_o - F_c$ omit map densities contoured at 3.0σ for the dipeptide and two Zn^{2+} ions are shown (Figure 2A). The $F_o - F_c$ and simulated annealing omit map of bound peptide and metal ions confirm their presence (Figure S1 of the Supporting Information). Also, the identity of Zn^{2+} is confirmed from metal content analysis by atomic absorption spectroscopy. Zn^{2+} ions and atoms of the peptide have thermal *B* factors between 35 and 40 Å^2 that are similar to the observed average *B* factors for the protein Cα atoms of 30 Å^2. The N-terminus of Gly-Cys dipeptide is in the proximity of both Zn^{2+} ions and active site residues of catalytic and lid domains.

Besides it being the first structure of a M20 family metallopeptidase in complex with a natural peptide, the several stereochemical features of the bound peptide−enzyme complex are unique. The dipeptide moiety is clamped tightly by a number of hydrogen bonds. Analyses of interactions between peptide and active site residues show that peptide is primarily recognized through its N- and C-termini (Figure 2B). The N-terminus of the peptide is bound to the S1 pocket, whereas the C-terminus is bound to the S1′ pocket. Both N- and C-termini of the Gly-Cys dipeptide are held tightly by a number of hydrogen bonds, contributed mostly by residues of catalytic (Asp137, His233, His450, and Ser422) and lid (Cys218, Arg348, and His385) domains. The N-terminus of the peptide is fixed by interactions with metal ions as well as with catalytic residues at the S1 site. The NH group of the glycine mediates two hydrogen bonds, one each to two zinc ions, and forms two more hydrogen bonds, one each to side chains of Asp137 and His450. In addition, the side chain Nε_2 atom of His233 from the other subunit also mediates a hydrogen bond to the NH group of the peptide. The C-terminus of the peptide is held by side chains of Arg348, Ser422, His385, and Cys218 at the S1′ site. Two NH groups (NH1 and NH2) of the guanidine group of Arg348 and the main chain amide of Ser422 hydrogen bond to the terminal carboxyl group of the peptide. Similarly, the side chain Nε_2 atom of His385 forms a weak interaction with the carboxylate of the peptide. Two hubs of interactions that hold the two ends of the dipeptide suggest that the binding mode is similar to that of the two-prong plug two-hole socket model. Additional interactions include weak interaction between the sulfhydryl group of the cysteine of the peptide and Cys218 at a distance of 3.8 Å and a hydrogen bond between two main chain carbonyl oxygens of Ser422 and the peptide, the carbonyl oxygen of the Gly-Cys peptide bond.

Stereochemical Features of the Active Site of the Dug1p−Peptide Complex. To improve our understanding of the unique structural features of the Dug1p−peptide complex, we superposed its structure with structures of related binuclear family peptidases that share a two-domain architecture. Structures of mouse CN2 (PDB entry 2ZOG),[6] PepV (PDB entry 1LFW),[10] caboxypeptidase G2 (PDB entry 1CG2),[12] PepT (PDB entry 1FNO),[31] and Sapep (PDB entry 3KHX)[32] were compared. The Dug1p structure is quite

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

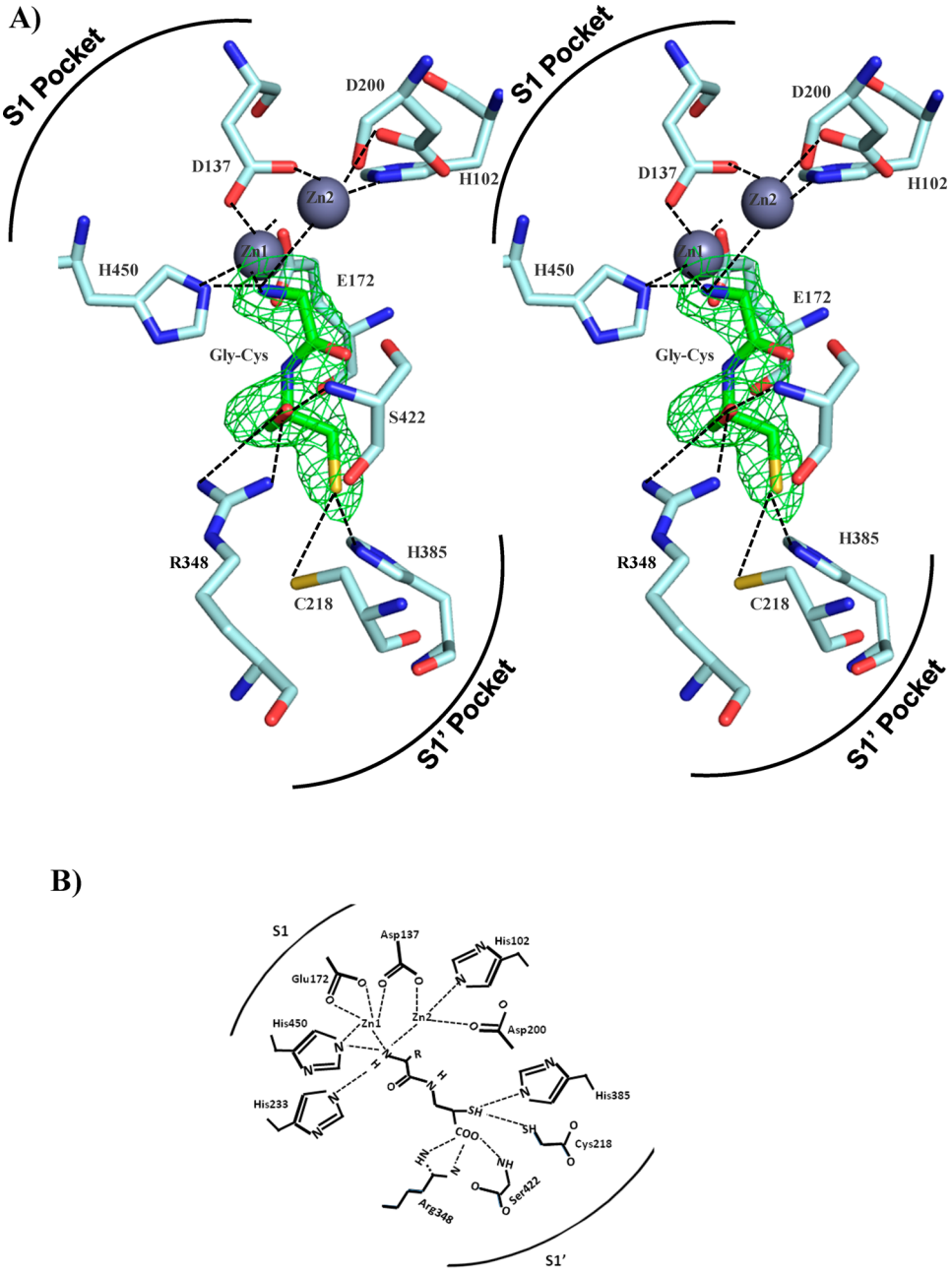


Figure 2. View of the Dug1p active site in complex with catalytic zinc ions and Gly-Cys peptide. (A) Stereoview of the interaction of two Zn^{2+} ions (spheres, labeled Zn_1 and Zn_2) and peptide (green sticks) with active site residues. The $F_o - F_c$ map (green mesh), contoured at 3.0σ, confirms the presence of ligands at the active site. (B) Schematic diagram describing the details of interactions between peptide and active site pockets S1 and S1′.

similar to that of CN2 (overall rmsd of ~0.65 Å), and both catalytic (rmsd of ~0.7 Å) and lid (rmsd of ~0.6 Å) domains of the two structures superimpose well. Dug1p has a unique polyproline loop (residues Gln69−Asn77) that is highly flexible and adopts an alternative conformation compared to that of its counterpart (Lys68−Glu74) in the CN2 structure. Both loops extend into the entrance of the active site channel in their respective structures, suggesting that they may have a role in substrate entry.

Interestingly, the comparison of the active site of the Dug1p−peptide complex with that of the CN2−bestatin complex reveals several conspicuous differences. Conformational features and positions of catalytically important active site residues and metal ions differ between these two structures. Both zinc ions in the CN2 structure maintain more or less octahedral coordination. In contrast, coordination geometries of two metal ions in the Dug1p structure are different. The first Zn^{2+} ion, named Zn_1, has trigonal-bipyramidal coordination and is in contact with His450, Asp137, Glu172 (two hydrogen bonds), and the NH group of the dipeptide. The second Zn^{2+} ion, termed Zn_2, acquires tetrahedral geometry, and four ligands (Asp137, Asp200, His102, and NH group of the dipeptide) coordinate Zn_2 (Figure 3A,B). Two carboxylate oxygens of the Asp137 side chain and the NH group of the dipeptide interact with both Zn^{2+} ions. The position of catalytic Zn_2, relative to Zn_1, is different from respective positions of metal ions in other metallopeptidases.[6,10−12] Compared to carnosinase (PDB entry 2ZOF), the catalytic Zn_2 has moved

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

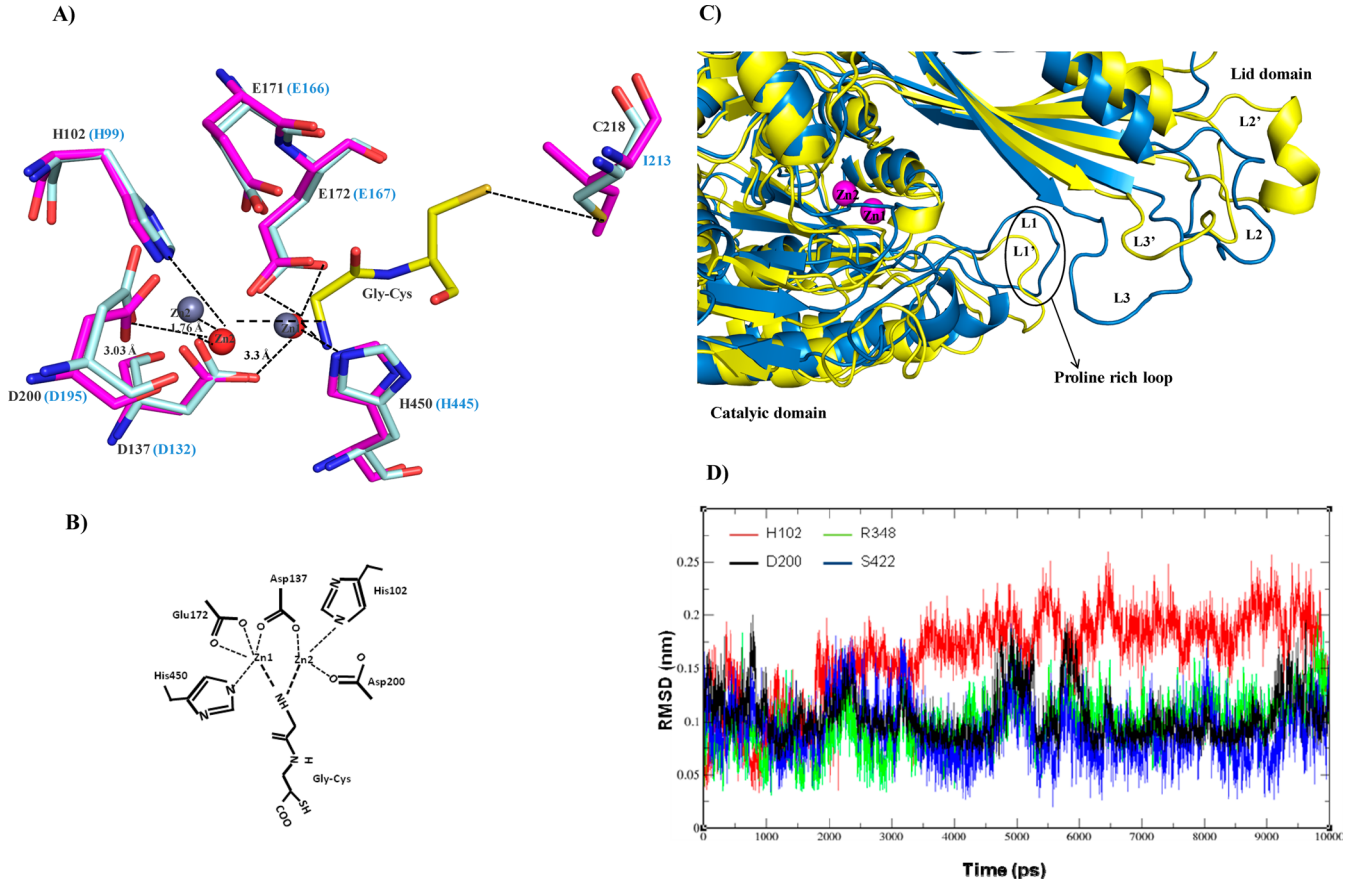


Figure 3. Structural analyses of the Dug1p active site. (A) Overview of the comparison of active sites of Dug1p (sticks, cyan) and mouse CN2 (sticks, light magenta) showing differences in coordination geometries of Zn^{2+} ions (Dug1p, red spheres; CN2, purple spheres) and conformational changes of active site residues. Zn_2 moves 1.7 Å into the active site. Asp137, Glu171, and Asp200 side chains orient differently to promote dipeptide binding. (B) Schematic representation of coordination of Zn_1 and Zn_2 within the Dug1p active site. Dotted lines represent noncovalent interactions. (C) Molecular dynamics simulations reveal open and closed states of Dug1p. Superimposition of open (yellow) and closed (sky blue) states. The L3 loop moves away from the catalytic domain, and distances between L3 of the lid domain and the polyproline loop of the catalytic domain are shown. (D) Cα rmsd values for active site residues H102 (red), D200 (black), R348 (green), and S422 (blue).

~1.7 Å into the active site and its coordination changed from octahedral to tetrahedral geometry. This movement is accompanied by the loss of hydrogen bonds with OD_1 of Asp200 and increased the distance between $N\varepsilon_2$ of His102 and Zn_2, leading to weak interaction between these two residues. The side chains of Glu171, Asp200, and His102 ligate Zn_2 in CN2. Peptide binding has altered side chain orientations of Glu171 and Asp200. We hypothesize that peptide binding-mediated dynamics may have weakened the interaction between ligating residues and Zn_2, facilitating the movement of Zn_2 toward the bound peptide. The side chain of Asp200 has flipped away from Zn_2, resulting in the loss of one hydrogen bond between Zn_2 and Asp200. It is possible peptide binding has altered the conformation of Asp200, and thus, Zn_2 has moved away from His102, closer to the peptide. Although the purpose of the movement of Zn_2 toward the catalytic center is not immediately obvious from the structure, we propose that Zn_2 may play a role in stabilizing the bound peptide. On the other hand, conformational features of residues at the S1′ site do not differ in either structure, suggesting that the S1 subsite is more dynamic than the S1′ site.

In summary, the unique polyproline loop, L2 and L3 loops of the lid domain, and active site residues at the S1 site showed conformational differences, indicative of dynamics. To gain more information about dynamics, we performed all-atom molecular dynamic simulations using the structure of Dug1p resolved in this study as the initial template. We simulated the system for 10 ns and analyzed the trajectories of simulation. Principal component analyses (PCA) of coordinates of Dug1p from a 10 ns simulated trajectory by diagonalization of the covariance matrix of the atomic fluctuations yielded eigenvalues and eigenvectors that describe collective modes of conformational fluctuations. Results show the entire lid domain moves away from the catalytic domain, on average ~2.9 Å (calculated for CA atoms for residues 204−412). The two loops, L2 and L3, of the lid domain undergo large conformational changes along with a central twisted β-sheet of the lid domain and move outward by 1.9 Å, indicating that the structure of the Dug1p−peptide complex is in the closed state (Figure 3C). Further, we analyzed the dynamics of active site residues (Asp200 and His102) that interact with metal ions and two other residues, Arg348 and Ser422, that interact with the C-terminus of the peptide. Results show that His102 shows high flexibility as compared to dynamics of three other residues (Figure 3D). These results support our structural analyses that showed that the lid domain and S1 site of Dug1p are more dynamic, and results of the simulation rationalize the movement of Zn_2, which moved away from His102 and toward the catalytic center.

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

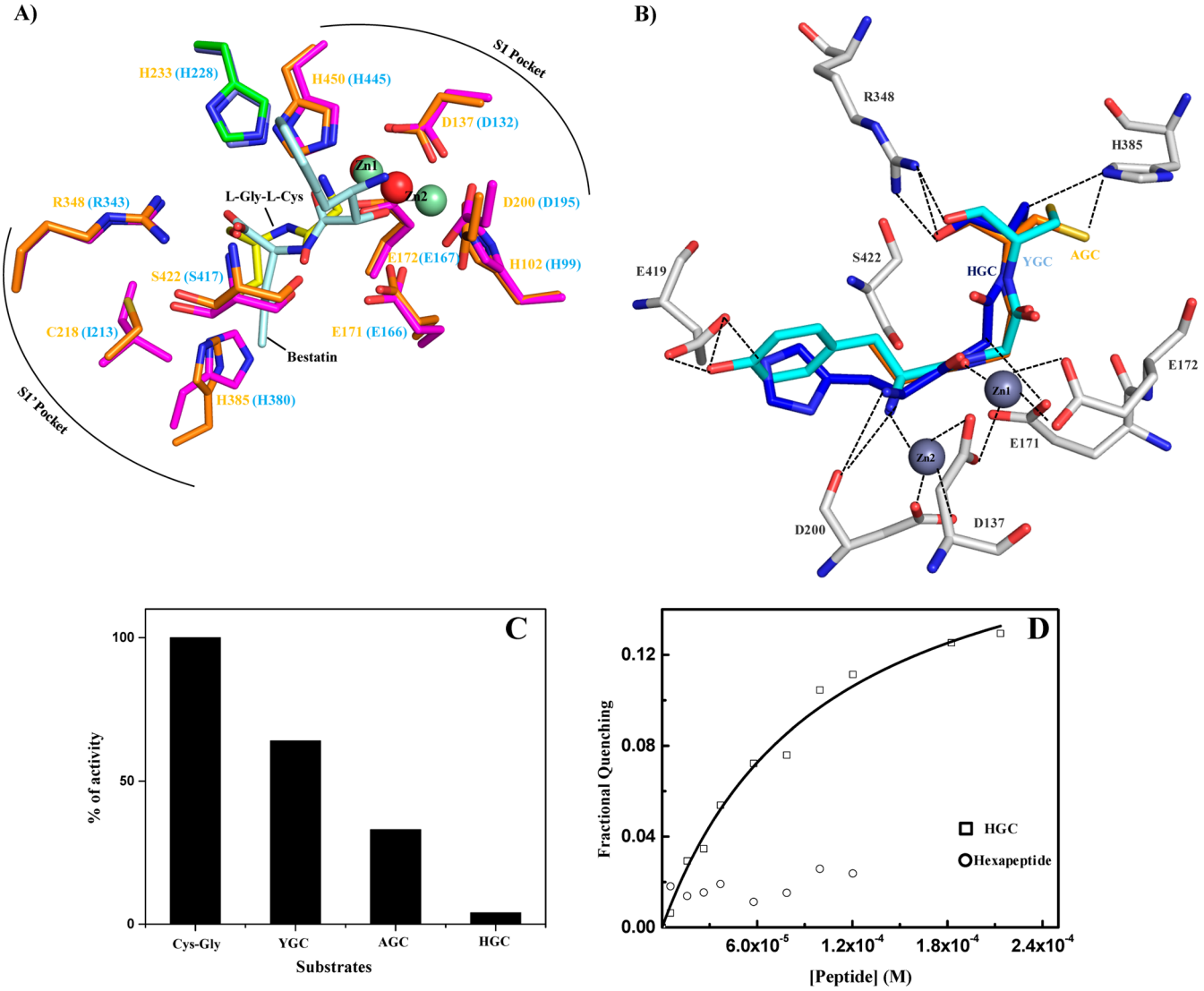


Figure 4. Comparison of inhibitor and tripeptide recognition. (A) Active site of Dug1P (sticks, orange) and CN2 (sticks, magenta) are superimposed to show that backbone of the peptide (yellow) follows that of bestatin (cyan). Zinc ions are shown as spheres (Dug1p, red; CN2, pale green). (B) Modeling and fitting of Tyr-Gly-Cys (YGC, cyan), Ala-Gly-Cys (AGC, green), and His-Gly-Cys (HGC, blue) peptides into the active site and positioning of the target carbonyl group of the peptide bond with respect to Zn_1. (C) Tripeptide hydrolysis by Dug1p. Percentage activities normalized to Cys-Gly hydrolysis are shown. (D) Binding isotherms of the interaction of Dug1p with catalytically inactive His-Gly-Cys tripeptide and Gly-Cys-Gly-Cys-Gly-Cys hexapeptide. Fractional quenching of tryptophan fluorescence upon peptide binding is plotted vs peptide concentration. The solid line represents the best fit to simple 1:1 binding model as described in eq 1.

Comparison of Peptide and Inhibitor Binding Poses. Superimposing both structures reveals that the backbone of Gly-Cys peptide coincides well with that of bestatin until the O2 hydroxyl, and thereafter, the amino group of bestatin takes a sharp turn (Figure 4A). Both bestatin and peptide exhibit similar hydrogen bonding at the S1′ pocket. Carboxyl termini of both bestatin and Gly-Cys peptide are fixed by the backbone amide group of Ser422 and the guanidyl group of Arg348, whereas the backbone carbonyl oxygen of Ser422 positions N1 of bestatin and the carbonyl oxygen of the Gly-Cys peptide bond. Although Gly-Cys peptide mimics the real substrate at the S1′ pocket, the N-terminus of bestatin penetrates into the S1 pocket that is left unoccupied in the structure of the Dug1p−dipeptide complex. O2 of bestatin extends into the S1 pocket in the carnosinase structure and maintains the proximity to Zn_2, mimicking the pose of the substrate carbonyl, to be cleaved. On the other hand, the carbonyl of the Gly-Cys peptide bond moves away from Zn_2 and interacts with the OH group of Ser422 at the S1′ pocket, suggesting that the Gly-Cys peptide bond is not the target for hydrolysis.

The S1 pocket of the CN2 active site is occupied by the headgroup of bestatin, but in the Dug1p structure, the S1 pocket is largely unoccupied. This suggested that a tripeptide would fit into the active site with its N-terminal amino acid fitting into the unoccupied S1 pocket. To test this, we modeled three tripeptides, X-Gly-Cys (X = Tyr, His, and Ala), into the active site by keeping all interactions at the S1′ pocket fixed (Figure 4B). Interestingly, modeled tripeptide skeletons overlap with that of bestatin, and the carbonyl of the Tyr-Gly peptide bond is positioned in the proximity of Zn_1. The interaction of this scissile carbonyl of the Tyr-Gly peptide bond with Zn_1 is consistent with the general mechanism. Zn_1 is anticipated to

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

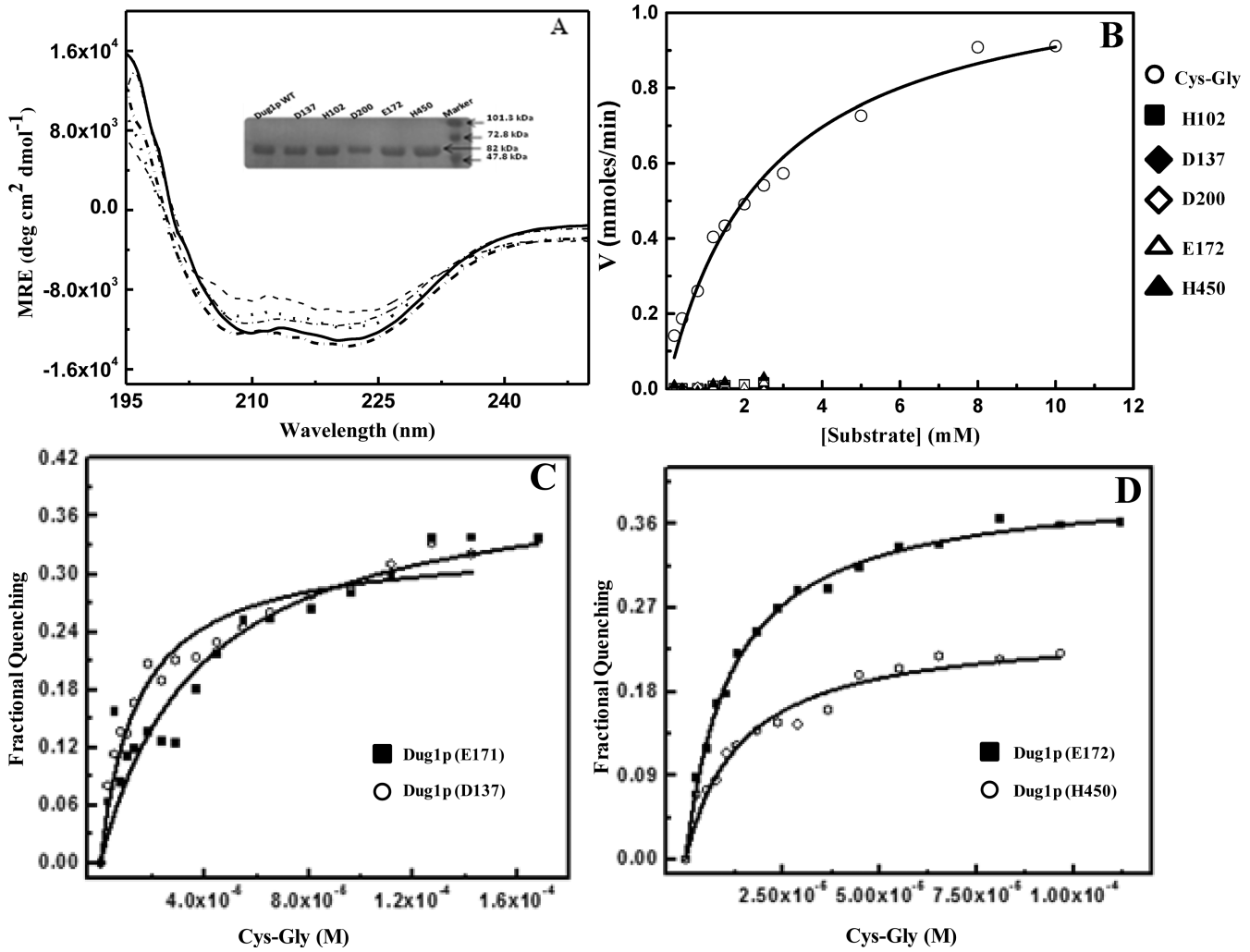


Figure 5. Characterization of active site mutants. (A) CD spectra of active site mutants confirm nativelike structures. The inset shows SDS−PAGE of mutants with markers. (B) Kinetics of Cys-Gly hydrolysis by wild type Dug1p and active site mutants (indicated by symbols). Initial velocities are plotted vs substrate concentration, and data were fit to Michaelis−Menten model as described in eq 4. (C and D) Analyses of Cys-Gly binding of active site mutants (as indicated by symbols). Fractional quenching upon substrate binding is plotted vs substrate concentration. All binding assays were conducted in binding buffer, and concentrations of active site mutants were kept constant (~1.0 × 10^{-7} M).

polarize the carbon atom of the target peptide bond and render the peptide bond susceptible to nucleophilic attack.[13,14] Therefore, we tested the possibility that these designed tripeptides could be hydrolyzed by Dug1p. Hydrolysis of tripeptides was monitored by mass spectrometry methods as described previously.[22] Dug1p was able to cleave Tyr-Gly-Cys and Ala-Gly-Cys tripeptides but showed no or very little cleavage activity toward His-Gly-Cys (HGC) (Figure 4C). We tested the possibility that HGC can still bind to Dug1p but cannot be hydrolyzed. We used fluorescence quenching experiments to monitor the binding of HGC, and results show that HGC can bind Dug1p with an estimated dissociation constant (K_d) of ~0.12 mM (Figure 4D). Although HGC tripeptide is resistant to proteolytic degradation by Dug1p, this tripeptide forms the enzyme−peptide complex, supporting our earlier proposal that hydrolysis might be the rate-limiting step in this case.[3] Therefore, our results demonstrate that Dug1p binds tripeptides but hydrolyzes selectively. To improve our understanding of the specificity of binding, we used a hexameric peptide with the sequence Gly-Cys-Gly-Cys-Gly-Cys and examined the binding. Dug1p does not bind this hexapeptide, suggesting that Dug1p shows length-dependent binding specificity (Figure 4D).

Catalytically Inactive Mutants Bind Substrate and Small Peptides. To further support structural analyses and confirm catalytic residues, we designed several point mutations (H102, D137, D200, E172, and H450) based on the Dug1p structure. All mutants were expressed and purified to homogeneity. Purified mutant proteins show nativelike CD spectra, indicating that point mutation of any one of the catalytic residues has not significantly altered the secondary or tertiary structure (Figure 5A). Dug1p has an extra polyproline stretch that is not present in CN2, and our structure shows that this motif may form an extended loop at the active site gate. Therefore, to test the role of this loop, we deleted this motif. However, this deletion mutant was not soluble under our purification conditions, and hence, no assays with this deletion mutant were conducted.

First, we tested the catalytic activities of mutants using Cys-Gly as the substrate. None of the point mutants showed any detectable catalytic activity compared to the wild type, indicating that all five mutants are catalytically inactive (Figure 5B). Dug1p shows comparable kinetics (K_M ~ 2.4 mM, and

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

V_{max} ~ 1.1 μmol/min), as reported previously.[22] Next, we examined the substrate−peptide binding properties of active site mutants to understand the effect of active site mutation on substrate recognition. Again, quenching of intrinsic tryptophan fluorescence upon peptide binding is monitored as a function of peptide concentration. The binding isotherms were normalized to fractional saturation, and equilibrium binding constants were obtained as described in Materials and Methods. All active site mutants that lack catalytic activity bind the cognate substrate Cys-Gly (Figure 5C,D). Binding affinities of active site mutants are very similar and are listed in Table 2. The average equilibrium dissociation constant (23.0 μM) estimated for active site mutants is comparable to the equilibrium dissociation constant (~6.0 μM) estimated for wild type Dug1p from the earlier study.[3] We further tested the possibility that indiscriminate peptide binding but selective hydrolysis may compromise the Dug1p activity in the presence of a "nonsubstrate" peptide like the HGC tripeptide used in this study. The competitive titration activity study shows that the Cys-Gly hydrolysis activity of is reduced to ~45% of the original activity in the presence of HGC tripeptide at concentrations equivalent to its equilibrium dissociation constant.

Table 2. Binding Affinities of Dug1p and Active Site Mutants

peptide	K_d (M)	ΔG (kcal mol^{-1})
HGC	1.18×10^{-4}	−5.1 (Dug1P)
Cys-Gly	1.96×10^{-5}	−6.3 (D200)
Cys-Gly	1.41×10^{-5}	−6.6 (H450)
Cys-Gly	3.70×10^{-5}	−6.0 (E172)
Cys-Gly	1.40×10^{-5}	−6.6 (D137)

Mass Spectrometry Analyses of Dug1p. To test further whether the bound dipeptide could be the hydrolytic product of a tripeptide (X-Gly-Cys) from which the N-terminal X residue is cleaved, we modeled three tripeptides that mimic bestatin. We performed MALDI analyses of purified native Dug1p and Dug1p incubated with YGC tripeptide under assay conditions. The MALDI data of purified Dug1p showed two peaks, one prominent peak belonging to a molecular mass of ~53837 Da (expected mass of ~53842 Da) and a shoulder peak corresponding to a molecular mass of ~54034 Da (Figure 6A,B). Because the protein is purified to homogeneity, the peak with the increased molecular mass may be interpreted as the Dug1p−ligand complex. The molecular mass of the Gly-Cys dipeptide is ~178 Da, and therefore, if Dug1p is purified with Gly-Cys bound, the net molecular mass would be equivalent to ~54021 Da, which is similar to the observed mass. Interestingly, when Dug1p is incubated with YGC peptide for 2 h, the shoulder peak corresponding to a molecular mass of ~54034 Da increases slightly, suggesting that hydrolyzed dipeptide may be still bound to the active site (Figure 6B). It is possible that release of the cleaved dipeptide product is the rate-limiting step in the hydrolysis. However, detailed studies are necessary to support this hypothesis.

DISCUSSION

We report here structural and mechanistic features of recently discovered M20 family metallopeptidase Dug1p from *S. cerevisiae*. Dug1p hydrolyzes Cys-Gly dipeptides with high specificity but binds a variety of small peptides, di- and tripeptides.[1−3] M20 family peptidases, including Dug1p and carnosinases, play crucial roles in a wide variety of cellular processes, and therefore, enzymes of the M20 family have attracted much attention from the perspective of protein engineering, tailoring substrate specificity, and the design of peptide inhibitors.[33,34] However, recognition of diverse peptide substrates by M20 family peptidases and a lack of structural information about peptide recognition principles present challenges for the design of peptidase specific inhibitors. The structure of Dug1p in complex with a natural dipeptide and biochemical analyses presented here provide the first insights into the stereochemical features of a natural peptide bound at the catalytic site. This structure advances our understanding of how Dug1p and other closely related M20 family members recognize peptide substrates.

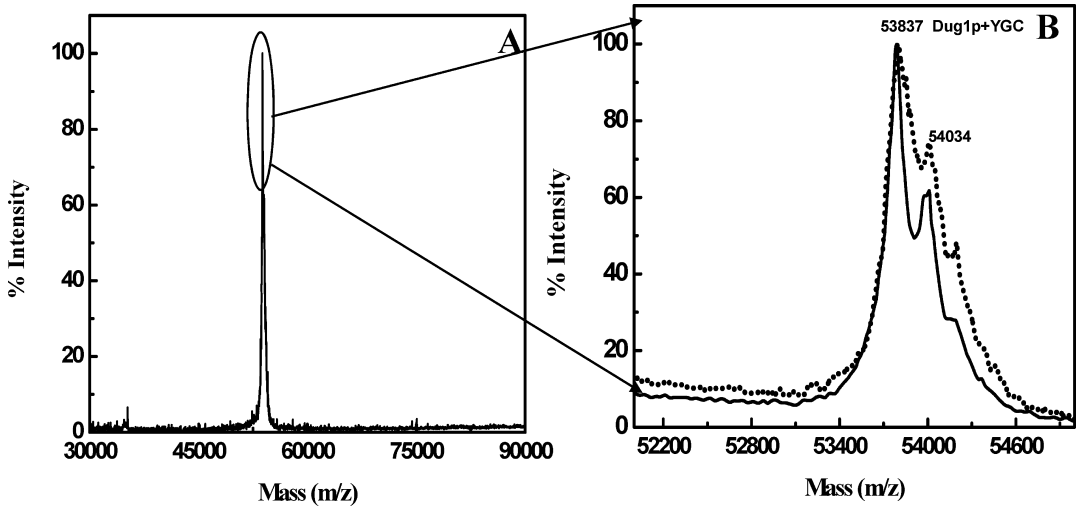


Figure 6. MALDI analyses of Dug1p. (A) The solid line represents the Dug1p peak with an average molecular mass of ~53837.9 Da. (B) Close-up of panel A. The solid line represents the MALDI profile for size exclusion-purified Dug1p, and the second peak is observed at 54034 Da. Overlay of MALDI profiles of Dug1p and Dug1p incubated with YGC peptide (dashed). Both profiles show a significant fraction of higher-molecular mass species; the *x*-axis is the *m*/*z* ratio and the *y*-axis the percent intensity. The expanded profile shows that ~50−65% of higher-molecular mass species are present.

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

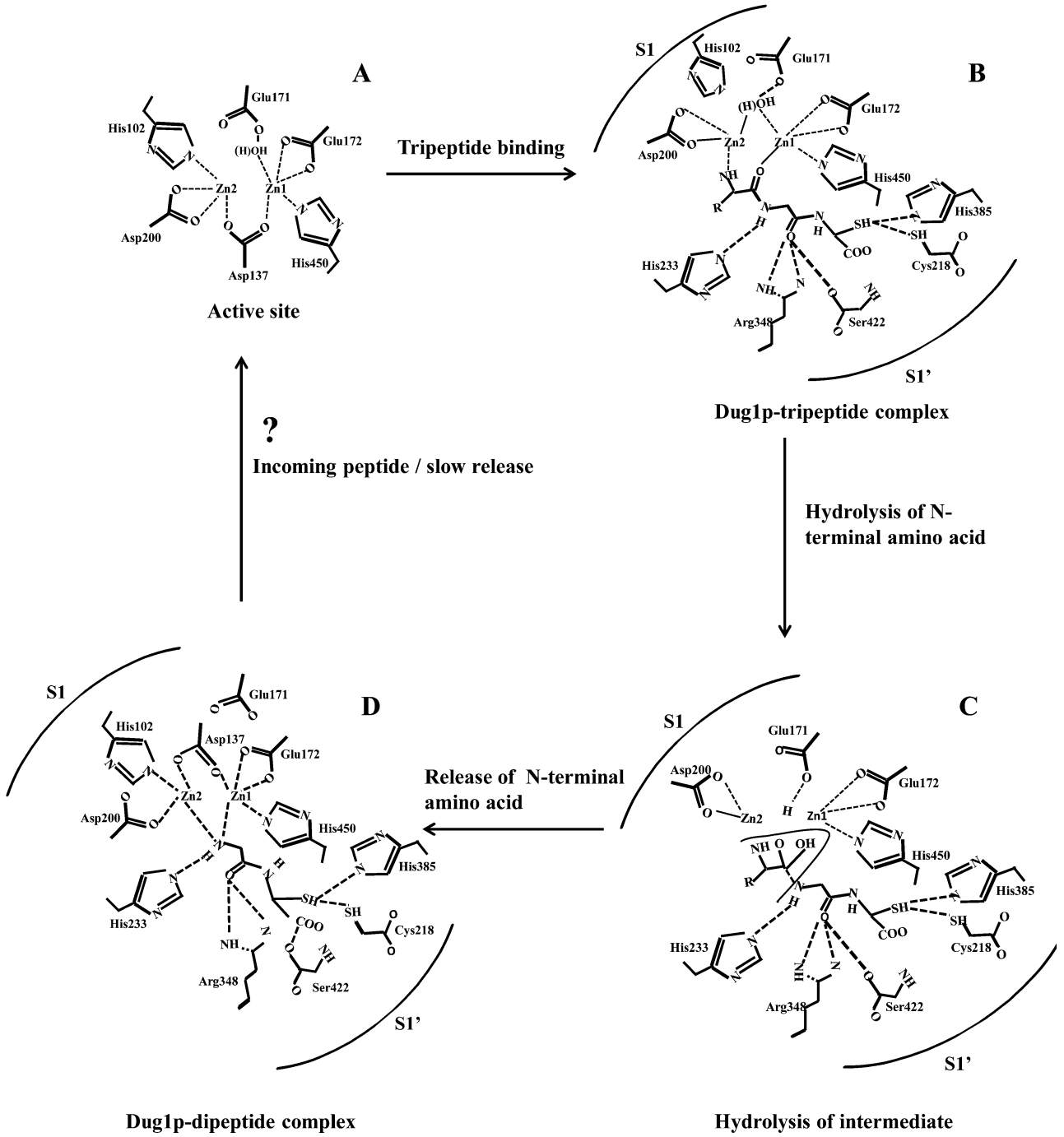


Figure 7. Schematic diagram depicting the plausible catalytic cycle of Dug1p-mediated hydrolysis. (A) View of the active site in the free state. (B) View of the active site residue after a tripeptide bound. Note that Asp200, His102, and Zn_2 are shown to be in different coordinations in the bound state. (C) Activation of the catalytic water molecule, hydrolysis of the target peptide bond, and release of the amino-terminal product. (D) View of Gly-Cys dipeptide, the possible product left at the Dug1p active site. The forward step from panel D to panel A is rate-limited, and the mechanism of product release is not known.

The most notable observation is that peptides are recognized through their N- and C-termini, with minimal contributions from the side chain. The stereochemical features of interactions between the peptide and active site residues explain how Dug1p binds a variety of small peptides in a sequence-independent manner.[3] The mode of peptide binding is similar to the "two-prong plug two-hole socket" model.[35] The N-terminus of the peptide fits into the catalytic S1 binding pocket, whereas the C-terminus of the peptide is firmly bound to the S1′ pocket. Point mutation of any of the catalytically important residues of the S1 pocket (Asp137, Asp200, Glu172, and His450) does not compromise peptide binding, supporting the structural observation and also indicating that peptides are recognized through a noncooperative mode. Further, the tunnel-like, large, and solvent accessible active site provides sufficient room for accommodating small peptides of different lengths and sequences. Together, structural and biochemical analyses reported here explain how Dug1p and other M20 family members with large catalytic pockets recognize diverse peptides.

dx.doi.org/10.1021/bi501263u | *Biochemistry* 2014, 53, 7870−7883

The next important question is how Dug1p and other M20 family members attain specificity toward their respective substrate peptides. Binding-induced conformational changes that promote substrate hydrolysis are considered one of the key features of enzyme-mediated catalysis. The same mechanism is also proposed for Dug1p-mediated hydrolysis of peptides.[3] To obtain evidence of active site remodeling, we compared active site features of Dug1p and CN2 structures. Positions and conformational features of metal ions and catalytic residues differ between these two structures, suggesting that peptide binding and inhibitor binding influence active sites differently. Because all catalytic residues are highly conserved between Dug1p and CN2 and both share very high degrees of sequence and structure similarity, differences in conformational properties between two active sites can be attributed to binding-mediated effects of dipeptide and bestatin, respectively. More structures of M20 family enzymes in complex natural peptides will improve our understanding of the conformational dynamics of the active site.

Two main features of the Dug1p structure support the notion that the bound peptide could be the product, generated from hydrolysis of tripeptides. The similarity between binding poses of Gly-Cys and bestatin, which mimics a tripeptide and the vacant space at the S1 pocket where the N-terminus of the bestatin binds in the CN2 structure, suggests that a tripeptide could be easily accommodated into the Dug1p active site. According to the well-accepted canonical mechanism, the carbonyl of the target peptide bond should interact with one of the metal ions.[14] However, the carbonyl oxygen of the Gly-Cys peptide bond faces away from both metal ions and interacts with the main chain carbonyl of Ser422. These features suggest that the dipeptide does not mimic the catalytic substrate. Results of tripeptide hydrolysis and MALDI analyses encourage us to propose that the bound peptide could be the cleaved product of a tripeptide. The increased mass, matching the molecular mass of Gly-Cys peptide-bound Dug1p, supports this assumption. The C-terminal cysteine of the dipeptide is held tightly by Ser422 and Arg348, and its side chain sulfhydryl group makes a weak interaction with the side chain of Cys218. It is possible that peptides with a C-terminal cysteine may have longer residential times within the Dug1p active site. *E. coli* produces many metal-chelating peptides that belong to the cysteine rich peptide (CP) group that have multiple Gly-Cys dipeptide motifs.[36,37] Degradation of metal binding CPs by endogenous proteases and peptidases would produce a load of Cys-Gly or Gly-Cys dipeptide motifs or other small peptides containing these motifs.

Comparison of the structure of Dug1p with known structures of metallopeptidases and proteinases and results of our biochemical studies allow us to propose a catalytic mechanism that is consistent with the stereochemical features of the structure presented here. The proposed catalytic mechanism is captured in the scheme depicted in Figure 7. The tripeptide enters the active site with its N-terminus locked into the S1 pocket and the C-terminus positioned by residues in the S1′ pocket. The binding mode of the peptide is similar to the two-prong plug and two-socket binding mode.[35] Peptide binding triggers conformational changes that position the target peptide bond for hydrolysis, if the peptide is a potential substrate. We describe a four-step hydrolysis model in the following order (Figure 7A–D). (A) In the unbound state or before peptide binding, the catalytic water molecule is hydrogen bonded to the carboxylate of Glu171 as well as to both metal ions. (B) The incoming peptide latches onto S1 and S1′ binding pockets. The carbonyl oxygen of the target peptide bond is positioned to be in close contact with Zn_1, whereas Zn_2 interacts with the N-terminus of the peptide. Residues at the S1′ pocket, Arg348 and Ser422, fix the peptide carboxylate, while His233 from the other subunit forms a hydrogen bond to the second peptide bond, providing more stability. (C) This step consists of events related to dynamics at the active site and hydrolysis: dynamics of Asp200, Zn_2, and His102 results in the movement of Zn_2 toward peptide and positioning of the N-terminus so that the carbonyl of the first target peptide bond is positioned in the proximity of Zn_1 and the carbonyl carbon is polarized. The sequence is hydrolysis of the target peptide bond by the activated nucleophilic hydroxyl ion, generated by abstraction of protein from the catalytic water by Glu171. (D) In the next step, the cleaved amino-terminal residue is released and Gly-Cys dipeptide is left in the active site. However, by a canonical mechanism, the cleaved C-terminal fragment (Gly-Cys) should exit first, and then in the next step, the N-terminal X residue of X-Gly-Cys tripeptide is released. However, the structure of the Dug1p−dipeptide complex shows that Gly-Cys is tightly bound at the S1′ and S1 sites. In the canonical mechanism proposed earlier, the exit of Gly-Cys would be facilitated by the protonation of the N-terminal amide nitrogen by Glu171.[14] However, stereochemical features of the bound Gly-Cys dipeptide suggest that Glu171 cannot donate a proton because the configuration of the cleaved N-terminal amide is misaligned and the side chain of Glu171 is 4.4 Å from the peptide N-terminus. Therefore, the exit of the cleaved X residues first from the active site is more consistent with the structure of the Dug1p−dipeptide complex observed here (Figure 7D). The proposed model is supported by the observation that the cleaved Gly-Cys dipeptide retains most of its prehydrolytic interactions in the S1′ pocket. Because most conformational dynamics are restricted to the S1 pocket, the cleaved N-terminal residue may lose its prehydrolysis contacts and may leave the active site first. Although the last step is not supported by results of this study, stereochemical features of the Dug1p active site adequately describe here the first three steps. Nonetheless, this structure supports a mechanism in which the amino-terminal moiety of the cleaved peptide, which is bound at the S1 site, should leave first. This observation is also supported by an earlier structural study that showed that matrix metalloproteinase-12 (MMP-12) functions in a similar manner. On the basis of structural snapshots, the study proposed that the C-terminus of the substrate peptide (Pro-Gln-Gly/Ile-Ala-Gly) binds tightly, and upon hydrolysis of Gly/Ile peptide bond, the Pro-Gln-Gly segment leaves the active site first.[38] The other new mechanistic information, other than understanding the principles of peptide recognition, gained from this study is that the role of Zn_2 might be in positioning the substrate peptide within the active site by interacting with its N-terminus. The role of Zn_2 is dictated by the substrate binding and subsequent dynamics of active site residues as described here. The role of Zn_2 was not clear from earlier studies, and its dynamics within the active site was not anticipated.

Structural similarity between Dug1p and other M20 family members and Cys-Gly-degrading peptidases suggests that these enzymes may share common substrate recognition features and are likely to operate through the common four-step mechanism proposed here. The Gly-Cys dipeptide binds in an extended conformation with its backbone trajectory superimposing well with that of bestatin in the CN2 structure and of other

inhibitors bound to different structures. This observation supports the notion that competitive inhibitors mimic natural substrates or products when bound to their respective active sites, and the structure provides peptide specific stereochemical features that would be useful for designing natural peptide inhibitors. The binding mode of the peptide conforms to stereochemical requirements of inhibitors, suggesting that optimization of the S1′ binding pocket parameters is necessary for designing substrates and inhibitors. This is an important observation because stereochemical features of the bound peptide would serve as a "real" template for the design of natural peptide inhibitors. Natural peptide inhibitors are more demanding as they have low immunogenicity and high specificity.[39]

In summary, the crystal structure of recently discovered Dug1p has been resolved in complex with a dipeptide. The structure explains roles of metal ions and active site residues in peptide recognition. Structural analyses of Dug1p and related family members, modeling, and biochemical studies revealed that the S1 pocket, where catalysis occurs, is dynamic and the S1′ pocket serves as the anchor to stabilize the peptide at the active site. Results also reveal the tripeptide hydrolytic potential of Dug1p that was not known previously. Molecular determinants of peptide recognition described here are likely to be shared by M20 family members that possess a large active site cleft and share a high degree of structural similarity. Further, similarity between binding poses of the peptide and known metallopeptidase inhibitors suggests that the structure of the Dug1−peptide complex can serve as a template for the design of natural peptide inhibitors against M20 metallopeptidases.

■ ASSOCIATED CONTENT

Supporting Information

Simulated annealing omit map of peptide and two zinc ions (Figure S1). This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Council of Scientific and Industrial Research (CSIR), Institute of Microbial Technology, G. N. Ramachandran Protein Centre, Sector 39A, Chandigarh 160036, India. Telephone: 91-172-6665474. Fax: 91-172-2690585. E-mail: skumaran@imtech.res.in.

Author Contributions

A.K.S. and M.S. contributed equally to this work.

Funding

A.K.S., M.S., V.K.P., V.S., B.G.L., and M.M. received research fellowships from CSIR, India, and the Department of Biotechnology (DBT), India. This work was supported by the Department of Science and Technology (DST) and CSIR, India.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We gratefully acknowledge access to in-house X-ray and three-dimensional structural biology facilities at IMTECH. We are grateful to Prof. A. K. Bachhawat for his gift of Dug1p plasmid and discussions.

■ ABBREVIATIONS

IPTG, isopropyl β-thiogalactopyranoside; Ni-NTA, nickel-nitrilotriacetic acid; GSH, glutathione; CD, circular dichroism; HPLC, high-performance liquid chromatography; ESI-MS, electrospray ionization mass spectrometry; MALDI, matrix-assisted laser desorption ionization; MD, molecular dynamics; PCR, polymerase chain reaction.

■ REFERENCES

(1) Ganguli, D., Kumar, C., and Bachhawat, A. K. (2007) The alternative pathway of glutathione degradation is mediated by a novel protein complex involving three new genes in *Saccharomyces cerevisiae*. *Genetics 175*, 1137−1151.

(2) Kaur, H., Kumar, C., Junot, C., Toledano, M. B., and Bachhawat, A. K. (2009) Dug1p Is a Cys-Gly peptidase of the γ-glutamyl cycle of *Saccharomyces cerevisiae* and represents a novel family of Cys-Gly peptidases. *J. Biol. Chem. 284*, 14493−14502.

(3) Kaur, H., Datt, M., Ekka, M. K., Mittal, M., Singh, A. K., and Kumaran, S. (2011) Cys-Gly specific dipeptidase Dug1p from *Saccharomyces cerevisiae* binds promiscuously to di-, tri-, and tetra-peptides: Peptide-protein interaction, homology modeling, and activity studies reveal a latent promiscuity in substrate recognition. *Biochimie 93*, 175−186.

(4) Hanson, H. T., and Smith, E. L. (1949) Carnosinase: An enzyme of swine kidney. *J. Biol. Chem. 179*, 789−801.

(5) Teufel, M., Saudek, V., Ledig, J. P., Bernhardt, A., Boularand, S., Carreau, A., et al. (2003) Sequence identification and characterization of human carnosinase and a closely related non-specific dipeptidase. *J. Biol. Chem. 278*, 6521−6531.

(6) Unno, H., Yamashita, T., Ujita, S., Okumura, N., Otani, H., Okumura, A., et al. (2008) Structural basis for substrate recognition and hydrolysis by mouse carnosinase CN2. *J. Biol. Chem. 283*, 27289−27299.

(7) Rawlings, N. D., Waller, M., Barrett, A. J., and Bateman, A. (2012) MEROPS: The database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res. 42*, D503−D509.

(8) Sauerhofer, S., Yuan, G., Braun, G. S., Deinzer, M., Neumaier, M., Gretz, N., et al. (2007) L-Carnosine, a substrate of carnosinase-1, influences glucose metabolism. *Diabetes 56*, 2425−2432.

(9) Pulido-Cejudo, G., Conway, B., Proulx, P., Brown, R., and Izaguirre, C. A. (1997) Bestatin-mediated inhibition of leucine aminopeptidase may hinder HIV infection. *Antiviral Res. 36*, 167−177.

(10) Jozic, D., Bourenkow, G., Bartunik, H., Scholze, H., Dive, V., Henrich, B., et al. (2002) Crystal structure of the dinuclear zinc aminopeptidase PepV from *Lactobacillus delbrueckii* unravels its preference for dipeptides. *Structure 10*, 1097−1106.

(11) Stec-Niemczyk, J., Pustelny, K., Kisielewska, M., Bista, M., Boulware, K. T., Stennicke, H. R., et al. (2009) Structural and functional characterization of SplA, an exclusively specific protease of *Staphylococcus aureus*. *Biochem. J. 419*, 555−564.

(12) Roswell, S., Paupit, R. A., Tucker, A. D., Melton, R. G., Blow, D. M., and Brick, P. (1997) Crystal structure of carboxypeptidase G2, a bacterial enzyme with applications in cancer therapy. *Structure 5*, 337−347.

(13) Chang, C. Y., Hsieh, Y. C., Wang, T. Y., Chen, Y. C., Wang, Y. K., Chiang, T. W., et al. (2010) Crystal structure and mutational analysis of aminoacylhistidine dipeptidase from *Vibrio alginolyticus* reveal a new architecture of M20 metallopeptidases. *J. Biol. Chem. 285*, 39500−39510.

(14) Wilcox, D. E. (1996) Binuclear Metallohydrolases. *Chem. Rev. 96*, 2435−2458.

(15) Holz, R. C., Bzymek, K. P., and Swierczek, S. I. (2003) Co-catalytic metallopeptidases as pharmaceutical targets. *Curr. Opin. Chem. Biol. 7*, 197−206.

(16) Otwinowski, Z., Borek, D., Majewski, W., and Minor, W. (2003) Multiparametric scaling of diffraction intensities. *Acta Crystallogr. A59*, 228−234.

(17) Kabsch, W. (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr. D66*, 133−144.
(18) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr. 40*, 658−674.
(19) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., et al. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. D67*, 235−242.
(20) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D66*, 486−501.
(21) DeLano, W. L. (2010) *The PyMOL Molecular Graphics System*, version 1.3r1, Schrödinger, LLC, New York.
(22) Pandya, V., Ekka, M. K., Dutta, R. K., and Kumaran, S. (2011) Mass spectrometry assay for studying kinetic properties of dipeptidases: Characterization of human and yeast dipeptidases. *Anal. Biochem. 418*, 134−142.
(23) Gaitonde, M. K. (1967) A spectrophotometric method for the direct determination of cysteine in the presence of other naturally occurring amino acids. *Biochem. J. 104*, 627−633.
(24) Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL Workspace: A web-based environment for protein structure homology modeling. *Bioinformatics 22*, 195−201.
(25) Berendsen, H. J. C., Van der Spoel, D., and Van Drunen, R. (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Comput. Phys. Commun. 91*, 43−56.
(26) Jorgensen, W. L., and Tirado-Rives, J. (1988) The OPLS Force Field for Proteins. Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc. 110*, 1657−1666.
(27) Wu, Y., Tepper, H. L., and Voth, G. A. (2006) Flexible simple point-charge water model with improved liquid-state properties. *J. Chem. Phys. 124*, 024503.
(28) Berendsen, H. J. C., Postma, J. P. M., Van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys. 81*, 3684−3690.
(29) Hess, B., Bekkar, H., Berendsen, H. J. C., and Johannes, J. G. E. M. (1997) LINCS: A linear constraint solver for molecular simulations. *J. Comput. Chem. 18*, 1463−1472.
(30) Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: An N.log (N) method for Ewald sums in larger systems. *J. Chem. Phys. 98*, 10089−10092.
(31) Håkansson, K., and Miller, C. G. (2002) Structure of peptidase T from *Salmonella typhimurium*. *Eur. J. Biochem. 269*, 443−450.
(32) Girish, T. S., and Gopal, B. (2010) Crystal structure of *Staphylococcus aureus* metallopeptidase (Sapep) reveals large domain motions between the manganese-bound and apo-states. *J. Biol. Chem. 285*, 29406−29415.
(33) Harbut, M. B., Velmourougane, G., Reiss, G., Chandramohandas, R., and Greenbaum, D. C. (2008) Development of bestatin-based activity-based probes for metallo-aminopeptidases. *Bioorg. Med. Chem. Lett. 18*, 5932−5936.
(34) Mucha, A., Drag, M., Dalton, J. P., and Kafarski, P. (2010) Metallo-aminopeptidase inhibitors. *Biochimie 92*, 1509−1529.
(35) Waksman, G., Shoelson, S. E., Pant, N., Cowburn, D., and Kuriyan, J. (1993) Binding of a high affinity phosphotyrosyl peptide in the src SH2 domain: Crystal structures of the complexed and peptide-free forms. *Cell 72*, 779−790.
(36) Kotrba, P., Doleckova, L., de Lorenzo, V., and Ruml, T. (1999) Enhanced bioaccumulation of heavy metal ions by bacterial cells due to surface display of short metal binding peptides. *Appl. Environ. Microbiol. 65*, 1092−1098.
(37) Pazirandeh, M., Wells, B. M., and Ryan, R. L. (1998) Development of bacterium-based heavy metal biosorbents: Enhanced uptake of cadmium and mercury by *Escherichia coli* expressing a metal binding motif. *Appl. Environ. Microbiol. 64*, 4068−4072.
(38) Bertini, I., Calderone, V., Fragai, M., Luchinat, C., Maletta, M., and Yeo, K. J. (2006) Snapshots of the reaction mechanism of matrix metalloproteinases. *Angew. Chem., Int. Ed. 45*, 7952−7955.
(39) Mason, J. M. (2010) Design and development of peptides and peptide mimetics as antagonists for therapeutic intervention. *Future Med. Chem. 2*, 1813−1822.