Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structure-based elucidation of the regulatory mechanism for aminopeptidase activity

Hai Minh Ta,[a] Sangsu Bae,[b] Seungsu Han,[c] Jihyuck Song,[a] Tae Kyu Ahn,[d] Sungchul Hohng,[b,e] Sangho Lee[c]* and Kyeong Kyu Kim[a]*

[a]Department of Molecular Cell Biology, School of Medicine, Samsung Biomedical Research Institute, Sungkyunkwan University, Suwon 440-746, Republic of Korea, [b]Department of Physics and Astronomy and National Center for Creative Research Initiatives, Seoul National University, Seoul 151-742, Republic of Korea, [c]Department of Biological Sciences, Sungkyunkwan University, Suwon 440-746, Republic of Korea, [d]Department of Energy Science, Sungkyunkwan University, Suwon 440-746, Republic of Korea, and [e]Biophysics and Chemical Biology, Seoul National University, Seoul 151-742, Republic of Korea

Correspondence e-mail: sangholee@skku.edu, kyeongkyu@skku.edu

Received 11 December 2012
Accepted 8 May 2013

PDB References: PepS, closed form, 4icq; open form, 4icr; substrate-bound form, 4ics

The specificity of proteases for the residues in and length of substrates is key to understanding their regulatory mechanism, but little is known about length selectivity. Crystal structure analyses of the bacterial aminopeptidase PepS, combined with functional and single-molecule FRET assays, have elucidated a molecular basis for length selectivity. PepS exists in open and closed conformations. Substrates can access the binding hole in the open conformation, but catalytic competency is only achieved in the closed conformation by formation of the S1 binding pocket and proximal movement of Glu343, a general base, to the cleavage site. Hence, peptides longer than the depth of the binding hole block the transition from the open to the closed conformation, and thus length selection is a prerequisite for catalytic activation. A triple-sieve interlock mechanism is proposed featuring the coupling of length selectivity with residue specificity and active-site positioning.

1. Introduction

In the catabolic pathways of proteins, a polypeptide is sequentially degraded into individual amino acids *via* oligopeptides (Goldberg & Dice, 1974). In this process, various peptidases with different substrate selectivities are involved in a hierarchical order. In bacteria, proteasome-like proteases, such as HslUV, ClpAP, Lon and FtsH, function in an ATP-dependent manner to yield peptide fragments of about ten amino acids each (Kisselev *et al.*, 1998; Saric *et al.*, 2004; De Mot *et al.*, 1999). These oligopeptides are subsequently digested by tricorn proteases or aminopeptidases, leading to the final release of free individual amino acids (Tamura *et al.*, 1996, 1998).

Tight control of proteolytic activity is essential to prevent the undesired cleavage of proteins and to maintain cellular homeostasis and signalling. Since this control is achieved mostly by substrate selectivity, exploring such selectivity is a prerequisite for a comprehensive understanding of protein catabolism. Aminopeptidases select their substrates by filtering for both the amino-acid residues in and the length of substrates. However, although a plethora of biochemical and structural studies have been performed on residue specificity (Lowther & Matthews, 2002), little is known about length selectivity. Therefore, further comprehensive study is necessary to investigate the substrate selectivity of peptidases from the perspective of both residues and length.

Proteolytic activity can also be regulated by controlling access of the substrate to the active site. While many

peptidases have active sites that are located in accessible clefts, some peptidases restrict substrate access to the active site in either an unfoldase-assisted manner, as in ClpP coupled with either ClpA or ClpX (Kim & Kim, 2008; Sauer *et al.*, 2004), or *via* narrow entry channels through which the substrate enters the active sites, as exemplified by PepA from *Streptococcus pneumoniae* (Kim *et al.*, 2010). Substrate access can also be regulated by a domain movement. For example, HtrA proteases mobilize the 'lid' motif to toggle between 'open' and 'closed' states to control substrate access (Kim *et al.*, 2003; Kim & Kim, 2008). Some bacterial aminopeptidases, such as AmpT, have also been proposed to control the access of substrates by conformational change (Odintsov *et al.*, 2005*b*). AmpT has two conformational states, 'open' and 'closed', and only the 'open' state is thought to be accessible, based on the observation that five promoters with a range of conformations are observed in the asymmetric unit of the *Thermus thermophilus* AmpT crystal structure. However, the low-resolution crystal structure of AmpT (3.7 Å) prevented further detailed structural characterization of substrate recognition or of the activation mechanism following substrate entry. Despite the intensive structural characterization of aminopeptidases, their regulation mechanisms, which are accompanied by domain movements, have never been confirmed at the single-molecule level.

In this study, we investigate the catalytic activity and regulation mechanism of PepS (Kissinger *et al.*, 1999), a metalloaminopeptidase from *S. pneumoniae*. PepS belongs to the M29 family in clan MQ (Rawlings *et al.*, 2008) and exhibits residue specificity for peptides containing arginine or aromatic amino acids at the N-terminus and a length selectivity of 2–9 residues (Kissinger *et al.*, 1999). Physiologically, PepS and its homologues support bacterial survival. For example, PepS from *S. thermophilus* provides phenylalanine and arginine, which stimulate growth of the CNRZ302 strain (Fernandez-Espla & Rul, 1999). Furthermore, antisense RNA against AmpS from *Staphylococcus aureus*, a PepS homologue, causes a growth defect (Ji *et al.*, 2001), and a recent study suggested that PepS may also be involved in peptidoglycan metabolism (Thomas *et al.*, 2010).

By performing structural, functional and single-molecule examinations, we elucidated the molecular basis of the substrate-length selectivity of aminopeptidase. We further propose a concerted regulatory mechanism for aminopeptidase, in which the catalytic activity is controlled by a combination of the following three factors: active-site positioning, substrate-length selectivity and residue specificity.

Table 1
Data-collection and refinement statistics.

Values in parentheses are for the highest resolution shell.

Conformation	Open†	Closed	
Mutation	E343D	None	E343D
Ligand	None	None	Trp-Gly
Data collection			
Space group	$P2_12_12$	$P4_32_12$	$P4_32_12$
Unit-cell parameters (Å, °)	a = 94.3, b = 185.3, c = 59.2, $\alpha = \beta = \gamma$ = 90	a = 126.6, b = 126.6, c = 139.0, $\alpha = \beta = \gamma$ = 90	a = 126.4, b = 126.4, c = 139.3, $\alpha = \beta = \gamma$ = 90
Resolution (Å)	50–2.17	50–2.25	50–1.97
R_{sym} or R_{merge} (%)	10.5 (25.1)	8.3 (48.5)	8.3 (42.4)
$\langle I/\sigma(I)\rangle$	20.0 (8.0)	46.0 (5.9)	42.1 (5.1)
Completeness (%)	99.6 (100)	100 (99.6)	99.6 (100)
Multiplicity	7.2 (7.3)	14.3 (14.6)	9.3 (8.4)
Refinement			
Resolution (Å)	50–2.17	50–2.25	50–1.97
No. of reflections	55649 (3961)	54281 (3932)	79807 (5806)
R_{work}/R_{free}	0.229/0.300	0.199/0.246	0.225/0.257
No. of atoms			
Protein	6262	6366	6364
Ligands/ions	0/4 Zn, 4 cacodylate	0/4 Zn	38/4 Zn
Waters	457	246	512
B factors (Å^2)			
Protein	36.6	33.6	28.5
Ligands/ions	—/31.7	—/35.8	35.5/33.6
Waters	45.9	40.0	42.1
R.m.s. deviations			
Bond lengths (Å)	0.025	0.022	0.018
Bond angles (°)	1.93	1.87	1.76

† The data set was collected from one crystal.

2. Experimental procedures

2.1. Expression and purification of wild-type and mutant PepS

S. pneumoniae PepS was cloned into pVFT1S (Korean Patent 10-0690230) and the resulting plasmid was transformed into *Escherichia coli* strain BL21(DE3) for protein expression. Cells were grown in LB medium to an OD_{600} of 0.6 prior to induction with 1 m*M* isopropyl β-D-1-thiogalactopyranoside and were then grown for a further 24 h at 291 K. The harvested cells were resuspended in 50 m*M* Tris–HCl pH 8.0, 500 m*M* NaCl, 20 m*M* imidazole, disrupted by sonication and centrifuged at 20 000 rev min^{-1} for 40 min. The supernatant was applied onto an Ni–NTA column (GE Healthcare) and PepS was eluted with a linear gradient of 50 m*M* to 1.0 *M* imidazole. Fractions containing PepS were dialyzed against 25 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl, 1 m*M* DTT and were then treated with TEV protease to remove the N-terminal His tag. The resulting solution was applied onto an Ni–NTA column to remove uncleaved PepS. The final purification step was performed by size-exclusion chromatography using a Superdex 200 column (GE Healthcare) equilibrated with 25 m*M* Tris–HCl pH 8.0, 100 m*M* NaCl. PepS mutants were prepared by site-directed mutagenesis using the QuikChange

kit (Stratagene) and were purified as described for wild-type PepS.

2.2. Crystallization and data collection

All crystallization experiments were performed using the microbatch method by mixing 1 µl reservoir solution and 1 µl protein solution at 295 K. Crystals of closed ligand-free PepS were grown in 2 *M* $MgSO_4$, 0.1 *M* sodium acetate pH 4.6. Crystals of $PepS^{E343D}$–Trp-Gly were obtained using the same conditions as for the closed PepS. A fivefold molar excess of substrate (Trp-Gly) was mixed with $PepS^{E343D}$ before crystallization. Crystals of open $PepS^{E343D}$ were obtained in 10% 2-propanol, 0.1 *M* cacodylate pH 6.5, 0.2 *M* zinc acetate, and attempts were made to cocrystallize open $PepS^{E343D}$ with Trp-Gly as above. Despite the presence of the substrate Trp-Gly in the crystallization drops, the crystals of open $PepS^{E343D}$ did not contain the substrate.

Diffraction data for crystals of the two conformations of PepS and of the complex with the substrate Trp-Gly were collected at resolutions of 1.97–2.25 Å using a wavelength of 1.0000 Å with a Bruker AXS PROTEUM 300 CCD detector on beamline 6B at Pohang Light Source, Korea and an ADSC Quantum 210 CCD detector on beamline AR-NW12 at the Photon Factory, Japan. Prior to data collection, crystals were cryoprotected by the addition of 20–25%(*v*/*v*) glycerol and were flash-cooled in a cold nitrogen stream. All diffraction images were processed, integrated and scaled using *HKL*-2000 (Otwinowski & Minor, 1997). The data-collection statistics are summarized in Table 1.

2.3. Structure determination and refinement

The structure of closed ligand-free PepS was solved by molecular replacement using *EPMR* (Kissinger *et al.*, 1999), with the *S. aureus* AmpS structure (PDB entry 1zjc; Odintsov *et al.*, 2005*a*) as a search model. Interactive model building was performed with *O* (Jones *et al.*, 1991) and *Coot* (Emsley & Cowtan, 2004). Structural models were iteratively refined with *REFMAC* (Murshudov *et al.*, 2011) and *CNS* (Brünger *et al.*, 1998). The crystals of the $PepS^{E343D}$–Trp-Gly complex were

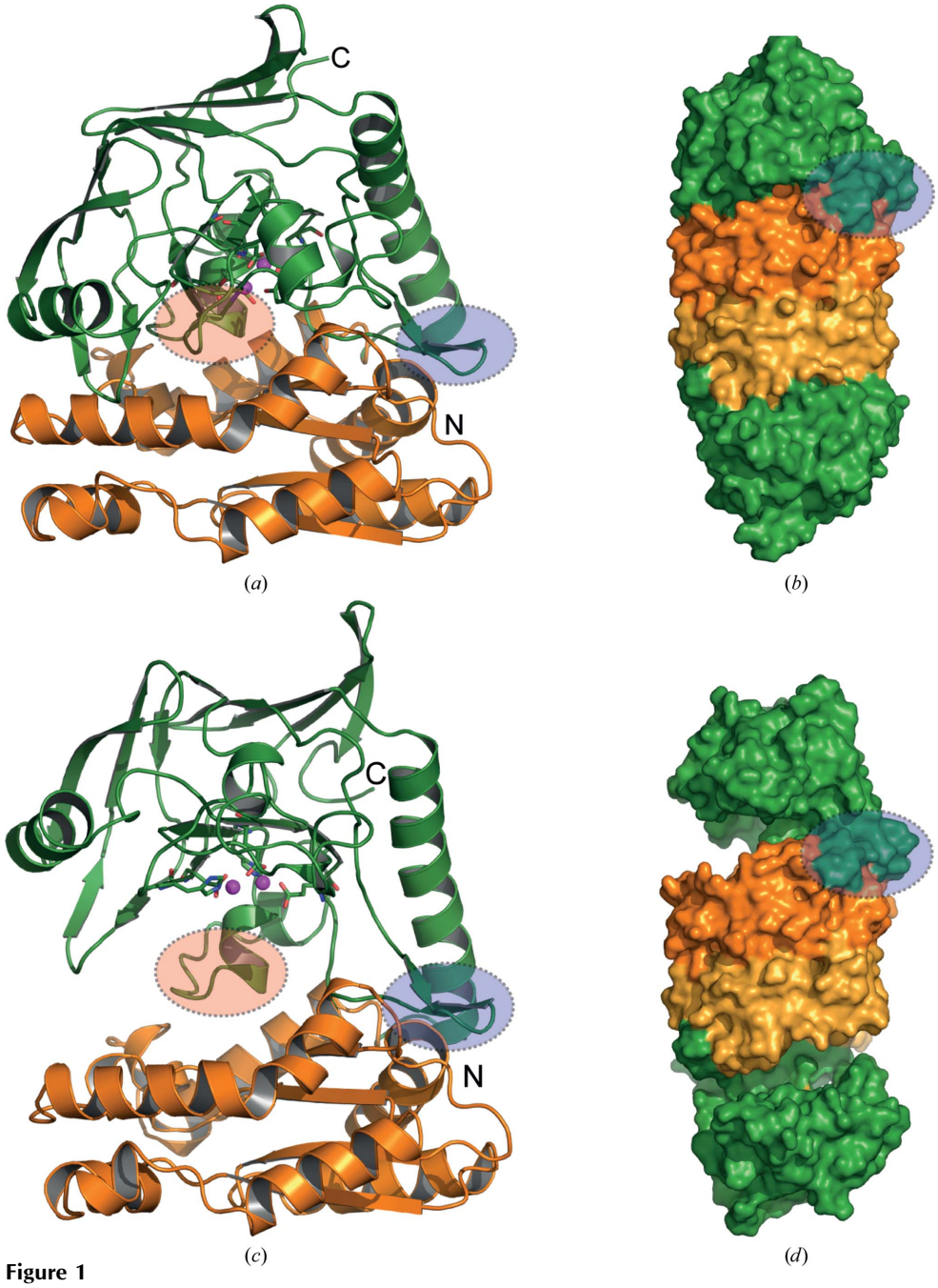

Figure 1
Structures of the two conformations of PepS. (*a*) Crystal structure of PepS in the closed conformation. The N-terminal domain (residues 1–182) is shown in orange and the C-terminal domain (residues 183–413) is shown in green. The hinge (residues 235–252) and Glu343 (residues 322–347) regions are indicated by blue and pink ovals, respectively. The N- and C-termini are indicated. (*b*) Surface representation of the dimeric arrangement of PepS in the closed conformation. The hinge regions are indicated by the blue oval. The colour scheme is the same as that in (*a*) but the second subunit is lightly coloured to distinguish it from the first. (*c*) Crystal structure of $PepS^{E343D}$ in the open conformation. (*d*) Surface representation of the dimeric arrangement of PepS in the open conformation.

found to be isomorphous to the closed native PepS crystals, and thus the bound substrate was introduced into the $F_o - F_c$ difference electron density and was refined together with PepS using *REFMAC* (Murshudov *et al.*, 2011). The N- and C-terminal domains of open PepS were located with *MOLREP* (Vagin & Teplyakov, 2010) using the N- and C-terminal domains of closed ligand-free PepS as separate search models. Inspection of the $2F_o - F_c$ electron-density map indicated that two loop regions, the Glu343 region (residues 322–347) and the hinge region (residues 243–255), needed to be rebuilt. Automatic model building of the Glu343 region and the hinge region was performed with *ARP*/*wARP* (Morris *et al.*, 2003). The final structural models were validated with *PROCHECK* (Laskowski *et al.*, 1993). Refinement statistics are given in

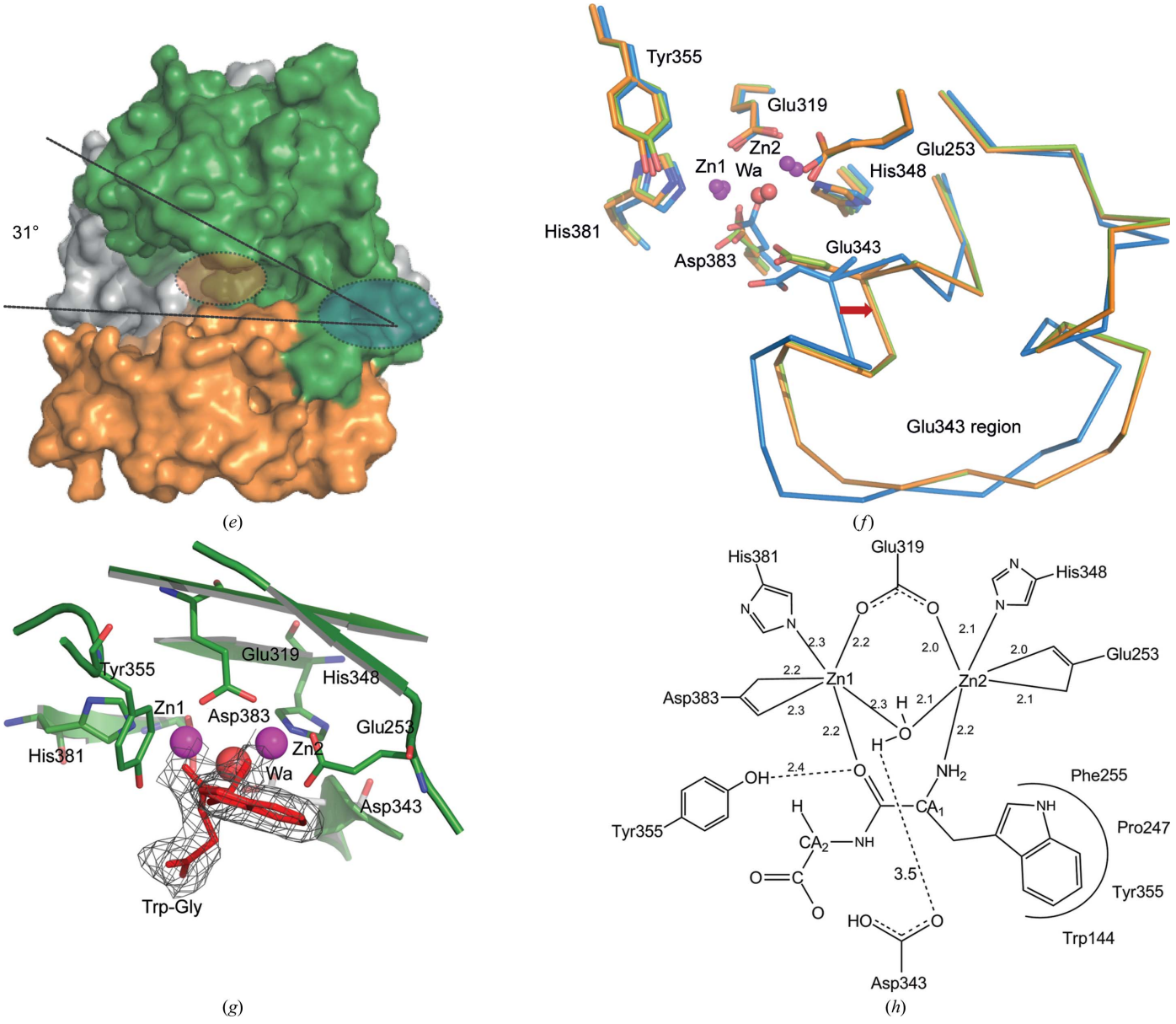

Figure 1 (continued)
(*e*) Superimposition of the closed and open conformations using the N-terminal domain as a reference. The N-terminal domains are indicated in orange (closed) and light orange (open) and the C-terminal domains are coloured grey (closed) and green (open). Domain movement is indicated by the angle between the two lines connecting C^{α} atoms in the hinge region (residue 244) and helix α11 (residue 308). (*f*) Superimposition of the active sites and the Glu343 regions of the three crystal structures (closed, open and substrate-bound). The active-site residues are presented as stick models and the Zn atoms and the bridging water molecules are presented as spheres. C^{α} traces of the Glu343 region from the five crystal structures are also shown. The colouring scheme is as follows: ligand-free closed form, green; open form, blue; substrate-bound form, orange; Zn atoms, magenta; bridging water molecules, orange. Note that Glu343 is replaced by Asp in the open and substrate-bound forms ($PepS^{E343D}$). The red arrow indicates dislocation of the C^{α} atom of Glu343 from the open form to the closed form (2.6 Å). (*g*) $PepS^{E343D}$ with the substrate Trp-Gly (red sticks) bound to the active site in the closed conformation. The $2F_o - F_c$ map is shown in grey, two zinc ions are shown in magenta and the bridging water labelled 'Wa' is shown in red. (*h*) Active-site configuration of the $PepS^{E343D}$–Trp-Gly complex in the closed conformation. Active-site residues, metal ions and the bridging water are labelled. Numbers indicate hydrogen-bonding distances (in Å). The tryptophan side chain of the substrate is specifically recognized by the S1 recognition pocket, which is depicted as an open circle and is formed by a series of hydrophobic residues including Phe255, Pro247, Tyr355 and Trp144. See also Table 1 and Supplementary Fig. S1.

Table 1. All structure figures were generated using *PyMOL* (DeLano, 2002).

2.4. Activity assay

The peptidase activity of PepS was monitored spectroscopically on a GeneQuant spectrophotometer (GE Healthcare). All experiments were performed in reaction buffer (20 m*M* Tris–HCl pH 8.0) using purified PepS and substrate peptides, which were purchased from Sigma and BioFD&C, respectively. The activity assay was performed based on a ninhydrin reaction, which was modified from a published method (Rosen, 1957). Briefly, 10 µl 15 m*M* peptide was mixed with 80 µl reaction buffer and the mixture was incubated at 310 K for 5 min with gentle mixing for full dissolution of the peptides. After incubating the mixture with 10 µl 2 µ*M* purified PepS for 0, 10, 20 and 30 min at 310 K, 200 µl 2% ninhydrin solution (Sigma) and 500 µl reaction buffer were added to the mixture. The resulting solution was boiled for 5 min and the absorbance at 570 nm was measured. The reaction buffer was used as a blank for spectrophotometry. For kinetic analysis, 0.02 µ*M* purified PepS was mixed with varying concentrations of substrates of different lengths (WG, WG3, WG7 and WG10). Curve fitting was performed using *Prism* (GraphPad Software Inc.).

2.5. Single-molecule FRET experiments

For dye labelling of PepS, we engineered a triple mutant PepS$^{C180S/C320S/S366C}$, in which the native cysteines (Cys180 and Cys320) and Ser336 were replaced by serines and a cysteine, respectively, and incubated the mutant with Maleimide Mono-Reactive Cy3 and Cy5 dyes (GE Healthcare) in a 1:5 ratio (300 µ*M*) for 2 h at room temperature. To remove free dyes after labelling, we used size-exclusion chromatography with a 2 ml Sephacryl (Sigma–Aldrich) bead column. Although a cysteine in each monomer was stochastically labelled with either Cy3 or Cy5, single-molecule measurements allowed us to identify PepS dimers which were labelled with one donor and one acceptor. Structural analysis suggested that a fluorophore attached to Ser366 would manifest distance changes specific to the two conformations (72 and 54 Å between C^{α} atoms in the open and closed forms of the PepS dimer, respectively). His-tagged PepS$^{C180S/C320S/S366C}$ was immobilized on the quartz surface *via* streptavidin–biotin, which were coupled with X-nitrilotriacetic acid (Biotin-X-NTA) to remove nonspecific binding (Supplementary Fig. S2*a*[1]). To prevent the nonspecific binding of enzymes, cleaned quartz microscope slides (Finkenbeiner) and cover slips (VWR) were coated with polyethylene glycol (mPEG 5000; Laysan Bio Inc.) and biotinylated PEG (biotin-PEG 5000; Laysan Bio Inc.) in a 40:1 ratio. A narrow channel was made between the cleaned slide and the cover slip using double-sided adhesive tape. After treating the quartz surface with 0.2 mg ml^{-1} streptavidin (Invitrogen) for 5 min and 0.5 mg ml^{-1} Biotin-X-NTA (Anaspec) for 20 min, dye-labelled His-PepS$^{C180S/C320S/S366C}$ (1.2 n*M*) was introduced into the channel. After 10 min incubation and removal of floating proteins, single-molecule fluorescence images were obtained in a wide-field total-internal-reflection fluorescence microscope with 100 ms time resolution using an electron-multiplying charge-coupled device camera (iXon DV887ECS-BV, Andor Technology) and a homemade program written in Visual C++ (Bae *et al.*, 2011). Measurements were performed in imaging buffer containing 25 m*M* Tris–HCl pH 8.0 and an oxygen-scavenger system to slow photobleaching (Roy *et al.*, 2008). The ALEX technique was used to determine whether the enzyme was labelled with Cy3 and Cy5 (Lee *et al.*, 2010), and only molecules labelled with both dyes were used for data analysis. All experiments were performed at 298 K.

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: MH5086). Services for accessing this material are described at the back of the journal.

3. Results

3.1. Structure of PepS in two conformations and active sites

Initially, we determined the crystal structure of PepS, which consists of an N-terminal dimerization domain and a C-terminal catalytic domain (Fig. 1*a* and Table 1). The PepS dimer has an elongated shape (Fig. 1*b*) with an extensive interface (Supplementary Fig. S1*a*), similar to AmpS (Odintsov *et al.*, 2005*a*) and AmpT (Odintsov *et al.*, 2005*b*) in the M29 family, in which the sequence is relatively well conserved (Supplementary Fig. S1*b*). The metal-binding site of PepS reveals that Glu343 is located close (2.5 Å) to a water molecule that coordinates two zinc ions, suggesting that Glu343 acts as a general base by extracting protons from the bridging water (Supplementary Figs. S1*c* and S1*d*). The PepS^{E343D} mutant exhibits little catalytic activity (Supplementary Fig. S1*e*), confirming the role of Glu343 as a general base, as well as the structural significance of the region surrounding Glu343 in achieving full catalytic activity. The crystal structure of mutant PepS^{E343D} (Figs. 1*c* and 1*d* and Table 1) shows that the C-terminal domain swings 31° away from the wild-type (WT) structure, centred on the 'hinge' region (Fig. 1*e*). This suggests that PepS can exist in at least two distinct conformations. We designated the WT structure the 'closed' form and the PepS^{E343D} structure the 'open' form. Domain movement of aminopeptidase has previously been proposed on observing five molecules with a range of conformations in the asymmetric unit of the AmpT crystal (Odintsov *et al.*, 2005*b*). However, the low-resolution crystal structure of AmpT (3.7 Å) prevented a more detailed structural characterization of the domain movement and its functional implications. The domain movement of PepS does not cause any major conformational change within each domain (Supplementary Fig. S1*f*). However, local conformational changes are observed in the hinge region and a loop surrounding Glu343 in the active site (the 'Glu343 region'; Fig. 1*f* and Supplementary S1*f*). The Glu343 region dislocates upon transition from the closed to the open conformation, with the following characteristics: (i) a noticeable change in the backbone conformation of the Glu343 region between the

open and the closed forms, with a C^{α} shift of 2.6 Å in Glu343, and (ii) no significant change in the residues and the bridging water molecule coordinating the Zn atoms during the conformational change. Although it has previously been proposed that the closed form represents the active state of M29-family enzymes (Odintsov *et al.*, 2005*a*), in this study the repositioning of the Glu343 region in the active site was first identified structurally (Fig. 1*f*) and subsequently supported by functional assays (Supplementary Fig. S1*e*). Based on the open and closed structures, we hypothesized that substrate accessibility can be conferred by conversion from the closed to the open conformation and that catalytic activity is ensured by the position of the Glu343 region in the closed conformation.

To further explore the positioning of the Glu343 region in the active site of PepS, we determined the crystal structure of $PepS^{E343D}$ with substrate (Trp-Gly) in the closed conformation (Fig. 1*g* and Table 1). The structure of the $PepS^{E343D}$–Trp-Gly complex provided a detailed structural snapshot of substrate binding in the closed conformation. Although the mutation of Glu343 to Asp causes a shortening of the side chain of residue 343 by one methylene group, Asp343 still points to the bridging water at a distance of 3.1 Å, supporting the hypothesis that Glu343 functions as a general base in the closed conformation (Fig. 1*h*). However, the dislocation of Glu343 away from the active site in the open conformation prevents Glu343 from activating the bridging water molecule. Interestingly, the $PepS^{E343D}$–Trp-Gly complex structure contains a substrate and a nucleophilic water at the same time, representing a snapshot before hydrolysis. In the closed structure, five residues from the C-terminal domain (Ser234, Ala235, Pro247, Phe255 and Tyr355) and one residue from the N-terminal domain (Trp144) constitute the S1 binding pocket by encompassing the P1 residue of the substrate (Supplementary Figs. S1*g* and S1*h*). Consequently, it is expected that the open structure cannot form a complete S1 binding pocket. These results clearly indicate that the closed conformation represents the active state of PepS and that the open conformation reflects the inactive state. Collectively, structural analysis of the open and closed structures of PepS implies that substrate access to the immature active site is allowed in the open conformation, but that substrate hydrolysis is only catalyzed in the mature active site, which is only completed in the closed conformation.

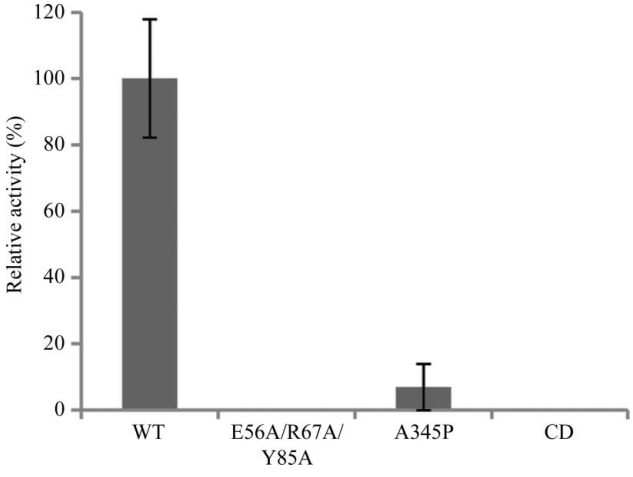

Figure 2
Catalytic activities of wild-type PepS (WT) and PepS mutants against the substrate WG. The activities of the mutants are indicated relative to WT, which was set as 100%. CD represents the C-terminal domain truncation mutant. E56A/R78A/Y85A indicates a monomeric mutant, $PepS^{E56A/R78A/Y85A}$, in which three key residues in the dimeric interface are mutated to Ala. CD and E56A/R78A/Y85A showed no detectable catalytic activity.

Table 2
Kinetic parameters.

Representative data are presented here from multiple data sets. n.d., not determined.

Substrate	WG	WG3	WG7	WG10
V_{max}† ($\mu M\ s^{-1}$)	2.286	2.470	2.142	n.d.
h‡	1.576	1.560	1.896	n.d.
K_H‡ (mM)	1.298	1.113	0.588	n.d.
R^2§	0.9967	0.9946	0.9949	n.d.

† The reaction rate V is defined as the change in the concentration of Ruhemann's purple, which is produced by ninhydrin and free tryptophan, for the first 10 min of the reaction. ‡ Data were fitted according to an allosteric sigmoidal model with the Hill equation $v = V_{max}[S]^h/(K_H + [S]^h)$, where h is the Hill coefficient and K_H is the half-maximal activity constant, which is conceptually equivalent to K_m in Michaelis–Menten kinetics (Sauro, 2012). § Quantitative indication of the goodness of fit, as calculated by the equation $R^2 = 1.0 - (wSS_{model}/wSS_{horizontal})$, where wSS_{model} is the weighted sum-of-squares of the residuals from the model and $wSS_{horizontal}$ is the weighted sum-of-squares of the residuals from the null-hypothesis model (GraphPad Tutorial; http://www.graphpad.com/guides/prism/6/curve-fitting/index.htm). Calculation of the kinetic parameters was performed using *Prism* (GraphPad Software Inc.)

3.2. Functional studies

Structural investigations of catalytic activity were complemented by functional assays of wild-type and mutant PepS (Fig. 2). A monomeric mutant $PepS^{E56A/R78A/Y85A}$ was made by mutating three key residues in the dimeric interface (Glu56, Arg78 and Tyr85) and its oligomerization state was confirmed by size-exclusion chromatography. In addition, Ala345, a key residue in the hinge region, was mutated to proline to restrict the conformational conversion between the open and the closed forms. Catalytic activity was deteriorated in the monomeric mutant $PepS^{E56A/R78A/Y85A}$ and the hinge mutant $PepS^{A345P}$ (Fig. 2), reinforcing the notion that the dimeric arrangement and conversion between the open and closed forms are essential for function. Moreover, the catalytic activity of the C-terminal domain only mutant ('CD') is completely abolished, which is consistent with the structural findings that both the N-terminal and the C-terminal domains are required for full activity.

3.3. Structural basis of substrate-length selectivity

The structural basis of the substrate-length selectivity of PepS was clearly revealed in this study. PepS has a unique substrate-binding site consisting of a long substrate-binding hole in the N-terminal domain and an S1 binding pocket in the C-terminal domain (Fig. 3*a*). The substrate-binding hole is 20 Å in length and 10 Å in diameter on average, with the widest part in the middle being about 20 Å. A hole of these dimensions is capable of accommodating substrates of up to eight residues, assuming an extended conformation, which is

consistent with a previous biochemical report (Fernandez-Espla & Rul, 1999). A model substrate with the sequence WGAAAAAA fits nicely into the substrate-binding site, with the first residue properly located in the S1 binding pocket and the remaining residues in the substrate-binding hole (Fig. 3*b*). To investigate the length selectivity of PepS from another aspect, we performed an activity assay with substrate peptides of various lengths: a dipeptide (amino-acid sequence WG), a tetrapeptide (WG_3), an octapeptide (WG_7) and an undecapeptide (WG_{10}) (Fig. 3*c*). Substrates up to octapeptides are cleaved by PepS to some extent, but cleavage of WG10 by PepS is mostly blocked, which is consistent with the size limit imposed by the substrate-binding hole. Kinetic studies have revealed that PepS exhibits a sigmoidal kinetic behaviour, implying that allostery may be involved in its reaction mechanism (Fig. 3*c* and Table 2). The kinetic data were successfully fitted to the Hill equation (Copeland, 2000), with Hill coefficient values of 2 being determined for WG, WG3 and WG7, reflecting the dimeric arrangement. K_H, the half-maximal activity constant, which is analogous to K_m in the Hill equation (Sauro, 2012), appears to decrease in the order WG, WG3, WG7 (Table 2). Such a systematic decrease in K_H values with increasing length of substrate peptides supports the idea that the longer substrates within the size limit can fit better into the substrate-binding hole. Therefore, the substrate-binding hole present in the N-terminal domain functions as a length filter, largely (if not entirely) limiting the length of the C-terminal side of the substrate to seven residues owing to structural constraints. In the dimeric PepS arrangement the substrate-binding holes from the two monomers are spatially connected. However, it is highly unlikely that a substrate spans both monomers because the substrate would have to bend abruptly. Therefore, dimerization seems to be necessary for catalytic activity but is not required for substrate filtering (Fig. 2).

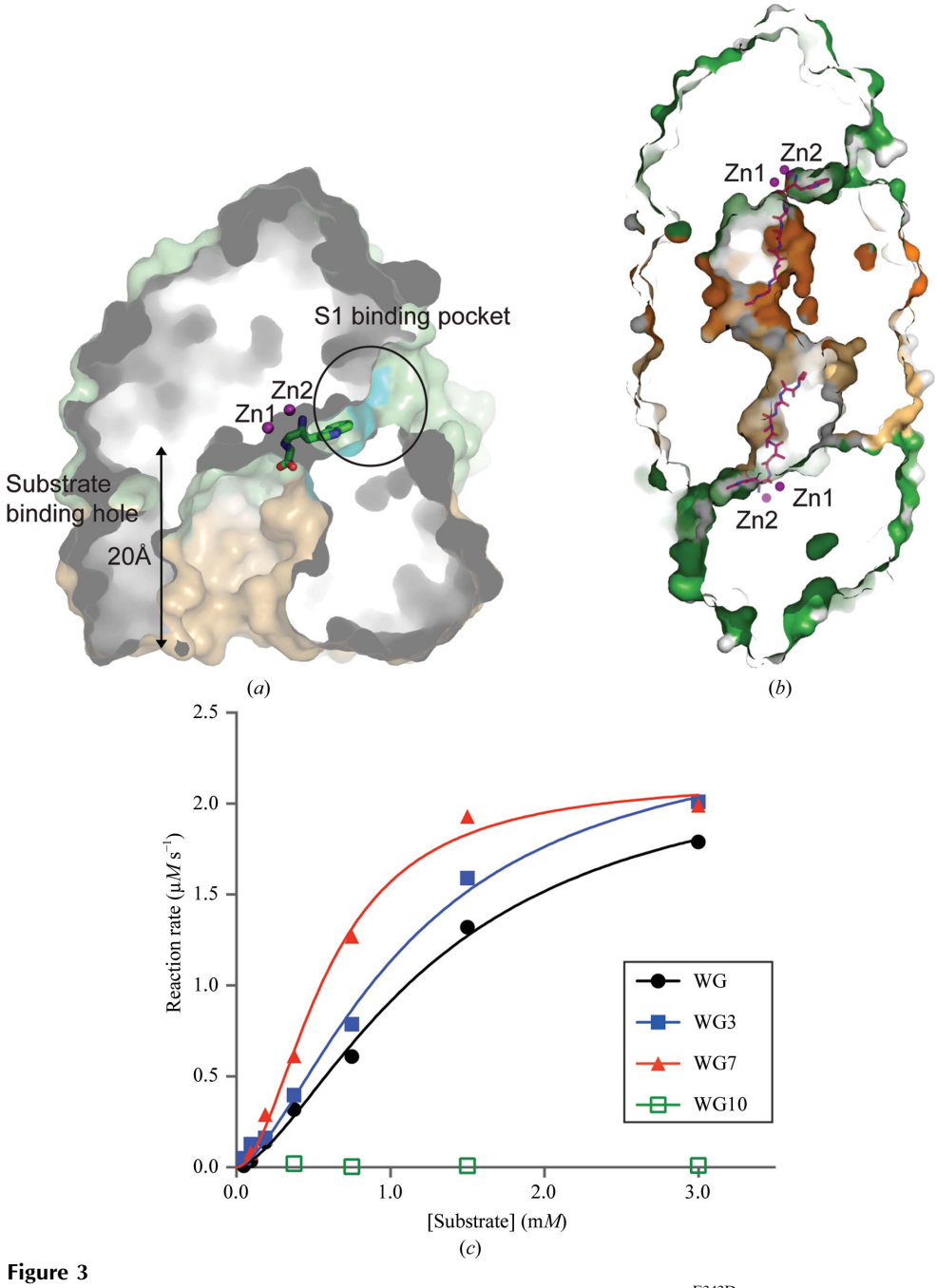

Figure 3
Substrate recognition of PepS, as revealed by the crystal structure of a PepS^{E343D}–Trp-Gly complex. (*a*) Surface representation of the substrate-binding hole and the S1 binding pocket of PepS. The two zinc ions are shown as magenta spheres and the substrate Trp-Gly is depicted as a green stick model. Residues constituting the S1 binding pocket are indicated by a black circle and are coloured cyan. The distance shown indicates the length of the substrate-binding hole from Zn1 to the bottom of the hole. (*b*) Modelling of an octapeptide (WGAAAAAA) in the substrate-binding site of the PepS dimer. (*c*) Kinetic analysis of PepS at 0.02 μ*M* with substrates of various lengths. Data were fitted using a sigmoidal fit routine in *Prism* (GraphPad Software Inc.) based on the Hill equation (Copeland, 2000). See also Table 2.

The crystal structure of the PepS^{E343D}–Trp-Gly complex also provides a molecular basis for the residue selectivity of PepS, with the S1 binding pocket preferring aromatic residues and arginine in the P1 position (Fernandez-Espla & Rul, 1999). The tryptophan side chain of the substrate Trp-Gly is well suited to van der Waals interactions with the

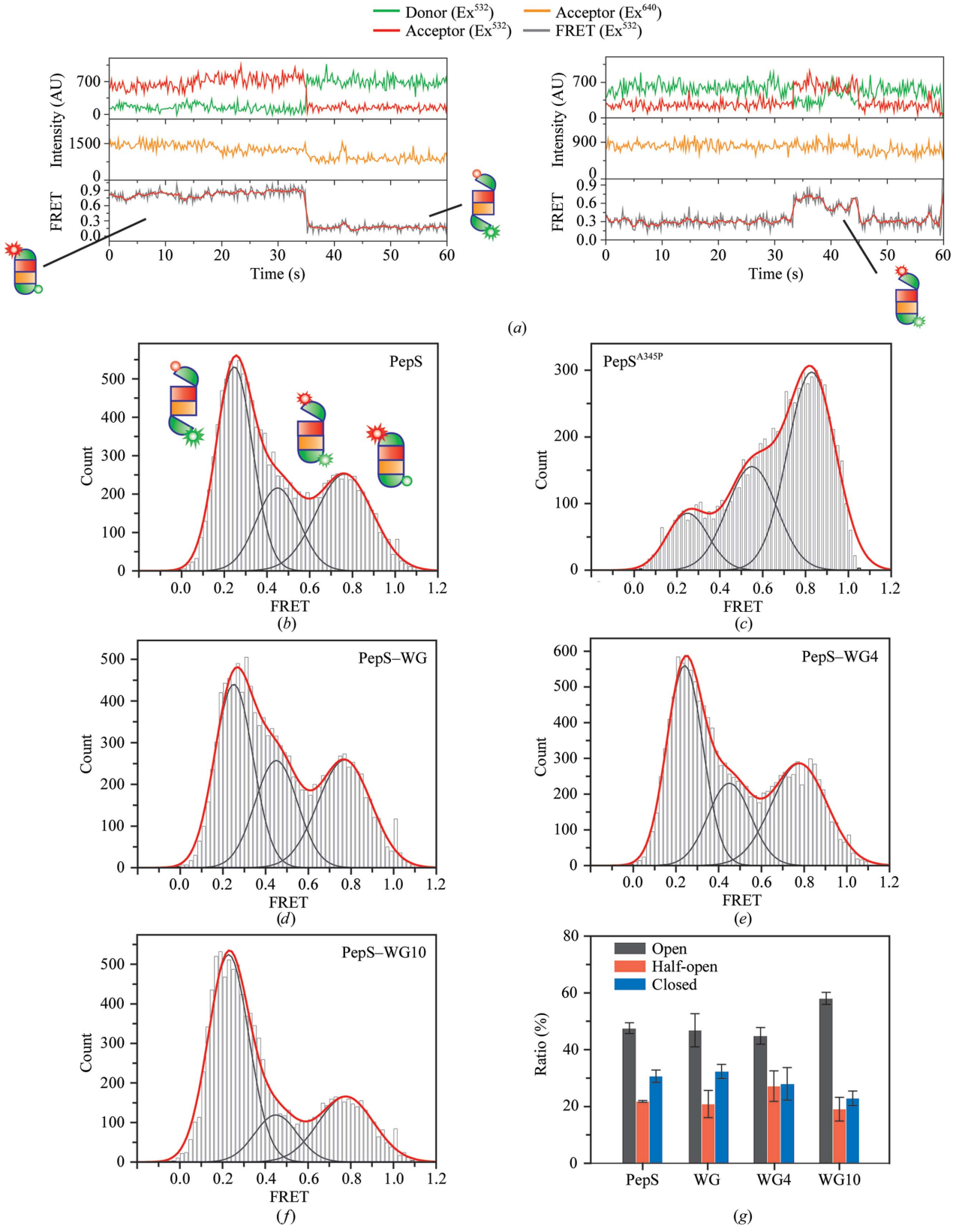

Figure 4

Single-molecule studies of PepS. (*a*) Representative time traces of dual-labelled PepS (green, Cy3 intensity at green excitation; red, Cy5 intensity at green excitation; orange, Cy5 intensity at red excitation; grey with red, FRET efficiency). Direct observation of three distinct conformations (open/open, closed/open and closed/closed) by single-molecule FRET. (*b–f*) FRET histograms of wild-type PepS (*b*), PepS^{A345P} (*c*) and PepS incubated with WG (*d*), WG4 (*e*) or WG10 (*f*). PepS was incubated with the substrates (20 µ*M*) for 1 h before data acquisition. All histograms are well fitted by the sum of three Gaussian functions, with FRET peaks at 0.23 (open/open), 0.45 (open/closed) and 0.75 (closed/closed). (*g*) The population distribution of the three species under different incubation conditions. The error bars from three separate measurements are indicated. See also Supplementary Fig. S2.

hydrophobic residues (Supplementary Fig. S1*g*). The potential rotameric arrangement of Glu233 suggests that the N^{ε} atom of arginine could electrostatically interact with Glu233, whereas lysine, which has a shorter side chain than arginine, would hinder the formation of such an interaction, thus accounting for the P1 specificity (Fernandez-Espla & Rul, 1999). Taken together, these data show that the substrate-binding site is composed of a length filter and that the S1 binding pocket can accommodate substrates with both length selectivity and residue specificity.

To probe the substrate-length selectivity of PepS in solution, single-molecule FRET experiments were performed using PepS with two fluorophores (Cy3 and Cy5) in each subunit (Supplementary Fig. S2). PepS exists in at least three distinct conformations with FRET values of 0.23, 0.45 and 0.75, corresponding to distances between the two fluorophores of 71, 60 and 48 Å, respectively (Fig. 4*a*). These distance values, which were derived from FRET experiments, are in excellent agreement with those predicted from structural analysis: the conformations with FRET values of 0.23 and 0.75 are consistent with the open and closed forms, which were observed in the crystal structures, while the third conformation with a FRET value of 0.45 corresponds to the 'half-open' form in which only one subunit of the dimer assumes the open conformation (Fig. 4*a*). Histogram analysis reveals that the open form accounts for 45% of the total protein, followed by the closed form (33%) and the half-open form (22%), suggesting that PepS favours the open form in solution (Fig. 4*b*). Interestingly, this preference is reversed in the $PepS^{A345P}$ mutant, in which conformational conversion is restricted and the equilibrium is shifted towards the closed form (Fig. 4*c*), which is consistent with the reduced catalytic activity of $PepS^{A345P}$ (Fig. 2). Next, the distribution of the three conformations of PepS was monitored upon the addition of substrates of different lengths (Figs. 4*d*–4*f*). Two substrates (WG and WG4) which were cleaved in the functional assay caused little change in the distribution (Figs. 4*d*, 4*e* and 4*g*), indicating that short substrates can be cleaved and released when considering longer time scales for the FRET measurement (200 ms) than enzymatic turnover. In contrast, WG10 drives the population of PepS conformations towards the open form (45% for PepS only *versus* 60% with WG10; Figs. 4*f* and 4*g*). It appears that a substrate longer than the size of the substrate-binding hole traps the protein in the open form, thereby preventing the transition from the open to the closed form, which is a prerequisite for turning on the catalytic activity of PepS.

4. Discussion

Based on crystallographic data, functional assays and single-molecule data, we propose a catalytic mechanism for PepS termed the 'triple-sieved interlock mechanism', featuring coupling of active-site positioning, residue specificity and length selectivity to regulate the catalytic activity (Fig. 5; for details, refer to Supplementary Fig. S3). 'Clogging' at any one of these three sieves leads to failed substrate hydrolysis. Substrate binding, which is allowed by passing the residue and length filters, as well as proper positioning of the Glu343 region in the active site, ensures catalytic cleavage. Therefore, peptides longer than eight residues can bind to the hole but cannot be cleaved, owing to incomplete formation of the active site. Length-dependent activation of catalytic activity has been studied using the aminopeptidase ERAP1 (Kochan *et al.*, 2011; Nguyen *et al.*, 2011). In this case, it was proposed that the binding of longer peptides to the regulatory site close to the catalytic site stabilizes the closed conformation, leading to increased enzymatic activity (Kochan *et al.*, 2011; Nguyen *et al.*, 2011). Therefore, substrate length is the limiting factor for the regulation of catalytic activity of ERAP1. However, the substrate-length selectivity observed in PepS differs from that in ERAP1 in that the length filter of PepS works directly with the active-site positioning in the absence of a regulatory site. Moreover, while the regulatory site of ERAP1 is not clearly identified and the positioning of the active site is not defined owing to a lack of structural information, the three sieves in PepS are well characterized by their crystal structures, and their

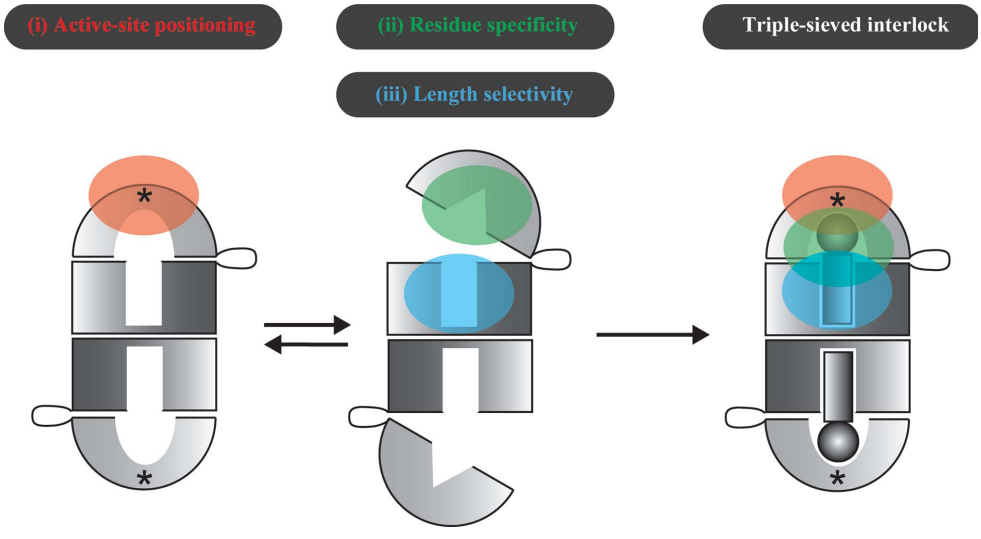

Figure 5
A triple-sieved interlock mechanism for the regulation of PepS catalytic activity. The triple sieve comprises (i) active-site positioning, in which domain movement and local conformational change including the Glu343 region allow the complete formation of the active site, (ii) residue specificity, conferred by the S1 binding pocket, and (iii) length selectivity, conferred by the substrate-binding hole. Catalysis occurs only when a substrate passes all three sieves. The asterisk indicates the activated catalytic site. The N-terminal domain is coloured dark grey, the C-terminal domain light grey and the substrate black. See also Supplementary Fig. S3.

interlocking mechanism is supported by the single-molecule studies. In bacteria, catabolic processing of peptides is performed by a hierarchical ordering of various peptidases, and substrate selection using a length filter is essential in order to discriminate each step in the protein-degradation hierarchy (Dougan *et al.*, 2002; Gonzales & Robert-Baudouy, 1996). Consequently, failure in length selection would lead to an interruption of the catabolic pathways in bacteria, thereby disrupting the fundamental physiology. Indeed, deletion of the *pepS* gene has been reported to inhibit bacterial cell growth and to alter lysozyme resistance (Ji *et al.*, 2001; Thomas *et al.*, 2010). The triple-sieved interlock mechanism proposed for PepS would ensure that bacterial protein degradation occurs in an orderly and proper fashion, thus preserving physiological function.

In summary, structural analysis of PepS reveals that the combined substrate-binding hole and pocket inside PepS imposes restrictions on how a substrate is selected based on its length and P1 residues. Structural interpretation of length selectivity is well supported by functional assays with various mutants and substrates, as well as by single-molecule FRET analysis of PepS in solution. Furthermore, our study revealed that substrate selectivity, in concert with the active-site positioning of the Glu343 region, plays a crucial role in regulating the catalytic activity of PepS. The triple-sieved interlock system of PepS based on active-site positioning, residue specificity and length selectivity is necessary to ensure the cleavage of substrates with the correct length and residues, and thus is likely to be present in other peptidases that exhibit substrate-length selectivity. The crystal structures of PepS combined with functional assays and single-molecule FRET studies provide new insights into how substrate binding and cleavage by aminopeptidases is controlled, thus expanding the repertoire of peptidase regulatory mechanisms.

This work was supported by the Basic Science Research Program (2011-0023402) and the Pioneer Research Center Program (NRF-2012-0009565) through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning and a grant from the Woo Jang Chun Special Project (PJ009106) funded by the Rural Development Administration, Republic of Korea (SL), the Ubiquitome Research Program (M10533010002-06N3301-00210), a National Research Foundation of Korea (NRF) grant (2011-0028878) and the Next-Generation BioGreen21 Program (SSAC PJ008107) (KKK) and the Creative Research Initiatives (Physical Genetics Laboratory, 2009-0081562) (SH).

References

Bae, S., Kim, D., Kim, K. K., Kim, Y.-G. & Hohng, S. (2011). *J. Am. Chem. Soc.* **133**, 668–671.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Copeland, R. A. (2000). *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis.* New York: Wiley-VCH.

DeLano, W. L. (2002). *PyMOL*. http://www.pymol.org.

De Mot, R., Nagy, I., Walz, J. & Baumeister, W. (1999). *Trends Microbiol.* **7**, 88–92.

Dougan, D. A., Mogk, A. & Bukau, B. (2002). *Cell. Mol. Life Sci.* **59**, 1607–1616.

Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.

Fernandez-Espla, M. D. & Rul, F. (1999). *Eur. J. Biochem.* **263**, 502–510.

Goldberg, A. L. & Dice, J. F. (1974). *Annu. Rev. Biochem.* **43**, 835–869.

Gonzales, T. & Robert-Baudouy, J. (1996). *FEMS Microbiol. Rev.* **18**, 319–344.

Ji, Y., Zhang, B., Van Horn, S. F., Warren, P., Woodnutt, G., Burnham, M. K. R. & Rosenberg, M. (2001). *Science*, **293**, 2266–2269.

Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kim, D. Y. & Kim, K. K. (2008). *J. Mol. Biol.* **379**, 760–771.

Kim, D. Y., Kim, D. R., Ha, S. C., Lokanath, N. K., Lee, C. J., Hwang, H.-Y. & Kim, K. K. (2003). *J. Biol. Chem.* **278**, 6543–6551.

Kim, D. Y., San, B. H., Moh, S. H., Park, H., Kim, D. Y., Lee, S. & Kim, K. K. (2010). *Biochem. Biophys. Res. Commun.* **391**, 431–436.

Kisselev, A. F., Akopian, T. N. & Goldberg, A. L. (1998). *J. Biol. Chem.* **273**, 1982–1989.

Kissinger, C. R., Gehlhaar, D. K. & Fogel, D. B. (1999). *Acta Cryst.* D**55**, 484–491.

Kochan, G., Krojer, T., Harvey, D., Fischer, R., Chen, L., Vollmar, M., von Delft, F., Kavanagh, K. L., Brown, M. A., Bowness, P., Wordsworth, P., Kessler, B. M. & Oppermann, U. (2011). *Proc. Natl Acad. Sci. USA*, **108**, 7745–7750.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Lee, S., Lee, J. & Hohng, S. (2010). *PLoS One*, **5**, e12270.

Lowther, W. T. & Matthews, B. W. (2002). *Chem. Rev.* **102**, 4581–4608.

Morris, R. J., Perrakis, A. & Lamzin, V. S. (2003). *Methods Enzymol.* **374**, 229–244.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Nguyen, T. T., Chang, S.-C., Evnouchidou, I., York, I. A., Zikos, C., Rock, K. L., Goldberg, A. L., Stratikos, E. & Stern, L. J. (2011). *Nature Struct. Mol. Biol.* **18**, 604–613.

Odintsov, S. G., Sabala, I., Bourenkov, G., Rybin, V. & Bochtler, M. (2005*a*). *J. Biol. Chem.* **280**, 27792–27799.

Odintsov, S. G., Sabala, I., Bourenkov, G., Rybin, V. & Bochtler, M. (2005*b*). *J. Mol. Biol.* **354**, 403–412.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Rawlings, N. D., Morton, F. R., Kok, C. Y., Kong, J. & Barrett, A. J. (2008). *Nucleic Acids Res.* **36**, D320–D325.

Rosen, H. (1957). *Arch. Biochem. Biophys.* **67**, 10–15.

Roy, R., Hohng, S. & Ha, T. (2008). *Nature Methods*, **5**, 507–516.

Saric, T., Graef, C. I. & Goldberg, A. L. (2004). *J. Biol. Chem.* **279**, 46723–46732.

Sauer, R. T., Bolon, D. N., Burton, B. M., Burton, R. E., Flynn, J. M., Grant, R. A., Hersch, G. L., Joshi, S. A., Kenniston, J. A., Levchenko, I., Neher, S. B., Oakes, E. S., Siddiqui, S. M., Wah, D. A. & Baker, T. A. (2004). *Cell*, **119**, 9–18.

Sauro, H. M. (2012). *Enzyme Kinetics for Systems Biology.* Seattle: Ambrosius Publishing.

Tamura, N., Lottspeich, F., Baumeister, W. & Tamura, T. (1998). *Cell*, **95**, 637–648.

Tamura, T., Tamura, N., Cejka, Z., Hegerl, R., Lottspeich, F. & Baumeister, W. (1996). *Science*, **274**, 1385–1389.

Thomas, S., Besset, C., Courtin, P. & Rul, F. (2010). *J. Appl. Microbiol.* **108**, 148–157.

Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.