**SUPPLEMENTARY MATERIAL**

# Structure-based elucidation of the regulatory mechanism for the aminopeptidase activity

Hai Minh Ta[1], Sangsu Bae[2], Seungsu Han[3], Jihyuck Song[1], Sungchul Hohng[2], Sangho Lee[3,*] and Kyeong Kyu Kim[1,*]

[1]Department of Molecular Cell Biology, School of Medicine, Samsung Biomedical Research Institute, Sungkyunkwan University, Suwon 440-746, Korea

[2]Department of Physics, Seoul National University, Seoul 151-742, Korea

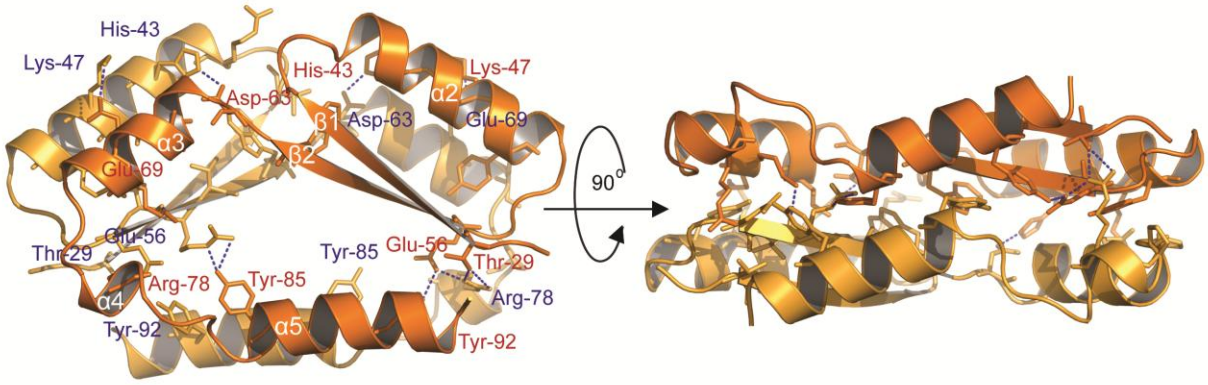
[3]Department of Biological Sciences, Sungkyunkwan University, Suwon 440-746, Korea

Correspondence should be addressed to: S.L. (sangholee@skku.edu) or K.K.K. (kkim@med.skku.ac.kr).

## Inventory of Supplementary Material

Figure S1

(*a*)


His-43
Lys-47
His-43
Lys-47
Asp-63
α2
β1
Asp-63
Glu-69
α3
β2
Glu-69
90°
Thr-29
Glu-56
Tyr-85
Glu-56
Thr-29
Arg-78
Tyr-85
α4
Arg-78
Tyr-92
α5
Tyr-92

(*b*)

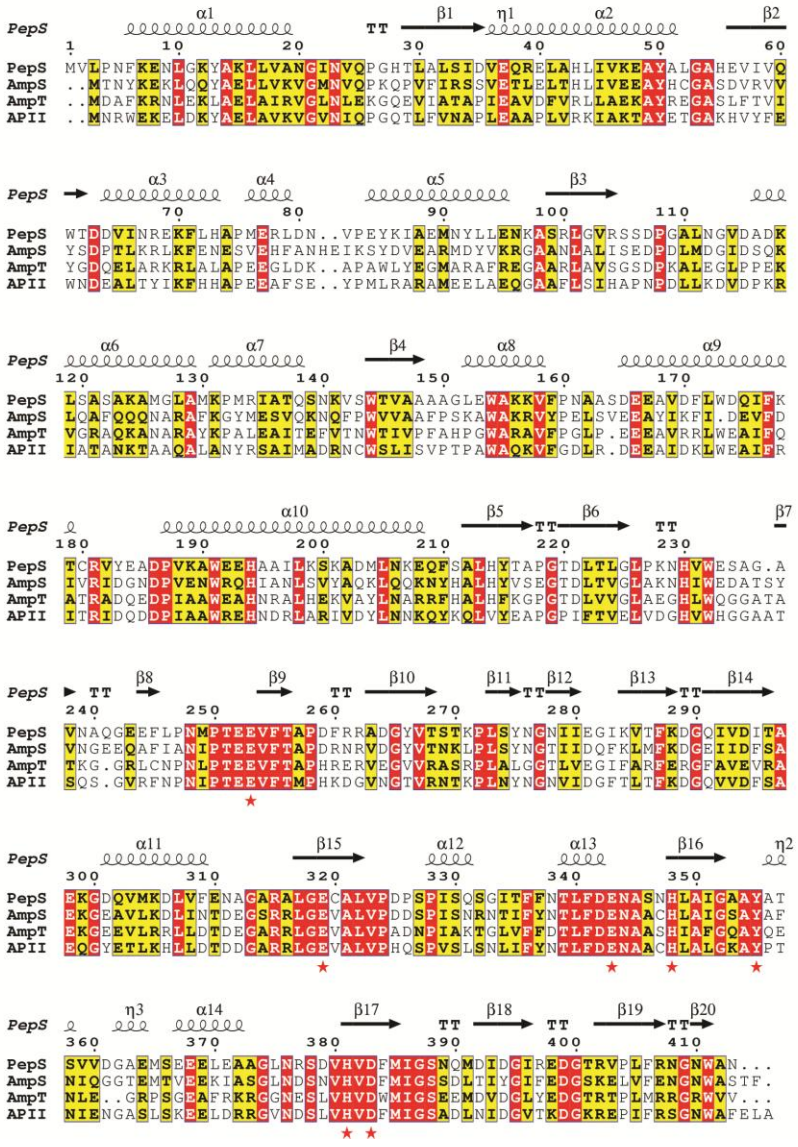
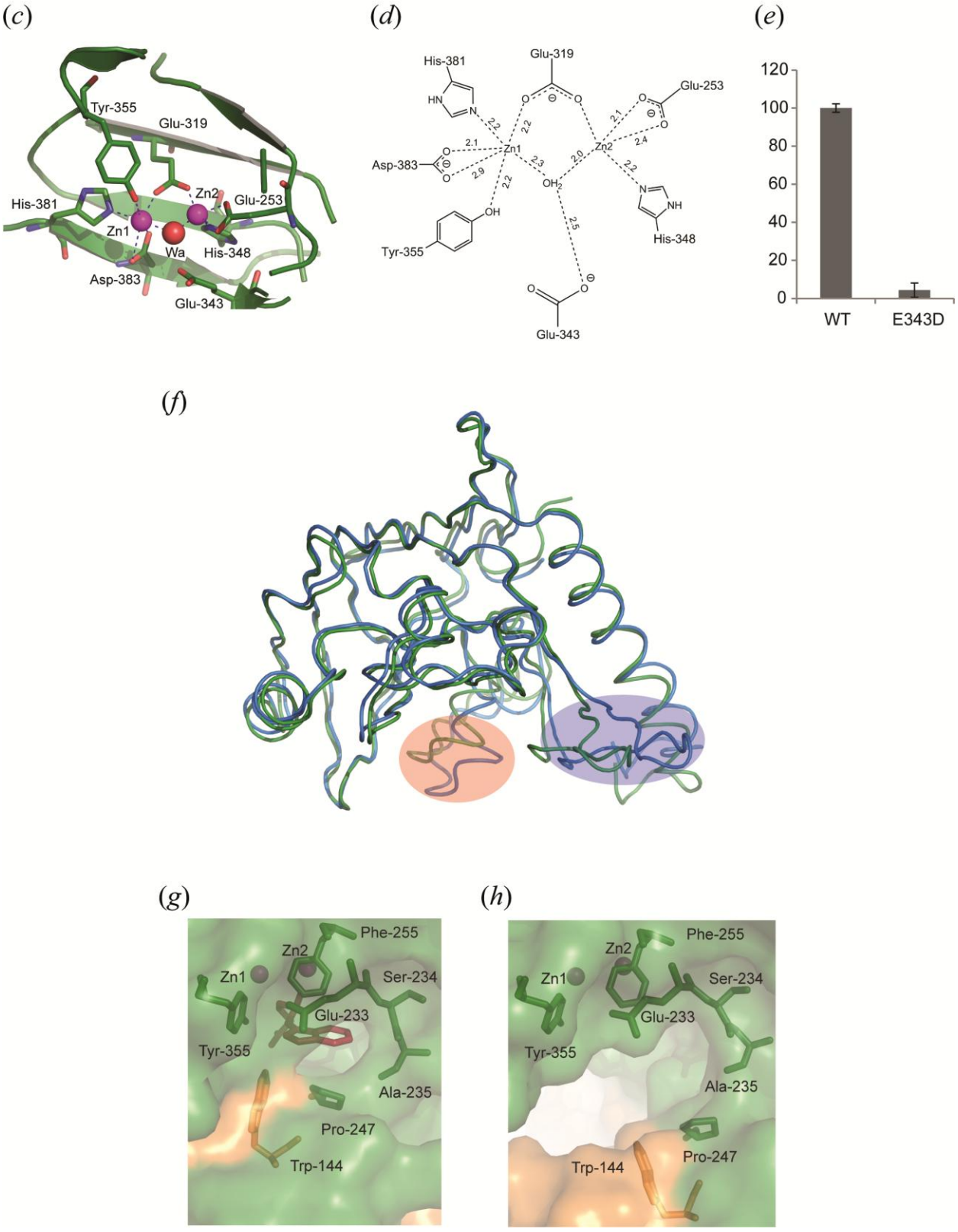

PepS
PepS MVLPNFKENLGKYAKLLVANGINVQPGHTLALSIDVEQRELAHLIVKEAYALGAHEVIVQ
AmpS ..MTNYKEKLQQYAELLVKVGMNVQPKQPVFIRSSVETLELTHLIVEEAYHCGASDVRVV
AmpT ..MDAFKRNLEKLAELAIRVGLNLEKGQEVIATAPIEAVDFVRLLAEKAYREGASLFTVI
APII ..MNRWEKELDKYAELAVKVGVNIQPGQTLFVNAPLEAAPLVRKIAKTAYETGAKHVYFE
PepS WTDDVINREKFLHAPMERLDN..VPEYKIAEMNYLLENKASRLGVRSSDPGALNGVDADK
AmpS YSDPTLKRLKFENESVEHFANHEIKSYDVEARMDYVKRGAANLALISEDPDLMDGIDSQK
AmpT YGDQELARKRLALAPEEGLDK..APAWLYEGMARAFREGAARLAVSGSDPKALEGLPPEK
APII WNDEALTYIKFHHAPEEAFSE..YPMLRARAMEELAEQGAAFLSIHAPNPDLLKDVDPKR
PepS LSASAKAMGLAMKPMRIATQSNKVSWTVAAAAGLEWAKKVFPNAASDEEAVDFLWDQIFK
AmpS LQAFQQQNARAFKGYMESVQKNQFPWVVAAFPSKAWAKRVYPELSVEEAYIKFI.DEVFD
AmpT VGRAQKANARAYKPALEAITEFVTNWTIVPFAHPGWARAVFPGLP.EEEAVRRLWEAIFQ
APII IATANKTAAQALANYRSAIMADRNCWSLISVPTPAWAQKVFGDLR.DEEAIDKLWEAIFR
PepS TCRVYEADPVKAWEEHAAILKSKADMLNKEQFSALHYTAPGTDLTLGLPKNHVWESAG.A
AmpS IVRIDGNDPVENWRQHIANLSVYAQKLQQKNYHALHYVSPGTDLTVGLAKNHIWEDATSY
AmpT ATRADQEDPIAAWEAHNRALHEKVAYLNARRFHALHFKGPGTDLVVGLAEGHLWQGGATA
APII ITRIDQDDPIAAWREHNDRLARIVDYLNNKQYKQLVYEAPGPIFTVELVDGHVWHGGAAT
PepS VNAQGEEFLPNMPTEEVFTAPDFRRADGYVTSTKPLSYNGNIIEGIKVTFKDGQIVDITA
AmpS VNGEEQAFIANIPTEEVFTAPDRNRVDGYVTNKLPLSYNGTIIDQFKLMFKDGEIIDFSA
AmpT TKG.GRLCNPNLPTEEVFTAPHRERVEGVVRASRPLALGGTLVEGIFARFERGFAVEVRA
APII SQS.GVRFNPNIPTEEVFTMPHKDGVNGTVRNTKPLNYNGNVIDGFTLTFKDGQVVDFSA
PepS EKGDQVMKDLVFENAGARALGECALVPDPSPISQSGITFFNTLFDENASNHLAIGAAYAT
AmpS EKGEAVLKDLINTDEGSRRLGEVALVPDDSPISNRNTIFYNTLFDENAACHLAIGSAYAF
AmpT EKGEEVLRRLLDTDEGARRLGEVALVPADNPIAKTGLVFFDTLFDENAASHIAFGQAYQE
APII EQGYETLKHLLDTDDGARRLGEVALVPHQSPVSLSNLIFYNTLFDENAACHLALGKAYPT
PepS SVVDGAEMSEPELEAAGLNRSDVHVDFMIGSNQMDIDGIREDGTRVPLFRNGNWANA...
AmpS NIQGGTEMTVEEKIASGLNDSNVHVDFMIGSSDLTIYGIFEDGSKELVFENGNWASTF.
AmpT NLE..GRPSGEAFRKRGGNESLVHVDWMIGSEEMDVDGLYEDGTRTPLMRRGRWVV...
APII NIENGASLSKEELDRRGVNDSLVHVDFMIGSADLNIDGVTKDGKREPIFRSGNWAFELA

Figure S1 (continued)



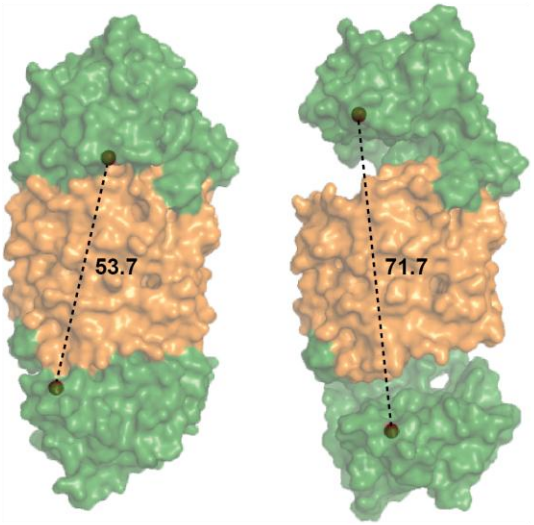
**Supplementary Figure S1. Structural analysis of PepS, related to Figure 1**

(***a***) Dimer interface of PepS. Top view looking down at the elongated dimer (**Left**) and side view obtained by rotating the top view by 90 degrees (**Right**). The first monomer on the top (orange) and the second monomer on the bottom (light orange) are shown as ribbon representations. Secondary structural elements are indicated in white. Residues involved in dimer interactions are labeled in red for the first monomer and in blue for the second monomer. Hydrogen bonds are displayed as blue dotted lines. (***b***) Multiple sequence alignment of clan MQ aminopeptidases: PepS from *Streptococcus pneumoniae*, AmpS from *Staphylococcus aureus*, AmpT from *Thermus thermophilus* and APII from *Geobacillus stearothermophilus*. Multiple sequence alignment was performed using ClustalW2 (Larkin *et al.*, 2007). Secondary structures of PepS are indicated: coil, α-helix; arrow, β-strand; η, $3_{10}$ helix; and T, β-turn. The active site residues are indicated by red stars. (***c***) Ribbon representation of the active site of ligand-free PepS in the closed conformation. Two zinc ions (magenta), the bridging water labeled "Wa" (red), and residues coordinating the metal ions (green) are shown. (***d***) Active site configuration of the ligand-free PepS in the closed conformation. Active site residues, metal ions, and the bridging water are labeled. The numbers indicate hydrogen bonding distances. Dotted lines indicate either metal coordination or hydrogen bonds. (***e***) Catalytic activities of wild-type (WT) and $PepS^{E343D}$ mutant. Catalytic activity of the $PepS^{E343D}$ mutant is given relative to the activity of WT PepS, which was set to 100%. (***f***) Cα superimposition of the C-terminal domains in the closed (green) and open (blue) conformations. The hinge (residues 235-252) and Glu-343 (residues 322-347) regions are indicated by blue and pink ovals, respectively. (***g***) The S1 binding pocket in the closed conformation consists of four hydrophobic residues (Phe-255, Tyr-355, Pro-247, and Trp-144) and is enclosed by the Cα backbone of three residues (Glu-233, Ser-234 and Ala-235). The S1

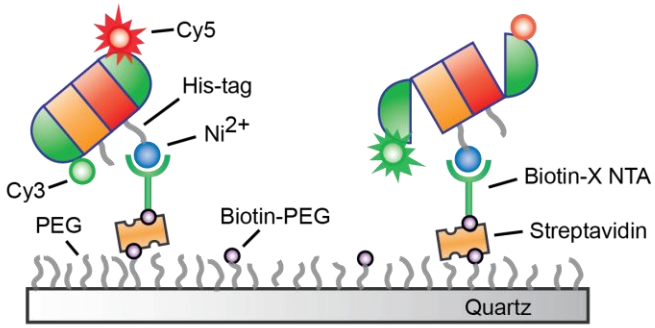
recognition residues from the N-terminal and C-terminal domains are colored orange and green, respectively. The Trp-Gly substrate bound to the active site is shown in purple and the two zinc ions in magenta. (***h***) The S1 binding pocket in the open conformation. The residues in the S1 recognition pocket and the zinc ions are labeled as in panel (***g***). The C-terminal domain was superposed to that in the closed conformation for comparison purpose.

Figure S2

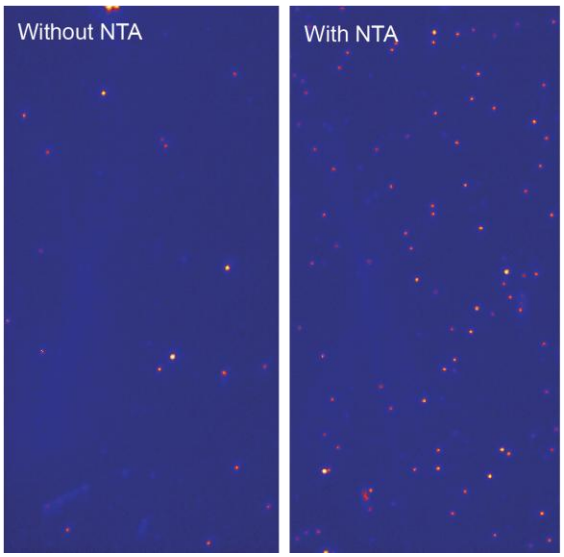
(*a*)




(*b*)



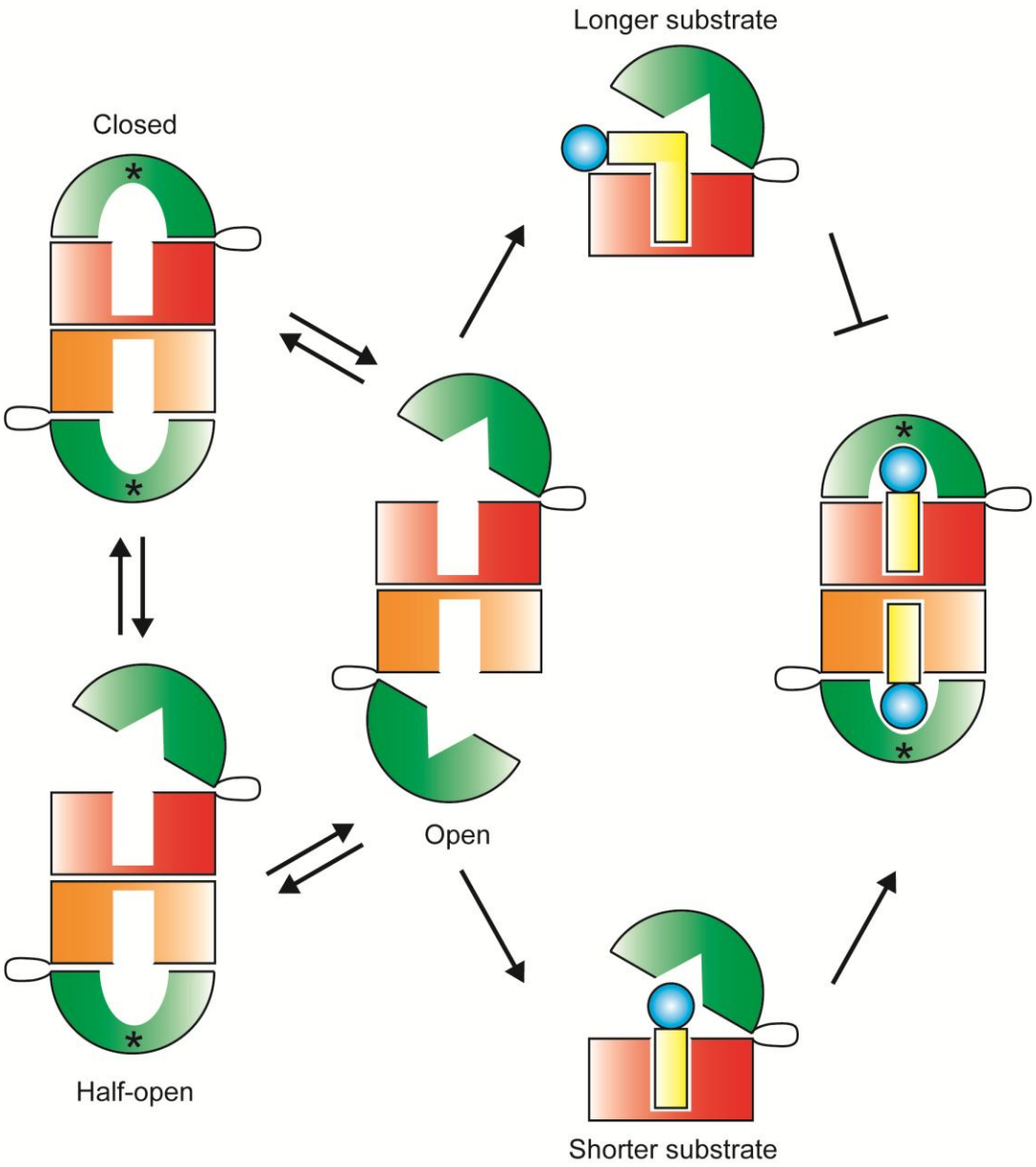

(*c*)



**SupplementaryFigure S2. Single molecule FRET measurements, related to Figure 4**

(***a***) Position of Ser-366 which was mutated to cysteine for labeling with Cy3 and Cy5. Cα atoms on Ser-366 are shown as red spheres. Distances between two Cα atoms on Ser-366 in dimeric arrangement are labeled in Å for closed (left) and open (right) conformations. (***b***) Schematic diagram of single-molecule FRET experiments. (***c***) Comparison of single-molecule fluorescence images of PepS (1.2 nM, 10 min incubation) on a quartz surface without (left) or with (right) X-nitrilotriacetic acid (X-NTA) treatment. Cy5 was excited with a red laser (640-nm). The number of spots increased more than 5-fold with X-NTA treatment.

Figure S3



**Supplementary Figure S3. Detailed explanation of the triple-sieved interlock mechanism, related to Figure 5**

In the absence of a substrate, three conformations of PepS – closed, half-open, and open forms – are presumed to be in dynamic equilibrium. In the presence of substrate, the open conformation is adopted to accommodate the substrate ("Open"). Since the substrate-binding hole is relatively straight, peptides with an extended conformation are likely to be accommodated without a large energy penalty. A triple sieve mechanism is proposed to regulate the activation of PepS ("Closed"). This involves: (i) active site positioning in which the Glu-343 region is repositioned close to the bridging water for proton abstraction along with domain movement; (ii) residue specificity for P1 residues conferred by the relocation of the S1 binding pocket residues to the active position; and (iii) length selectivity regulated by the substrate-binding hole. Substrates longer than the length of the substrate-binding hole can access the active site, but do not permit the conformational change to the active closed form due to steric clashes during domain movement ("Longer substrate"). In contrast, shorter substrates allow the conformational change and catalytic activation ("Shorter substrate"). The active closed form of PepS performs hydrolysis of the bound substrate. Finally, the products are released when PepS switches from the closed to the open conformation and the enzyme is ready for another cycle. The N-terminal domain is colored green and the C-terminal domain dark orange and light orange to emphasize the dimeric architecture of PepS. Substrate peptides are depicted in yellow, and the circle represents the N-terminal amino acid.

**Supplementary Reference**

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. & Higgin, D. G. (2007). *Bioinformatics* **23**, 2947-2948.