博士論文（要約）

Improving cold-activity of phytase by empirical and rational design

(経験的および合理的設計による フィターゼの低温高活性化)

Manami WADA

和田 愛未

Abbreviations

AppA	pH 2.5 acid phosphatase
HAP	histidine acid phosphatase
PCR	polymerase chain reaction
SASA	solvent accessible surface area
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
WT	wild type

本要約にて除外した箇所

Abbreviations（以下の内容に関わる部分）

Chapter 1（全文）

Chapter 2（全文）

Chapter 3 のうち、以下を除外。

3.1 Introduction（10－13 行目）

3.2 Materials and methods

3.2.2（全文）

3.3 Results

3.3.2（全文）

3.4.2（全文）

Chapter 4（全文）

Chapter 3: Improving AppA cold-activity by rational design 1

(Ala/Gly scanning of active site loop)

3.1 Introduction

The enzyme assay of Ala/Gly scanning mutants was carried out at 60°C (optimum temperature of AppA), to get insight into the contribution of each residue from the active site loop to AppA activity. As the movement of the loop governs substrate entry to the catalytic pocket and the enzyme reaction rate, unraveling the role of each residue in the loop upon functional expression would assist in elucidating the mechanism of AppA function. Furthermore, such work could provide valuable information to efficiently enhance AppA activity. Although Ala mutation is common for mutational analysis studies to negate the effect of the original residue's side chain, Gly scanning was also carried out to investigate the roles of Ala residues within the loop.

3.2 Materials and methods

3.2.1 Ala/Gly scanning of AppA active site loop

Plasmids and strains

The *appA* gene was amplified from *E. coli* BL21(DE3) genome extracted using nexttec 1-step DNA Isolation Kit for Bacteria (nexttec GmbH, Hilgertshausen, Germany). Using the purified genome, *appA* was amplified by PCR using a forward primer appA_F and reverse primer appA_AS (see Table 1) and was inserted into the vector plasmid pETDuet-1 (Merck Millipore, Darmstadt, Germany) at the *Nde*I and *Avr*II restriction sites, yielding the plasmid pETDuet-AppA. His6-tag was inserted at the C-terminus of *appA* for protein purification and named pETDuet-AppA+His. This plasmid was utilized to add His6-tag and construct Ala and Gly mutants by the protocols described below using the primers listed in Table 1.

The addition of His6-tag to pETDuet-AppA plasmid was carried out as follows. First,

primers for His6-tag introduction were prepared by amplifying two short primers containing complimentary linker and His6-tag sequences (AppA+His_F, 5'-cat acc ggc gtg cag ttt ggg tag cag cgg tca tca tca tca tca tca t-3', AppA+His_R, 5'-gca gcc tag gtt agt cag tta atg atg atg atg atg atg acc gct gct acc-3') by PCR [98°C for 10 sec, 1 cycle; 98°C for 10 sec, 68°C for 5 sec, 72°C for 30 sec, 30 cycles] using KOD-Plus-NEO DNA polymerase (Toyobo, Osaka, Japan). The PCR product was purified using FastGene Gel/PCR Extraction Kit (Nippon Genetics, Tokyo, Japan) and used as primers to introduce His6-tag in the subsequent PCR [94°C for 2 min, 1 cycle; 98°C for 10 sec, 68°C for 4 min, 30 cycles] using the protocols of the QuikChange site-directed mutagenesis kit (Agilent Technologies, Santa Clara, CA, USA). Methylated template plasmids were digested by *Dpn*I restriction enzyme (Takara Bio, Shiga, Japan). The reaction solution was then used to transform *E. coli* JM109 competent cells (Nippon Gene, Tokyo, Japan) with pETDuet-AppA+His plasmid. The plasmid was extracted from the cultivated cells using FastGene Plasmid Mini Kit (Nippon Genetics, Tokyo, Japan). Finally, *E. coli* BL21(DE3)pLysS competent cells (BioDynamics Laboratory, Tokyo, Japan) were transformed with pETDuet-AppA+His plasmid for protein expression. All primers used in this study were synthesized by Eurofins Genomics (Tokyo, Japan).

E. coli cell cultivation

Auto-induction medium was used for cultivation of *E. coli* BL21(DE3)pLysS competent cells transformed with plasmids constructed above (Studier *et al.* 2005). Cultivation was initiated by inoculating colonies into 96-well plates with non-inducing ZYM-505 medium (100 μL per well). After the overnight cultivation at 37°C, 10 μL from each well was transferred to a new well in 96-well plate with 150 μL inducing ZYM-5052 medium. Each strain was cultivated using 3 wells for one sample. Cells were harvested in 1.5 mL tubes after 20 h of cultivation, washed twice with Tris-HCl buffer (50 mM Tris-HCl, pH 7.4), and kept at –80°C until use.

Phytase enzymatic assay

Phytase assay was carried out based on the molybdenum blue method (Shivange *et al.* 2012). Cells were first thawed on ice, and 200 μL Tris-HCl buffer (pH 7.4) was added to each sample to suspend the cells. Following the addition of 100 μL lysozyme solution (10 mg/mL in Tris-HCl buffer, pH 7.4), cells were lysed by incubating at 37°C for 1 h. Supernatant of cell lysates was obtained by centrifuging the lysed cells (4°C, 20,600 ×*g*, 30 min).

The amount of the soluble form of AppA contained in the supernatant was quantified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using 10% gels. Coomassie Brilliant Blue-stained gel images were acquired using a Gel Doc EZ imager (Bio-Rad, Hercules, CA, USA). Quantification of AppA band was performed using the ImageLab software (Bio-Rad).

Phytase reaction was started by mixing 10 μL diluted supernatant of cell lysates (200 times dilution) and 70 μL substrate solution (1 mM sodium phytate in 250 mM sodium acetate buffer, pH 4.5) in 96-well plates, and incubated at 60°C for 15 min. The enzyme reaction was quenched by adding 70 μL trichloroacetic acid to the reaction solution. Following the addition of 70 μL color developing solution (0.54% ammonium molybdate, 2.16% ascorbic acid in 0.34 M sulfuric acid), the 96-well plate was placed under 50°C for 15 min for phosphate-molybdenum complex formation. An absorbance at 820 nm was measured (Synergy HTX multi-mode reader, BioTek, Vermont, USA) and used to calculate the activity for each sample using the KH_2PO_4 standard curve. Finally, the activity was standardized by SDS-PAGE band intensity of the soluble form of AppA contained in lysate supernatant. The activity measurements were carried out five times, and means ± standard errors are shown. For each experiment, three samples of the wild type were used (total of fifteen samples for wild type).

The apparent AppA activity was determined using the KH_2PO_4 standard curve and the 820 nm absorbance from the result of the AppA assay mentioned earlier. The activity was expressed in U, which means the amount of AppA required to liberate 1 μmol of PO^{4-} in one minute (Fig. 18). To standardize the loading amount in SDS-PAGE, the band intensity of *E. coli* in each mutant lane was divided by that in the WT lane. This was named a “correction factor” and used to normalize the soluble amount of AppA in the supernatant fraction by dividing an AppA band intensity (Vol (Int) in the ImageLab software; Fig. 19, Fig. 20 (Top)) with the corresponding correction factor. The resultant number is the relative amount of soluble AppA (Fig. 20 (Bottom)). Finally, the AppA activity was standardized by the soluble amount of AppA and expressed as a relative activity (%) compared to WT (Fig. 21).

3.3 Results

3.3.1 Ala/Gly scanning of 11 residues in the AppA active site loop

Overview

To investigate the role of each residue's role in AppA active site loop, 19 Ala and/or Gly mutants were designed as follows. Two Ala residues within the loop (A21 and A25) were replaced by only Gly, and one residue initially Gly (G18) was substituted with Ala. Other 8 positions were replaced with both Ala and Gly. In total, 19 mutants were designed and constructed by site-directed mutagenesis. *E. coli* BL21(DE3) pLysS competent cells were transformed by mutant and wild type plasmids. Proteins were expressed by cultivating the cells. Phytase activity assay was conducted using the supernatant of cell lysate under the AppA optimal conditions (pH 4.5, 60°C). After the enzyme reaction using sodium phytate as substrate, the 820 nm absorbance corresponding to the blue complex of product (PO_4^{3-}) and molybdenum was measured. Phytase activity was calculated using the absorbance (Fig. 18), SDS-PAGE band intensities were detected (Fig. 20(Top)) and used to standardize the activity by the relative amount of soluble AppA, calculated as described in the methods section (Fig. 19(Bottom)). This means that each value of activity was standardized by the AppA expression level. From the result of SDS-PAGE band analysis, no obvious difference in AppA expression levels were found among wild type and mutant AppA. Thus, a possibility of AppA activity heavily influenced by its expression level can be ignored. This analysis method enabled us to evaluate all the 19 mutants and wild type in a speedy manner.

As a result of activity assay, more than 90% of activity loss was confirmed from eight mutants (R16A, R16G, H17A, H17G, V19G, R20A, R20G. and T23G) (Fig. 21). Besides, when Ala and Gly mutants at the same position were compared, V19G was found to lost 50% more activity compared to V19A.

Correlation analysis

To understand the mutational effects caused by Ala/Gly scanning, correlations between the activity and several chemical/structural characteristics of each mutation site, such as solvent accessible surface area (SASA), B-factor, number of interactions with substrate and other structural units in AppA, distance from the substrate, and degree of conservation was investigated (Fig. 26, 27). SASA values were calculated by GetArea server using the crystal

structures of AppA with substrate (PDB ID: 1DKQ) and without substrate (PDB ID: 1DKN) to inspect structural differences induced by substrate-binding. B-factor value for Ca atoms of each site was obtained from the same PDB entries. Correlation analyses provided no clear correlations between SASA/B-factor and AppA activity, whether the substrate was present or not. However, there was a weak positive correlation between B-factor and relative activity under the presence of substrate (Fig. 27(C)). This infers those residues that become less flexible when substrate is bound might play an important role in function expression. In other words, those sites that maintain mobility regardless of the existence of substrate may not contribute much to the function, and thus can be potential further mutagenesis sites.

The number of contacts between each residue from the active site loop and the substrate or other parts of AppA was counted by the softwares "ligand-protein contacts" and "contacts of structural units" using the AppA structure with substrate (Sobolev *et al.* 1999). The number of interactions differed from 4 to 24, where the two Arg in the conserved motif (R16 and R20) having more contacts (24 and 23, respectively) compared to others. A strong negative correlation between the activity and the number of interactions of each loop residue and other structural units of AppA was obtained from the correlation analysis, displaying an overall tendency to lose more activity with the increase of interactions (Fig. 26(A)).

To obtain the distance from each mutational site to substrate, PyMol software (Molecular Graphics System, Schrodinger, LLC) was utilized to measure the distance between the side chain of each loop residue and any of the six phosphate groups on the substrate (phytic acid). The correlation analysis using all the mutants yielded a remarkably high correlation coefficient. In spite of this trend, the V19G mutant was distant from the regression line (Fig. 26(B)). With the V19G excluded, a further increase in the correlation was observed.

Sequence conservation among HAPs

Five residues from the AppA active site loop (R16, H17, G18, R20, P22) are conserved as active site motif (RHGXRXP) among the superfamily AppA belongs to (HAPs; histidine acid phosphatases). In addition, a preceding site-directed mutagenesis work revealed R16 and H17 to be substrate-binding and catalytic site, respectively. The result of AppA activity assay shows a complete (or almost complete) loss of activity among conserved charged residues (Arg and His), while non-charged ones (Gly and Pro) preserved activity of 50% or above (Fig. 21).

To compare the degree of conservation values for each site in AppA active site loop, the number of AppA homolog sequences with the same amino acid at a certain site was divided by the total number of sequences used. This time, 62 amino acid sequences from diverse species were acquired from the NCBI Conserved Domain Database within the protein category named “HP_HAP_like” (phosphatases having a His residue as active site) which AppA belongs to (Fig. 28). For the simplicity of analysis, one enzyme from each species were selected to calculate the degree of conservation. However, AppA2, which is another phytase from *E. coli*, was included to the analysis as an exception for the contrast with AppA. The correlation analysis of the degree of conservation and the relative activity of Ala/Gly produced a weak correlation. The reason for this result may attribute to the fact that mutations to the conserved residues did not necessarily end up in the reduction in activity and mutations to some of the non-conserved residues diminished remarkable fractions of the activity.

There was no obvious correlation observed in the AppA activity and the degree of conservation, some sequence variation patterns were found as a result of the multiple sequence alignment (Fig. 27(E) and 29). It turned out that the T23 site, which lies in the AppA amino acid sequence right after the conserved active site motif RHGXRXP, is partially conserved. In contrast, the positions A21, K24, A25, and T26 had varying amino acids in different species. Moreover, there was a strong tendency at the equivalent position of V19 to have either Ala or a residue with negative charge (Glu or Asp), and each residue was found in 19, 14, and 13 AppA homologs, respectively. Having Val at residue 19 like AppA was a rather unusual case in the 62 AppA homologs tested (only 5 out of 62 homologs). Moreover, the fact that the A21 position, which is a variable residue in the active site motif, has bulky hydrophobic amino acids (Tyr and Phe) in more than half the histidine phosphatases used for

3.4 Discussion

3.4.1 Ala/Gly scanning of 11 residues in the AppA active site loop

Mutations at conserved sites in the active site loop

Ala and Gly scanning mutagenesis was performed targeting AppA active site loop to gain knowledge on the function of each residue. AppA enzymatic assay was carried out to determine AppA activity, and the correlation analysis using the mutant activity and multiple chemical/structural properties for each mutation site was carries out.

As stated in the result section, Ala and/or Gly mutation on charged conserved sites resulted in significant decreases (Fig. 21). The reason for this great activity loss of charged conserved amino acids (R16, H17, and R20) is clear. Both R16 and R20 forms multiple contacts with the scissile 3-phosphate of substrate and are located very close to the substrate in the crystal structure (Fig. 2(B)). With the positively charged side chains eliminated, it would not be possible for AppA to hold the phosphate groups of substrate with negative charges. On the other hand, H17 gets phosphorylated during the AppA catalytic reaction. Thus, the positive charge of H17 is also not allowed to be occupied by aliphatic amino acids.

Despite the fact that G18 and P22 are both conserved, the effect on G18 and P22 mutants caused by Ala and/or Gly mutations were not significant (Fig. 21). This could be explained by the fact that G18 and P22 make no contact with the substrate and few contacts with other structural parts of AppA (Fig. 2(B)), which suggests they are conserved for maintain the structure of loop. The original residue at G18 and P22 positions give AppA the best activity, however, those positions can be substituted by other amino acids while retaining partial activity.

Mutations at non-conserved sites in the active site loop

When non-conserved sites were mutated, the AppA mutants did not severely lose the activity in general (Fig. 21). Notably, mutations on A25 and T26 barely affected their activity. According to the mutagenesis study by Wu *et al.* (2014), 79% of activity compared to wild type was detected in the A25F mutant. Our result on A25G mutant is comparable to this data. On the other hand, T23 and K24 mutants significantly decreased their activity, showing their importance in AppA catalytic activity. In addition, T23 mutants forfeited larger fraction of activity compared to K24 mutants (Fig. 21), possibly due to its more interactions with the substrate at a closer proximity (Fig. 2(B)). There is one water-mediated contact between T23 and the substrate (Lim *et al.*, 2000), which could lead to the stabilization of AppA structure by hydration. Besides, K24 is not conserved, while T23 is partly conserved (Fig. 29) despite it does not belong to the conserved active site motif. Taken together, the T23 site can be regarded to play important roles for both the structure and function. Thus, by mutating T23 and K24, AppA activity could be greatly affected depending on which amino acid to substitute with. Zhang *et al.* (2016) reported that K24E mutant displayed approximately 40% activity at 60°C. Exchanging K24 with a residue with smaller side chain volume may not be effective in improving activity at 60°C due to the loosening of the substrate-binding pocket.

When the pairs of Ala/Gly mutants at a certain position in the active site loop were compared, V19A and V19G had the most notable difference in the activity loss (Fig. 21; 29% and 95% activity loss, respectively). The result from the correlation analysis supports this fact: V19G mutant lost much more activity compared to other mutants with the comparable number of contacts with substrate or the distance from the substrate. On the AppA amino acid sequence, V19 is located between conserved residues G18 and R20. Since H17 undergoes phosphorylation by a phosphate group from the substrate during the AppA catalytic reaction, it can be assumed that the neighboring sites would need to render some space for H17. Nevertheless, under an elevated temperature like 60°C, which is the optimal temperature for AppA, two consecutive Gly residues (G18 and Gly mutation to V19) would provide more than enough conformational mobility to the active site loop, resulting in an inefficient catalysis/binding. Comparing AppA amino acid sequence with 62 other homologs, the position equivalent to V19 is most frequently replaced by either Asp or Glu (Fig. 29). As there was no homolog with Gly at the V19 position, this site presumably requires some bulkiness for the amino acid to be placed, and it does not necessarily need to be charged or hydrophobic.

Acknowledgements

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Munehito Arai, for letting me join his lab three years back. Also, I am thankful for his generosity in letting me start up a new research project. He never forgot to keep me stay on the right track of research by giving me proper insights.

A constant assistance provided by Asst. Prof. Dr. Yuuki Hayashi helped me go through the time-consuming optimization of experiment conditions of the directed evolution based on his professional knowledge on the subject.

Many special thanks to all my friends from both inside and outside of Arai Laboratory, they always encouraged to keep my faith in finishing the PhD course.

Last not but least, I want to send my warmest gratefulness to my family for caring me all the time.

References

Crooks GE, Hon G, Chandonia JM, *et al.* (2004) WebLogo: a sequence logo generator. *Genome research*, 14(6): 1188-1190.

Larkin MA, Blackshields G, Brown NP, *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics.* 23(21): 2947-2948.

Shivange, AV, Serwe, A, Dennig, A, *et al*. (2012) Directed evolution of a highly active Yersinia mollaretii phytase. *Applied microbiology and biotechnology*, 95(2), 405-418.

Sievers, F, Wilm, A, Dineen, D, *et al*. (2011) Fast, scalable generation of high‐quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology*, 7(1), 539.

Sobolev, V, Sorokine, A, Prilusky, J, *et al*. (1999) Automated analysis of interatomic contacts in proteins. *Bioinformatics (Oxford, England)*, 15(4), 327-332.

Studier, FW. (2005) Protein production by auto-induction in high-density shaking cultures. *Protein expression and purification*, 41(1), 207-234.

Zhang, J, Liu, Y, Gao, S, *et al*. (2016) Site-directed mutagenesis and thermal stability analysis of phytase from Escherichia coli. *Biochemical and Biophysical Research Communications*, 9(3), 357-365.

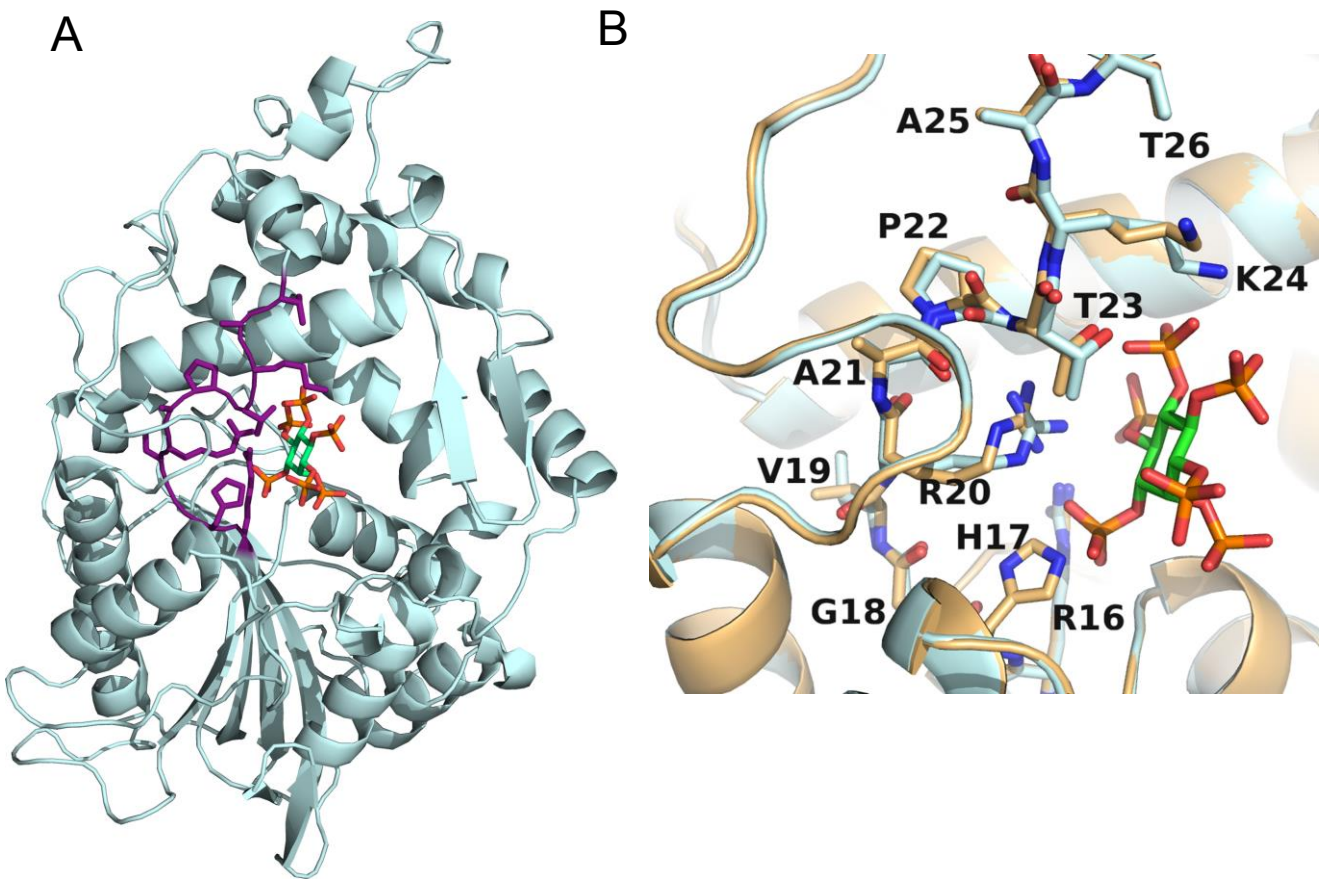

Fig. 2. Structure of *E. coli* phytase AppA. (A) Overall structure of AppA. The active site loop is shown in purple, and the substrate (phytic acid) is shown in green. (B) The enlarged view of AppA around its active site loop (sticks) and the substrate (green sticks). The AppA structures in the substrate-bound form (light cyan) and in the free form (gold) are overlaid. Blue, red, and orange denote the nitrogen, oxygen, and phosphorous atoms, respectively. The figures were drawn using the PyMOL Molecular Graphics System, Schrodinger, LLC.

Table 1. List of primers used in this study.

Primer name	Primer sequence (5'→ 3')
appA_F	agagcatatgaaagcgatcttaatcccatttttatctcttctgattcc
appA_AS	ctctcctaggttagtcagttacaaactgcacgccggtatgcgt
R16A_F	gtgtggtgattgtcagtgcgcatggtgtgcgtgctcc
R16A_R	ggagcacgcacaccatgcgcactgacaatcaccacac
R16G_F	tgtggtgattgtcagtggccatggtgtgcgtgctcc
R16G_R	ggagcacgcacaccatggccactgacaatcaccaca
H17A_F	tgtggtgattgtcagtcgtgcgggtgtgcgt
H17A_R	acgcacacccgcacgactgacaatcaccaca
H17G_F	tgtggtgattgtcagtcgtggcggtgtgcgtgctccaacc
H17G_R	ggttggagcacgcacaccgccacgactgacaatcaccaca
G18A_F	gattgtcagtcgtcatgcggtgcgtgctccaaccaa
G18A_R	ttggttggagcacgcaccgcatgacgactgacaatc
V19A_F	cagtcgtcatggtgcgcgtgctccaaccaa
V19A_R	ttggttggagcacgcgcaccatgacgactg
V19G_F	gtcagtcgtcatggtggccgtgctccaa
V19G_R	ttggagcacggccaccatgacgactgac
R20A_F	cagtcgtcatggtgtggcggctccaaccaagg
R20A_R	ccttggttggagccgccacaccatgacgactg
R20G_F	gtcgtcatggtgtgggcgctccaaccaa
R20G_R	ttggttggagcgcccacaccatgacgac
A21G_F	gtcatggtgtgcgtggcccaaccaaggcc
A21G_R	ggccttggttgggccacgcacaccatgac
P22A_F	tcatggtgtgcgtgctgcgaccaaggccacg
P22A_R	cgtggccttggtcgcagcacgcacaccatga
P22G_F	atggtgtgcgtgctggcaccaaggccacgc
P22G_R	gcgtggccttggtgccagcacgcacaccat
T23A_F	ggtgtgcgtgctccagcgaaggccacgcaact
T23A_R	agttgcgtggccttcgctggagcacgcacacc
T23G_F	tgcgtgctccaggcaaggcc
T23G_R	ggccttgcctggagcacgca
K24A_F	gcgtgctccaaccgcggccacgcaactgat
K24A_R	atcagttgcgtggccgcggttggagcacgc

K24G_F	tgtgcgtgctccaaccggcgccacgcaactgatgc
K24G_R	gcatcagttgcgtggcgccggttggagcacgcaca
A25G_F	tgctccaaccaagggcacgcaactgatgc
A25G_R	gcatcagttgcgtgcccttggttggagca
T26A_F	ctccaaccaaggccgcgcaactgatgcagg
T26A_R	cctgcatcagttgcgcggccttggttggag
T26G_F	gtgctccaaccaaggccggccaactgatgcaggatgt
T26G_R	acatcctgcatcagttggccggccttggttggagcac
AppA+His_F	cataccggcgtgcagtttgggtagcagcggtcatcatcatcatcatcat
AppA+His_R	gcagcctaggttagtcagttaatgatgatgatgatgaccgctgctacc

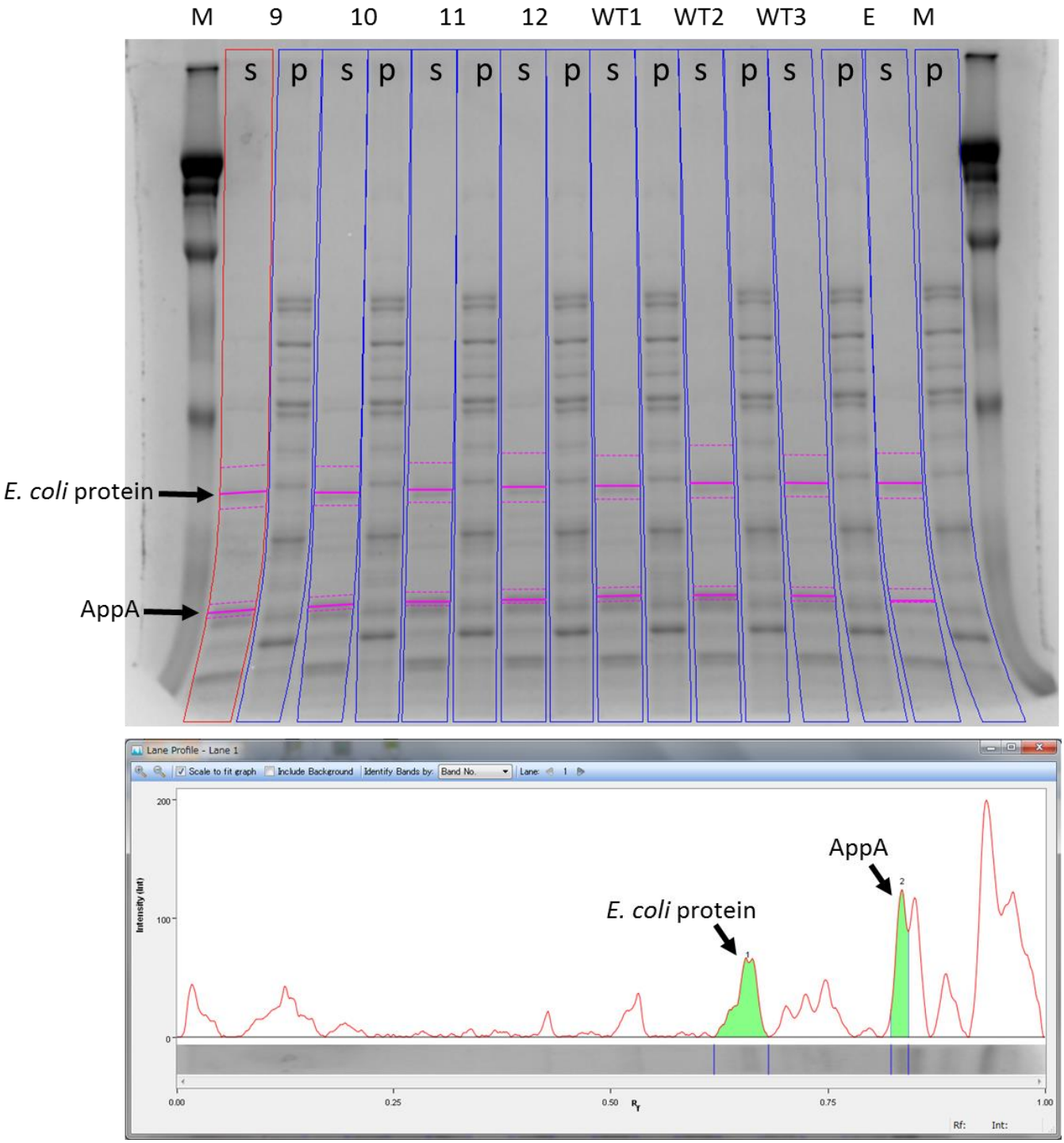

Fig. 4. An example of SDS-PAGE analysis of the *E. coli* protein (for the calculation of "correction factor" of the sample loading amount) and AppA bands. (Top): The gel image on the Image Lab software. The lane boundaries are shown by a rectangle, and the band boundaries are shown by the pink dotted lines. The *E. coli* protein and AppA bands are marked with black arrows. (Bottom): The image of the "Lane Profile" on the Image Lab software. The band boundaries can be adjusted by moving the blue lines beneath the band boundaries.

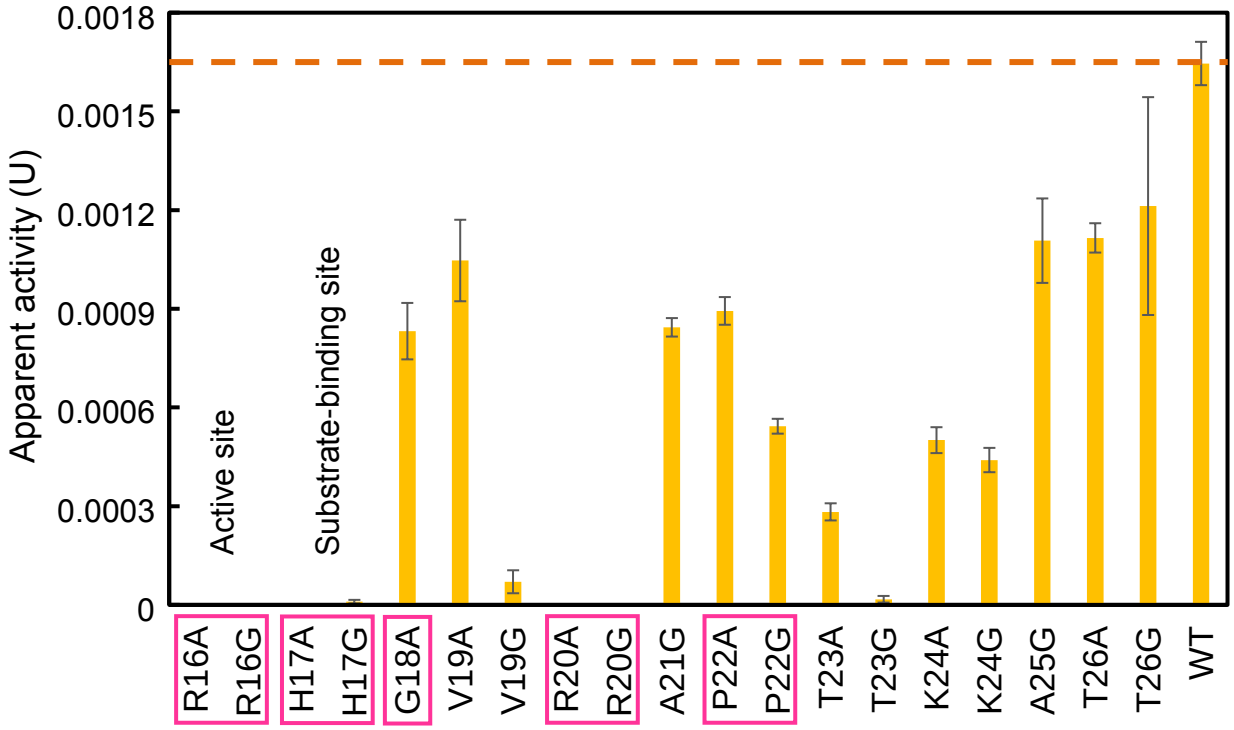

Fig. 18. Apparent AppA activity of mutant and wild type AppA. Supernatant of lysate from each sample was used to carry out the AppA enzymatic assay. The apparent AppA activity was calculated from the 820 nm absorbance of product-molybdenum complex using the KH_2PO_4 standard curve. The activity measurements were carried out five times, and means ± standard errors are shown. The assay was carried out at 60°C. The orange broken line shows the wild type activity.

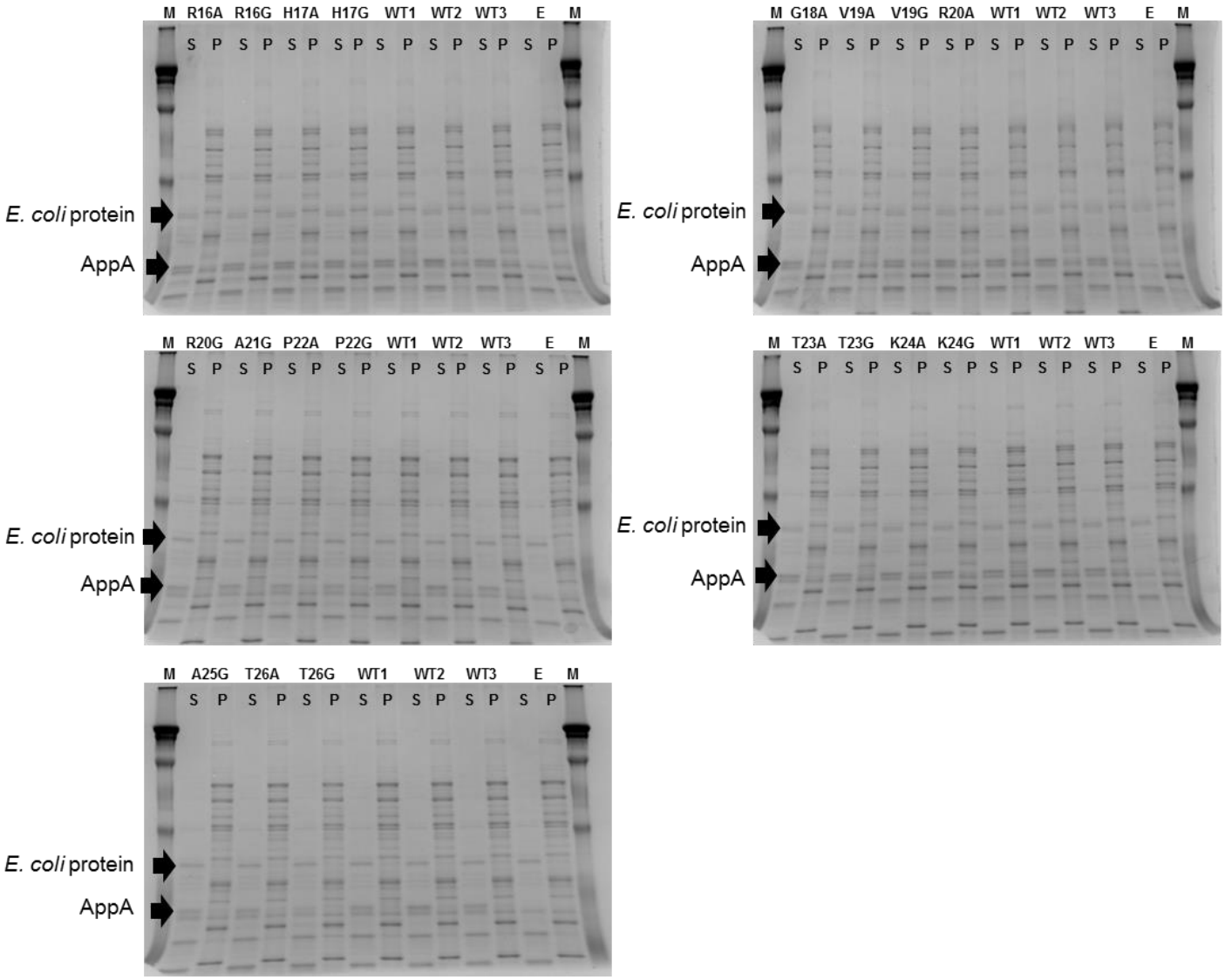

Fig. 19. SDS-PAGE of AppA and Ala and/or Gly mutants of the active site loop. S, supernatant of cell lysate. P, pellet of cell lysate. M, molecular weight marker. E, pETDuet-1 (the vector plasmid; negative control). The bands shown with black arrows are proteins from the expression strain *E. coli* BL21(DE3)pLysS used to calculate the "correction factors" to standardize the applied sample amount, and the AppA band.

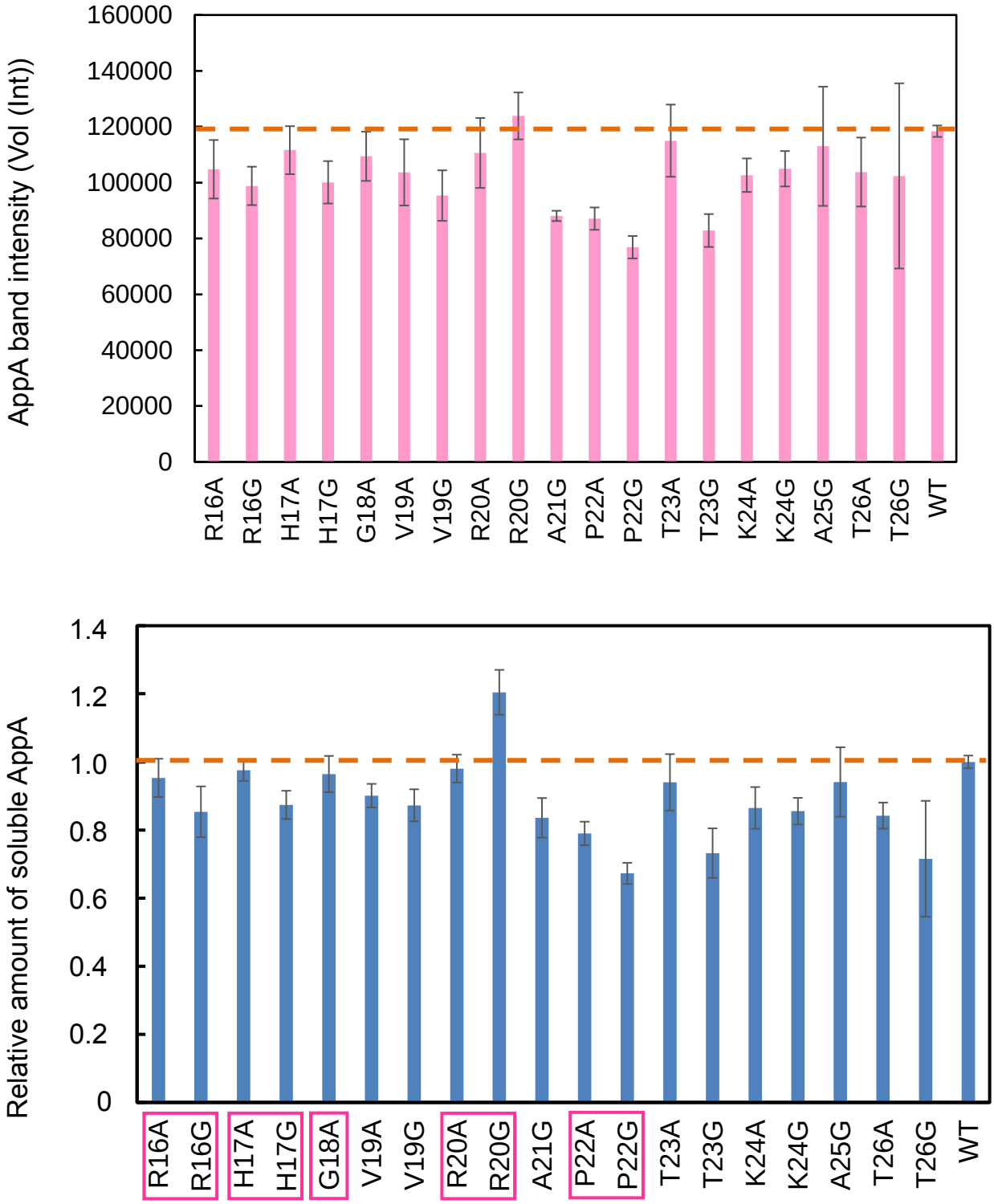

Fig. 20. SDS-PAGE analyses. Top: AppA band intensity values by analyzing the AppA band using the ImageLab software (Bio-Rad). Bottom: Relative amount of soluble AppA in the supernatant of lysate used for the measurement of the apparent AppA activity in Fig. 18. The orange broken lines show the wild type values.

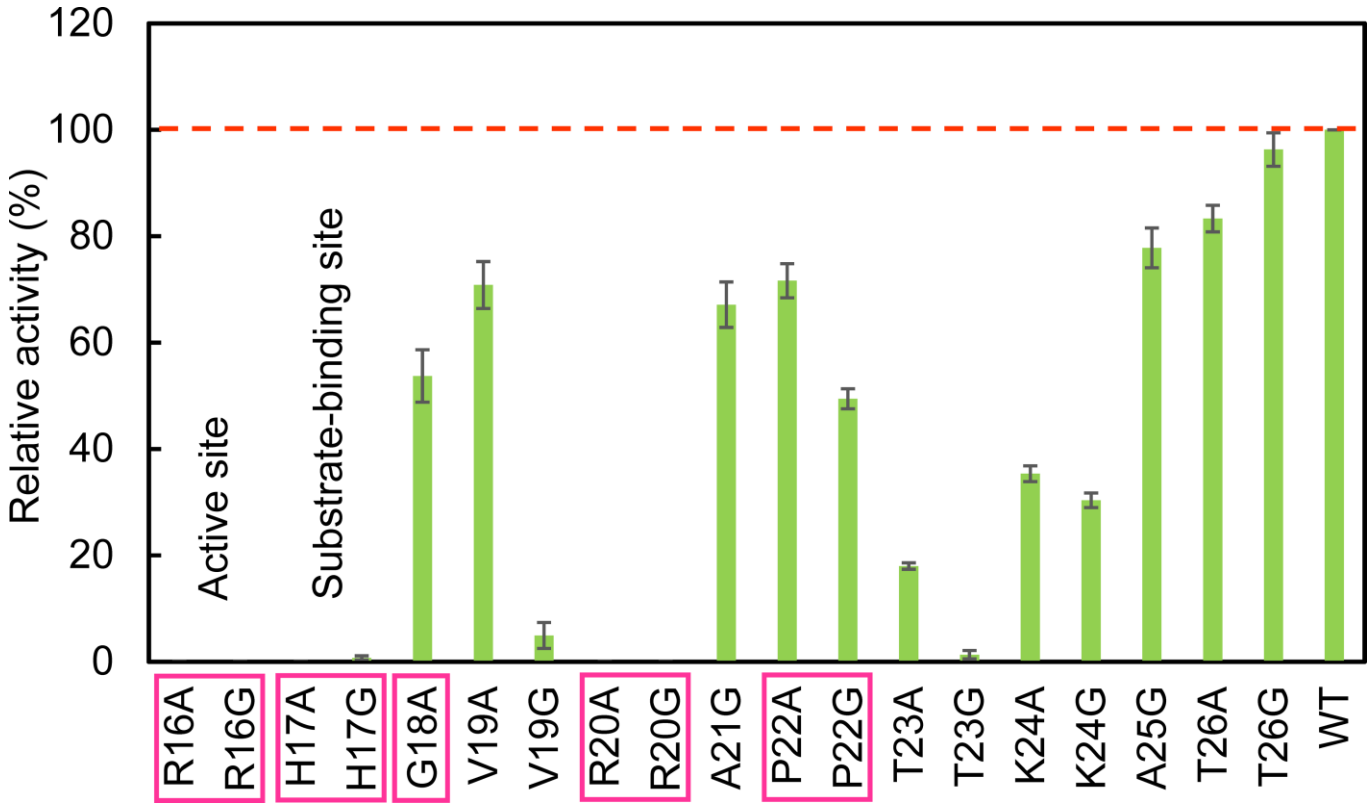

Fig. 21. Relative activity of Ala and Gly mutants compared with wild-type AppA. Mutants of conserved residues are boxed, and those of active site and substrate-binding site are designated. The broken line shows the activity of wild type (100%). WT denotes wild type. The activity measurements were carried out five times, and means ± standard errors are shown. The assay was carried out at 60°C.

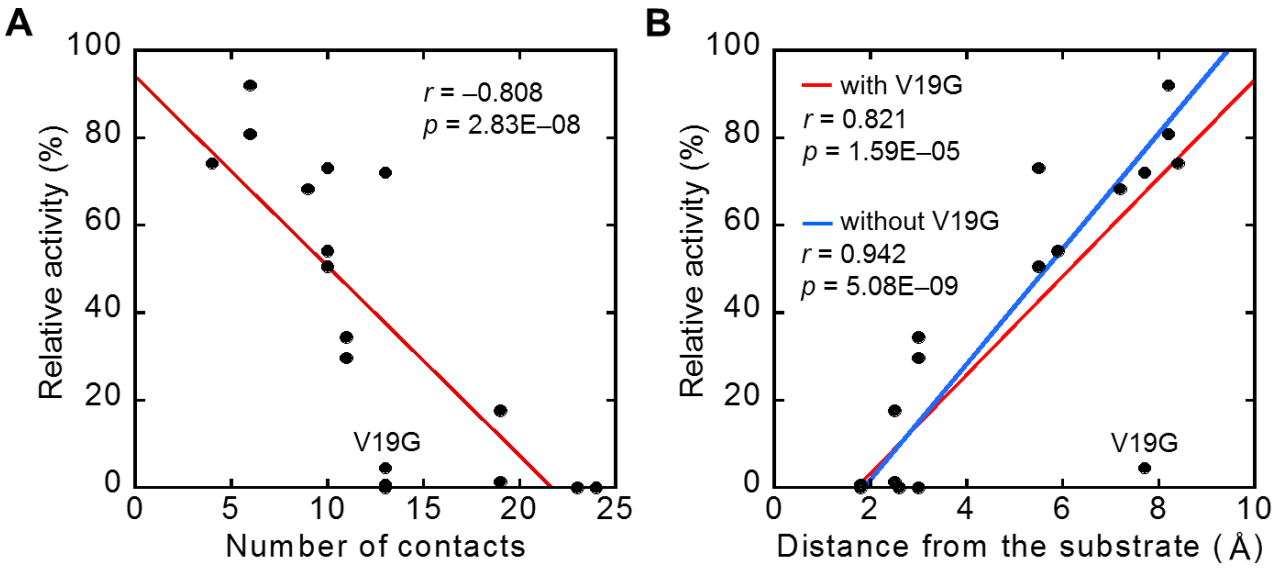

Figure 26. Correlation plots of relative activity of a mutant with (A) the number of contacts of each mutation site with substrate and other structural units in AppA, and (B) the distance from substrate to each mutation site. In both panels, a red continuous line shows the regression line. In (B), a blue continuous line shows the regression line obtained using the data without V19G.

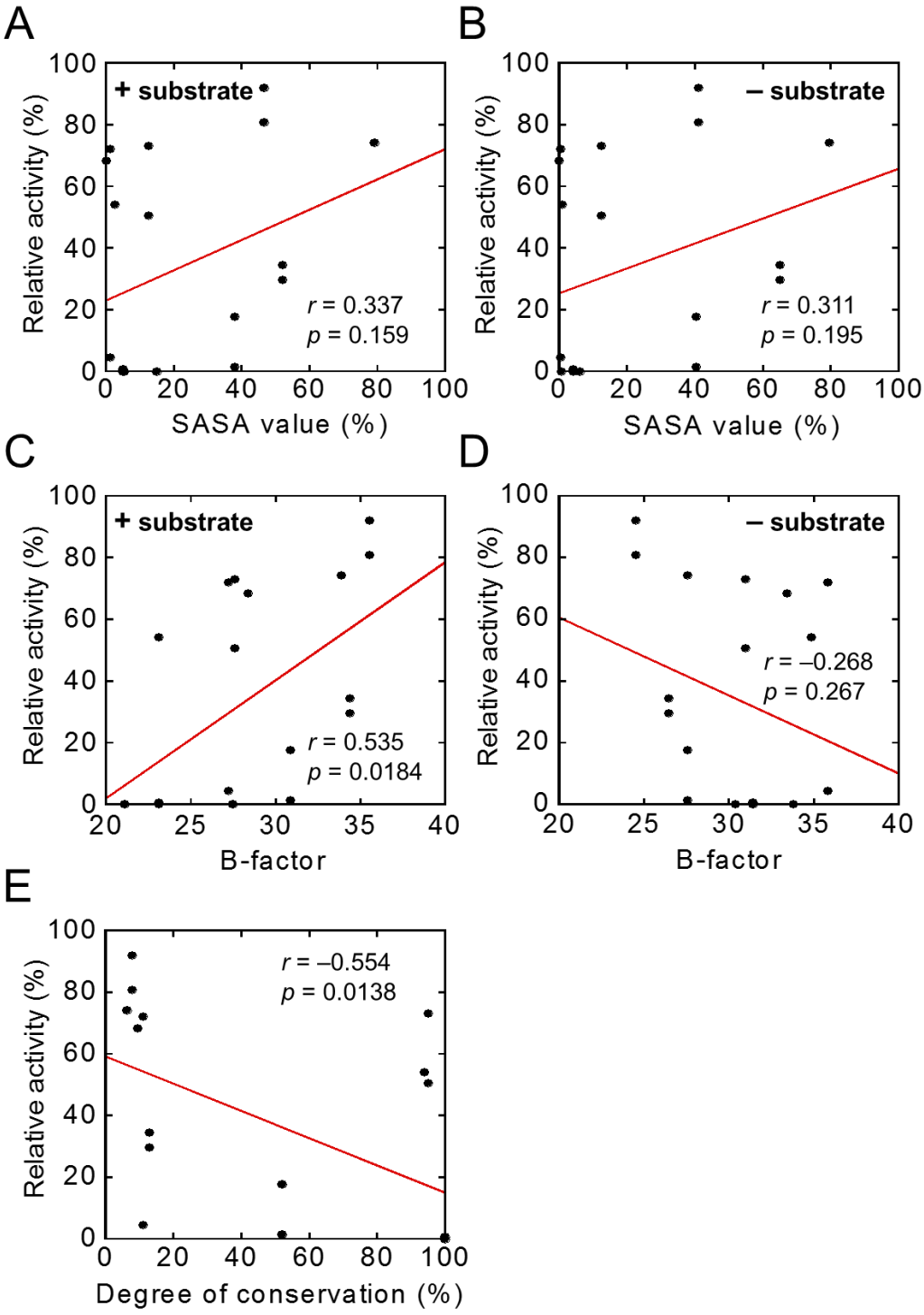

Fig. 27. Correlation analysis. (A-D) Correlation plots of relative activity of a mutant with SASA (A, B), B-factor (C, D), and (E) the degree of conservation of each mutation site. Crystal structures of AppA with substrate (A, C) and without substrate (B, D) were used for the calculation of SASA and B-factor. In all panels, a red continuous line shows the regression line.

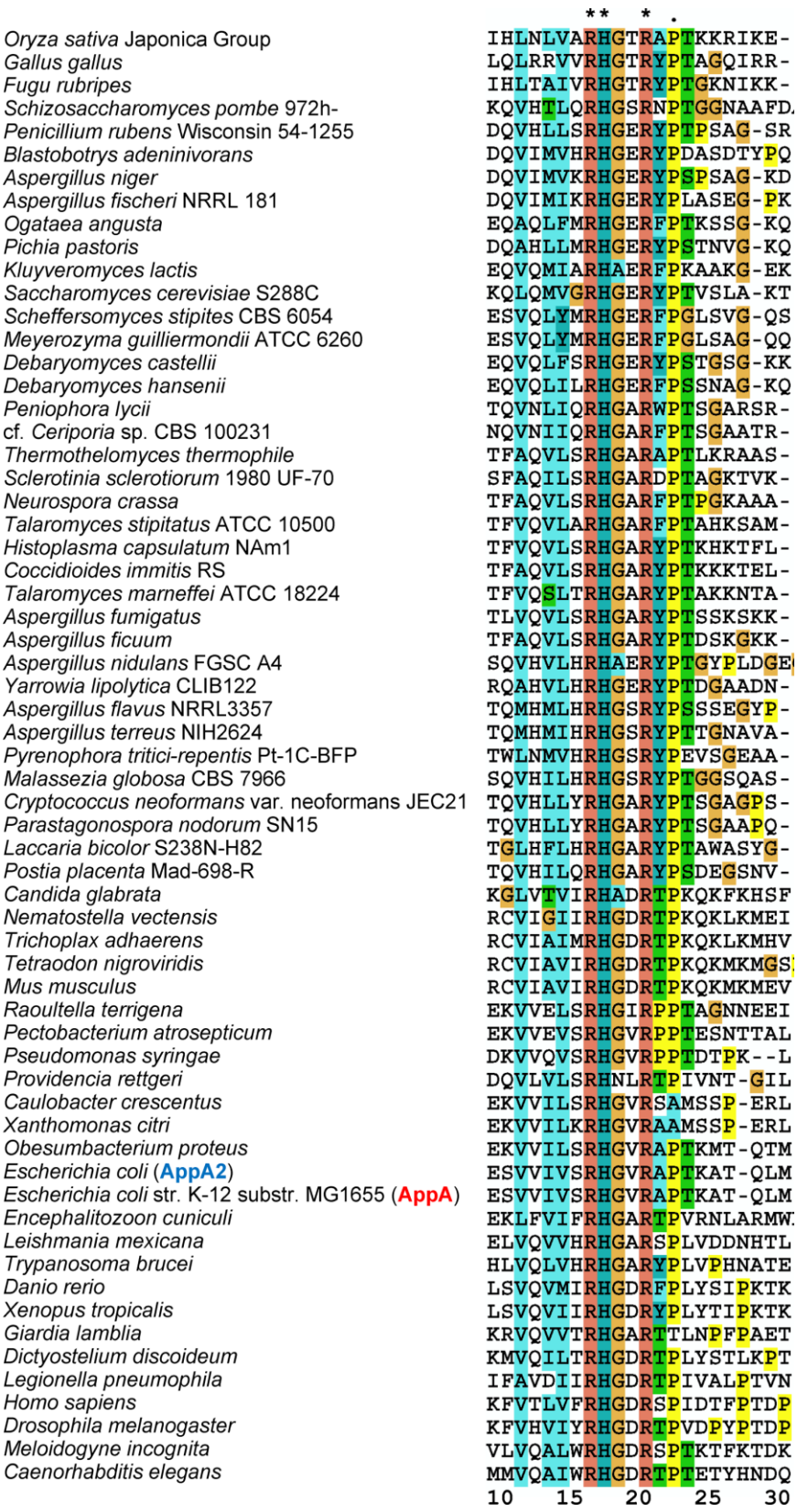

Fig. 28. Multiple sequence alignment of the active site loop of AppA and its 62 homologs. The alignment was performed using Clustal Omega (Sievers *et al*. 2011), and the figure was generated using Clustal X (Larkin *et al*. 2007). "*" indicates fully conserved residues, and "." indicates partially conserved residues. The numbers at the bottom are residue numbers in AppA sequence.

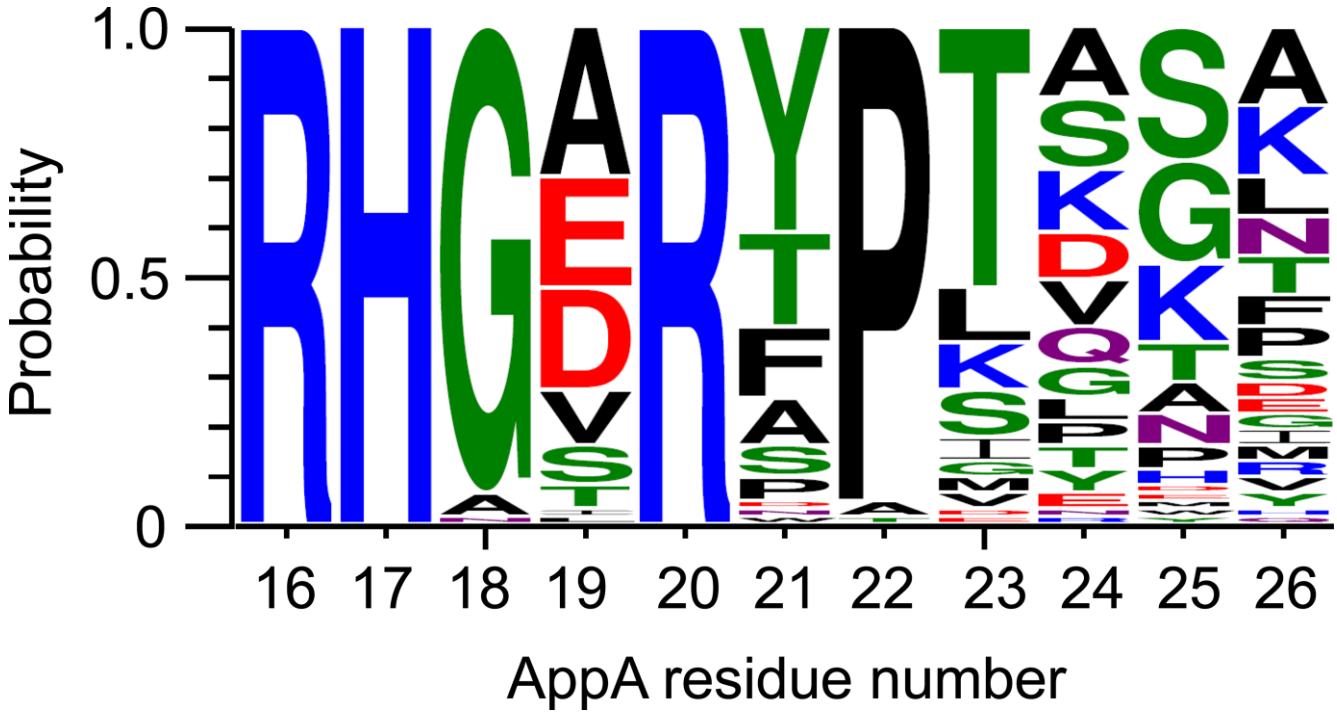

Fig. 29. Sequence logo showing the degree of conservation of active site loop residues based on the multiple sequence alignment of 63 HAPs (Fig.28). The horizontal axis shows the residue number in AppA, and the vertical axis shows the probability of a residue in a position. The larger the logo of a residue, the more likely the residue is placed in a certain position in the multiple alignment. The figure was created by WebLogo application (Crooks *et al.* 2004).