

TITLE

Elucidating the structure-function of *Helicobacter pylori* N-carbamoylputrescine amidase and to design structure-based inhibitor

INTRODUCTION

Helicobacter pylori is a gram-negative motile bacterium that colonizes the human stomach in an acidic environment. This pathogen infects nearly 60.3% of the global population (1). Once the bacterium enters the host, it evades the surrounding acidic environment by over-expressing the enzyme urease, which hydrolyzes urea to ammonia and carbon dioxide. The ammonia thus formed is known to elevate the pH of the microenvironment (2, 3). With the help of its flagellar motion and helical shape, the bacteria can then penetrate the gastric mucus barrier for colonization (4, 5). Persistent colonization can lead to host tissue damage resulting in peptic ulcers, duodenal ulcers, chronic gastritis, and gastric cancers (6–8). Therefore, this bacterium is classified as a Group 1 carcinogen by the World Health Organization (9).

Generally in prokaryotes and eukaryotes, arginine metabolism leads to the formation of polycationic compounds, which are crucial for cell growth and proliferation (10–12). In *H. pylori*, this metabolism is also important for acid resistance (2). Bacterial arginase utilizes host-derived L-arginine to produce ornithine and urea (13). The latter is then hydrolyzed by bacterial urease to produce ammonia, thus providing acid resistance to the pathogen (2). The bacterial arginase also competes with the host macrophages' inducible nitric oxide synthase for the common substrate L-arginine, limiting the host's arginine pool and thereby reducing NO production, an important component of innate immunity (14, 15). This modulation of the host immune system helps the bacterium to colonize near gastric epithelial cells (16). The L-arginine can also be hydrolyzed by bacterial arginine decarboxylase (ADC) to produce agmatine and carbon dioxide (17). The agmatine is then converted into N-carbamoylputrescine (NCP) by agmatine deiminase (18). In other bacteria and plants, the product NCP is known to be hydrolyzed by the enzyme N-carbamoylputrescine amidase (CPA) to produce putrescine, ammonia, and carbon dioxide (**Figure 1**) (19–21). However, the gene encoding CPA in *H. pylori* has thus far only been annotated.

In plants and other bacteria, agmatine is converted into putrescine by a two-step reaction, whereas in humans, agmatinase directly converts agmatine into putrescine (22). In general, putrescine is further metabolized by downstream enzymes that lead to the formation of polycationic compounds such as spermidine and spermine (11). In the host, these polycationic compounds can also be synthesized by the enzyme ornithine decarboxylase using the substrate ornithine produced through the arginase pathway (**Figure 1**) (23). However, the absence of ornithine decarboxylase in *H. pylori* suggests that arginine metabolism by ADC is the only pathway through which polycationic compounds can be produced. Thus, in plants and bacteria, including *H. pylori*, CPA appears to be a crucial enzyme for the formation of putrescine, the precursor of polycationic compounds. It has been reported that methylglyoxalbis(cyclopentylamidino)hydrazine, a known enzyme inhibitor for the synthesis of polycationic compounds reduces the growth of *H. pylori* when the bacteria is cultured in a medium containing the derivative (24). This observation further suggests the importance of polycationic compounds in bacterial proliferation and growth. In addition, when the ADC gene is knockout in *H. pylori*, the bacteria fails to grow (17), suggesting the gene is important for bacterial proliferation. This also implies that in *H. pylori*, the formation of polycationic compounds occurs

only through arginine metabolism via the ADC pathway. Overall, these data indicate that CPA is an attractive target for developing new therapeutics against *H. pylori*. It should be noted that with the exception of CPA, the enzymes involved in the *H. pylori* polyamine biosynthesis pathway (through arginine metabolism by ADC) have been previously studied (13, 17, 18).

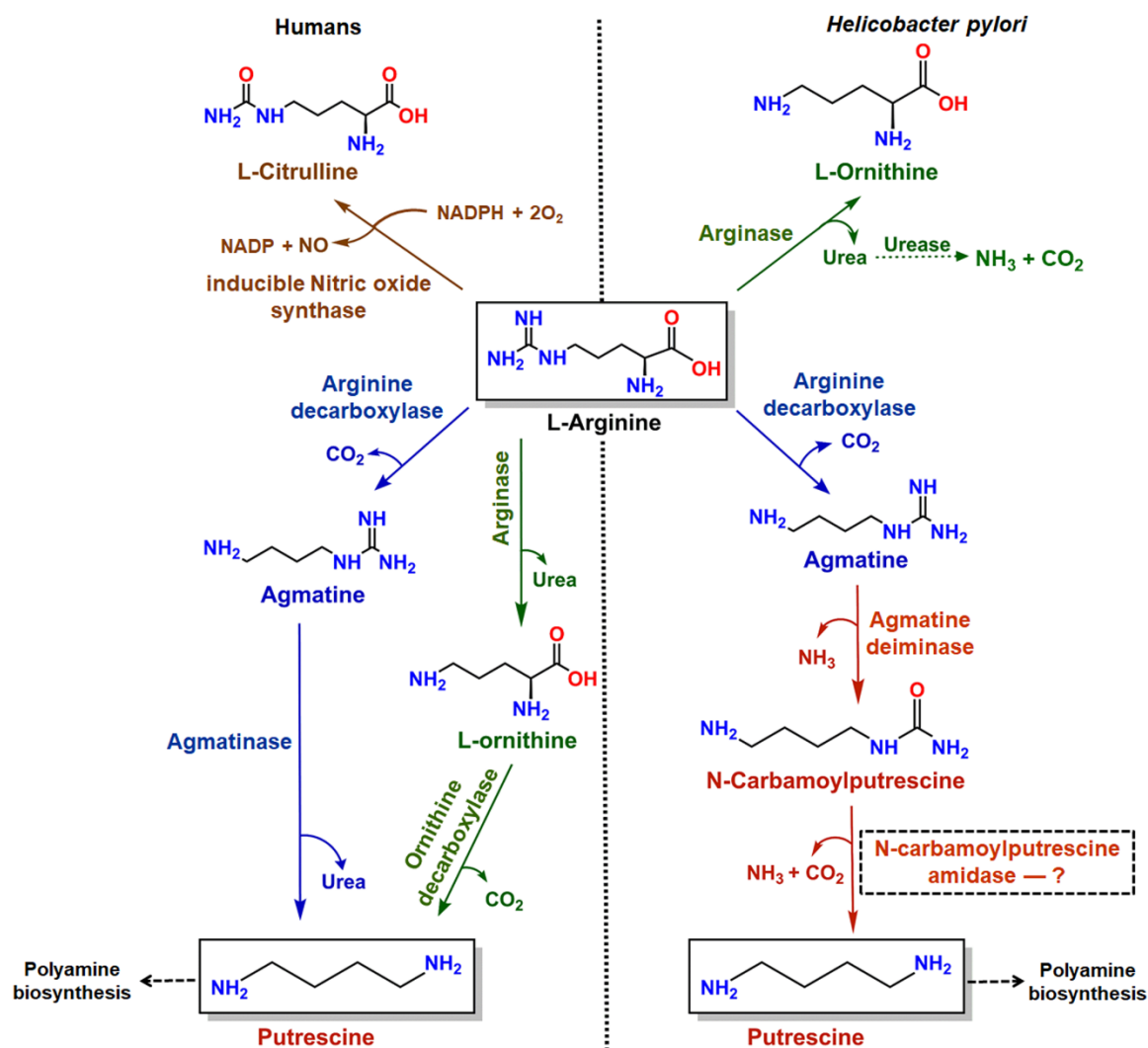


Figure 1. An outline of putrescine synthesis through L-arginine metabolism in humans and *H. pylori*. In humans, putrescine can be produced by two independent pathways using the enzymes arginase and arginine decarboxylase. In *H. pylori*, putrescine formation through arginase is not yet known.

As previously mentioned, the CPA gene in this bacterium has been annotated. Although analogous genes are present in other bacteria and plants, little is known with respect to their enzyme function (19–21). N-carbamoylputrescine amidase has been classified as a member of the nitrilase family, which hydrolyzes the non-peptide carbon-nitrogen bond (19, 25). Among the known CPAs, the crystal structure of the *Medicago truncatula* (model legume) homolog is

available (PDB 5H8I). Based on this crystal structure in complex with putrescine and a comparison of this structure with that of other known nitrilases, a mechanism for substrate hydrolysis by the plant CPA has been proposed using a set of putative catalytic residues (19).

Among the known *Helicobacter* gastric pathogens, *H. pylori* shows the highest incidence of infection (26). In this study, we report that *H. pylori* CPA hydrolyzes NCP to produce putrescine, ammonia, and carbon dioxide, thus establishing the missing link between arginine metabolism and polyamine biosynthesis in this bacterium. We discovered a unique tryptophan cluster located near the catalytic site and elucidated its role in catalytic function, which involves the formation of oligomers and their catalytically competent conformation. We also demonstrated the role of a set of three residues at the active site for enzyme function. Overall, this study provides new insights into the role of a conserved aromatic cluster involved in *H. pylori* CPA function, which will be useful for understanding the function of homologous enzyme. Additionally, the knowledge gained from this study can be exploited to design small molecule inhibitors against this *H. pylori* enzyme.

OBJECTIVES

The following objectives have been proposed for this study:

1. To find out whether the polyamine biosynthesis in *H. pylori* occurs through the metabolism of N-carbamoylputrescine by its CPA.
2. If so, to determine its three-dimensional structure, to identify the active site residues and also to study the mechanism of substrate hydrolysis.
3. To elucidate the role of a conserved tryptophan cluster located near the active site in catalytic function.
4. To screen and identify small molecule inhibitors against this enzyme

MATERIALS AND METHODS

Cloning and mutagenesis

Cloning of *H. pylori* CPA

The *aguB* gene (HP_RS03690) that encodes for CPA was amplified by Polymerase Chain Reaction (PCR) method from the genomic DNA of *H. pylori* 26695 strain using the appropriate forward and reverse primers (**Table S1**) having *Bam*HI/ *Xho*I restriction sites. The PCR product thus obtained was cloned into the bacterial expression vector, pPROEx-HTb that has a 6X-His tag. The clone was confirmed by sequencing.

Mutagenesis

The mutants of CPA were generated from the wild-type plasmid using HF-phusion polymerase (ThermoFisher Scientific) and *Pfu* Turbo (Agilent) according to the manufacture's protocol. Appropriate forward and reverse primers, listed in **Table S1** were used and the positive variants were identified using DNA sequencing.

Over-expression and purification of the recombinant proteins

The plasmid containing the CPA gene was transformed into *E. coli* BL21 (DE3) cells. A small volume of the starter culture was grown at 37°C, overnight in a sterile 2xTY medium containing 100 µg/mL of ampicillin. The culture was then transferred to a large volume of the ampicillin containing medium which was grown at 37°C until the value of OD₆₀₀ was 0.6. This was then induced with 0.5 mM isopropyl-b-D-thiogalactopyranoside (final concentration) and further grown at 30°C for 16 hours. The cells were harvested and resuspended in 50 mM Tris buffer (pH 8) which contained 500 mM NaCl, 1 mM phenylmethylsulfonyl fluoride (Sigma-Aldrich), 0.5% NP40, 0.05% β-mercaptoethanol and 10% glycerol at 4°C. The cell suspension was incubated with 1 mg/ml of lysozyme for 15 min and sonicated at 4°C for bursts of total 6 min. The lysate was centrifuged at 3500 × g for 60 min and the supernatant obtained was applied to a column containing 1 mL of Ni-NTA agarose (Qiagen). The column was then washed with a buffer containing 50 mM Tris (pH 8), 150 mM NaCl, 10 mM imidazole and 10% glycerol at 4°C. The elution of the wild-type protein was performed with an elution buffer (50 mM Tris, 100 mM NaCl and 10% glycerol, pH 7.2) containing 250 mM imidazole. Finally, size-exclusion chromatography was performed to purify the protein. HiLoadSuperdex 75 16/60 column (GE Healthcare) connected to AKTA FPLC system (GE Healthcare) was equilibrated with 50 mM potassium phosphate buffer (containing 10% glycerol, pH 7.2). The protein was collected, concentrated and stored at -80°C with 10% glycerol. Protein concentrations were measured spectrophotometrically by Bradford method. The mutants were also purified using this procedure.

Activity assay

The activity assays of the wild-type and mutants were carried out by the Berthelot and o-phthaldialdehyde method (OPA) (20, 27, 28). Briefly, 5 µM of the protein was applied to a reaction buffer containing 50 mM potassium phosphate buffer (pH 7.2). The substrate, NCP (NCP-Cl salt, custom synthesized from Merck) was added to the reaction mix (final concentrations, 250 µM for Berthelot method and 750 µM for OPA method) and incubated at 37°C for 1 min. The reaction was stopped using 5 mM sodium hypochlorite and 150 mM sodium hydroxide for the Berthelot method and 100 mM HCl for the o-phthaldialdehyde method. Finally, 60 mM sodium salicylate, 4 mM sodium nitroprusside and 0.15 mM EDTA were applied and ammonia formation was then

estimated at 698 nm, spectrophotometrically. For putrescine estimation, a mixture of OPA and β -mercaptoethanol was added to the reaction mixture and the complex formed was estimated at 340 nm using a spectrophotometer (UV450, Shimadzu, Japan).

For the steady-state kinetic assays, 5 μ M of the protein was used and the substrate concentration was varied at a final concentration of 0.05-2 mM. The rate of ammonia formed was estimated, spectrophotometrically.

The enzyme inhibition assay was performed using 5 μ M of the protein, 1.5 mM of NCP and 1 mM of the inhibitors. The rate of ammonia formed was estimated, spectrofluorometrically.

Circular dichroism measurements

Circular dichroism (CD) spectra of the wild-type and mutant proteins were recorded in the far-UV range in a spectropolarimeter (Chirascan, Applied Photophysics, Leatherhead, U.K.). The studies were performed at 25°C using 2 μ M protein in 50 mM potassium phosphate buffer, pH 7.2 placed in a 1-mm path length cuvette.

Steady-state fluorescence

Intrinsic tryptophan fluorescence measurements were carried out with 2 μ M protein in a spectrofluorometer (Fluoromax 4, Horiba, Japan) at an excitation wavelength of 295 nm. The emission wavelength was recorded between 310 to 450 nm. The monochromator slit width of 3 nm was used throughout. The buffer for the activity assay was used for this experiment.

Analytical size-exclusion chromatography

High performance liquid chromatography (HPLC) system (Perkin Elmer) was used to perform the analytical size-exclusion chromatography (SEC) measurements. A Phenomenex S-2000 column was pre-equilibrated with 50 mM potassium phosphate buffer, pH 7.2 at a flow rate of 1 mL/min and 30 μ M of 15 μ L protein was injected into it. A standard curve was generated from the elution volumes of proteins with known molecular mass ranging from 29 to 443 kDa. The molecular mass of the wild-type and mutant proteins were then determined from this curve.

Size-exclusion chromatography coupled with multi-angle light scattering

The absolute molecular mass of the wild-type and mutant proteins was determined using HPLC coupled with multi-angle light scattering (SEC-MALS) measurements. 60 μ M of the protein was injected into a WTC-030S5 column (Wyatt technology), which was connected to an HPLC system (LC-2030 Plus, Shimadzu). The eluant was then passed through an 18-angle (22.5 to 147°) DAWN light scattering instrument (Wyatt, Germany) and Optilab refractive index detector (Wyatt, Germany). The wavelength used for the laser in DAWN and light source in Optilab was 658 nm. ASTRA 7.3.2 software (Wyatt, Germany) was used to determine the molecular weight of the proteins using the light scattering and refractive index.

Molecular docking of wild-type protein with substrate NCP

The crystal structure of a carbon-nitrogen hydrolase from *H. pylori* G27 (PDB: 6MG6) was used as a template to construct the model structure of *H. pylori* 26695 CPA using MODELLER version 10.0 (University of California San Francisco, San Francisco, CA, USA) (29). The CPA structure was subjected to energy minimization using the GROMOS force field in order to remove conformational strain (30). The quality of the model was assessed using PROCHECK (EMBL-EBI, Hinxton, Cambridgeshire, UK) (31). The substrate, NCP was downloaded in 2D-structure data file from Pubchem Database. Substrate was prepared by energy minimization of 1000 steps

with the Ligprep module of Schrodinger (32). The active site was defined by a 10 Å radius sphere around the residues, Glu43, Lys115, and Cys152 of the *H. pylori* 26695 CPA pocket. Molecular Docking studies was performed with GOLD 5.1 software (The Cambridge Crystallographic Data Centre, Cambridge, UK) (33). Protein was kept as rigid and the number of Genetic Algorithm (GA) runs was set to 10 run per ligand with the default search algorithm parameters. Docking poses were evaluated and ranked according to the Gold score. Docking results were visualized by Discovery Studio (34). The Molecular dynamics simulations were performed using Gromacs-5.1.4 software. The MD snapshots were collected at the interval of 10 ps and analyzed by Gromacs utility programs.

RESULTS

The *aguB* gene encodes *H. pylori* CPA

The *aguB* gene (HP_RS03690) of the *H. pylori* 26695 strain is annotated for CPA. With an aim to check whether this enzyme is a CPA and to establish a missing link for the enzyme that catalyzes the conversion of NCP to putrescine through the ADC pathway in *H. pylori*, the *aguB* gene was cloned, over-expressed, and purified. On the basis of sodium dodecyl sulfate-polyacrylamide gel electrophoresis, its molecular mass was estimated to be near 35 kDa (theoretical mass of *H. pylori* CPA including the tag = 36.42 kDa).

N-carbamoylputrescine amidase is known to convert the substrate NCP to putrescine, ammonia, and carbon dioxide (20). With an anticipation whether the purified recombinant *H. pylori* protein exhibits CPA activity, we performed activity assays using a fixed concentration of substrate with varying concentrations of the enzyme (pH 7.4) at 37 °C, as previously described (20, 28, 35). We first estimated ammonia formation spectrophotometrically by measuring a colored complex at 698 nm, which increased with increasing concentrations of the enzyme (**Figure 2A**). We also measured the formation of other product putrescine, which was found to increase with increasing enzyme concentration (**Figure 2B**). These results suggested that the recombinant *H. pylori* protein was catalytically active. The data also implies that the enzyme is a CPA.

To kinetically characterize the enzyme, we performed steady-state kinetic experiments. In this assay, the substrate was used in excess over the enzyme so that multiple enzyme turnover kinetics could be studied. The assay was performed with a fixed concentration of enzyme and with varying concentrations of substrate. The data were fitted to a hyperbolic equation to obtain the kinetic parameters K_m (apparent affinity of the substrate to the enzyme) and k_{cat} (catalytic constant), which yielded the values $610 \pm 53 \mu\text{M}$ and $29 \pm 1 \text{ min}^{-1}$, respectively (**Figure 2C**). Among the known CPAs, the *Arabidopsis thaliana* homolog is kinetically characterized, with K_m and k_{cat} values reported to be $135 \mu\text{M}$ and 172 min^{-1} , respectively (20). Comparing these two kinetic parameters between the *H. pylori* CPA and the *Arabidopsis thaliana* homolog suggests the *H. pylori* enzyme has a 4.8-fold lower apparent substrate binding affinity and a 5.9-fold lower catalytic constant than its *A. thaliana* counterpart resulting in a nearly 32-fold decrease in catalytic efficiency (k_{cat}/K_m). Given the low sequence identity (~ 40%) of the proteins in these two organisms, the large decrease in the *H. pylori* enzyme's catalytic efficiency relative to the homolog could be associated with structural variation and/or with a difference in oligomer formation. The apparent substrate binding affinity for *H. pylori* CPA in sub-millimolar is not surprising because the arginine metabolic enzymes belonging to this pathogen (arginase and arginine decarboxylase) have evolved with the K_m values in millimolar (8.4 and 3.4 mM for arginase and arginine

decarboxylase, respectively) (16, 36–40). This could be relevant from an evolutionary perspective because *H. pylori* is a slow growing organism.

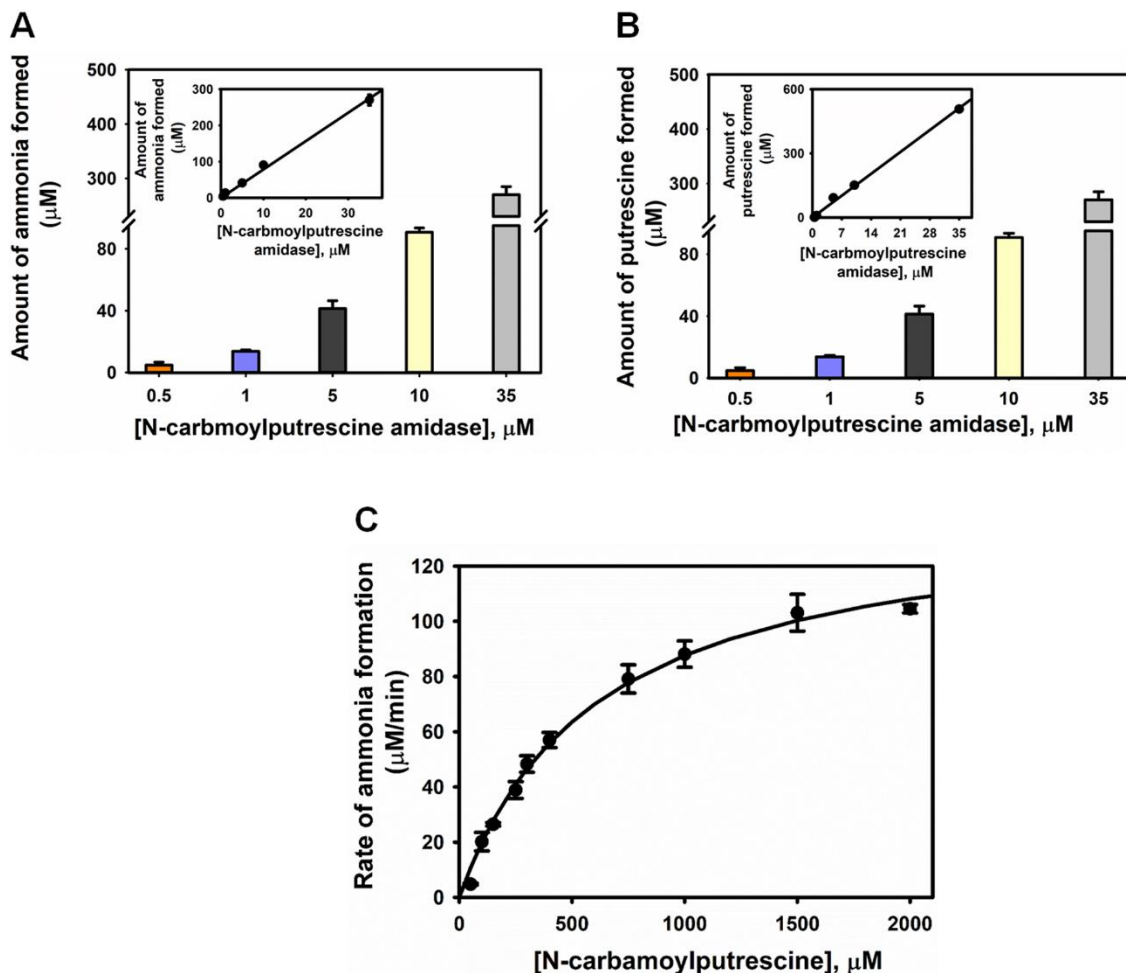


Figure 2. A-B) Plots of amount of product formed with increasing concentrations of *H. pylori* CPA. A) represents the ammonia formation and B) represents the putrescine production. The linear dependence of these products with varying concentrations of the enzyme is displayed in the inset. CPA stands for N-carbamoylputrescine amidase. C) Steady-state kinetic experiments using wild-type CPA. 5 μM of the protein was used and the substrate concentration was varied as shown in the figure. The data were fitted to a hyperbolic equation, $\text{rate} = k_{\text{cat}} \times [\text{E}_0] \times [\text{NCP}] / (K_{\text{m}} + [\text{NCP}])$ to determine the k_{cat} and K_{m} values. The quality of the fit was determined by a theoretical line drawn through the experimental data points. The data presented in these three plots are the average \pm SEM of three independent measurements.

A four-tryptophan cluster near the active site is vital for catalytic function

The 3D-structure of the *H. pylori* 26695 enzyme, which is not yet available, is essential to obtaining a structural insight into its function. Efforts were made to crystallize the protein and determine its 3D-structure. Although several crystals were found, only low-diffraction data sets were collected. Hence, we could not determine the structure. However, the crystal structure of a C-N hydrolase from the *H. pylori* strain G27 has recently been reported (PDB ID: 6MG6). A

sequence comparison of the *H. pylori* G27 C-N hydrolase with the *H. pylori* 26695 CPA (**Figure 3**) revealed that these two proteins share a high degree of identity (98%), indicating the *H. pylori* G27 protein is expected to be a CPA. In addition, the crystal structure of a plant homolog (*M. truncatula*, PDB: 5H8I) has been reported (19), but the plant protein displays low sequence identity with its *H. pylori* 26695 counterpart (40%; **Figure 3**). Therefore, we generated a model structure of an *H. pylori* 26695 CPA dimer using the crystal structure of the *H. pylori* G27 homolog (where the enzyme is a dimer) as a template. Despite the structure of the *H. pylori* G27 enzyme being available, no studies have been carried out to elucidate its function. Because CPA has been classified as a member of the nitrilase family, the location of the active site in the plant protein structure has been speculated on the basis of available nitrilases crystal structures. The plausible catalytic residues suggested are Glu48, Lys121, and Cys158 (19). However, their role in catalytic function has not been demonstrated. A sequence comparison of the plant CPA with other homologs revealed that the putative catalytic residues of the plant protein are conserved in the bacterial homologs including in *H. pylori* 26695 (**Figure 3**). An overlay of the crystal structure of the plant protein with the model structure of the *H. pylori* 26695 homolog reveals that the positioning of the putative catalytic residues in the *H. pylori* enzyme is comparable. This suggests that the analogous residues of the plant protein in the *H. pylori* homolog is expected to have similar function (study on the catalytic residues of the *H. pylori* enzyme is described later).

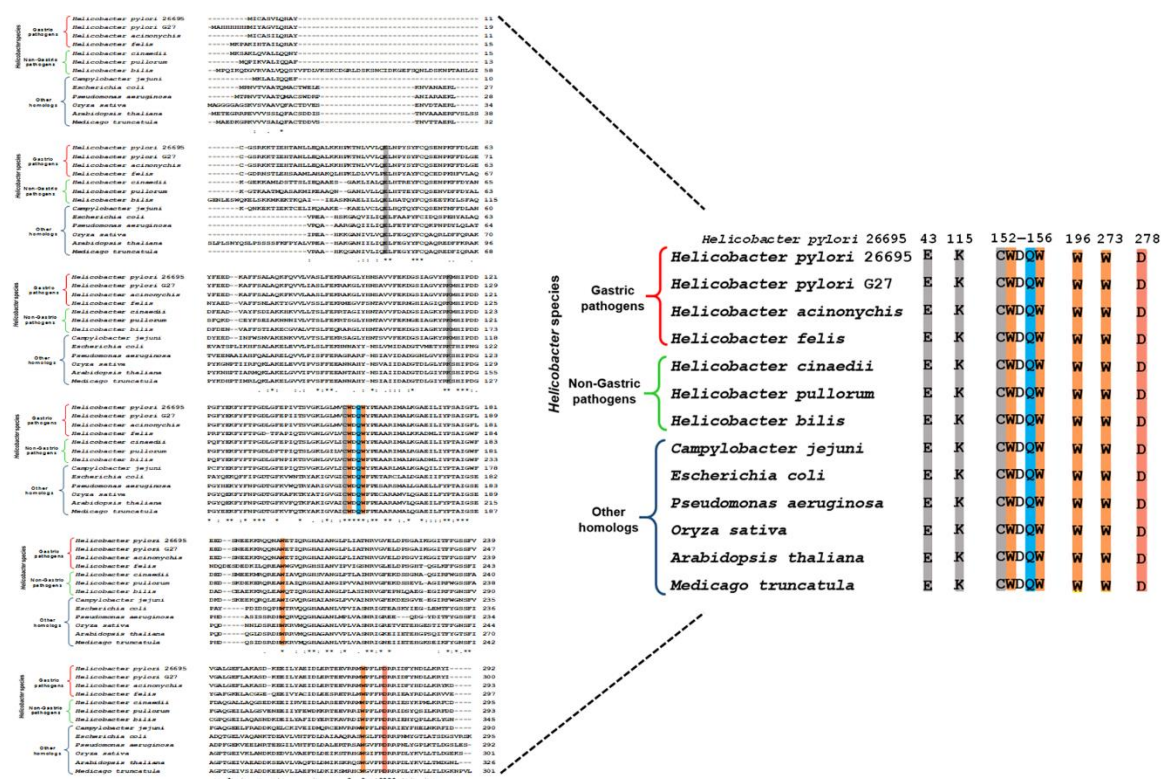


Figure 3. Sequence alignment of *H. pylori* 26695 CPA (accession number: HP_0757) with CPA from *H. pylori* G27 (accession number: B5Z7B9_HELPG), *H. acinonychis* (accession number: WP_104744497.1), *H. felis* (accession number: WP_199764868.1), *H. cinaedii* (accession number: WP_115025891.1), *H. pullorum* (accession number: WP_054194887.1), *H. bilis* (accession number: WP_077390008.1), *Campylobacter jejuni* (accession number: EAL1790463.1), *Escherichia coli* (accession number: MBL0992402.1), *Pseudomonas aeruginosa* (accession number: MBL0992402.1), *Oryza sativa* (accession number: MBL0992402.1), *Arabidopsis thaliana* (accession number: MBL0992402.1), and *Medicago truncatula* (accession number: MBL0992402.1).

(accession number: CRQ07783.1), *Oryza sativa* (accession number: EAY86155.1), *Arabidopsis thaliana* (accession number: NP_565650.1), and *Medicago truncatula* (accession number: G7ITU5_MEDTR). The catalytic residues Glu43, Lys115 and Cys152 are highlighted in gray. The four tryptophans shown in orange are fully conserved. The interacting residues, Gln155 (highlighted in cyan) and Asp278 (highlighted in red) of the tryptophans are also conserved. These residues of CPA among the organisms are enlarged at the right for clarity.

We wanted to explore the region beside these putative catalytic residues and whether any nearby amino acids could influence enzyme function. Thus, we examined the structure of the *H. pylori* 26695 CPA dimer and found a cluster containing four tryptophans (Trp153, Trp156, Trp196 and Trp273) located near the active site (**Figure 4**). A sequence alignment of *H. pylori* CPA with the other homologs revealed that these four tryptophans in the *H. pylori* enzyme are fully conserved in the homologs of plants and other bacteria (**Figure 3**). This suggests that the four tryptophans may have been selected during the evolution of the enzyme to perform a specific function. For many enzymes, it has been suggested that conserved residues clustered close to an active site may play a role in their structure and/or function (41, 42). In the *H. pylori* CPA dimer structure, Trp153, Trp156, and Trp196 are located in one of the monomers, whereas Trp273 is located in the other monomer (**Figure 4**). To explore the role of these residues in catalytic function, we individually mutated them to Ala and over-expressed and purified the variants. Interestingly, compared to wild-type, the activity of these four mutants was significantly reduced (1%, 2%, 3% and 6% of wild-type activity for the Trp196Ala, Trp156Ala, Trp273Ala and Trp153Ala mutants, respectively) with substrate concentration, which was 3-fold higher than the K_m value of the wild-type protein (**Figure 5A**). This suggests that the four tryptophans are individually crucial for catalytic function, with Trp196, Trp156, and Trp273 playing a slightly greater role than Trp153. We also performed steady-state kinetic assays for these four mutants. However, because they displayed such low catalytic activity, their kinetic parameters could not be reliably determined. Nevertheless, the data clearly suggest the importance of the cluster containing four tryptophans in catalytic function.

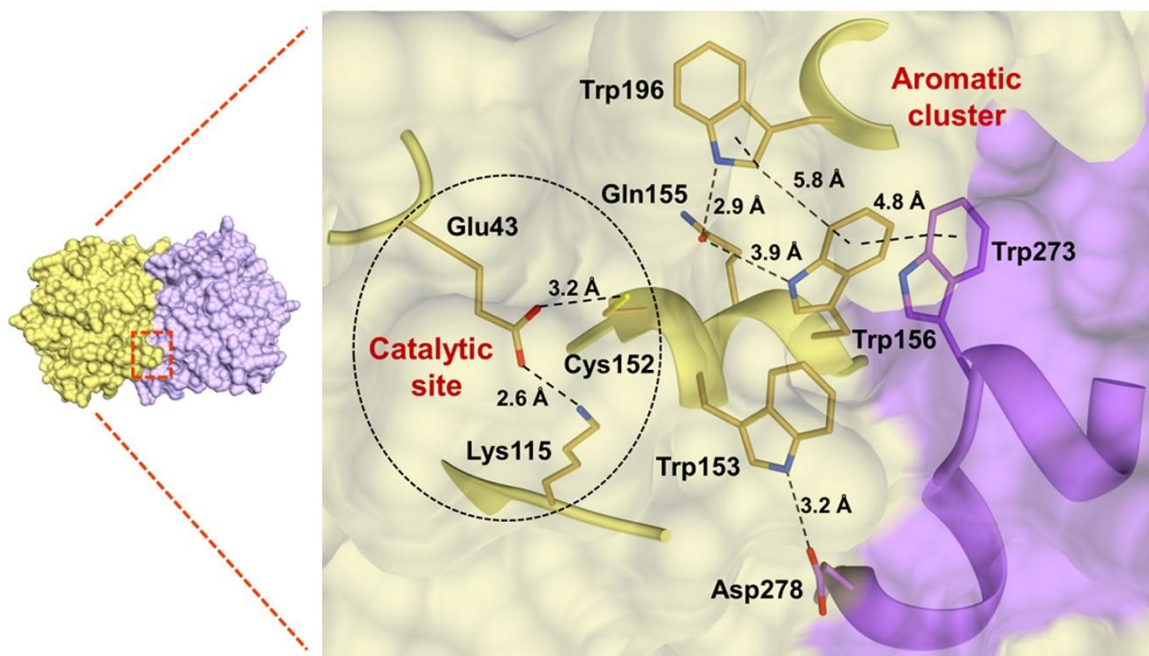
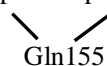


Figure 4. The left panel shows the model structure of *H. pylori* 26695 CPA. The site containing the catalytic residues and tryptophan cluster is shown in a red dotted box. The magnified view shows a cluster containing four tryptophans located close to the catalytic site. Out of them, Trp153, Trp156 and Trp196 are from the same monomer (coloured in khaki), while Trp273 originates from the other monomer (shown in purple). The indole -NH of Trp153 is at a distance of 3.2 Å from the carboxylate group of Asp278 located in the other monomer. The side chains of Trp156 and Trp273 from two different monomers are in close proximity. The distance of 4.8 Å was calculated from the centre of the aromatic rings. The indole -NH of Trp196 makes an H-bonding contact with the -CONH₂ group of Gln155 within the same monomer. The figure shows the network of interactions



involving . The -CONH₂ of Gln155 is at a distance of 3.9 Å from the indole -NH of Trp156. The distance between the two side chains of Trp156 and Trp196 was 5.8 Å (taken from the centre of the rings).

The cluster has a role in catalytic function through hexamer formation

The *M. truncatula* CPA exists as an octamer, whereas the *H. pylori* G27 homolog is a dimer. This suggests that CPA enzymes in different organisms can exist in different oligomers. It is not known whether the oligomer of these homologs is essential for catalytic function. To check whether the *H. pylori* 26695 homolog exists as an oligomer, and if so, whether these tryptophans play a role in catalytic function through oligomer formation, analytical size-exclusion chromatography (SEC) was performed with the wild-type and mutant proteins (**Figure 5B**). The wild-type protein eluted with a single peak corresponding to a molecular mass of nearly 230 kDa (theoretical molecular mass of hexamer including the tag= 218.5 kDa). This was close to the mass of a hexamer, suggesting that the *H. pylori* 26695 enzyme exists as a hexamer (**Table 2**). Unlike wild-type, the Trp153Ala and Trp273Ala mutants eluted with a peak that corresponds to a dimer. However, the Trp156Ala mutant eluted with a slightly higher retention volume that corresponds to a molecular mass which is also close to a dimer. The higher retention volume observed in the Trp156Ala mutant may be due to a change in conformation of the dimer through the disruption of interactions between Trp156 and other nearby residues. We also found that the Trp196Ala mutant eluted with a broader peak (peak maxima corresponding to a dimer) than the other Trp to Ala mutants, indicating that the Trp196 to Ala mutation leads to the presence of different oligomers in fast equilibrium. Nevertheless, these results clearly suggest the individual role of these four tryptophans in hexamer formation. We also found that the tryptophans are individually crucial for catalytic function. Taken together, we suggest that the significant loss in catalytic activity observed in these mutants is due to the loss of their hexamer formation.

Aromatic and H-bonding interactions of the tryptophans are essential for function through hexamer formation

We next wanted to obtain insight into the role of these tryptophans in catalytic function through hexamer formation. In the *H. pylori* 26695 CPA structure, Trp156 of one monomer is in close proximity with Trp273 of the other monomer. Similarly, Trp153 (same monomer as Trp156) makes contact with the carboxylate group of Asp278 of the other monomer through its indole -NH (distance 3.2 Å) and forms an H-bond. The indole ring of Trp196 makes H-bonding contact (distance 2.9 Å) with the -CONH₂ group of Gln155 within the same monomer (**Figure 4**).

H-bonding interaction between Trp153 and Asp278: To verify whether the significant decrease in catalytic activity found in the Trp153Ala variant (6% of wild-type activity) is due to the disruption of Trp153-Asp278 contact, we prepared the Trp153Phe (hydrophobicity at position 153 is reasonably restored) and Asp278Ala mutants. As expected, the Trp153Phe variant showed a considerable decrease in catalytic activity (8% of wild-type), as did the Asp278Ala mutant (6% of wild-type; **Figure 5A**). These results clearly suggest that rather than the hydrophobicity of Trp at position 153, it is the H-bonding interaction of Trp153 with Asp278 that plays a vital role in catalytic function. The observation of a dimer in the Trp153Ala mutant is likely to be because of the disruption of this Trp153-Asp278 interaction. To support this, we performed similar SEC assays with the Asp278Ala mutant, which also only formed dimers. To verify this further, we performed similar studies with the Trp153Phe variant, and it also eluted as a dimer (**Figure 5B**, **Table 2**). These results clearly suggest the importance of Trp153 in hexamer formation through its H-bonding interaction with Asp278 which in turn affects enzyme function.

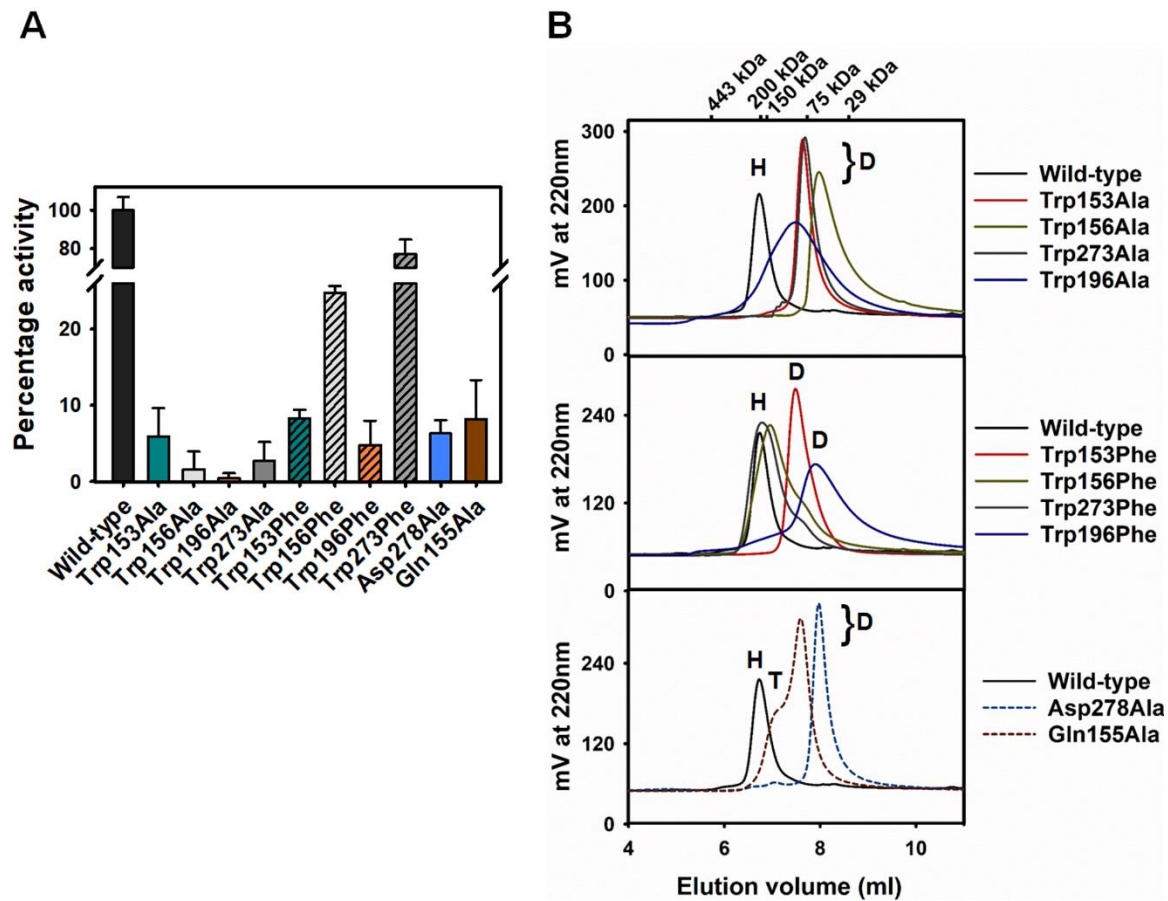


Figure 5. A) The percentage activity of the wild-type and its mutant proteins. The experiment was performed with 5 μ M protein at a substrate concentration which is three times higher than the K_m value of the wild-type enzyme. The values represented here are the average \pm SEM of three independent measurements. B) SEC profiles of wild-type and its mutants. H, T and D represents hexamer, tetramer and dimer, respectively.

Aromatic contact of Trp273-Trp156: To test whether the contact of Trp273 with Trp156 plays a role in catalytic function, we first prepared a Trp273Phe mutant, which reasonably restores the hydrophobicity of Trp at position 273. Interestingly, the activity of this mutant was much higher than that of the Trp273Ala variant (77% vs 3% of wild-type activity for the Trp273Phe and Trp273Ala mutants, respectively; **Figure 5A**). This suggests that the aromatic-aromatic contact between Trp273 and Trp156 plays a crucial role in catalytic function. To further support this, we next prepared a Trp156Phe mutant. Surprisingly, the activity of this variant was lower than that of the Trp273Phe mutant (24% vs 77% of wild-type activity for the Trp156Phe and Trp273Phe mutants, respectively; **Figure 5A**). This raises the possibility that Trp156 is in close contact with other nearby residue(s) in addition to its interaction with Trp273 (see the discussion). Together, these results indicate that the interaction of Trp273 with Trp156 plays a vital role in catalytic function.

To examine whether this aromatic contact plays a role in hexamer formation, we performed SEC measurements with the Trp156Phe and Trp273Phe mutants. Unlike the Trp156Ala and Trp273Ala variants (which only formed dimers), the corresponding Phe mutants displayed a mixture of dimer and hexamer, with hexamer being the major form (**Figure 5B, Table 2**). Thus, these results clearly suggest that the aromatic-aromatic contact between Trp156 and Trp273 is essential for catalytic function through hexamer formation. The lack of complete restoration of hexamer formation in these two Trp to Phe mutants is likely due to the less effective aromatic-aromatic contact between Trp and Phe vs Trp and Trp.

H-bonding contact of Trp196-Gln155: To verify the drastic decrease in catalytic activity observed in the Trp196Ala mutant (1% of wild-type activity) is likely due to the disruption of Trp196-Gln155 contact, we prepared a Trp196Phe variant, which also showed a significant decrease in catalytic activity compared to wild-type (5% of wild-type activity; **Figure 5A**). This result indicates that the above contact plays a vital role in catalytic function. To support this further, we prepared another mutant, Gln155Ala, which displayed activity similar to that of the Trp196Phe variant (Gln155Ala mutant showed 8% of wild-type activity; **Figure 5A**). The slightly greater loss of catalytic activity observed in the Trp196Ala mutant versus the Trp196Phe variant (1% vs 5% of wild-type activity for the Trp196Ala and Trp196Phe mutants, respectively) could be due to a local structural perturbation from the substitution of a bulky Trp with a small Ala residue at position 196. Taken together, these results clearly suggest that Trp196 plays a critical role in catalytic function through its interaction with Gln155.

It should be noted that Trp196 is not located at the monomer-monomer interface, but rather is present near Trp156, which is important for hexamer formation. (**Figure 4**). To our surprise, Trp196Ala mutant also formed dimers, suggesting a role for Trp196 in hexamer formation as well. As noted, Gln155 is crucial for catalytic function through its interaction with Trp196. Thus, to explore whether the Gln155 to Ala mutation has an impact on hexamer formation, we performed similar studies with the Gln155Ala mutant (**Figure 5B, Table 2**). Interestingly, this variant showed a mixture of dimers and tetramers with dimers being the predominant form, indicating a role for Gln155 in hexamer formation. The presence of tetramers in this variant may be due to a loss of interaction between Gln155 and surrounding residues other than Trp196 (see the discussion). The SEC data for the Gln155Ala mutant indicated that Gln155 plays a role in hexamer formation through its interaction with Trp196. To verify this, we performed similar measurements with the Trp196Phe variant, where the hydrophobicity of Trp at position 196 is reasonably restored. This variant mainly formed dimers with a smaller proportion of hexamers (**Figure 5B, Table 2**). These results clearly suggest that the contact between Trp196 and Gln155 is crucial for hexamer

formation, and the significant loss of catalytic activity observed in the Trp196Ala and Gln155Ala mutants is likely due to the defect in their hexamer formation. Altogether, the data of Trps and their interacting residues mutants suggest that hexamer of the wild-type protein is essential for enhanced catalytic function.

We also performed steady-state kinetic measurements for these Trp to Phe mutants and estimated their kinetic parameters (**Table 1**). In these mutants, the k_{cat} values were primarily decreased, with only a marginal effect on the magnitude of K_{m} compared to that of wild-type. However, for the Trp196Phe variant, the kinetic parameters could not be reliably determined because this mutant displayed significantly less catalytic activity (5% of wild-type). These data are in agreement with the results of the activity assays for the Trp to Phe mutants and further support the role of the individual tryptophans in catalytic function through hexamer formation.

Although our studies show the importance of the tryptophans, Gln155, and Asp278 in hexamer formation, on the basis of *H. pylori* CPA's dimer structure one might anticipate that individual mutations of these residues would lead to the formation of monomers. However, the dimerization of these mutants suggests that the hexamer formation is also mediated by the residues of other monomer-monomer interfaces.

Table 1. Steady-state kinetic parameter for the *H. pylori* wild-type and mutant CPA proteins

Protein	k_{cat} (min ⁻¹)	K_{m} (μM)	$k_{\text{cat}}/K_{\text{m}}$ (min ⁻¹ . μM ⁻¹)
Wild-type	28 ± 1	610 ± 53	0.046
Trp153Phe	3 ± 0.5	769 ± 246	0.004
Trp156Phe	6 ± 1	650 ± 230	0.009
Trp273Phe	26 ± 5	549 ± 150	0.047

Note: The values represented are the average ± SEM of three independent experiments.

Tryptophans have a marginal impact on secondary and local tertiary structure

We wanted to understand the separate roles of these four tryptophans in the secondary structure. To investigate this, we recorded the CD spectra of the Trp to Ala mutants in the far-UV region and compared the results with that of the wild-type (**Figure 6A**). The CD intensities of the four mutants were similar to that of wild-type, indicating that the four Trp residues do not play individual roles in the secondary structure.

To examine the effect of these Trp to Ala substitutions on local tertiary structure, we performed intrinsic tryptophan fluorescence measurements with the wild-type and the Trp to Ala mutants. It may be noted that this enzyme contains only these four tryptophans. As the emission maximum of free tryptophan is known to be approximately 355 nm, the 335 nm fluorescence emission maximum of the wild-type protein indicated that the tryptophan microenvironment in the

wild-type protein is non-polar (**Figure 6B**). The emission maxima of the four Trp to Ala mutants were blue shifted by 2-3 nm from that of wild type (332-333 vs 335 nm). This indicates that the microenvironment experienced by the tryptophans in the mutant proteins may be slightly more non-polar than that in the wild-type. Overall, the results of the CD and fluorescence measurements suggest that the individual tryptophans have a negligible effect on the protein's secondary and local tertiary structure.

It should be noted that the mutations of the residues interacting with the tryptophans, Gln and Asp to Ala at positions 155 and 278, respectively, also did not affect the secondary structure (**Figure 6A**). In the fluorescence assays, the emission maximum of the Gln155Ala mutant was similar to that of wild-type. However, the Asp278Ala variant was red shifted by nearly 3 nm, indicating that the microenvironment around the tryptophans may have become slightly polar upon the mutation of Asp278 to Ala (**Figure 6B**). These results suggest that the residues interacting with the tryptophans (Gln155 and Asp278) have also a marginal effect on the overall structure of the protein.

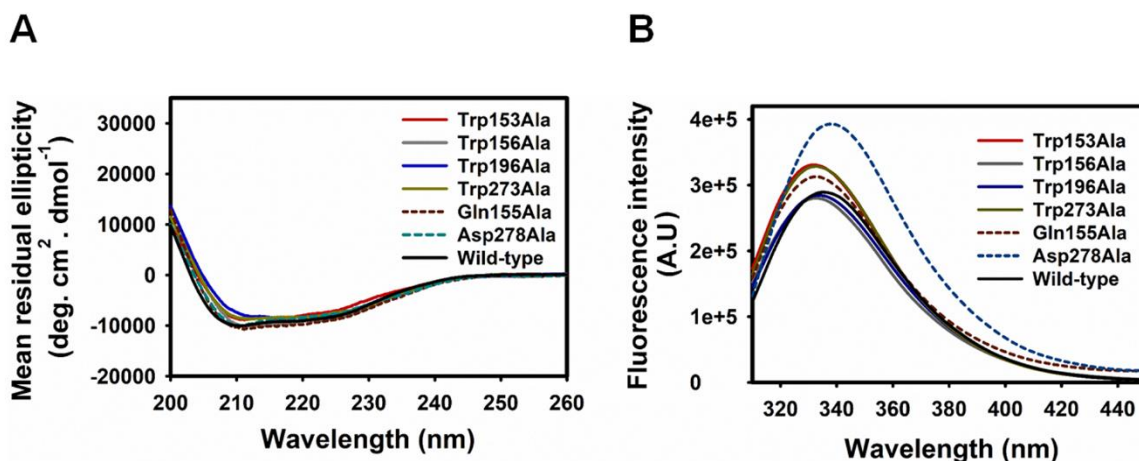


Figure 6. A) Circular dichroism measurements in the far-UV region and B) Tryptophan fluorescence emission spectra of the wild-type and the mutant proteins. The emission spectra were recorded upon excitation at 295 nm. The data shown here are the average of three independent scans.

Glu43, Lys115, and Cys152 are crucial for catalytic function

As previously described, the putative catalytic residues of the plant CPA are conserved in the *H. pylori* homolog. These residues in the *H. pylori* enzyme are Glu43, Lys115, and Cys152 (**Figure 7A**). However, neither the role of these three residues in *H. pylori* CPA nor the role of the analogous residues in plants and bacterial homologs has been demonstrated. To verify this in the *H. pylori* enzyme, we prepared three individual mutant constructs, Glu43Ala, Lys115Ala, and Cys152Ala, which were over-expressed and purified similar to the wild-type enzyme. The activity assays with these mutants showed they exhibited no detectable enzymatic activity, even with a substrate concentration 3-fold higher than the K_m value of the wild-type enzyme (**Figure 7B**). This suggests that Glu43, Lys115 and Cys152 are individually crucial for catalytic function.

To check the individual roles of these three residues in the secondary structure, the circular dichroism (CD) spectra of the wild-type and mutant proteins were examined (**Figure 7C**). The spectra of the three mutants were similar to those of wild-type, indicating that the Glu43, Lys115

and Cys152 residues do not play individual roles in the secondary structure. To examine the effect of these three residues on local tertiary structure, we performed fluorescence assays with the mutants (**Figure 7D**). The fluorescence emission maxima of the mutant proteins were similar to that of wild-type, suggesting the individual mutations of the three residues did not affect the tryptophan microenvironment. However, the change in intensity could also due to variation in the positioning of a surrounding protein matrix containing glutamic acid, aspartic acid, histidine, and/or cysteine, as these residues are known to quench the tryptophan fluorescence (43, 44). Overall, the CD and fluorescence results suggest that the three residues have a negligible effect on the protein's secondary and local tertiary structures.

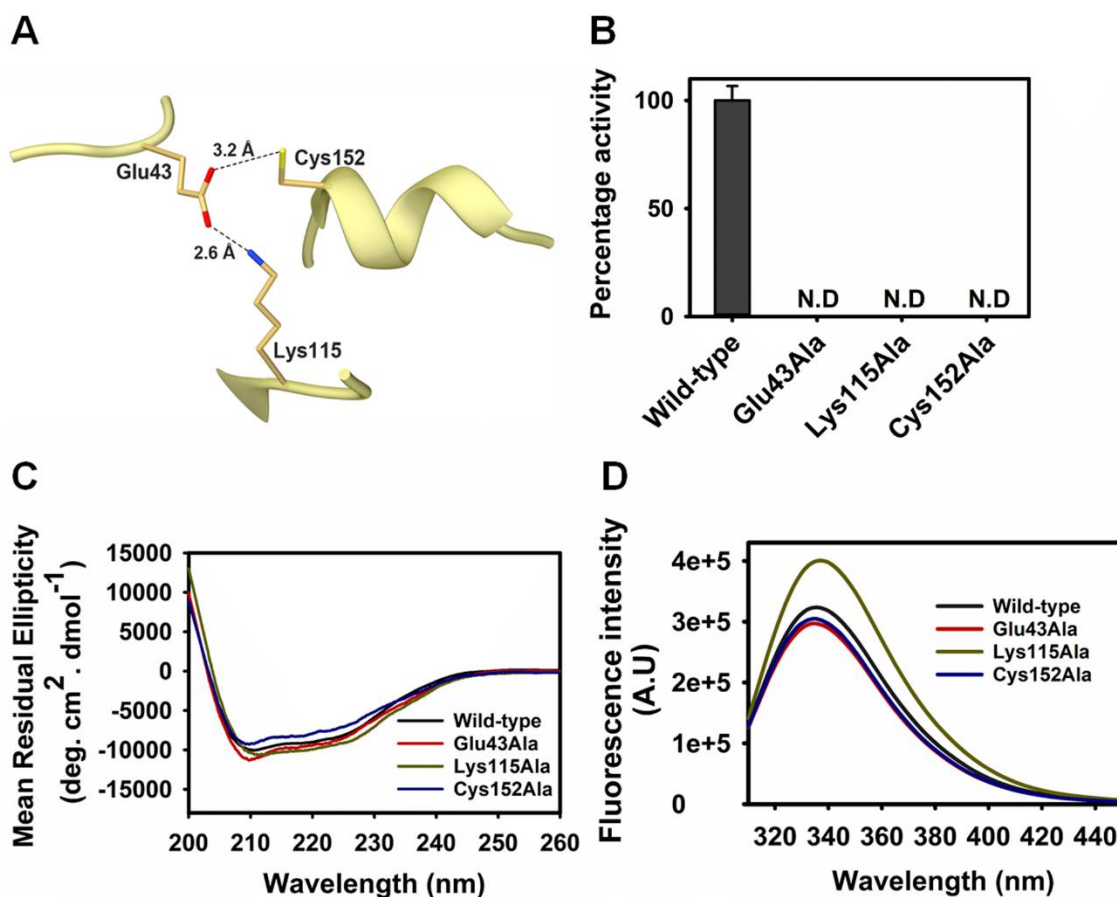


Figure 7. A) The *H. pylori* 26695 CPA structure showing the site containing the putative catalytic residues, Glu43, Lys115 and Cys152. They are in close proximity to each other. B) The percentage activity of the wild-type and mutants. N.D. represents no detectable activity. Wild-type activity $\leq 0.05\%$ was considered to be of no detectable activity. The average \pm SEM values of three independent measurements are represented here. C) Circular dichroism spectroscopy and D) Intrinsic tryptophan fluorescence studies of the wild-type and mutant proteins. The data presented in C) and D) are the average of three independent scans.

Glu43, Lys115, and Cys152 play a catalytic role

To check whether the loss of catalytic function in these three mutants was due to an alteration in substrate binding or related to a defect in catalysis, substrate binding measurements were

performed using intrinsic tryptophan fluorescence using their Ala mutants. Since they do not exhibit catalytic activity, the experiment was performed with a fixed concentration of the mutant proteins, but with varying concentrations of substrate. The change in fluorescence intensities with increasing substrate concentrations was fitted to the binding equation, $\Delta F = \Delta F_{\max} \times [\text{Substrate}] / (K_d + [\text{Substrate}])$ to determine the K_d , where ΔF and ΔF_{\max} represent the change in fluorescence intensity and maximum fluorescence, respectively, and K_d is the dissociation constant (**Figure 8A-C**). The K_d values for the Glu43Ala, Lys115Ala, and Cys152Ala mutants were 559 ± 34 , 125 ± 8 and 596 ± 56 μM , respectively. The K_d value for the wild-type protein could not be reliably estimated, because it hydrolyzes the substrate during the course of the titration experiment. Moreover, a non-hydrolyzable substrate analog is not commercially available. However, the K_m value of the wild-type protein is 610 ± 53 μM . Assuming this K_m value is comparable to the K_d value of the wild-type protein, these results indicate that the individual substitutions of the three residues with Ala did not hamper substrate binding. The large decrease in the K_d value for the Lys115Ala mutant suggests that the Lys115 to Ala mutation may have altered the local structure to favor substrate binding. Thus, the loss of catalytic activity is due to a defect in catalysis.

As already shown, the tryptophan cluster is essential for catalytic function through hexamer formation. Glu43, Lys115 and Cys152 are located close to this cluster. Therefore, we wanted to check whether the mutation of these residues to Ala have an impact on hexamer formation. Similar to wild-type, the three mutants (Glu43Ala, Lys115Ala and Cys152Ala) also eluted with peaks close to the size of a hexamer (**Figure 8D, Table 2**). This suggests that the individual mutations of the three residues to Ala did not affect hexamer formation. Collectively, these results suggest that Glu43, Lys115, and Cys152 play a catalytic role. The data further supports the importance of wild-type hexamer in catalytic function.

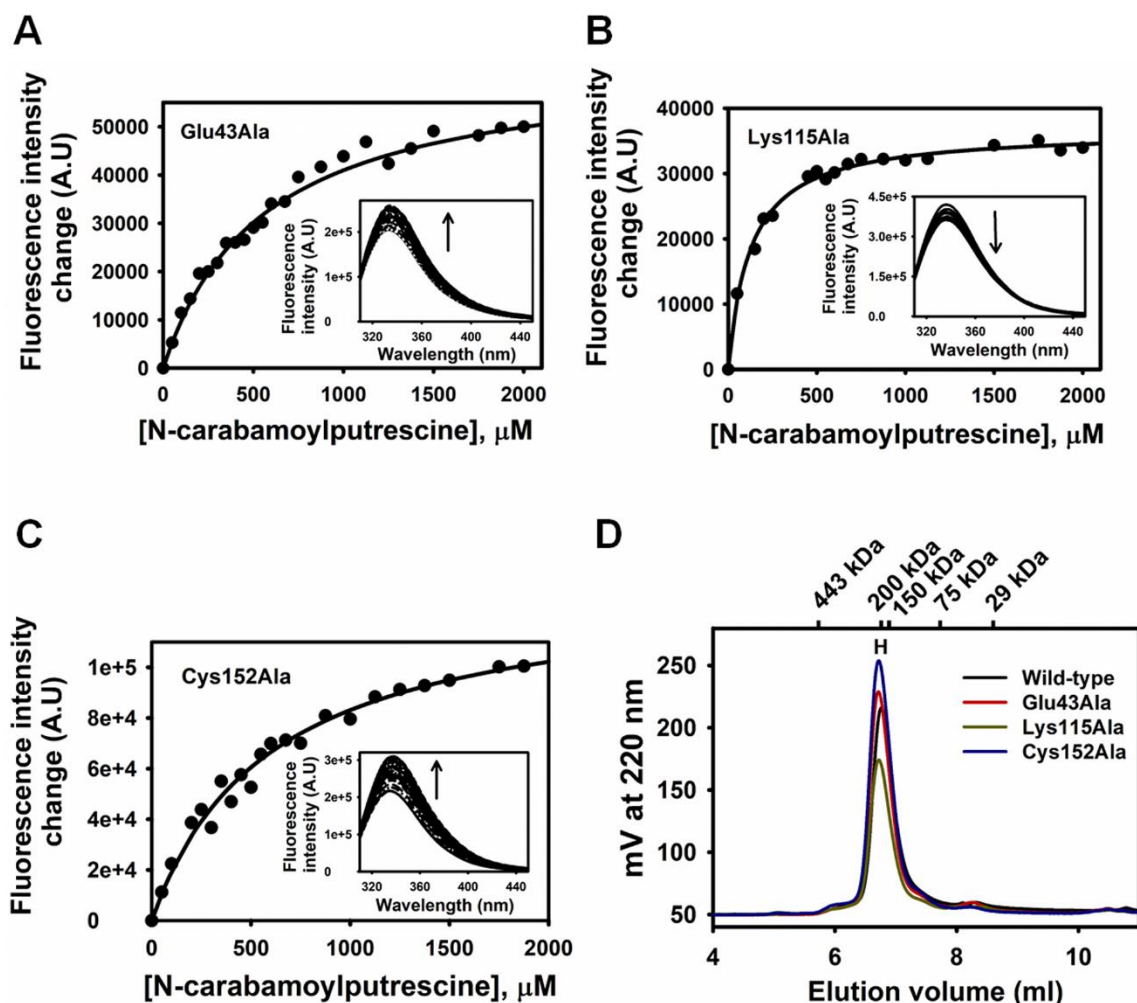


Figure 8. A-C) shows the change in tryptophan fluorescence of the mutants at 333 nm with increasing concentrations of the substrate, NCP, upon excitation at 295 nm using 2 μM protein. The data were fitted to a binding equation, $\Delta F = (\Delta F_{\text{max}} \times [\text{NCP}]) / (K_d + [\text{NCP}])$ to determine the K_d value, where ΔF and ΔF_{max} is the change and maximum change in fluorescence intensity, respectively. The inset shows the change in fluorescence intensity upon addition of the substrate. Three independent measurements were performed to obtain the average K_d value, but the plots shown in A-C is the representation of one measurement. D) Size-exclusion chromatography profiles of the wild-type and mutant proteins. Similar to wild-type, the mutants eluted as a hexamer, denoted as H.

SEC-MALS studies support the role of tryptophan cluster in hexamer formation

To further support the hexamer of wild-type protein and the role of tryptophan cluster in this oligomer formation, we performed analytical SEC coupled with multi-angle light scattering (SEC-MALS) measurements of the wild-type protein and Trp to Ala mutants (**Figure 9A, Table 2**). The molecular mass of the wild-type protein was estimated as $m_{\text{MALS}} = 213 \text{ kDa}$, which is in close agreement with the theoretical mass of a hexamer ($m_{\text{calculated}} = 218.5 \text{ kDa}$), further validating our finding that *H. pylori* CPA exists as a hexamer. The molecular masses of the Trp153Ala, and Trp273Ala mutants were estimated as 74, and 75 kDa, respectively, which strongly agrees with

the theoretical mass of a dimer ($m_{\text{calculated}} = 72.8$ kDa). However, Trp156 which interacts with Trp273, its mutant (Trp156Ala) displayed a higher retention time with a molecular mass of 65 kDa as compared to the Trp273Ala variant, which is in agreement with the SEC experiments. This suggests a different dimer conformation, possibly because of the loss of interaction of Trp156 with other nearby residues upon substitution of a bulky Trp with a less hydrophobic Ala. In the SEC assays, the Trp196Ala mutant eluted with a broader peak (peak maxima corresponding to dimers) than the other mutants, suggesting it forms different oligomers that are in fast equilibrium (**Figure 5A**). To check this, we performed SEC-MALS studies on the Trp196Ala variant, which showed a mixture of expanded dimer, trimer, and undecamer species ($m_{\text{MALS}} = 92, 128, 410$ kDa; the 92 kDa species was considered an expanded dimer, because the theoretical molecular mass of a dimer is 72.8 kDa), with dimer being the major form (**Figure 9A, Table 2**). The differences in the elution profiles of the mutant in the SEC-MALS studies versus the SEC measurements could be due to the difference in the column matrices used in the two experiments. The observation of a small fraction of undecamer in the Trp196Ala mutant may be due to the increased concentration of the protein used in the SEC-MALS experiments in comparison to the SEC studies. This result also indicates the role of Trp196 in maintaining the hexamer integrity possibly through its interaction with both Gln155 and Trp156 (see discussion).

The Gln155Ala mutant existed as a mixture of dimers and tetramers ($m_{\text{MALS}} = 71$ and 155 kDa, respectively), which is in agreement with the SEC studies. However, the difference in their proportion in these two experiments could be because of the variation in the protein concentration. These data also suggest that the oligomers of the Gln155Ala mutant are in dynamic equilibrium with each other. Similar to Trp153Ala mutant, its interacting residue mutant (Asp278Ala) eluted as a dimer ($m_{\text{MALS}} = 71$ kDa) (**Figure 9B, Table 2**). Overall, the SEC-MALS studies further validate our observations on the role of the tryptophan cluster in hexamer formation.

We also performed the similar experiments with the Glu43Ala, Lys115Ala, and Cys152Ala variants (**Figure 9C, Table 2**). The molecular masses of each of these mutants were determined as $m_{\text{MALS}} = 234, 239$ and 237 kDa, respectively, which is slightly higher than that of the wild-type. This could be due to the expanded form of a hexamer upon mutation of these individual residues. Our findings thus suggest that Glu43, Lys115, and Cys152 individually play a role in maintaining a compact hexamer.

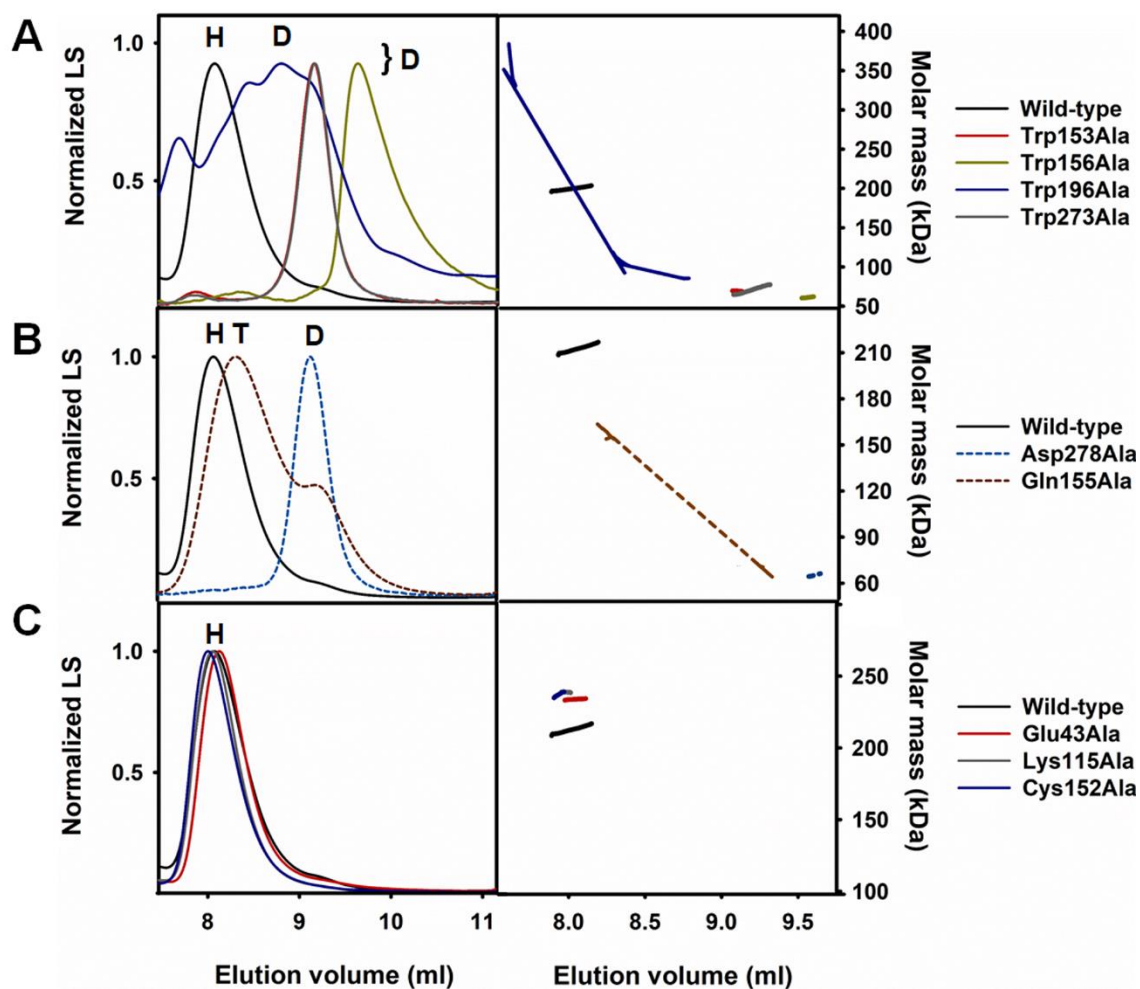


Figure 9. The left panel displays SEC-MALS profiles of wild-type CPA and its mutants. LS stands for light scattering. H, T and D represent hexamer, tetramer and dimer, respectively. The right panel shows the molecular weight traces of the corresponding proteins as obtained from MALS experiment. A flat trace indicates homogeneity as observed in cases of wild-type (black), Trp153Ala (magenta), Trp156Ala (yellow) and Trp273Ala (gray) mutants in Figure 9A. Molecular weight traces that are angled downwards suggest heterogeneity which could be due to dynamic equilibrium between different oligomers. This is visible in case of Trp196Ala (blue, Figure 9A) and Gln155Ala (brown, Figure 9B) variants. Asp278Ala (cyan, Figure 9B) and the three catalytic site mutants, Glu43Ala (red), Lys115Ala (green) and Cys152Ala (blue) in Figure 9C show peak homogeneity (45).

Table 2. Oligomeric states of the *H. pylori* wild-type and mutant proteins determined from SEC and SEC-MALS studies

Protein	Oligomeric state	
	SEC	SEC-MALS
Wild-type	Hexamer	Hexamer
Trp153Ala	Dimer	Dimer
Trp156Ala	Dimer	Dimer
Trp196Ala	Dimer	Expanded dimer (major), Trimer, Dodecamer
Trp273Ala	Dimer	Dimer
Asp278Ala	Dimer	Dimer
Gln155Ala	Tetramer, Dimer (major)	Tetramer (major), Dimer
Glu43Ala	Hexamer	Expanded hexamer
Lys115Ala	Hexamer	Expanded hexamer
Cys152Ala	Hexamer	Expanded hexamer

Note: SEC stands for size-exclusion chromatography and SEC-MALS stands for size-exclusion chromatography coupled with multi-angle light scattering.

Structure-based virtual screening, molecular docking and identification of small molecule inhibitor against *H. pylori* CPA

As previously discussed, CPA in *H. pylori* is a crucial intermediate enzyme, which establishes the missing link between arginine metabolism and polyamine biosynthesis through the ADC pathway in the bacterium. Moreover, the gene encoding for CPA (*aguB*) is absent in humans. Thus, we performed a structure-based virtual screening for the search of small molecule inhibitors against the *H. pylori* enzyme. We carried out docking studies of about 1,35,000 compounds (<http://www.lifechemicals.com/>) using the *H. pylori* 26695 CPA structure. The catalytic site and its surrounding region were targeted for the protein-ligand interaction. The compounds with favorable binding energy (< -8 kcal/mol) were selected for further studies.

Enzyme inhibition assay

To examine whether the shortlisted compounds can inhibit the *H. pylori* CPA activity, the activity assay of the enzyme was individually performed in the presence of these molecules and compared the results to that of the free protein. The rate of ammonia production was found to decrease by nearly 80% in the presence of 1000 μM of compound C3 (**Figure 10A**). However, at the same concentrations, compounds C4 and C5 failed to show any noticeable decrease in the activity. C1 and C2 individually displayed minor decrease in enzyme function (~20%). These data suggests that out of the five molecules, C3 can significantly inhibit the enzyme activity. Therefore, this inhibitor was used for further studies.

Determination of IC_{50} value of C3 against *H. pylori* CPA

We performed the experiment to determine the IC_{50} value. The activity was found to decrease with increasing inhibitor concentration with a nearly 80% maximum inhibition (**Figure 10B**). The IC_{50} value for C3 was estimated to be 100 μM . The micromolar range IC_{50} value suggests that C3 has moderate affinity towards the *H. pylori* enzyme.

Determination of nature of inhibition and estimation of K_i

To obtain an insight into the inhibitor binding site, we determined the nature of inhibition. **Figure 10C** shows that mechanistically this inhibition is turned to be competitive, and the K_i value was determined to be 20 μM . The micromolar range of the further supports our observation that compound C3 has moderate affinity towards *H. pylori* CPA.

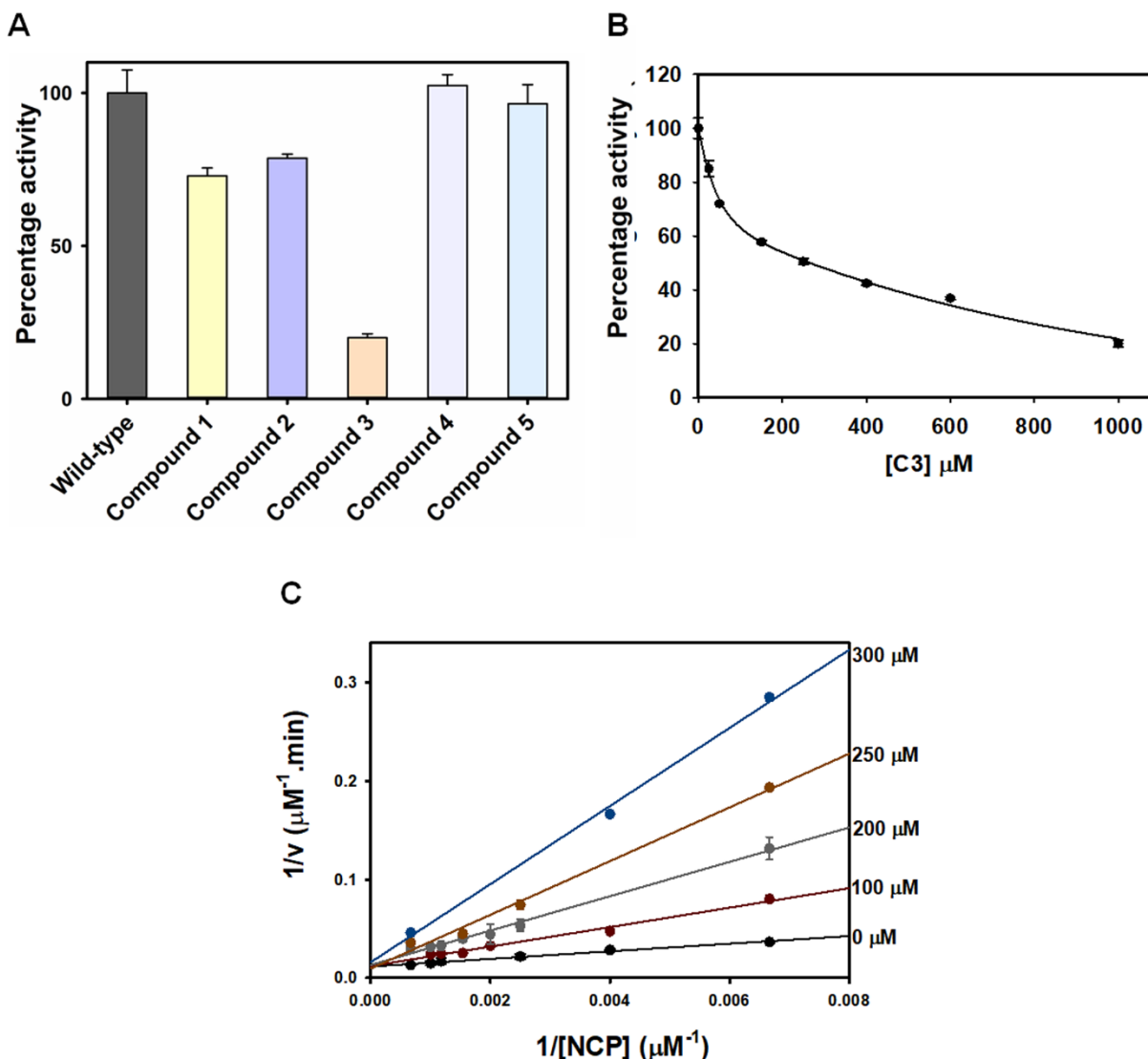


Figure 10. A) The percentage activity of the *H. pylori* CPA in the presence of 1000 μM of different compounds. The wild-type represents activity of the protein in absence of the molecules. B) A plot of percentage activity vs C3 concentrations to determine the IC_{50} value. C) Line-weaver Burk plot of the protein in presence of different concentrations of NCP and at various inhibitor concentration. The average \pm SEM values of three independent measurements are displayed in all the plots.

Molecular dynamics simulation studies

We next wanted to explore the region where the inhibitor can bind to the enzyme. We thus performed molecular dynamics simulations studies of the enzyme-inhibitor complex using the lowest energy structure obtained from molecular docking studies. The RMSD values of the complex was found to be stable during the course of simulation (**Figure 11A**). We analysed the 200 ns structure, which reveals that C3 binding site is located in a pocket that consists of residues, Cys152, Trp153 and Trp156. The aromatic ring of C3 is in contact with the side chain of Trp153 at a distance of 4.4 Å (**Figure 11B**). This could be the reason for reduced enzyme activity upon inhibitor binding.

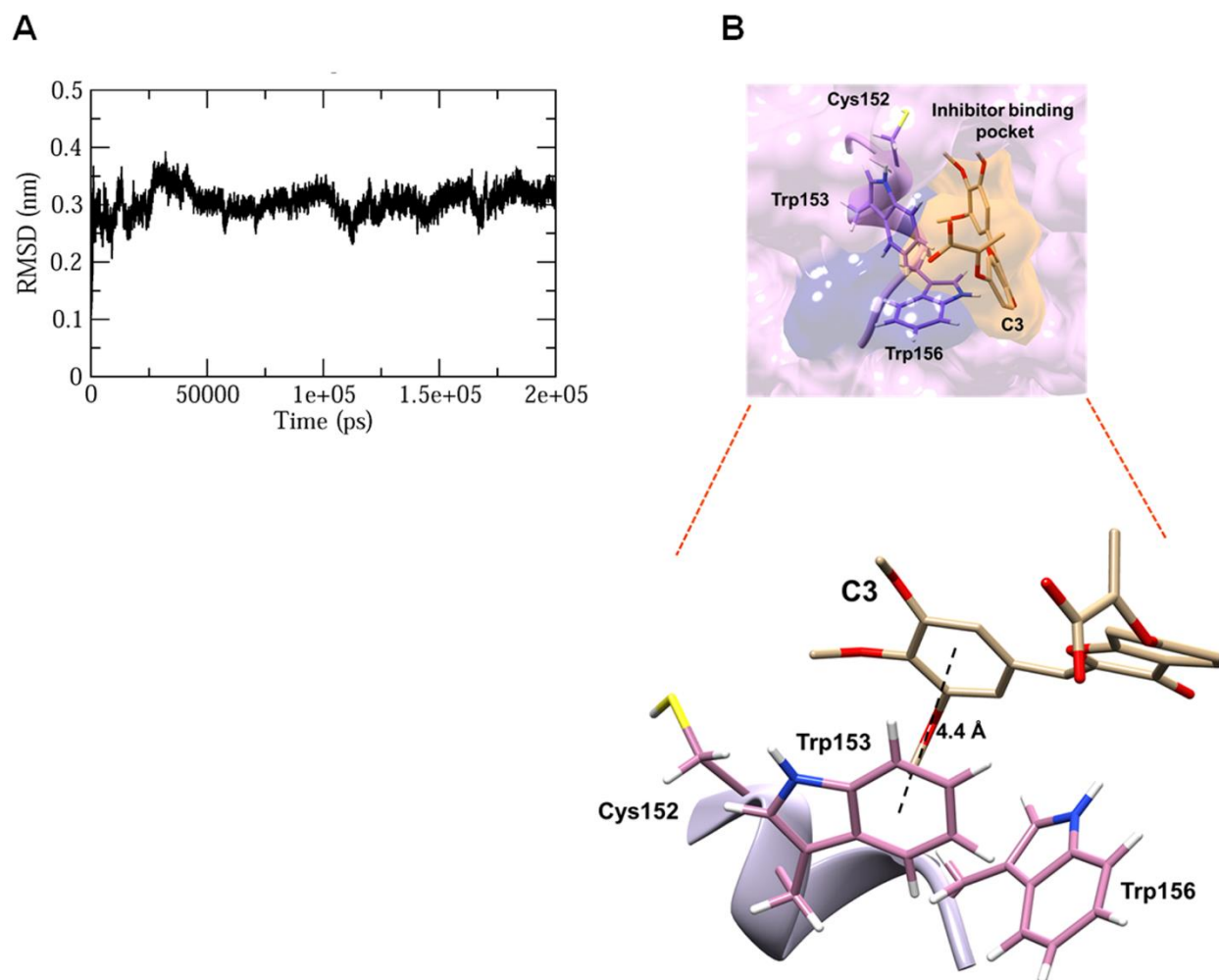


Figure 11. A) Root mean square deviation (RMSD) of *H. pylori* CPA in complex with the inhibitor C3 obtained during the MD simulations. B) The top panels show the inhibitor-docked CPA structure. C3 is found to bind in a pocket present near the catalytic Cys152. The enlarged view of this is displayed at the bottom panel. The aromatic ring of C3 is at a distance of 4.4 Å from the side chain of Trp153.

DISCUSSION

As previously mentioned, among the known *Helicobacter* gastric pathogens, *H. pylori* exhibits the highest incidence of infection (26). Although there are several studies on the polyamines and polyamine biosynthesis pathways in *H. pylori*, none have explored the metabolism of NCP to putrescine by CPA in this bacterium, despite it being a crucial intermediate step of polyamine synthesis. The present study not only establishes that CPA is the missing link for polyamine synthesis in *H. pylori*, but also provides new insights into its function, particularly with regard to the highly conserved tryptophan cluster located in the vicinity of the active site.

Our data clearly suggest a catalytic role for Glu43, Lys115 and Cys152. The next question to understand is how these three residues are crucial to catalysis. In the protein structure, these three residues are in close contact. As described above, the carboxylate group of Glu43 makes contact with the side chains of both Lys115 and Cys152 at distances of 2.6 and 3.2 Å, respectively. These three corresponding residues in the plant protein structure are similarly positioned. On this basis, we suggest that the network of interactions involving these three residues suitably positions the catalytic machinery so that substrate hydrolysis can occur. Thus, Glu43, Lys115 and Cys152 may act as a catalytic triad. To understand a possible mechanism of the hydrolysis reaction, we also performed docking studies on the *H. pylori* enzyme with NCP as the substrate (**Figure 12A**). The binding energy of the substrate to the enzyme was -6.3 kcal/mol. The docked structure revealed that the substrate binding site is located in a pocket with its carbonyl group 2.9 Å away from the -SH side chain of Cys152. This distance seems to favor a nucleophilic attack by Cys152 on the substrate amide group. On the basis of the positioning of the three catalytic residues, the location of the substrate in the docked *H. pylori* CPA structure, and the proposed mechanism of substrate hydrolysis by the *M. truncatula* homolog (19), we suggest an outline of substrate hydrolysis by the *H. pylori* enzyme as follows (**Figure 12B**). As found, Glu43 interacts with Cys152. This is likely to activate the -SH group of Cys152 for a nucleophilic attack on the carbamoyl moiety of the substrate for the formation of a covalent oxyanion intermediate, which could be stabilized by the adjacent positively charged Lys115 (distance between the Lys115 side chain and the carbamoyl group of the substrate is 3.2 Å). This will lead to the release of ammonia as the first product. The S-acyl reaction intermediate (second intermediate) is then hydrolyzed to form a third species. Finally, this is then further hydrolyzed to yield putrescine and carbon dioxide. However, additional studies are required to validate this mechanism.

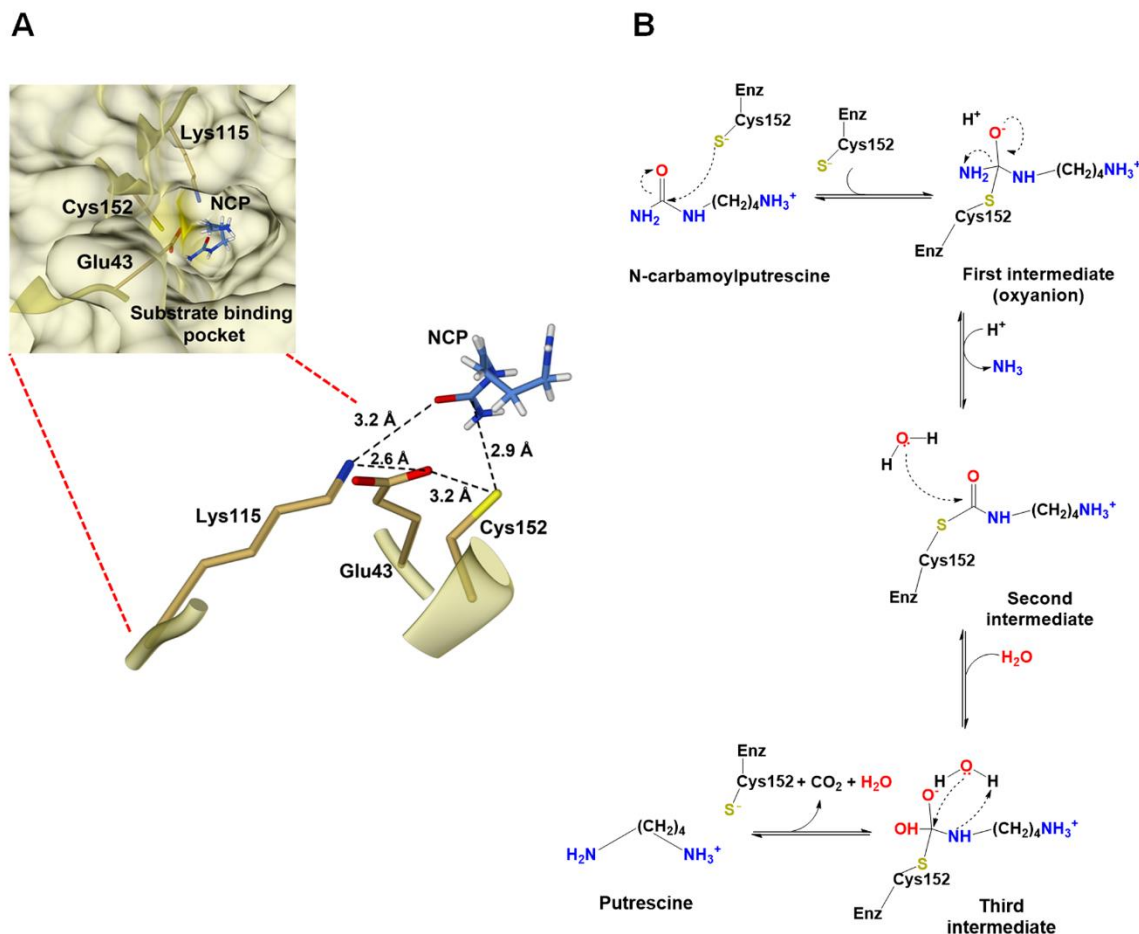
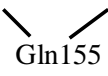


Figure 12. A) The substrate-bound docked structure of *H. pylori* 26695 CPA reveals that the substrate is in a binding site, which is shown at the top panel. The bottom panel shows the magnified view of the structure, where the -SH group of Cys152 is at a distance of 2.9 Å from the carbonyl group of the substrate, NCP. NCP also lies at a distance of 3.2 Å from the side chain of Lys115. B) An outline of the proposed mechanism of substrate hydrolysis by the *H. pylori* enzyme based on the docked structure and the known structure of *Medicago truncatula* CPA in complex with its product, putrescine (PDB: 5H8I).

The H-bonding and aromatic interactions of the tryptophans in the *H. pylori* enzyme have been shown to play a vital role in catalytic function through hexamer formation. Why does the hexamer exhibit significantly greater catalytic activity than the dimer? When looking the kinetic parameters of the Trp to Ala/Phe mutants, which exist as a dimer, only those of the Trp153Phe variant could be determined (the other mutants exhibited considerably less activity). We also found that the apparent substrate binding affinity of this mutant dimer was similar to that of the wild-type hexamer (**Table 1**). Taken together, these results suggest that the significantly lower catalytic activity observed in the Trp153Phe mutant dimer is not because of an alteration in substrate binding, rather due to a defect in catalysis. Thus, it appears that the organization of the three dimers in the hexameric wild-type protein generates a conformation that correctly positions the catalytic machinery for efficient catalysis.

The aromatic-aromatic contact between Trp156 and Trp273 at the monomer-monomer interface was found to be important for hexamer formation. Given that similar amounts of hexamer were observed, why then did the Trp156Phe mutant exhibit considerably less catalytic activity than the Trp273Phe variant (24% vs 77% of wild-type activity for the Trp156Phe and Trp273Phe mutants, respectively)? In the wild-type protein structure, the indole -NH of Trp156 is 3.9 Å away from the side chain carbonyl oxygen in Gln155. The Trp156 side chain is also nearly 5.8 Å away from the indole ring of Trp196 (**Figure 4**). These distances should be too far for Trp156 to make effective contact with both Gln155 and Trp196. However, it is possible that in the hexameric organization of the protein, these distances are decreased by the movement of individual Trp156-Trp273 and/or Gln155-Trp196 units towards each other, allowing Trp156 to make effective contact with both Gln155 and Trp196 without affecting interactions between Trp156 and Trp273.

Trp196-Trp156-Trp273

This may form a network of interactions involving  , which stimulates the catalytic function (**Figure 4**). Thus, our study seems to provide the new insight that hexamer formation alone is not sufficient for enhanced catalytic activity. Rather, this described network of interactions along with hexamer formation is essential to generating a catalytically competent conformation.

It may be noted that the three analogous tryptophans of *H. pylori* CPA (Trp153, Trp156 and Trp273) in the *M. truncatula* homolog were speculated to be involved in substrate recognition based on the crystal structure information. However, the individual mutations of these residues to Phe in the *H. pylori* enzyme did not hamper the apparent substrate binding affinity, suggesting that the analogous tryptophans in the plant homolog may have evolved to perform their role in catalytic function through oligomerization instead of substrate binding.

As reported, the four *H. pylori* CPA tryptophans, which are fully conserved in plant and other bacterial homologs, are important for catalytic function through hexamer formation. We also reported on the importance of two residues (Gln155 and Asp278) that interact with the tryptophans for catalytic function. Interestingly, these two residues of the *H. pylori* enzyme are also similarly conserved in the plant and other bacterial homologs (**Figure 3**). It should be noted that the Trps and their interacting residues (Gln155 and Asp278) do not have much impact on either the secondary or local tertiary structures. These observations suggest that the Gln155, Asp278, and four-tryptophan aromatic cluster in *H. pylori* have been selected during the evolution of the enzyme, not only to form a hexamer, but also to make it catalytically efficient. This perhaps explains why CPA in *H. pylori* has evolved with these six specific residues located near the active site to perform its catalytic function. In the plant protein structure, the analogous tryptophans and their interacting residues are also similarly positioned. Thus, the corresponding six residues in the homologs of other bacteria and plants including *M. truncatula* may have similar function.

Our study establishes how *H. pylori* CPA metabolizes the NCP derived from the formation of agmatine from arginine, which was a missing link in polyamine biosynthesis in this bacterium. We discovered a unique aromatic cluster containing four tryptophans as well as Gln155 and Asp278 residues that are specifically conserved and located near the catalytic site; each is vital for hexamer formation and hence catalytic function. We have also demonstrated the catalytic role of a set of fully conserved residues (Glu43, Lys115, and Cys152) and thus, they may act as a catalytic triad through the involvement of their network of interactions. The knowledge gained in this study should aid in understanding the function of CPA in other homologs. In addition, we targeted the

site containing catalytic residues and its vicinity to design small molecule inhibitor(s) specific to the *H. pylori* enzyme. Although the study identifies a new inhibitor with moderate affinity to this enzyme, there is a tremendous scope for designing inhibitors with greater efficacy based on the structure of C3.

IMPACT OF THE RESEARCH IN THE ADVANCEMENT OF KNOWLEDGE OR BENEFIT TO MANKIND

In India, more than 20 million people are accounted to suffer from peptic ulcer disease due *H. pylori* infections. At present, various therapeutic methods are adopted however, increasing resistance to these procedures contribute to poor eradication rates. In the WHO's list of antibiotic resistant priority pathogen, *H. pylori* falls into the high priority category (46). Thus, there is an urgent need to investigate new antibiotics for the treatment of *H. pylori* infection.

Helicobacter pylori CPA is a crucial intermediate enzyme for polyamine biosynthesis in the bacterium. The absence of CPA in humans makes its presence in *H. pylori* an attractive target for therapeutics. We identified a new and novel inhibitor of the bacterial enzyme having moderate affinity. It specifically binds to the site containing the catalytic residues and its nearby region. This novel inhibitor can be further exploited to design compounds with greater efficacy based on its structure. We also unearthed a four-tryptophan cluster as well as Gln155 and Asp278 that are conserved and located near the catalytic site. They are individually vital for enzyme function through hexamer formation. We elucidated the catalytic role of Glu43, Lys115 and Cys152, and they may act as a catalytic triad. These findings could help in understanding the CPA function in other homologs. Therefore, our study offers an opportunity to design small molecule inhibitors not only against the *H. pylori* enzyme but also against other bacterial CPAs.

LITERATURE REFERENCES

1. Hooi, J.K.Y., W.Y. Lai, W.K. Ng, M.M.Y. Suen, F.E. Underwood, D. Tanyingoh, P. Malfertheiner, D.Y. Graham, V.W.S. Wong, J.C.Y. Wu, F.K.L. Chan, J.J.Y. Sung, G.G. Kaplan, and S.C. Ng. 2017. Global Prevalence of *Helicobacter pylori* Infection: Systematic Review and Meta-Analysis. *Gastroenterology*. 153:420–29.
2. McGee, D.J., F.J. Radcliff, G.L. Mendz, R.L. Ferrero, and H.L.T. Mobley. 1999. *Helicobacter pylori* rocF is required for arginase activity and acid protection in vitro but is not essential for colonization of mice or for urease activity. *J. Bacteriol.*
3. Hovey, J.G., E.L. Watson, M.L. Langford, E. Hildebrandt, S. Bathala, J.R. Bolland, D. Spadafora, G.L. Mendz, and D.J. McGee. 2007. Genetic microheterogeneity and phenotypic variation of *Helicobacter pylori* arginase in clinical isolates. *BMC Microbiol.* 7:1–15.
4. Eaton, K.A., S. Suerbaum, C. Josenhans, and S. Krakowka. 1996. Colonization of gnotobiotic piglets by *Helicobacter pylori* deficient in two flagellin genes. *Infect. Immun.* 64:2445–58.
5. Yoshiyama, H., and T. Nakazawa. 2000. Unique mechanism of *Helicobacter pylori* for colonizing the gastric mucus. *Microbes Infect.* 2:55–60.
6. Marshall, B.J., and J.R. Warren. 1984. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet*. 323:1311–1315.
7. Uemura, N., S. Okamoto, S. Yamamoto, N. Matsumura, S. Yamaguchi, M. Yamakido, K. Taniyama, N. Sasaki, and R.J. Schlemper. 2001. *Helicobacter pylori* Infection and the Development of Gastric Cancer. *N. Engl. J. Med.* 345:784–789.
8. Correa, P. 1992. Human Gastric Carcinogenesis: A Multistep and Multifactorial Process—First American Cancer Society Award Lecture on Cancer Epidemiology and Prevention. *Cancer Res.* 52:6735–6740.
9. 1994. Schistosomes, liver flukes and *Helicobacter pylori*. In: IARC monographs on the evaluation of carcinogenic risks to humans. Lyon, France: World Health Organization, International Agency for Research on Cancer.
10. Michael, A.J. 2016. Polyamines in eukaryotes, bacteria, and archaea. *J. Biol. Chem.* 291:14896–14903.
11. Michael, A.J. 2016. Biosynthesis of polyamines and polyamine-containing molecules. *Biochem. J.* 473:2315–2329.
12. Gogoi, M., A. Datey, K.T. Wilson, and D. Chakravorty. 2016. Dual role of arginine metabolism in establishing pathogenesis. *Curr. Opin. Microbiol.* 29:43–48.
13. McGee, D.J., J. Zabaleta, R.J. Viator, T.L. Testerman, A.C. Ochoa, and G.L. Mendz. 2004. Purification and characterization of *Helicobacter pylori* arginase, RocF: Unique features among the arginase superfamily. *Eur. J. Biochem.* 271:1952–1962.

14. Morris, S.M. 2010. Arginine: Master and commander in innate immune responses. *Sci. Signal.* 3:1–5.
15. Nathan, C., and M.U. Shiloh. 2000. Reactive oxygen and nitrogen intermediates in the relationship between mammalian hosts and microbial pathogens. *Proc. Natl. Acad. Sci. U. S. A.* 97:8841–8848.
16. Dutta, A., M. Mazumder, M. Alam, S. Gourinath, and A.K. Sau. 2019. Metal-induced change in catalytic loop positioning in *Helicobacter pylori* arginase alters catalytic function. *Biochem. J.* 476:3595–3614.
17. Valenzuela, M., A. Cáceres, O. Almarza, D. Bravo, S. Soto, O. Cerda, and H. Toledo. 2014. Characterization of the Arginine Decarboxylase Gene (ORF HP0422, speA) Involved in Acid Tolerance in *Helicobacter pylori*. *Helicobacter.* 19:182–193.
18. Jones, J.E., C.P. Causey, L. Lovelace, B. Knuckley, H. Flick, L. Lebioda, and P.R. Thompson. 2010. Characterization and inactivation of an agmatine deiminase from *Helicobacter pylori*. *Bioorg. Chem.* 38:62–73.
19. Sekula, B., M. Ruszkowski, M. Malinska, and Z. Dauter. 2016. Structural investigations of N-carbamoylputrescine amidohydrolase from *Medicago truncatula*: Insights into the ultimate step of putrescine biosynthesis in plants. *Front. Plant Sci.* 7:350.
20. Piotrowski, M., T. Janowitz, and H. Kneifel. 2003. Plant C-N hydrolases and the identification of a plant N-carbamoylputrescine amidohydrolase involved in polyamine biosynthesis. *J. Biol. Chem.* 278:1708–1712.
21. Nakada, Y., Y. Jiang, T. Nishijyo, Y. Itoh, and C.D. Lu. 2001. Molecular characterization and regulation of the aguBA operon, responsible for agmatine utilization in *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 183:6517–6524.
22. Iyer, R.K., H.K. Kim, R.W. Tsoa, W.W. Grody, and S.D. Cederbaum. 2002. Cloning and characterization of human agmatinase. *Mol. Genet. Metab.* 75:209–218.
23. Dufe, V.T., D. Ingner, O. Heby, A.R. Khomutov, L. Persson, and S. Al-Karadaghi. 2007. A structural insight into the inhibition of human and *Leishmania donovani* ornithine decarboxylases by 1-amino-oxy-3-aminopropane. *Biochem. J.* 405:261–68.
24. Takaji, S., H. Hibasami, I. Imoto, Y. Hashimoto, K. Nakao, N. Ikemura, Y. Taguchi, and K. Nakashima. 1997. Growth inhibition of *Helicobacter pylori* by a polyamine synthesis inhibitor, methylglyoxal bis(cyclopentylamidino)hydrazine. *Lett. Appl. Microbiol.* 25:177–180.
25. Bork, P., and E. V. Koonin. 1994. A new family of carbon-nitrogen hydrolases. *Protein Sci.* 3:1344–1346.
26. Solnick, J. V., and D.B. Schauer. 2001. Emergence of diverse *Helicobacter* species in the pathogenesis of gastric and enterohepatic diseases. *Clin. Microbiol. Rev.* 14:59–97.
27. Patton, C.J., and S.R. Crouch. 1977. Spectrophotometric and Kinetics Investigation of the

Berthelot Reaction for the Determination of Ammonia. *Anal. Chem.* 49:464–69.

28. Qi, X., W.F. Wang, J. Wang, J.L. Yang, and Y.P. Shi. 2018. Highly selective colorimetric detection of putrescine in fish products using o-phthalaldehyde derivatization reaction. *Food Chem.* 259:245–250.
29. Šali, A., and T.L. Blundell. 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* 234:779–815.
30. Hess, B., C. Kutzner, D. Van Der Spoel, and E. Lindahl. 2008. GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4:435–447.
31. Laskowski, R.A., J.A.C. Rullmann, M.W. MacArthur, R. Kaptein, and J.M. Thornton. 1996. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR.* 8:477–486.
32. Didi-Huberman, G., G. Boehm, G. Brandstetter, B. Stiegler, M. Sedlaczek, and K. Stierle. 2018. Protein Preparation Wizard; Epik. LigPrep. *Schrödinger Release 2018-2.* 71–76.
33. Jones, G., P. Willett, R.C. Glen, A.R. Leach, and R. Taylor. 1997. Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* 267:727–748.
34. Accelrys Software Inc. 2013. Discovery studio modeling environment, release 4.0. .
35. Davis, D.R., A.G. Fogg, D. Thorburn Burns, and J.S. Wragg. 1974. A colorimetric method for the determination of phenacetin and paracetamol part iii. Studies of the indophenol reaction: an alternative manual procedure for the determination of phenacetin and paracetamol and its application to the determination of other ph. *Analyst.* 99:12–18.
36. Srivastava, A., and A.K. Sau. 2010. Biochemical studies on *Helicobacter pylori* arginase: Insight into the difference in activity compared to other arginases. *IUBMB Life.* 62:906–915.
37. Srivastava, A., S.K. Meena, M. Alam, S.M. Nayeem, S. Deep, and A.K. Sau. 2013. Structural and functional insights into the regulation of *helicobacter pylori* arginase activity by an evolutionary nonconserved motif. *Biochemistry.* 52:508–519.
38. Alam, M., A. Srivastava, A. Dutta, and A.K. Sau. 2018. Biochemical and biophysical studies of *Helicobacter pylori* arginine decarboxylase, an enzyme important for acid adaptation in host. *IUBMB Life.* 70:658–669.
39. Dutta, A., D. Sarkar, P. Murarka, T. Kausar, S. Narayan, M. Mazumder, S.R.K. Ainavarapu, S. Gourinath, and A.K. Sau. 2021. An evolutionary non-conserved motif in *Helicobacter pylori* arginase mediates positioning of the loop containing the catalytic residue for catalysis. *Biochem. J.* 478:871–894.
40. George, G., M. Kombrabail, N. Raninga, and A.K. Sau. 2017. Arginase of *Helicobacter Gastric Pathogens* Uses a Unique Set of Non-catalytic Residues for Catalysis. *Biophys. J.* 112:1120–1134.

41. Matsui, E., J. Abe, H. Yokoyama, and I. Matsui. 2004. Aromatic Residues Located Close to the Active Center Are Essential for the Catalytic Reaction of Flap Endonuclease-1 from Hyperthermophilic Archaeon *Pyrococcus horikoshii*. *J. Biol. Chem.* 279:16687–96.
42. Ronnebaum, T.A., S.M. Gardner, and D.W. Christianson. 2020. An Aromatic Cluster in the Active Site of epi-Isozizaene Synthase Is an Electrostatic Toggle for Divergent Terpene Cyclization Pathways. *Biochemistry*. 59:4744–54.
43. Ghisaidoobe, A.B.T., and S.J. Chung. 2014. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: A focus on förster resonance energy transfer techniques. *Int. J. Mol. Sci.* 15:22518–22538.
44. Chen, Y., and M.D. Barkley. 1998. Toward understanding tryptophan fluorescence in proteins. *Biochemistry*. 37:9976–9982.
45. Sarkar, P., N. Akhavantabib, and S.D. Arcy. 2020. Comprehensive analysis of histone-binding proteins with multi-angle light scattering. *Methods*. 184:93–101.
46. Tshibangu-kabamba, E. *Helicobacter pylori* infection and antibiotic resistance — from biology to clinical implications. *Nat. Rev. Gastroenterol. Hepatol.* 0123456789.

Ditsa Sarkar

Ditsa Sarkar
Ph.D Scholar
Protein Engineering Lab
National Institute of Immunology
New Delhi-110067
India
Email: ditsasarkar@gmail.com