The gene cluster for agmatine catabolism of *Enterococcus faecalis*. Studies of recombinant putrescine transcarbamylase and agmatine deiminase and a snapshot of agmatine deiminase catalyzing its reaction

Running title: *E. faecalis* agmatine deiminase operon and structure

José L. Llácer,† Luis Mariano Polo,† Sandra Tavárez,† Benito Alarcón, Rebeca Hilario and Vicente Rubio*

Instituto de Biomedicina de Valencia (IBV-CSIC), C/ Jaime Roig 11, 46010-Valencia, Spain

Correspondent footnote

*Dr. Vicente Rubio

Instituto de Biomedicina de Valencia (IBV-CSIC)

C/ Jaime Roig 11

46010-Valencia, Spain

Phone: +34 96 339 17 72

Fax: +34 96 369 08 00

E-mail: rubio@ibv.csic.es

†J.L.L., L.M.P, & S.T. contributed equally to this work

Abstract

Enterococcus faecalis makes ATP from agmatine in three steps catalyzed by agmatine deiminase (AgDI), putrescine transcarbamylase (PTC) and carbamate kinase (CK). An antiporter exchanges putrescine for agmatine. We have cloned the *E. faecalis ef0732* and *ef0734* genes of the reported gene cluster for agmatine catabolism, overexpressing them in *Escherichia coli,* purifying the products, characterizing them functionally as PTC and AgDI, and crystallizing and X-ray diffracting them. The 1.65-Å-resolution structure of AgDI forming a covalent adduct with an agmatine-derived amidine reactional intermediate is described. We provide definitive identification of the gene cluster for agmatine catabolism and confirm that ornithine is genuinely a poor PTC substrate, suggesting that PTC (found here to be trimeric) evolved from ornithine transcarbamylase (OTC). N-(Phosphonoacetyl)-putrescine was prepared and shown to strongly (K_i=10nM) and selectively inhibit PTC and to improve PTC crystallization. We find that *E. faecalis* AgDI, which is committed to ATP generation, closely resembles the AgDIs involved in making polyamines, suggesting the recruitment of a polyamine-synthesizing AgDI into the AgDI pathway. The arginine deiminase (ADI) pathway of arginine catabolism probably supplied the genes for PTC and CK, but not for the agmatine/putrescine antiporter, and thus, the AgDI and ADI pathways are not related by a single *"en bloc"* duplication event. The AgDI crystal structure reveals a tetramer with a 5-blade propeller subunit fold, proves that AgDI closely resembles ADI despite lack of sequence identity, and explains substrate affinity, selectivity, and Cys357-mediated covalent catalysis. A three-tongued agmatine triggered gating opens or blocks access to the active center.

Introduction

In addition to the fermentation of carbohydrates, *Enterococcus faecalis* (formerly *Streptococcus faecalis*) is able to use arginine and the decarboxylated derivative thereof, agmatine, as an energy source for growth (8,10,45,48,49). Arginine and agmatine are metabolized via the arginine deiminase (ADI) and agmatine deiminase (AgDI) pathway, respectively. Both metabolic routes are very similar and include the sequential action of three enzymes (48,49) and one antiporter (11) that are analogous in the two pathways. Arginine and agmatine, respectively, are deiminated by ADI (EC 3.5.3.6) and AgDI (EC 3.5.3.12), yielding citrulline and carbamoyl putrescine, which are phosphorolyzed by ornithine transcarbamylase (OTC; EC 2.1.3.3) and putrescine transcarbamylase (PTC; EC 2.1.3.6), generating carbamoyl phosphate for use in ADP phosphorylation by pathway-specific carbamate kinase (CK; EC 2.7.2.2) isozymes, producing one ATP molecule (48,49). The resulting ornithine and putrescine are exchanged with external arginine or agmatine by an arginine/ornithine antiporter in one pathway and an agmatine/putrescine antiporter in the other pathway (11).

Possibly no microbial species has been more important for the biochemical characterization of the ADI and AgDI pathways than *E. faecalis*. It was in this microorganism where both pathways were originally demonstrated (20,24,45,50), the corresponding enzymatic steps characterized and shown to be coordinately induced by arginine or agmatine (48,49), respectively, the enzymes except AgDI purified (31,32,42,56), and CK (the ADI pathway isozyme) crystallized and its structure determined at atomic resolution (29,30). Despite the abundance of biochemical information, there was little genetic information on these routes in *E. faecalis* until we sequenced and determined the gene structure, organization and some regulatory features for the components of the ADI pathway (3). However, in the case of the AgDI pathway,

there was for very long time no other genetic information than the observation that three mutant strains of *E. faecalis* that were unable to use agmatine were devoid of either AgDI activity, PTC activity, or both (48). The loss in one mutant of the two enzymes and the triggering by agmatine of coordinated increases in the levels of AgDI and PTC appeared consistent with the physical association of the genes for these two enzymes within the same operon, as is the case for the genes for the ADI pathway (3,48,49). Only recently, after the identification in *Pseudomonas aeruginosa* of the gene *aguA* (38), encoding the AgDI that is involved in putrescine and polyamine biosynthesis in plants and microorganisms that decarboxylate arginine (2) (not the case of *E. faecalis*), a putative *aguA* gene was identified in the cariogenic organism *Streptococcus mutans* (17), and, by sequence similarity, in *E. faecalis* (gene *ef0734* of *E. faecalis* V583 genome, TIGR database; http://www.tigr.org). In both species this gene is preceded by the genes for a putative antiporter and for a transcarbamylase (in *E. faecalis* V583, genes *ef0733* and *ef0732*, respectively) and is followed by a putative gene for carbamate kinase (*ef0735*; Fig. 1A). Thus, this gene cluster would contain the genes for all the catalysts required for operation of the AgDI pathway, and, indeed, a polar disruption of the first gene in this cluster of *S. mutans* decreased strongly AgDI activity measured in permeabilized cells, as expected for the AgDI operon (17). Further, the amino acid sequence predicted to be encoded by *ef0732* coincides with the N-terminal sequence reported long ago for *E. faecalis* PTC (39,54).

Nevertheless, the ultimate test for ascribing specific functions to genes, the cloning of the gene, its expression and the purification and functional characterization of the corresponding gene product, has not been published for the genes for the AgDI pathway. As a consequence of an independent effort to identify and characterize the AgDI pathway genes, we describe here the cloning of the *E. faecalis ef0732* and *ef0734* genes, their overexpression in *Escherichia coli* and the purification of the corresponding protein products, the enzymatic characterization of these

products as PTC and AgDI, and their crystallization and X-ray analysis, reporting also the crystal structure at high resolution of *E. faecalis* AgDI containing a covalently bound derivative of agmatine at the active center. Our results not only confirm conclusively the nature of the operon, but they are the first that characterize functionally an AgDI committed to fermentative ATP production [all previously well characterized examples are involved in polyamine biosynthesis, and only one is bacterial, from *Pseudomonas aeruginosa* (23,37,59)] revealing also the structure of this enzyme during catalysis. For the other gene product studied here, PTC, previously characterized from a single source (56), *E. faecalis*, we demonstrate that the bisubstrate analog for this enzyme, N-(phosphonoacetyl)-putrescine (PAPU), is a highly selective and very powerful (K_i=10 nM) PTC competitive inhibitor, relative to carbamoyl phosphate, clarifying the substrate binding order in this enzyme. This inhibitor is proven here to be crucial for obtaining good diffracting crystals of PTC, opening the way for crystal structure determination and for clarification of the structural bases for PTC specificity for putrescine. The finding of clear structural similarities between the ADI and AgDI folds indicates that these enzymes, which do not exhibit significant sequence similarity, are homologous. On the basis of this finding we propose a potential mechanism for the evolutionary relations between the ADI and AgDI operons.

Materials and methods

Bacterial growth and characteristics. *E. faecalis* SD10 was grown overnight at 37ºC, without shaking, in medium A (49) supplemented with 25 mM glucose. This strain is highly similar to *E. faecalis V583*, judged from previous studies on the ADI operon (3), and also from the comparison of the present sequences determined here for genes *ef0732* and *ef0734*, which have revealed, of a total of 2130 bases, only 6 trivial base differences relative to the corresponding sequence of the *E. faecalis* V583 genome, none of them causing any amino acid change. Genomic DNA was isolated according to a standard procedure for bacteria (57).

Cloning and expression in *E. coli* of *ef0732* and *ef0734*. *ef0732* and *ef0734* were PCR-amplified from genomic DNA from *E. faecalis* SD10, by utilizing a high-fidelity thermostable DNA polymerase (Deep Vent; New England Biolabs) and the primer pairs 5'^{689512}AGGAGGAACACCATA*TGAAAAGAGATTAC*^{689540} and 5'^{690565}AATCAGTGGAAGCTTGGCCG*TTAAATGC*^{690538}, for *ef0732*, and 5'^{691969}GAACGAAAGCATA*TGGCTAAACGAATTG*^{691996} and 5'^{693106}ATCACTATTTTTGAATTCTGTTTCCCTC*C*^{693078}, for *ef0734*, where the first and third of these primers correspond to the coding strand and the second and fourth to the complementary strand, the superscript numbers are the coordinates in the TIGR database for the *E. faecalis* genome, the underlining indicates mutated bases, and cursive lettering identifies nucleotides belonging to the open reading frame to be amplified. These primers were designed to introduce a *Nde*I site at the initiator ATG codon and *Hin*dIII and *Eco*RI sites 6 or 11 nucleotides downstream of the stop codon of *ef0732* or *ef0734*, respectively. The PCR products, digested with *Nde*I-

130 *Hin*dIII or *Nde*I-*Eco*RI, were inserted directionally in the corresponding sites of plasmid pET-22b
131 behind the promoter recognized by T7 DNA polymerase. The resulting plasmids, isolated from
132 transformed *E. coli* DH5α cells grown in Luria-Bertani (LB) medium containing 0.1 mg ml^{-1} of
133 ampicillin, were mutated at the translation termination codon using the QuickChange™ site-
134 directed mutagenesis kit (from Stratagene) and the oligonucleotide pairs
135 $^{5'}$CAAAGCATTTCAGCGGCCAAGCTTG$^{3'}$ and $^{5'}$CTTGGCCGCTGAAATGCTTTGAGTG$^{3'}$
136 for *ef0732*, to replace the translation termination codon by serine and to introduce and extra G
137 after this mutated codon; and the pair
138 $^{5'}$GAACCAAAGCGCGTAGGAGGGAAACAGAATTCG$^{3'}$ and
139 $^{5'}$CTGTTTCCCTCCTACGCGCTTTGGTTCTTGTTG$^{3'}$ for *ef0734*, to introduce a G before the
140 translation termination codon. These mutations abolish termination at the normal stop codon and
141 place in frame the plasmid sequence for incorporating at the cloned protein C-terminus a linker
142 and a 6-His sequence. In this way, the *ef0732* and *ef0734* gene products include, respectively, the
143 16- and 24-amino acid C-terminal extensions SAAKLAAALEH_6, and
144 VGGKQNSSSVDKLAAALEH_6. The mutant plasmids, isolated from transformed DH5α cells
145 and confirmed by sequencing to carry the correct constructions, were used to transform *E. coli*
146 BL21(DE3) cells. After growth of the cells at 37°C (*ef0732*) or 30ºC (*ef0734*) in liquid LB
147 medium supplemented with 0.1 mg ml^{-1} of ampicillin until a turbidity at 600 nm of 0.6 to 0.7 was
148 attained, 0.1 mM isopropyl ß-D-thiogalactopyranoside (IPTG) was added, and the culture was
149 continued for 3-4.5 additional hours before the cells were harvested by centrifugation. All
150 subsequent purification steps were carried out at 0-4ºC.
151

Purification of the product of the cloned *ef0732* gene. The cells were suspended in 1/100 of the original culture volume of 20 mM K-phosphate, pH 7.4, containing 10 mM putrescine and 20 mM imidazole, they were broken by sonication (four pulses of 30 s each; MSE Soniprep 150 fitted with the standard probe), and the sonicate was centrifuged at 15,800 × g for 10 min. The supernatant was loaded onto a 5-ml His-trap Ni-affinity column (Amersham Biosciences) mounted on an ÄKTA fast protein liquid chromatography system (FPLC, Amersham Biosciences), equilibrated and run at 1 ml min^{-1} with 50 mM K-phosphate, pH 7.0, containing 20 mM imidazole. The column was washed with the same buffer until the optical absorption of the effluent returned to baseline, and then a 100-ml linear gradient of 20 to 500 mM imidazole in 50 mM K-phosphate, pH 7, was applied and 3-ml fractions were collected. Fractions containing the essentially pure protein (monitored by SDS-PAGE and Coomassie staining) were pooled, concentrated to ~20 mg ml^{-1} by centrifugal ultrafiltration (Amicon Ultra 30K device, from Millipore), 20 % (v/v) glycerol was added, and the protein was stored at -20ºC.

Purification of the product of the cloned *ef0734* gene. The purification was as for the product of the *ef0732* gene except for: 1) the utilization as cell suspension buffer, of 50 mM K-phosphate, pH 7, containing 1 mM dithiothreitol (DTT) and 50 mM phenylmethylsulfonyl fluoride; 2) the inclusion of 1 mM DTT in all the solutions; 3) the use with the His-trap step of a 25-ml gradient; and 4) the incorporation of two additional purification steps as follows. The fractions (2 ml each) of the first His-trap column step containing the purer protein (SDS-PAGE monitoring) were pooled, concentrated, and placed in 50 mM K-phosphate, pH 7.0, 1 mM DTT, 0.5 M NaCl, by repeated centrifugal ultrafiltration and then were subjected to repurification through the 5-ml His-trap column as in the first step except for the inclusion in all the solutions of 0.5 M NaCl. The fractions containing the purer protein were concentrated again and freed from imidazole by

centrifugal ultrafiltration, and were subjected to size-exclusion chromatography (~10 mg per injection to the column) on a Superdex200 HR 10/30 column (Amersham Biosciences) mounted on an ÄKTA FPLC system equilibrated and run at 0.25 ml min^{-1} using a solution of 50 mM K-phosphate, pH 7.0, 1 mM DTT and 0.5 M NaCl. The fractions containing the essentially pure protein were pooled, concentrated to ~20 mg ml^{-1}, and placed in Tris-HCl 50 mM pH 7.4, 0.5 M NaCl, 1 mM DTT, by centrifugal ultrafiltration, and were then supplemented with 10% (v/v) glycerol and stored at -20ºC.

Enzyme activity assays. AgDI and PTC activities were assayed at 37ºC by the production of carbamoyl putrescine, determined colorimetrically at 465 nm in an assay for ureido groups (40) based on the Archibald procedure (1). The color yield of carbamoyl putrescine in this color reaction (24,320 M^{-1} cm^{-1}) was estimated after complete conversion of agmatine to carbamoyl putrescine using a large excess of AgDI, and was found to be 25 % higher than the color yield of citrulline. The AgDI assay mixture (33) contained 50 mM EDTA brought to pH 7.8 with NaOH, 1 mg ml^{-1} bovine serum albumin (at the high dilutions used, the enzyme was unstable unless 1 mg ml^{-1} bovine serum albumin was added) and 5 mM agmatine (unless varied) or the compounds tested to replace agmatine (L-arginine, L-argininamide or arcaine). The PTC assay mixture contained 50 mM Tris-HCl pH 7, 0.1 mg ml^{-1} bovine serum albumin and 10 mM of both carbamoyl phosphate and putrescine (unless indicated). When varying the concentration of one substrate, the other was fixed at 10 mM. In both assays the amount of the enzyme was adjusted to assure that there was no consumption of >20% of any substrate, even at the low substrate concentrations used in the investigation of K_m values. The reactions were terminated after 5-15 min with 7% cold trichloroacetic acid, and the amount of carbamoyl putrescine was determined. Results at variable substrate concentrations were fitted to hyperbolae using the program

GraphPad Prism (GraphPad Software, San Diego, Calif.). One enzyme unit corresponds to the production of 1 μmol carbamoyl putrescine min^{-1}.

Analytical gel filtration chromatography. A Superdex 200HR (10/30) column was used, mounted on an ÄKTA FPLC, and equilibrated and eluted at 24°C, at a flow rate of 0.25 ml min^{-1}, with a solution of 50 mM Tris-HCl, pH.7.5, containing 0.15 M NaCl. The sample contained 0.1 mg of the protein of interest in 0.25 ml. Protein in the effluent was monitored by the optical absorption at 280 nm. A semilogarithmic plot of the molecular masses of marker proteins [from Amersham Biosciences or Sigma, or produced in our laboratory (14,28,44)] versus the distribution coefficient (K_d) for each protein was used for estimating the masses of AgDI and PTC. K_d values were calculated from the expression, $K_d = (V_e - V_0) / (V_i - V_0)$, taking V_0, V_i, and V_e as the volumes of elution of Blue Dextran, water (estimated by monitoring conductivity) and the protein of interest, respectively.

Growth of protein crystals and data collection by X-ray diffraction. The sparse-matrix sampling vapour-diffusion method (22) was used for crystallization tests carried out in hanging drops in multiwell plates using commercial kits (Crystal Screen I and II, from Hampton Research). The drops contained equal volumes (1-1.5 μl) of reservoir solution and of a 10 mg ml^{-1} solution of PTC or AgDI, prepared by repeated centrifugal ultrafiltration of the enzyme in 50 mM Tris-HCl, pH 7.45, containing also, in the case of AgDI, 1 mM dithiothreitol and 20 mM NaCl. Crystals of the two enzymes grew in about one week at 21ºC. The best PTC crystals were obtained in the presence of 430 μM PAPU, using a crystallization solution consisting of 125 mM $(NH_4)_2SO_4$, 17 % PEG 3.35K (Hampton Research) and 0.1 M Bis-Tris pH 5.5. The best AgDI

223 crystals were obtained in the presence of 5 mM agmatine, using as reservoir fluid 0.1 M Hepes,
224 pH 7.5, 1.5 M sodium chloride and 1.6 M ammonium sulfate. The crystals were harvested in the
225 corresponding crystallization solution supplemented with 15% (v/v) glycerol as cryoprotectant,
226 they were flash-cooled in liquid nitrogen, and were diffracted at 100 K (Oxford Cryo-Systems)
227 using synchrotron radiation (ESRF, Grenoble; beamline ID23-2 for PTC and BM-16 for AgDI).
228 The PTC and AgDI datasets, collected, respectively, to 3 and 1.65-Å resolution, were processed
229 and scaled with MOSFLM and SCALA [CCP4, (6)]. Table 1 gives the results of the data
230 collection as well as the spatial group and size of the cell for each of the proteins.
231
232 **Phasing, model building and refinement with the AgDI crystal data.** Molecular replacement
233 using MOLREP (55), utilizing as model the deposited (although not yet analyzed or reported)
234 structure at 2.9 Å of the subunit of AgDI from *Streptococcus mutans* (PDB accession number
235 2EWO), yielded a solution consisting of 8 subunits in the asymmetric unit. Rigid body and
236 restrained refinement were performed using REFMAC (36), alternating with graphic model-
237 building sessions with program Coot (12). B-factors and positional non-crystallographic
238 symmetry restraints were used and gradually released as refinement progressed. TLS (58) was
239 used in the last step of refinement. All the diffraction data were used throughout the refinement
240 process, except the 5% randomly selected data for calculating R_{free}. Refinement converged to a
241 final R value of 16.8% (R_{free}= 19.2%). The final model, at 1.65 Å resolution, consisted in the
242 chain spanning residues 2 - 367, 2-364, 2-368, 2-373, 1-368, 2-368, 2-367 and 2-366, for subunits
243 A, B, C, D, E, F, G and H, respectively. The model includes in all the subunits one molecule of
244 agmatine (as an amidine derivative, see results) covalently bound to Cys357. The stereochemistry

of the model, checked with PROCHECK (27) is reasonably good. Table 1 summarized the data on the refinement process and on the final model.

Other methods. Protein was assayed by the method of Bradford (5) using a commercial reagent from Bio-Rad, and bovine serum albumin as a standard. SDS-PAGE was carried out according to Laemmli (26). Sequence alignments were carried out with ClustalW (53), using defect values. Superposition of structures was carried out with program SSM (25). Buried surface areas were calculated using NACCESS (http://wolf.bms.umist.ac.uk/naccess). Figures of protein structures were generated using BOBSCRIPT (13), Raster3D (34), and Pymol, (http://pymol.sourceforge.net/).

Atomic coordinate and structure factors. The coordinates and structure factors are deposited in the Protein Data Bank (http://www.rcsb.org/) with the accession code 2J2T.

Materials. Purified recombinant *E. faecalis* ornithine transcarbamylase (3,31) (specific activity, 4,021 U mg^{-1}) was a gift of J. Sellés, from this laboratory. N-(Phosphoacetyl)-putrescine (PAPU) was prepared and purified as previously reported (41), and had the expected contents of phosphate [determined after hot acid digestion (4)] and free amino groups [assayed with ninhydrin (51) or by reverse-phase HPLC after orto-phtaldialdehyde derivatization (43); phosphoethanolamine was used as standard]. It yielded a mass (4700 Proteomic analyzer MALDI-TOF-TOF, from Applied Biosystems; CIPF, Valencia) of 212.05 Da (expected mass of the monopositive ion, 211.2 Da). Agmatine, putrescine, cadaverine, ornithine, carbamoyl phosphate and arcaine were from Sigma.

RESULTS

The agmatine catabolism gene cluster of *E. faecalis* V583.

Predicted genes *ef0732*, *ef0733*, *ef0734* and *ef0735* (Fig. 1), are on the same DNA strand of the *E. faecalis* V583 chromosome, separated by proposed intergenic distances of 66, 76 and 11 bp, and anotated in the current version of the TIGR database as the genes for putative ornithine transcarbamylase, an amino acid permease, a hypothetical conserved protein and a putative carbamate kinase, respectively. By analogy with the operon for arginine catabolism, in which the genes for arginine deiminase, ornithine transcarbamylase, carbamate kinase and the arginine/ornithine antiporter are designated *arc* (from arginine catabolism) *ABCD*, we will designate here the genes *ef0734, ef0732, ef0735 and ef0733* as *agc* (from agmatine catabolism) *ABCD*, respectively (Fig. 1A). No open reading frames have been identified on the same DNA strand within the 1,143 bases preceding *agcB* or in the 197 bases following *agcC*. The latter 197-base region hosts a predicted good, highly stable, protein-independent transcription terminator hairpin (terminator #851 for the *E. faecalis* V583 genome; TransTerm v 2.0 Beta program, http://www.cbcb.umd.edu/software/TransTerm) that may limit downstream the span of the transcriptional unit. A less stable terminator hairpin is predicted in the 66-residue intergenic region between *agcB* and *agcD* (terminator #850 for the *E. faecalis* genome), thus resembling the observation in the ADI operon of an internal hairpin of suboptimal stability after the gene for ornithine transcarbamylase, which only caused partial termination (3).

The identification in the TIGR database of the first codon of the open reading gene for *agcA* is in error: there are two more upstream inframe ATG codons, at 12 and 28 triplets from the proposed initiator ATG, of which the most upstream one is the genuine one, because 1) it is the only one that is preceded, 12 bases upstream, by a good Shine Dalgarno ribosomal binding sequence (AG<u>A</u>AGG; the base differing from the canonical sequence is underlined); 2) the

292 protein expressed from this ATG is a highly active AgDI (see below); 3) there is correspondence
293 between these 28 N-terminal residues and the N-terminal sequence of the AgDI from *P.*
294 *aeruginosa* (38); and 4) in the crystal structure of *E. faecalis* AgDI presented here, all these
295 residues except Met1 are well ordered and integrated into the enzyme crystal structure as
296 expected for a genuine portion of the natural enzyme. Therefore, this ATG is eight bases into the
297 preceding *agcD* gene, and thus *agcD* and *agcA* overlap.
298
299 **The product of the cloned *agcB* gene is genuinely putrescine transcarbamylase.**
300 The amino acid sequence encoded by the first gene of the cluster, *agcB* (ORF spanning
301 nucleotides 689526-690545 of the *E. faecalis* genome) has the same length (339 amino acids)
302 and exhibits 31% sequence identity with the ornithine transcarbamylase (OTC) encoded by the
303 *arcB* gene of the ADI operon of *E. faecalis* (3). The identity extends to the carbamoyl phosphate
304 and ornithine binding signature sequences ^{52}STRTR and ^{268}HCLP (the amino acid numbering
305 corresponds to the *agcB*-encoded protein sequence) and to 58 of the 85 residues that are totally
306 conserved in the anabolic and catabolic OTCs of *P. aeruginosa* and in the *arcB*-encoded *E.*
307 *faecalis* OTC (3). However, as might be expected if the product of the *agcB* gene were a
308 transcarbamylase that carbamylates a substrate different from ornithine (although not much
309 different, given the conservation of the ornithine signature), 11 of the 14 residues that are
310 invariant in these three OTCs but that are not conserved or conservatively replaced in the *agcB*
311 product, map in the C-terminal half of the enzyme, corresponding to the putative ornithine
312 domain of OTC. Further, the invariant SMG sequence of OTCs, which belongs to a mobile loop
313 that encircles the substituents around the ornithine C^{α} (47), is not conserved in the putative
314 product of *agcB*.

315 Cloning of *agcB* into the expression plasmid pET-22b(+) and overexpression of the gene
316 has confirmed that the corresponding protein product is PTC. The plasmid-encoded His_6-tagged
317 protein, overexpressed in BL21 (DE3) *E. coli* cells (see Materials and Methods), was produced in
318 large amounts in soluble form upon IPTG induction, and was purified to essential homogeneity
319 (Fig. 1B) in an approximate yield of 25 mg per liter of initial culture, by a simple procedure
320 based on the use of Ni affinity chromatography. The electrophoretic mobility of the purified
321 protein in SDS-PAGE (Fig. 1B) corresponded to a mass estimate of 40 kDa, in agreement with
322 the expected mass, deduced from the sequence, of 40,091 Da. Fourteen cycles of N-terminal
323 sequencing yielded the sequence MKRDYVTTETYTKE which includes the N-terminal Met, and
324 which corresponds to the amino acid sequence expected from the gene sequence. As previously
325 reported for genuine *E. faecalis* putrescine transcarbamylase, the protein appears to be a highly
326 stable trimer, as judged from its behavior, relative to other proteins of known mass, when
327 subjected to chromatography in a column of Superdex-200HR (Fig. 2).
328 Enzyme activity assays in the presence of 10 mM of both putrescine and carbamoyl
329 phosphate proved the recombinant protein to be a highly active putrescine transcarbamylase (Fig.
330 1, lower part of the figure) exhibiting comparable although somewhat higher specific activity
331 than the non-recombinant enzyme purified from *E. faecalis* [597 U mg^{-1}, versus 460 U mg^{-1} for
332 non-recombinant PTC (56)], and yielding K_m values for carbamoyl phosphate (58 ± 6 μM) and
333 putrescine (2.3 ± 0.3 mM) that also agree with prior determinations of the kinetic constants for
334 *E. faecalis* PTC (56). Furthermore, also according with prior results with PTC (56), the enzyme
335 exhibits some weak activity when 10 mM putrescine is replaced by either 10 mM ornithine or
336 cadaverine (6 and 9 %, respectively, of the activity observed with putrescine).
337

Phosphonoacetyl putrescine (PAPU) is a very potent and highly selective inhibitor of PTC.

Studies with aspartate and ornithine transcarbamylases demonstrated that phosphonoacetyl-L-aspartate (7) (PALA) and phosphoacetyl-L-ornithine (35) (PALO) are, respectively, highly potent inert bisubstrate inhibitors of these enzymes. Since these inhibitors have been successfully used in crystallization trials with these two enzymes (21,47) that led to the determination of their 3-D structures by X-ray diffraction, we reasoned that phosphonoacetyl putrescine (PAPU) might be also a very potent and highly specific inhibitor of PTC and if so it might help enzyme crystallization (see below). Although PAPU was synthesized previously (41), to our knowledge it has never been used with PTC. Fig. 3A shows that PAPU, at μM concentrations, is a very potent inhibitor of PTC, causing complete inhibition. In contrast, this compound, at the same concentrations, does not inhibit *E. faecalis* OTC, highlighting the selectivity of this inhibitor for PTC. The inhibition is non-competitive versus putrescine (Fig. 3B) and competitive versus carbamoyl phosphate (Fig. 3C), as expected if substrate binding in the PTC reaction is ordered, with carbamoyl phosphate binding first. From the slope of the plot of the apparent K_m for carbamoyl phosphate versus the concentration of PAPU (Fig. 3C), a K_i value can be estimated for PAPU of 10 nM. This low K_i value highlights the high affinity of the enzyme for this bisubstrate inhibitor, thus offering good opportunities for the preparation of PAPU-containing crystalline complexes of PTC that might shed structural light on substrate binding by, and specificity of, the enzyme.

Use of PAPU has allowed generation of PTC crystals suitable for X-ray analysis

To try to clarify the differences between PTC and OTC that justify the different specificities of these enzymes, we have initiated studies to determine the structure of PTC by X-

ray diffraction of protein crystals. Initial crystallization trials in the absence of substrates or inhibitors, or in the presence of putrescine, yielded crystals under some conditions, and some of these crystals were of sufficient size for diffraction studies, but they diffracted X-rays poorly (poorer than 4 Å resolution) even when synchrotron sources were used. The addition of PAPU to the crystallization drop dramatically improved the results of the crystallization trials, strongly suggesting that these new crystals contain bound PAPU. The crystals, having prismatic shape and ~0.3 mm maximal dimension (Fig. 1D, top panel), grew in about 1 week in the presence of 0.43 mM PAPU, using as crystallizants $(NH_4)_2SO_4$ and polyethylene glycol 3.35K (from Hampton). The crystals diffract X-rays (ESRF synchrotron ID-23-1 source) at 3 Å resolution, allowing determination of the space group of the crystal (Table 1), which is hexagonal $P6_322$, with a unit cell that would allow accommodating 2 or 3 enzyme subunits in the asymmetric unit, depending on whether 55% or 33% of the volume of the crystal is occupied by the solvent. We are presently in the process of searching for the phases by molecular replacement, using the structure of OTC from *Pyrococcus furiosus* (PDB file 1A1S) as search model.

Properties of the product of the *agcA* gene, agmatine deiminase

Using the most upstream ATG of the open reading frame (see first section of the Results), the coding region for the third gene of the cluster, *agcA,* spans nucleotides 691981-693078 of the *E. faecalis* genome. The predicted protein product has nearly identical length (365 versus 368 amino acids) and exhibits 54% sequence identity with respect to the AgDI encoded by the *aguA* gene of *P. aeruginosa* (38). In contrast, there is no significant identity (11.6 % identity, with 11 gaps) with the 408-residue sequence for the ADI of *E. faecalis* (3). Nevertheless, the Clustal W alignment of the AgDI and ADI sequences (not shown) aligns a cysteine residue that is near the C-termini of both sequences (Cys357 of AgDI) and which is conserved in both ADIs and AgDIs

385 and plays in both enzymes an analogous key catalytic role (see below our structural data on
386 AgDI).

387 The *agcA* gene, cloned from the most upstream ATG codon in the expression plasmid
388 pET-22b(+), triggered upon IPTG induction massive expression of the expected protein (Fig.
389 1C), in soluble form, as shown by the appearance of a large band in SDS-PAGE with a mass (46
390 kDa) corresponding, within experimental error, to the expected mass of the recombinant protein
391 (43,778 Da, including 2,589 extra Da due to the 24-residue C-terminal 6-His-containing
392 extension, VGGKQNSSSVDKLAAALEH_6). The extracts of the cells expressing the protein, but
393 not those transformed with the empty parental pET-22 plasmid, exhibited important AgDI
394 activity (Fig. 1, bottom) whereas ADI activity was nil in the same extracts. The recombinant
395 enzyme, purified by a combination of two Ni-affinity chromatography steps and a gel filtration
396 step, was obtained in high yield (~40 mg per liter of initial culture) in highly homogeneous form
397 (Fig. 1C) and was proven by gel filtration (and also by the crystal structure, see below) to be
398 tetrameric (Fig. 2). This is a substantial difference with respect to the AgDIs that are involved in
399 polyamine synthesis, which appear to be dimeric (23,37,59). Nevertheless, *E. faecalis* AgDI
400 resembles the well characterized polyamine synthesizing AgDIs of corn and *Arabidopsis thaliana*
401 (23,59) in the relatively low K_m value for agmatine (35 ± 3 μM, versus 12 and 110 μM for corn
402 and *A. thaliana* AgDIs, respectively), and in the similar magnitude of the activity at agmatine
403 saturation (22.3 ± 0.4 U mg^{-1} versus respective activities of 32 and 26 U mg^{-1} for corn and *A.*
404 *thaliana*). These results differ importantly from those for the only bacterial AgDI studied
405 biochemically, the AgDI of the arginine decarboxylase (ADC) pathway of *P. aeruginosa* (37), for
406 which a much larger K_m value (0.6 mM) and an ~4-fold-lower specific activity, relative to the *E.*
407 *faecalis* enzyme were reported. However, these differences may not be real, given the

methodological difficulties with colorimetric activity assays at low substrate concentrations and given the instability of AgDI upon large dilution in the assay solution (prevented in our case by adding 1 mg ml^{-1} bovine serum albumin). Similarly to all previous reports with AgDIs from other sources (23,37,60), the *E. faecalis* enzyme appears highly specific for agmatine, not using L-arginine (Fig. 1, lower panel), L-argininamide or arcaine (1,4 diguanidinobutane). Arcaine was reported to be a competitive inhibitor (relative to agmatine) of the corn (59) and cucumber (46) enzymes, with K_i values of ~3 and 7 μM, and we have found this compound to be also a competitive inhibitor of *E. faecalis* AgDI, with a K_i value of 28 ± 5 μM.

AgDI and ADI share the same basic fold.

AgDI monocrystals of up to 1 mm lenght (Fig. 1D, bottom panel) diffracted X-rays to 1.65 Å, allowing the determination of the crystal structure of the enzyme at atomic resolution. The asymmetric unit of the AgDI crystals (Table 1) contains eight subunits organized as two tetramers having identical structure. When surperposed, the rmsd for monomers is 0.17 Å (for 364 Cα atoms). Each of the monomers has a globular fold with approximate dimensions 53 × 45 × 40 $Å^3$. The monomer has a similar tertiary fold to that of the catalytic domain of ADI (Fig. 4A,B), although it lacks the five-helix bundle domain of this enzyme (9). Thus, the AgDI monomer has the fan-like structure with five blades that is a distinctive trait of ADI, and which results from a 5-fold pseudosymmetric structure in which each repeating element consists of a three-stranded mixed β-sheet and a helix in a ββαβ arrangement. Given the absence in AgDI of the five-helix bundle that distorts in ADI the fan-like structure, the AgDI monomer is closer to fivefold pseudosymmetry than the catalytic domain of ADI (Fig. 4A,B). Nevertheless, the first repeat diverges from the canonical structure of the repeat since it has two helices and the

arrangement βαβαβ, and it is flanked near the fivefold pseudosymmetry axis by the C-terminal strand running antiparallel to the other three strands. The lengths and amino acid sequences, however, vary considerably from one element to another (Fig. 4E) and from those in the catalytic domain of ADI (not shown), and, indeed, the superposition of AgDI with the catalytic domain of ADI (from subunit A of the *Mycoplasma arginini* enzyme, PDB file 1S9R) yields a relatively large rmsd (2.77 Å for 238 Cα atoms). A distinctive characteristic of the AgDI monomer fold is the existence of large loops emerging on the side of the fan corresponding to the C-end of the two parallel strands of the repeats (Fig. 4E and bottom part of Fig. 4F). Particularly large is the loop that emerges from the end of repeat 4, which includes a β hairpin (β15 and β16). This loop folds flat over the other loops and over a protruding long α helix that emerges from repeat 1 (helix 2, Fig. 4E) in the same direction as the loops. The presence of these loops and of the protruding α helix 2 renders highly different the two faces of the monomer that correspond to opposite edges of the repeats β sheets, and serves also the purposes of forming the active center and of providing interactions with the other subunits to form the tetramer (see below). Because of the presence of these loops on one side of the subunit, and also since the α helix of each repeat fills the space between adjacent repeats diverging from the pseudosymmetry axis, the subunit has a ball-like rounded shape (see each subunit in Fig. 4F and 4G).

The agmatine binding site justifies the high specificity of AgDI for its substrate.

A large mass of electron density not corresponding to the polypeptide chain and having an elongated shape was clearly visible (Fig. 4C) filling an internal cavity of the enzyme, and being connected to the density of the S atom of Cys357, the cysteine residue that is close to the enzyme C-terminus and that is conserved in both agmatine deiminase and arginine deiminase. The

electron density at 1.65 Å resolution fits a completely extended molecule of agmatine, with its Cξ atom (the C atom of the guanidinium group of agmatine) covalently linked with the S atom of Cys357. Agmatine is bound centrally (Fig. 4A), approximately along the five-fold pseudosymmetry axis near its exit from the loop-rich side of the fan, in a closed, elongated, and crowded cavity. The central position results in the involvement in the building of the site of elements connected to all five repeats of the subunit. Thus, the site is formed between the long loop of repeat 2 and the loops that connect repeats 5-to-1, 1-to-2, and 3-to-4. The cavity is closed at its entry by a three-tongued gate formed by the loop of repeat 2 and by the long loops connecting repeats 3-to-4 and 4-to-5 (seen lateraly in Fig. 4F). In our structure the closure is assured by mutual interactions between some residues of these loops, although it is clear that these loops have to retreat at the beginning and at the end of the catalytic cycle, to allow substrate binding and carbamoyl putrescine release. The extended substrate runs parallel to and makes extensive Van der Waals contacts with a straight stretch of three glycines (Gly351-Gly352-Gly353), making also a hydrogen bond between the agmatine amino group and the O atom of Gly351. These glycines are a part of the conserved sequence (G/A)GGNIHCITQQ(E/Q)P, which includes the catalytic cysteine (underlined) and which can be considered a signature of AgDI. The four-carbon portion of the molecule of agmatine is also surrounded by the indolic rings of the invariant Trp93 and Trp119 (Fig. 4C), which are like flat tiles that wall the substrate binding cavity, and by the methyl group of invariant Thr215. The agmatine amino group also makes a bond with the γ-COO^- of Glu214 (Fig. 4C), a residue that may play a key role in making the enzyme extremely selective against arginine, since it would not favor placing near it another negatively charged group as it would be the case for the carboxylate group of arginine. Anyway, the region that surrounds carbon 1 of agmatine is packed with predominantly hydrophobic

477 groups, leaving no room for a carboxyl or for any other group of substantial size and less so if the
478 group is polar and charged as in an α carboxylate. On the opposite end of the agmatine molecule,
479 around the guanidinium group, the invariant residues Asp96, His218 and Asp220 surround the
480 bound substrate and play catalytic roles to be described below (Fig. 4C and 4D).
481
482 **The covalent adduct provides a snapshot of AgDI catalyzing its reaction**
483 A close examination of the electron density around the Cξ atom of agmatine (Fig. 4D)
484 shows that this carbon is covalently linked with the S atom of Cys357 (C-S bond distance, 1.79
485 Å) and with two nitrogens (Nε and Nξ2). The Cξ in our structure is somewhat displaced from the
486 plane formed by its three covalent ligands (S, Nε and Nξ2 atoms) towards a water molecule (W1,
487 Fig. 4D) which is located at only 2.5 Å (a very short distance for non-bonded C and O atoms).
488 The water molecule is fixed by hydrogen bonds to one O atom of each of the two side-chain
489 carboxylates of Asp220 and Asp96, and to the δ1N atom of His218. In turn, the ε2N of His218 is
490 linked to the γ-COO^- of Glu157. A similar adduct was reported with ADI within a complex which
491 replicates essentially all of the details of the present complex, including the presence and the
492 interactions of the fixed water (9). This complex was interpreted to represent the covalent
493 amidino complex proposed long ago to be formed in the mechanism of ADI (9,16). Therefore,
494 the present complex is highly indicative of a common, very strictly conserved mechanism of
495 deimination by ADI and AgDI. This mechanism (Fig. 5) involves two tetrahedral intermediates
496 but only one amidino adduct. Nevertheless, the amidino adduct must exist first in the presence of
497 the leaving ammonia and, later on, in the presence of the attacking water. Thus, the W1 molecule
498 could also correspond to ammonia, and the present adduct may represent either the amidino
499 compound with the leaving ammonia or the same compound with the attacking water.

500

501 **The architecture of the AgDI tetramer**

502 In accordance with the conclusions derived from gel filtration data, AgDI is organized as
503 a tetramer. This tetramer has tetrahedral shape with the four subunits located in the vertices (Fig.
504 4G) and can be considered composed of two identical dimers, each of them (Fig. 4F) built by a
505 180º-rotation of the monomer around an axis that is approximately parallel to the 5-fold
506 pseudosymmetry axis. Thus, this dimer has the aspect of two fans in battery. The interactions in
507 this dimer are mediated by the elements of the first and second repeats, with the edge of the more
508 external β strand of the first repeat interacting with the C-terminal two turns of α helix 4, the
509 helix belonging to the second repeat. These interactions are generally hydrophobic in the core
510 region and polar towards the periphery, and the surface involved amounts to an average of 991 $Å^2$
511 per monomer (determined with a probe radius of 1.4 Å) or ~6.9% of the surface of each
512 monomer. The two subunits in the dimer leave a valley between them (Fig. 4F) in the loop-rich
513 face. It is this valley which is used for tetramer formation by having the valley of one dimer
514 interact in a crossed over way with the valley in the other dimer, so that the twofold axes of the
515 two dimers are coincident and the longest axes of the two dimers run in perpendicular directions
516 (Fig. 4G). One subunit (called A for the purpose of this description) of one dimer interacts with
517 the two subunits of the other dimer (called here C and D). The N-termini of helices 2 from A and
518 C interact mutually, and residues 286 to 289 of the long hairpin loop that connects repeat 4 and 5
519 of A interact with the outer surface of helix 3 and also with the N terminal turn of helix 2 of C,
520 and vice versa. The interactions between A and D are restricted to mutual hydrophobic contacts
521 between the long loop of repeat 2 of both subunits. The buried surface per monomer is 639 $Å^2$ for
522 the interactions between A and C, and only 252 $Å^2$ for those between A and D. Overall, each

523 monomer has in the tetramer a buried surface of 1927 $Å^2$, accounting for 13.5 % of its total
524 accessible surface area, justifying the stability of the enzyme tetramer that has been observed in
525 the present studies.

Discussion

By cloning the genes and by studying the expressed proteins, we provide here the most conclusive proof to date that the *E. faecalis agcB* and *agcA* genes encode two key enzymes of agmatine catabolism, PTC and AgDI, confirming and extending previous (17,18,39) but more indirect evidence for the identification of these genes. When initially purified from *E. faecalis,* PTC exhibited some (although low) activity with ornithine (56), and this is confirmed here with the recombinant, His-tag-purified enzyme, virtually completely excluding OTC contamination as the cause for this activity. Nevertheless, the relatively low specific activity of PTC compared with the activity of pure OTC (~5-fold higher; the low PTC activity is not due to the poly-His tail, since wild-type PTC isolated from *E. faecalis* has even somewhat less activity (56)) rendered desirable to confirm that this enzyme is a genuine PTC, what has been done here by demonstrating that this enzyme is powerfully inhibited by very low concentrations (K_i=10 nM) of the PTC-specific bisubstrate analog inhibitor PAPU, a compound that does not inhibit OTC at similar concentrations.

Since *E. faecalis* OTC and PTC share 31% sequence identity, these two enzymes either derive from a common ancestor of broad specificity, or, perhaps more likely since OTC cannot use putrescine, PTC may derive from OTC and may not have perfected yet discrimination between putrescine and ornithine, with the process of shifting specificity possibly having resulted in somewhat compromised catalytic efficiency. To discriminate between these possibilities and to clarify the determinants of specificity and catalytic efficiency, it would be important to compare the structures of PTC and OTC, a goal that is now at closer reach thanks to the use of PAPU, since we report here the production of X-ray diffracting PTC crystals generated in the presence of this bisubstrate inhibitor.

549 We have characterized also the protein product encoded by *agcA*, both functionally and
550 structurally, as AgDI. The structure of this enzyme closely resembles that of ADI, the enzyme
551 that catalyzes the same reaction except for the use of arginine as substrate, exhibiting the
552 characteristic five-blade propeller fold presented by the catalytic domain of ADI (9,15), with the
553 substrate, also similarly to ADI, binding in a deep, central, very tight cavity. We have found here
554 that AgDI, again similarly to ADI (9,16), makes a covalent substrate amidino adduct involving a
555 catalytic thiol group belonging to a conserved Cys residue that is close to the enzymes C-termini.
556 Thus, ADI and AgDI are homologous enzymes, although the lack of significant sequence identity
557 between them indicates a long period of divergence. The inability of each of these enzymes to use
558 the substrate of the other (3,37, and present results) further suggests that the separation between
559 ADI and AgDI occurred long ago, with enough time for optimization of substrate specificity.
560 The AgDI studied here, committed to making ATP fermentatively from agmatine (48),
561 exhibits ~50 % sequence identity (not shown) with the more widespread AgDIs, which belong to
562 the arginine decarboxylase (ADC) pathway and are involved in polyamine production (37,38)
563 (although this pathway can also serve for agmatine utilization as a carbon and nitrogen source, as
564 in *Pseudomonas aeruginosa* (19,52)). The most relevant difference is that ADC-pathway AgDIs
565 appear to be dimeric (23,37,59), whereas *E. faecalis* AgDI is tetrameric, although in fact it is a
566 dimer of dimers and thus even in this aspect it does not depart much from the characteristics of
567 the ADC pathway AgDIs. Since AgDIs neither exhibit cooperativity for the substrate or
568 regulatory properties (46), we presently have no indications that the degree of oligomerization of
569 AgDIs is important functionally.
570 Although not studied experimentally here, there can be little doubt that the product of
571 *agcC* (Fig. 1A) is a true CK, since it is only one amino acid shorter than and exhibits 49%
572 sequence identity (data not shown) with the CK of the *E. faecalis* ADI operon (30), an enzyme

573 for which the 3-D structure was determined (29). Since the CKs involved in arginine and
574 agmatine catabolism appear similar and there is no evidence of CK regulation by effectors (32),
575 one plausible reason for having two separate CK isozymes in each of these pathways may be to
576 facilitate concerted expression of all the genes of one or the other pathway. The important
577 sequence identity of these two CK isozymes indicates that their separation is not remote. In
578 contrast, the lack of significant sequence identity (14%) between the *E. faecalis arcD* and *agcD*
579 gene products (the putative arginine/ornithine and agmatine/putrescine antiporters) indicates
580 ancient divergence, the homology of these genes being supported by the similarity of the
581 polypeptide lengths (483 and 458 residues, respectively) and transmembrane helix predictions
582 (11-12 helices), and also by the analogous functions and substrates of the antiporters.

583 The comparison of the genes of the ADI and AgDI pathways contradicts the naïve view
584 that the two pathways might have arisen by a process of duplication of a complete ancient four-
585 gene cluster. As already indicated, the deiminase and antiporter components of both pathways
586 have evolved separately for much longer than the transcarbamylase and carbamate kinase
587 components, in contrast with the expectation for a common duplication event for all of the
588 elements of the gene cluster, followed by coevolution. Nevertheless, the genes for the
589 transcarbamylase and for CK may have duplicated simultaneously, given their similar degree of
590 conservation in one gene cluster relative to the corresponding genes in the other cluster and also
591 since in both clusters the transcarbamylase gene physically precedes the CK gene. Since
592 agmatine utilization cannot precede agmatine production, the close relation between the AgDI of
593 the ADC and AgDI pathways suggests that the latter may have derived from and ADC-pathway
594 gene for AgDI. The evolution of AgDI may have been initiated by its divergence from ADI to
595 serve the purpose of polyamine synthesis. Much more recently AgDI may have become
596 committed into a novel route of agmatine catabolism, made by recombining elements of the ADI

pathway with the arginine decarboxylase pathway element AgDI, and with an agmatine antiporter of obscure origin.

An important contribution of the present work is the clarification of the structure and, based on the structure, of the reactional and catalytic mechanism of AgDI. The high affinity of AgDI for agmatine is accounted by the extension and closeness of the interactions between substrate and enzyme, since agmatine is buried, fitting tightly a binding site where there is no empty space. The high specificity is justified by the negative charge at the entry of the site provided by Glu214, which would not fit the placement of the α-carboxylate group of arginine, and also by the very crowded environment where even minor volumes around the C1 of agmatine would be excluded. The relatively low k_{cat} for a hydrolase exhibited by AgDI (17 s^{-1} at 37ºC) may be justified by the deepness of the site, with the catalytic groups far into the subunit structure, an also by the existence of a gating mechanism at the entry of the site that has to open, close, and open again in each catalytic cycle, possibly rendering limiting substrate access or product release. Agmatine binding may trigger site closure, since the amino end of the substrate interacts with elements of the loops that contribute to the gating mechanism, particularly with Glu214. Since the formation of the covalent adduct with the thiol group of Cys357 should shorten somewhat the bound molecule, the closing mechanism could be described as "pulling the gate from the inside" by the covalently bound substrate. The importance of substrate binding for gate closure is supported by the observation of the deposited structure of *S. mutans* AgDI, which contains no bound agmatine, and where the largest loop involved in the gating mechanism is retracted and the site is more accessible. Our structure, containing the covalently bound substrate, is closed, and would have to be open at the end of the reaction. The simplest triggering mechanism to open the gate could be defined as "pushing the gate from the inside", whereby the increase in the volume

resulting from the coexistence of the ureido group in carbamoyl putrescine and the free thiol in Cys357 may result in some displacement of the molecule of the product towards the gate, particularly given the extreme narrowness of the site, which should not allow bending of the bound product.

Puzzlingly, in our crystal structure the enzyme has retained the covalent amidino adduct without progressing further along the reactional path. The corresponding analog for arginine has also been reported in ADI (9). The presence of the trapped intermediate strongly suggests that some component in the crystallization solution has stabilized the amidino intermediate, in fact resulting in enzyme inhibition. Perhaps ammonia, present in our solution at a concentration of 3.2 M, has resulted in the stabilization of the ammonia-containing amidino complex (W1 in our structure could equally be ammonia), blocking further reaction with water (Fig. 5). Whichever the mechanism, the observation of the complex has had the value of clarifying substrate binding and catalysis. The catalytic process involves centrally, as in the case of ADI (9,15), a charge relay system consisting of Glu157 and His218, which promotes formation of the tetrahedral intermediates by providing or withdrawing a proton; and Cys357 with its SH group being abnormally acidic, perhaps because of the presence of the guanidinium group of the substrate and possibly also by the inducing effect of the nearby (3Å) carboxylate of Asp96. We are presently undertaking studies to subject to experimental test, by site-directed mutagenesis, the roles proposed on the basis of the structure for catalysis of the reaction by these residues.

Acknowledgments

This work was supported by grant BFU2004-05159 of the Spanish Ministry of Education and Science. LM Polo is a fellow of CSIC-Banco de Santander and JL Llácer and S Tavárez of the Spanish Ministry of Education and Science. We thank the EU, ESRF and EMBL Grenoble for financial support for ESRF synchrotron X-ray data collection; the ESRF personnel for expert help; A Marina and M López for diffracting AgDI; JJ Calvete (IBV-CSIC, Valencia, Spain) for N-terminal sequencing; D Gigot (Université Livre de Bruxelles, Belgium) for advice and A Cantin and MA Miranda (ITQ-CSIC, Valencia, Spain) for help with the synthesis of PAPU; C Aguado (CIPF, Valencia, Spain) for MALDI-TOF mass spectrometry; and J Sellés, P Tortosa and L Osuna (IBV-CSIC, Valencia) for technical help.

Reference List

1. **Archibald, R.** 1944. Determination of citrulline and allantoin and demonstration of citrulline in blood plasma. J. Biol. Chem. **156**:121-141.
2. **Bagni, N. and A. Tassoni**. 2001. Biosynthesis, oxidation and conjugation of aliphatic polyamines in higher plants. Amino Acids **20**:301-317.
3. **Barcelona-Andres, B., A. Marina, and V. Rubio**. 2002. Gene structure, organization, expression, and potential regulatory mechanisms of arginine catabolism in *Enterococcus faecalis*. J. Bacteriol. **184**:6289-6300.
4. **Bartlett, G. R.** 1959. Phosphorus assay in column chromatography. J. Biol. Chem. **234**:466-468.
5. **Bradford, M. M.** 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. **72**:248-254.
6. **Collaborative Computational Project Number 4**. 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D. **50**:760-763.
7. **Collins, K. D. and G. R. Stark**. 1971. Aspartate transcarbamylase. Interaction with the transition state analogue N-(phosphonacetyl)-L-aspartate. J. Biol. Chem. **246**:6599-6605.
8. **Cunin, R., N. Glansdorff, A. Pierard, and V. Stalon**. 1986. Biosynthesis and metabolism of arginine in bacteria. Microbiol. Rev. **50**:314-352.
9. **Das, K., G. H. Butler, V. Kwiatkowski, A. D. Clark, Jr., P. Yadav, and E. Arnold**. 2004. Crystal structures of arginine deiminase with covalent reaction intermediates; implications for catalytic mechanism. Structure. **12**:657-667.

10. **Deibel, R. H.** 1964. Utilization of arginine as an energy source for the growth of *Streptococcus faecalis*. J. Bacteriol. **87**:988-992.
11. **Driessen, A. J., E. J. Smid, and W. N. Konings**. 1988. Transport of diamines by *Enterococcus faecalis* is mediated by an agmatine-putrescine antiporter. J. Bacteriol. **170**:4522-4527.
12. **Emsley, P. and K. Cowtan**. 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. **60**:2126-2132.
13. **Esnouf, R. M.** 1999. Further additions to MolScript version 1.4, including reading and contouring of electron-density maps. Acta Crystallogr. D. Biol. Crystallogr. **55**:938-940.
14. **Fernandez-Murga, M. L., F. Gil-Ortiz, J. L. Llacer, and V. Rubio**. 2004. Arginine biosynthesis in *Thermotoga maritima*: characterization of the arginine-sensitive N-acetyl-L-glutamate kinase. J. Bacteriol. **186**:6142-6149.
15. **Galkin, A., L. Kulakova, E. Sarikaya, K. Lim, A. Howard, and O. Herzberg**. 2004. Structural insight into arginine degradation by arginine deiminase, an antibacterial and parasite drug target. J. Biol. Chem. **279**:14001-14008.
16. **Galkin, A., X. Lu, D. Dunaway-Mariano, and O. Herzberg**. 2005. Crystal structures representing the Michaelis complex and the thiouronium reaction intermediate of *Pseudomonas aeruginosa* arginine deiminase. J. Biol. Chem. **280**:34080-34087.
17. **Griswold, A. R., Y. Y. Chen, and R. A. Burne**. 2004. Analysis of an agmatine deiminase gene cluster in *Streptococcus mutans* UA159. J. Bacteriol. **186**:1902-1904.
18. **Griswold, A. R., M. Jameson-Lee, and R. A. Burne**. 2006. Regulation and physiologic significance of the agmatine deiminase system of *Streptococcus mutans* UA159. J. Bacteriol. **188**:834-841.

19. **Haas, D., H. Matsumoto, P. Moretti, V. Stalon, and A. Mercenier**. 1984. Arginine degradation in *Pseudomonas aeruginosa* mutants blocked in two arginine catabolic pathways. Mol. Gen. Genet. **193**:437-444.

20. **Hills, G. M.** 1940. Ammonia production by pathogenic bacteria. Biochem. J. **34**:1057-1069.

21. **Huang, J. and W. N. Lipscomb**. 2004. Aspartate transcarbamylase (ATCase) of *Escherichia coli*: a new crystalline R-state bound to PALA, or to product analogues citrate and phosphate. Biochemistry **43**:6415-6421.

22. **Jancarik, J. and S. H. Kim**. 1991. Sparse matrix sampling: a screening method for crystallization of proteins. J. Appl. Cryst. **24**:409-411.

23. **Janowitz, T., H. Kneifel, and M. Piotrowski**. 2003. Identification and characterization of plant agmatine iminohydrolase, the last missing link in polyamine biosynthesis of plants. FEBS Lett. **544**:258-261.

24. **Jones, M. E. and F. Lipmann**. 1960. Chemical and enzymatic synthesis of carbamyl phosphate. Proc. Natl. Acad. Sci. U. S. A **46**:1194-1205.

25. **Krissinel, E. and K. Henrick**. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D. Biol. Crystallogr. **60**:2256-2268.

26. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227**:680-685.

27. **Laskowski, R. A., M. W. MacArthur, D. S. Moss, and J. M. Thornton**. 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Cryst. **26**:283-291.

28. **Marco-Marin, C., S. Ramon-Maiques, S. Tavarez, and V. Rubio**. 2003. Site-directed mutagenesis of *Escherichia coli* acetylglutamate kinase and aspartokinase III probes the catalytic and substrate-binding mechanisms of these amino acid kinase family enzymes and allows three-dimensional modelling of aspartokinase. J. Mol. Biol. **334**:459-476.

29. **Marina, A., P. M. Alzari, J. Bravo, M. Uriarte, B. Barcelona, I. Fita, and V. Rubio**. 1999. Carbamate kinase: New structural machinery for making carbamoyl phosphate, the common precursor of pyrimidines and arginine. Protein Sci. **8**:934-940.

30. **Marina, A., M. Uriarte, B. Barcelona, V. Fresquet, J. Cervera, and V. Rubio**. 1998. Carbamate kinase from *Enterococcus faecalis* and *Enterococcus faecium*. Cloning of the genes, studies on the enzyme expressed in *Escherichia coli*, and sequence similarity with N-acetyl-L-glutamate kinase. Eur. J. Biochem. **253**:280-291.

31. **Marshall, M. and P. P. Cohen**. 1972. Ornithine transcarbamylase from *Streptococcus faecalis* and bovine liver. I. Isolation and subunit structure. J. Biol. Chem. **247**:1641-1653.

32. **Marshall, M. and P. P. Cohen**. 1966. A kinetic study of the mechanism of crystalline carbamate kinase. J. Biol. Chem. **241**:4197-4208.

33. **Mercenier, A., J. P. Simon, D. Haas, and V. Stalon**. 1980. Catabolism of L-arginine by *Pseudomonas aeruginosa*. J. Gen. Microbiol. **116**:381-389.

34. **Merritt, E. A. and M. E. Murphy**. 1994. Raster3D Version 2.0. A program for photorealistic molecular graphics. Acta Crystallogr. D. Biol. Crystallogr. **50**:869-873.

35. **Mori, M., K. Aoyagi, M. Tatibana, T. Ishikawa, and H. Ishii**. 1977. N^{δ}–(phosphonacetyl) –L–ornithine, a potent transition state analogue inhibitor of ornithine carbamoyltransferase. Biochem. Biophy. Res. Commun. **76**:900-904.

36. **Murshudov, G. N., A. A. Vagin, and E. J. Dodson**. 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D. Biol. Crystallogr. **53**:240-255.

37. **Nakada, Y. and Y. Itoh**. 2003. Identification of the putrescine biosynthetic genes in *Pseudomonas aeruginosa* and characterization of agmatine deiminase and N–carbamoylputrescine amidohydrolase of the arginine decarboxylase pathway. Microbiology **149**:707-714.

38. **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.

39. **Naumoff, D. G., Y. Xu, N. Glansdorff, and B. Labedan**. 2004. Retrieving sequences of enzymes experimentally characterized but erroneously annotated : the case of the putrescine carbamoyltransferase. BMC. Genomics **5**:52.

40. **Nuzum, T. C. and P. J. Snodgrass**. 1976. Multiple Assays of the Five Urea Cycle Enzymes in Liver Homogenates, p. 325-349. *In* S. Grisolia, R. Báguena, and F. Mayor (ed.), The Urea Cycle. John Wiley & Sons, New York, U.S.A.

41. **Penninckx, M. and D. Gigot**. 1978. Synthesis and interaction with *Escherichia coli* L–ornithine carbamolytransferase of two potential transition–state analogues. FEBS Lett. **88**:94-96.

42. **Petrack, B., L. Sullivan, and S. Ratner**. 1957. Behavior of purified arginine desiminase from *S. faecalis*. Archives of Biochemistry and Biophysics **69**:186-197.

43. **Portoles, M. and V. Rubio**. 1986. High-performance liquid chromatographic assay of argininosuccinate: its application in argininosuccinic aciduria and in normal man. J. Inherit. Metab Dis. **9**:31-38.

44. **Ramon-Maiques, S., H. G. Britton, and V. Rubio**. 2002. Molecular physiology of phosphoryl group transfer from carbamoyl phosphate by a hyperthermophilic enzyme at low temperature. Biochemistry **41**:3916-3924.

45. **Roon, R. J. and H. A. Barker**. 1972. Fermentation of agmatine in *Streptococcus faecalis*: occurrence of putrescine transcarbamoylase. J. Bacteriol. **109**:44-50.

46. **Sakakibara, Y. and H. Yanagisawa**. 2003. Agmatine deiminase from cucumber seedlings is a mono-specific enzyme: purification and characteristics. Protein Expr. Purif. **30**:88-93.

47. **Shi, D., H. Morizono, X. Yu, L. Tong, N. M. Allewell, and M. Tuchman**. 2001. Human ornithine transcarbamylase: crystallographic insights into substrate recognition and conformational changes. Biochem. J. **354**:501-509.

48. **Simon, J. P. and V. Stalon**. 1982. Enzymes of agmatine degradation and the control of their synthesis in *Streptococcus faecalis*. J. Bacteriol. **152**:676-681.

49. **Simon, J. P., B. Wargnies, and V. Stalon**. 1982. Control of enzyme synthesis in the arginine deiminase pathway of *Streptococcus faecalis*. J. Bacteriol. **150**:1085-1090.

50. **Slade, H. D. and W. C. Slamp**. 1952. The formation of arginine dihydrolase by streptococci and some properties of the enzyme system. J. Bacteriol. **64**:455-466.

51. **Spies, J. R.** 1957. Colorimetric procedures for amino acids, p. 467-477. *In* Methods in Enzymology. Academic Press.

52. **Stalon, V. and A. Mercenier**. 1984. L-arginine utilization by *Pseudomonas* species. J. Gen. Microbiol. **130**:69-76.

53. **Thompson, J. D., D. G. Higgins, and T. J. Gibson**. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673-4680.

54. **Tricot, C., J. L. De Coen, P. Momin, P. Falmagne, and V. Stalon**. 1989. Evolutionary relationships among bacterial carbamoyltransferases. J. Gen. Microbiol. **135**:2453-2464.

55. **Vagin, A. and A. Teplyakov**. 1997. MOLREP: an Automated Program for Molecular Replacement. J. Appl. Cryst. **30**:1022-1025.

56. **Wargnies, B., N. Lauwers, and V. Stalon**. 1979. Structure and properties of the putrescine carbamoyltransferase of *Streptococcus faecalis*. Eur. J. Biochem. **101**:143-152.

57. **Wilson, K.** 2001. Preparation of Genomic DNA from Bacteria, p. 2.4.1-2.4.2. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.), Current Protocols in Molecular Biology. John Wiley & Sons, New York, USA.

58. **Winn, M. D., M. N. Isupov, and G. N. Murshudov**. 2001. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr. D. Biol. Crystallogr. **57**:122-133.

59. **Yanagisawa, H.** 2001. Agmatine deiminase from maize shoots: purification and properties. Phytochemistry **56**:643-647.

60. **Yanagisawa, H. and Y. Suzuki**. 1981. Corn agmatine iminohydrolase: purification and properties. Plant Physiol. **67**:697-700.

Figure legends

FIG. 1. Agmatine catabolism gene cluster and PTC and AgDI purification and crystallization. (A) Gene organization of the two strands of the sequenced region, with indications of the number of amino acid (aa) residues expected for each gene product and the length (in base pairs, bp) of the intergenic regions. Genes are given the identifyers of the TIGR database, together with the *agc* denominations given to them here. The positions of the three predicted stem-loops are indicated by the open circles, and a putative *cre* box preceding *agcB* is indicated with a grey rectangle.The gene in the opposite strand corresponds to a *luxR* regulator. (B) and (C) Coomassie-stained SDS-PAGE analyses of the various steps of the purifications of PTC and AgDI. The crude extracts are the postsonication supernatants. Panel B includes also a blank extract of *E. coli* BL21 cells transformed with the parental pET-22 plasmid carrying no gene insert, to highlight the differences with the extracts of cells transformed with the plasmids carrying the genes for PTC or for AgDI. Molecular weight marker proteins were from Sigma (Dalton Mark VII-L). Results of enzyme activity assays for PTC and AgDI are shown below the purification steps at which the activities were assayed. The results of the activites obtained when putrescine was replaced by 10 mM ornithine (OTC) or when agmatine was replaced by 5 mM arginine (ADI) are shown also. Values preceded by a < symbol are detection limits for assays giving no activity. (D) Crystals obtained of both enzymes that have been used for diffraction studies. The small horizontal bars denote 0.1 mm.

FIG. 2. Investigation of the oligomeric state of *E. faecalis* PTC and AgDI, using gel filtration. Semilogarithmic plot of molecular mass versus elution volume (expressed as K_d, see Materials and Methods) from the Superdex 200HR column. The closed circles correspond to the following

protein standards: cytochrome C (12.3 kDa), lactalbumin (14.2 kDa), carbonic anhydrase (29.0 kDa), ovalbumin (42.7 kDa), bovine serum albumin (66.4 kDa), the dimer of bovine serum albumin (132.9 kDa), *Pyrococcus furiosus* carbamate kinase (68.8 kDa), intact (97.1 kDa) and truncated (31.9 kDa) aspartokinase III of *E. coli*, alcohol dehydrogenase (146.8 kDa), aldolase (156.8 kDa), *Thermotoga maritima* N-acetyl-L-glutamate kinase (182.0 kDa), amylase (223.8 kDa), catalase (230.3 kDa), ferritin (440 kDa), and thyroglobulin (669 kDa). The triangle and square denote, respectively, the position of elution of the peaks of *E. faecalis* PTC and AgDI, assuming that PTC is a trimer (sequence-deduced mass, 120,273 Da) and AgDI is a tetramer (deduced mass, 165,412 Da).

FIG. 3. Effects of PAPU on *E. faecalis* PTC activity. (A) Inhibition and lack of inhibition of *E. faecalis* PTC and OTC, respectively, by PAPU. Activities are given as fractions of the activity in the absence of PAPU. (B) and (C) Influence of PAPU on the kinetic parameters of PTC for putrescine (Put), or on the K_m for carbamoyl phosphate (Carb-P_i).

FIG. 4. The structure of AgDI. (A) and (B), show, respectively, in ribbons representation, the monomer of *E. faecalis* AgDI and of the catalytic domain of *Mycoplasma arginini* ADI (PDB entry 1S9R without residues 75 – 148, which are not a part of the catalytic domain), both containing the covalently bound substrate in space-filling representation. α-Helices, β-sheets and loops are colored red, yellow and green, respectively. (C) Stereo view of the active site of AgDI, showing the ($2F_{obs} - F_{calc}$) density map contoured at 0.9 σ level, around the covalent adduct. The substrate is colored yellow, and the surrounding protein residues are colored grey. O, N and S atoms are colored red, blue and green, respectively. (D) Interatomic distances between the

852 catalytic protein residues and the substrate around the reactive carbon center. The interactions
853 with a fixed water molecule (W1) believed to be important in the mechanism are represented
854 also. The ($2F_{obs} - F_{calc}$) density map contoured at 0.75 σ for the covalent amidino complex is
855 shown. (E) Correspondence between the amino acid sequence and the secondary structure. Bars,
856 arrows and lines above the structure denote, respectively, α helices, β-strands and loops (only
857 long loops are depicted), numbered in ascending order from N to C terminus, and, when
858 belonging to a repeat, enclosed between parentheses and having a subscript that denotes the
859 repeat number. Open triangles under the sequence denote residues having decreased accessibility
860 upon the binding of agmatine. Circles denote decreased accessibility upon dimer (open) and
861 tetramer (shadowed) formation. The grey sequence backgrounds highlight residues that are
862 invariant in the AgDIs o *E. faecalis, Streptococcus mutans, Pseudomonas aeruginosa and*
863 *Arabidopsis thaliana* (Swissprot accession numbers, Q837U5, Q8DW17, Q9I6J9, Q8GWW7).
864 (F) Ribbon diagram of AgDI dimer viewed perpendicularly to the molecular twofold axis.
865 Coloring and substrate representation are as in (A). (G) Ribbon representation of the AgDI
866 tetramer viewed along one of the three twofold molecular axes. The two subunits of one and the
867 other dimer (as defined in the text) are shown in different shades of red or blue. Covalently
868 bound substrate is in space-filling representation.
869
870 FIG. 5. Proposed five-step mechanism for the AgDI reaction. Step 1 leads to the formation of the
871 first tetrahedral carbon center intermediate as a consequence of the attack by the activated thiol of
872 Cys357. Asp96 and the non-protonated primary N of the guanidinium group may induce
873 deprotonation of the thiol group. A proton is extracted by His218, which forms a charge relay
874 system with Glu 157. In step 2 the tetrahedral intermediate collapses to the trigonal amidino

intermediate, with liberation of ammonia. Asp96, Asp220 and His218 help stabilize the leaving ammonia and the positive charge development in the amidino group. In step 3 ammonia is replaced by water positioned for attack on the carbon center, interacting with the same groups as the ammonia. The intermediate revealed here by X-ray crystallography corresponds to one of the two complexes (either the ammonia or the water complex) with the amidino intermediate. Step 4 is the formation of the second tetrahedral carbon intermediate. His218 helps this step by abstracting one proton from water. The final step is the collapse of the tetrahedral intermediate to carbamoylputrescine and the regenerated thiol group.

883

884 TABLE 1. X-Ray Data and Structure Refinement Statistics
885

Parameter[a]	PTC	AgDI
Data statistics		
Space Group	$P6_322$	$P2_1$
Unit Cell	a = b = 118.6 Å c = 227.4 Å, α = β = 90º γ = 120º	a = 107.7 Å, b = 130.2 Å, c = 126.7 Å, α = γ = 90º β = 93.6º
Resolution range (Å)	76.25 - 3.00 (3.16-3.00) [b]	45.36-1.65 (1.74 –1.65)[b]
R_{sym} (%) overall [c]	16.3 (37.7) [b]	7.1 (37.1) [b]
Completeness (%)	98.3 (98.3) [b]	100 (100) [b]
I/σ[c]	21.9 (8.4) [b]	8.4 (2.0) [b]
Total / unique reflections	313,291 / 18,766	1,556,987 / 417,377
Refinement Statistics		
Resolution range (Å)		50.0 - 1.65
Polypeptide chains / amino acid residues		8 / 2934
Agmatine molecules		8 (as covalent adducts)
Protein atoms / water molecules		23,227 / 2,174
R-factor / R_{free} [d]		16.8 / 19.2
RMSD bonds (Å) / angles (º)		0.015 / 1.519
Ramachandran plot (%)		
(fav./all./gen.all./disall) [e]		87.6 / 11.7 / 0.7 / 0

886 [a] Abreviations: RMSD, root mean square deviation. fav., favored. all., allowed. gen.all.,

887 generously allowed. disall., disallowed.

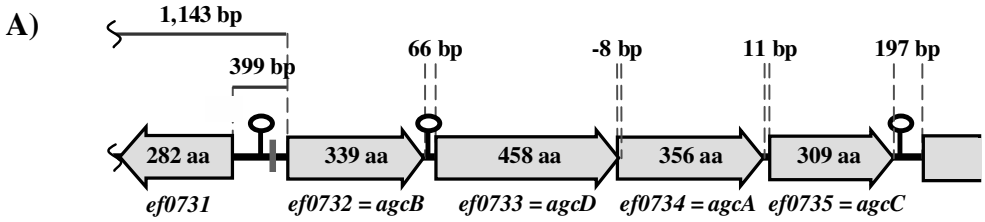
888 [b]Values in parenthesis are data for the highest resolution shell.

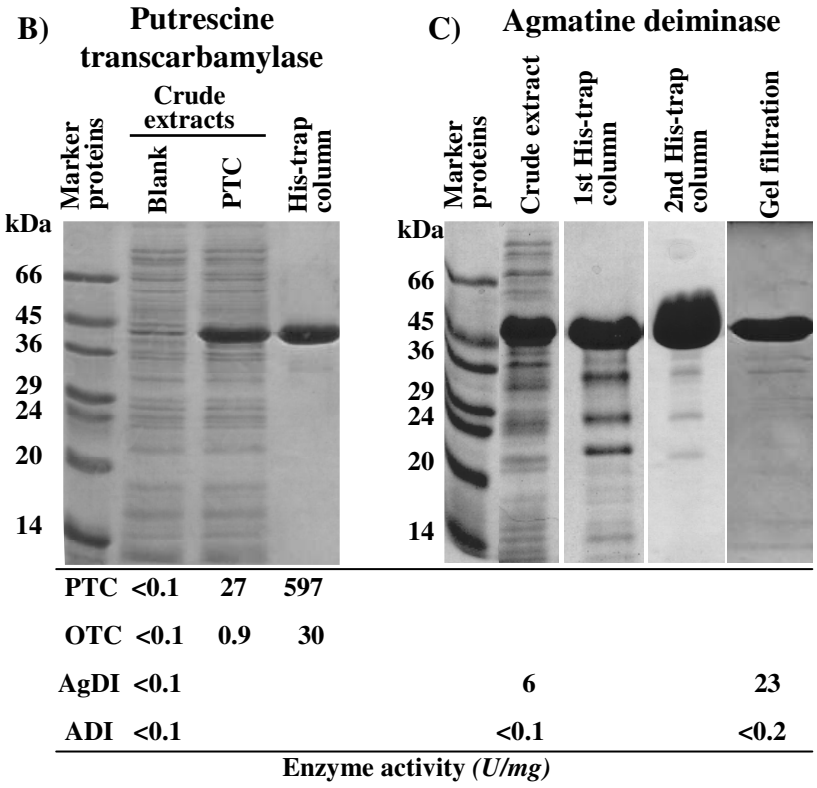
889 [c] $R_{sym} = \Sigma I - \langle I \rangle / \Sigma I$, where I is the observed intensity and $\langle I \rangle$ the average intensity. σ, standard

890 deviation.

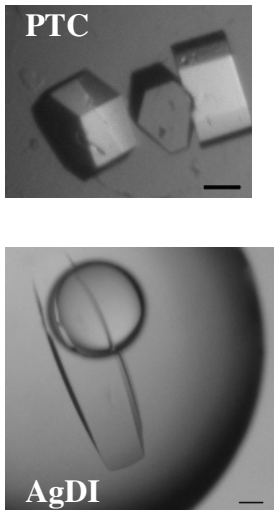
[d] R-factor = Σ_h ||F_{obs}| - |F_{calc}|| / Σ_h |F_{obs}|, where |F_{obs}| and |F_{calc}| are observed and calculated structure factors amplitudes for all reflections (R-factor). $R_{free,}$ R based on 5% of the data, withheld for the cross-validation test.

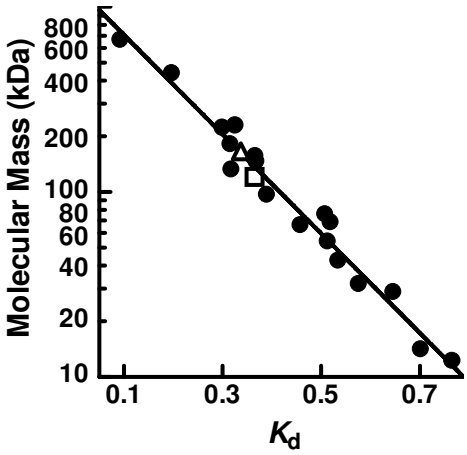
[e] Using PROCHECK (27)

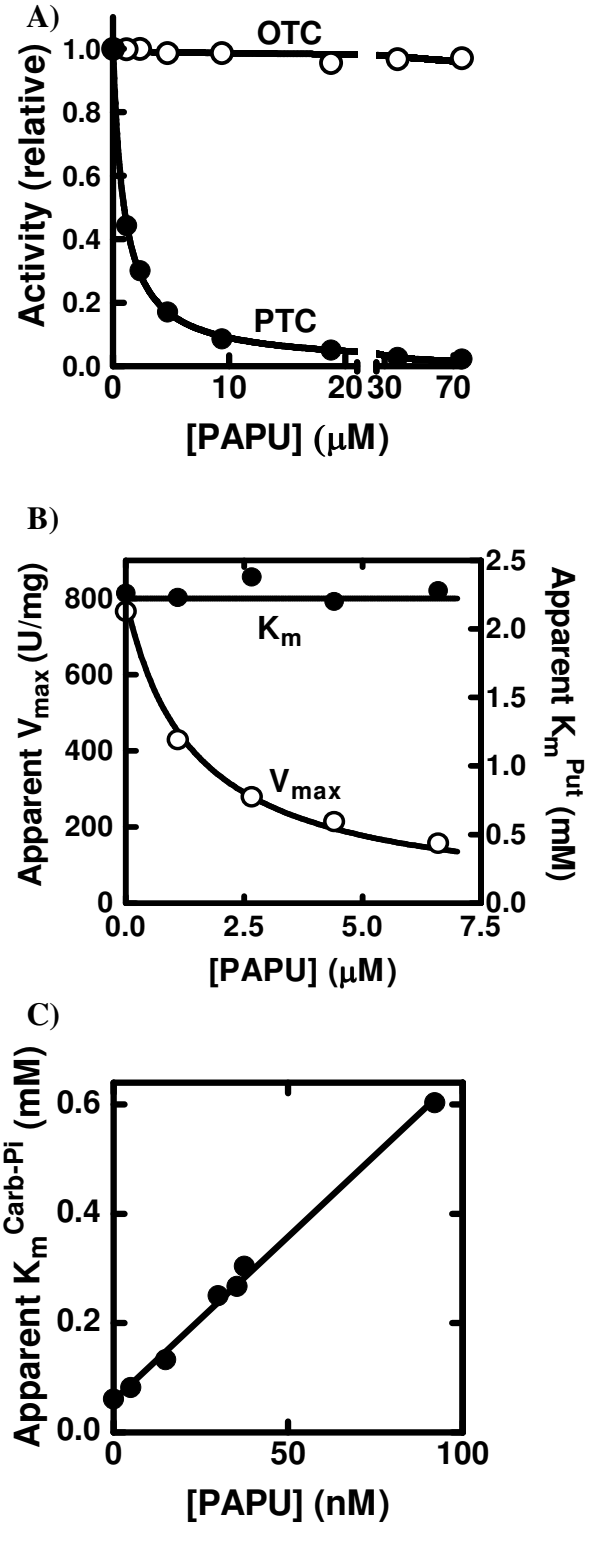


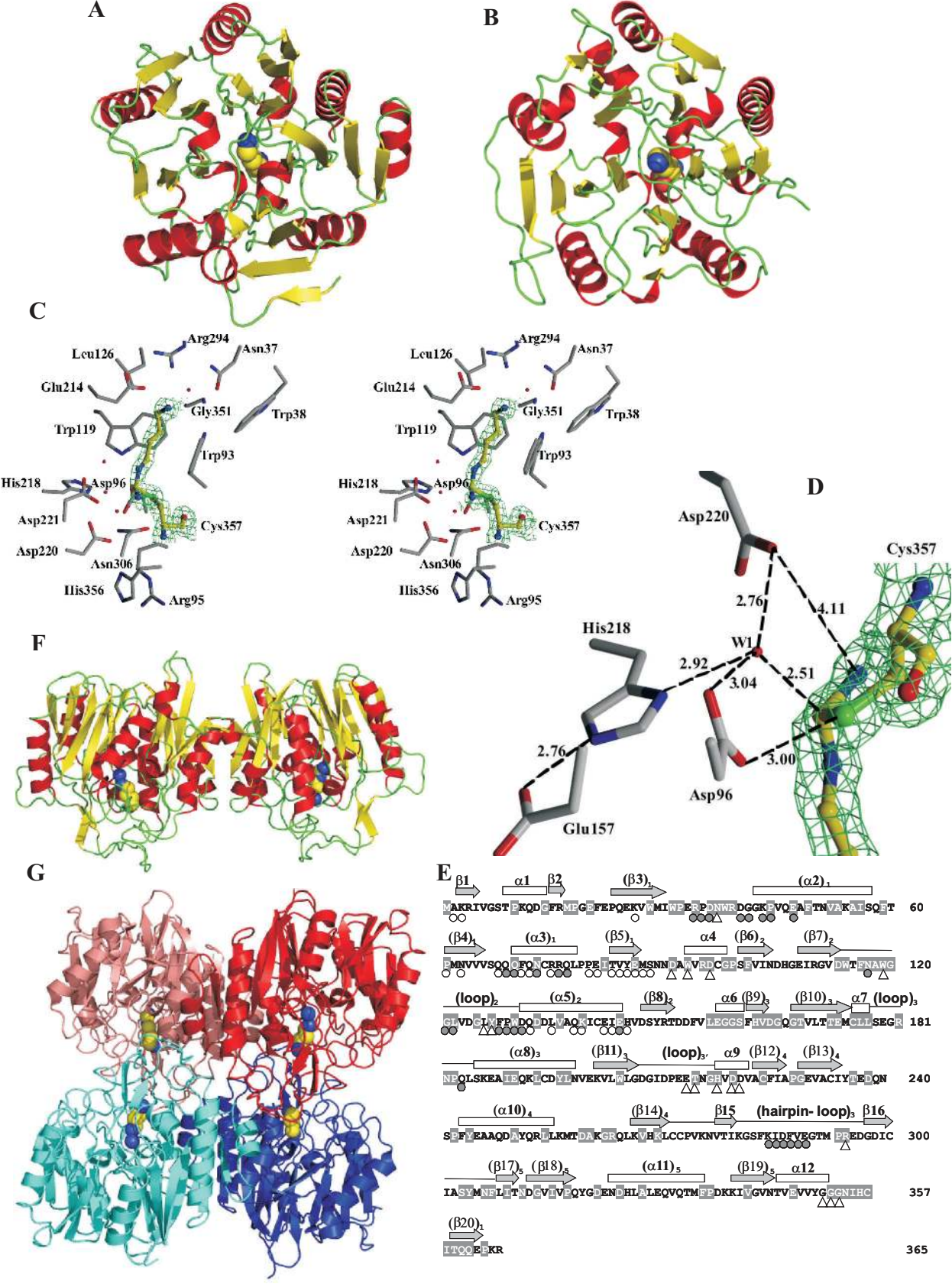


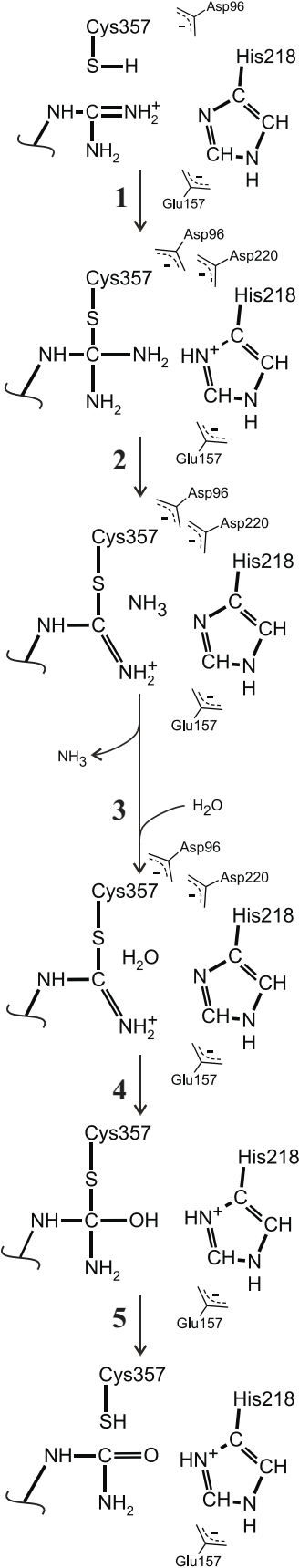
Enzyme activity *(U/mg)*						
PTC	<0.1	27	597			
OTC	<0.1	0.9	30			
AgDI	<0.1			6		23
ADI	<0.1			<0.1		<0.2




Molecular Mass (kDa)
800
600
400
200
100
80
60
40
20
10
0.1
0.3
0.5
0.7
K_d
A
D


A)
OTC
PTC
Activity (relative)
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
70
[PAPU] (μM)
B)
Apparent V_{max} (U/mg)
Apparent K_m^{Put} (mM)
K_m
V_{max}
800
600
400
200
0
2.5
2.0
1.5
1.0
0.5
0.0
0.0
2.5
5.0
7.5
[PAPU] (μM)
C)
Apparent $K_m^{Carb-Pi}$ (mM)
0.6
0.4
0.2
0.0
0
50
100
[PAPU] (nM)


A
B
C
Arg294
Asn37
Leu126
Glu214
Gly351
Trp38
Trp119
Trp93
His218
Asp96
Asp221
Cys357
Asp220
Asn306
His356
Arg95
D
Asp220
Cys357
2.76
4.11
His218
W1
2.92
3.04
2.51
2.76
3.00
Asp96
Glu157
F
G
E
β1
α1
β2
(β3)1
(α2)1
(β4)1
(α3)1
(β5)1
α4
(β6)2
(β7)2
(loop)2
(α5)2
(β8)2
α6
(β9)3
(β10)3
α7
(loop)3
(α8)3
(β11)3
(loop)3'
α9
(β12)4
(β13)4
(α10)4
(β14)4
β15
(hairpin- loop)3
β16
(β17)5
(β18)5
(α11)5
(β19)5
α12
(β20)1
60
120
181
240
300
357
365


Cys357
Asp96
His218
Glu157
1
Asp220
2
NH3
3
H2O
4
OH
5
SH