A single mutation in the active site swaps the substrate specificity of *N*-acetyl-L-ornithine transcarbamylase and *N*-succinyl-L-ornithine transcarbamylase

DASHUANG SHI,[1] XIAOLIN YU,[1] JUAN CABRERA-LUQUE,[1] TONY Y. CHEN,[2] LAUREN ROTH,[2] HIROKI MORIZONO,[1] NORMA M. ALLEWELL,[2] AND MENDEL TUCHMAN[1]

[1]Children's National Medical Center, The George Washington University, Washington, DC 20010, USA
[2]College of Chemical and Life Sciences, University of Maryland, College Park, Maryland 20742, USA

(RECEIVED March 30, 2007; FINAL REVISION May 14, 2007; ACCEPTED May 14, 2007)

Abstract

Transcarbamylases catalyze the transfer of the carbamyl group from carbamyl phosphate (CP) to an amino group of a second substrate such as aspartate, ornithine, or putrescine. Previously, structural determination of a transcarbamylase from *Xanthomonas campestris* led to the discovery of a novel *N*-acetylornithine transcarbamylase (AOTCase) that catalyzes the carbamylation of *N*-acetylornithine. Recently, a novel *N*-succinylornithine transcarbamylase (SOTCase) from *Bacteroides fragilis* was identified. Structural comparisons of AOTCase from *X. campestris* and SOTCase from *B. fragilis* revealed that residue Glu92 (*X. campestris* numbering) plays a critical role in distinguishing AOTCase from SOTCase. Enzymatic assays of E92P, E92S, E92V, and E92A mutants of AOTCase demonstrate that each of these mutations converts the AOTCase to an SOTCase. Similarly, the P90E mutation in *B. fragilis* SOTCase (equivalent to E92 in *X. campestris* AOTCase) converts the SOTCase to AOTCase. Hence, a single amino acid substitution is sufficient to swap the substrate specificities of AOTCase and SOTCase. X-ray crystal structures of these mutants in complexes with CP and *N*-acetyl-L-norvaline (an analog of *N*-acetyl-L-ornithine) or *N*-succinyl-L-norvaline (an analog of *N*-succinyl-L-ornithine) substantiate this conversion. In addition to Glu92 (*X. campestris* numbering), other residues such as Asn185 and Lys30 in AOTCase, which are involved in binding substrates through bridging water molecules, help to define the substrate specificity of AOTCase. These results provide the correct annotation (AOTCase or SOTCase) for a set of the transcarbamylase-like proteins that have been erroneously annotated as ornithine transcarbamylase (OTCase, EC 2.1.3.3).

Keywords: argF gene; transcarbamylase; *N*-succinylornithine; *N*-acetylornithine; arginine biosynthesis; protein engineering; crystal structure

Reprint requests to: Dashuang Shi, Children's Research Institute, Children's National Medical Center, 111 Michigan Avenue, NW, Washington, DC 20010-2970, USA; e-mail: dshi@cnmcresearch.org; fax: (202) 884-6014.

Abbreviations: ACDase, *N*-acetylcitrulline deacetylase; ANOR, *N*-acetylnorvaline; SNOR, *N*-succinylnorvaline; AODase, *N*-acetylornithine deacetylase; AOTCase, *N*-acetylornithine transcarbamylase; ATCase, aspartate transcarbamylase; CP, carbamyl phosphate; *m*-DAP, *meso*-diaminopime-late; OTCase, ornithine transcarbamylase; SOTCase, *N*-succinylornithine transcarbamylase; PTCase, putrescine transcarbamylase; WT, wild type.

Article published online ahead of print. Article and publication date are at http://www.proteinscience.org/cgi/doi/10.1110/ps.072919907.

The transcarbamylase family of enzymes catalyzes the transfer of a carbamyl group from carbamyl phosphate (CP) to an amino group of the second substrate. Aspartate transcarbamylase (ATCase) and ornithine transcarbamylase (OTCase) are the two best-known members of this family. Phylogenetic analysis for nearly 450 transcarbamylase sequences demonstrated that most of them are either ATCases or OTCases, forming two major branches in the phylogenetic tree (Naumoff et al. 2004). OTCases are essential enzymes that catalyze the carbamylation of

L-ornithine in the arginine biosynthetic pathway in lower organisms or in the urea cycle in mammals. Along the OTCase branch, several small clades exist between the root and OTCase branch. The work reported herein demonstrates that most of the transcarbamylase sequences in these clades are incorrectly annotated and belong to novel classes of transcarbamylase.

The canonical biosynthesis of arginine in microbes and plants proceeds from glutamic acid through eight enzymatic steps using acetylated intermediates in the first five steps (Cunin et al. 1986; Davis 1986; Slocum 2005). The recent discovery of *N*-acetylornithine transcarbamylase (AOTCase) indicates that some plant pathogens in the *Xanthomonadaceae* genus use a modified biosynthetic pathway to produce arginine (Shi et al. 2005a, 2006b; Morizono et al. 2006). In these organisms, two enzymes, *N*-acetylornithine deacetylase (AODase) and ornithine transcarbamylase (OTCase), are missing; instead, two novel enzymes, AOTCase and *N*-acetylcitrulline deacetylase (ACDase), replace them (Shi et al. 2005b, 2007). Thus, the order of deacetylation and transcarbamylation is reversed relative to the canonical pathway. When we compared the structure of AOTCase from *Xanthomonas campestris* with the structure of a transcarbamylase-like protein essential for arginine biosynthesis in *Bacteroides fragilis* (Shi et al. 2002), an additional novel enzyme, *N*-succinyl-transcarbamylase (SOTCase), was uncovered (Shi et al. 2006a). This discovery implies that *B. fragilis*, a clinically important human pathogen, and other bacteria with homologous enzymes may use succinyl intermediates rather than acetyl intermediates for arginine biosynthesis (Shi et al. 2006a). This finding provides interesting clues about the evolutionary relationships among enzymes in the arginine biosynthesis pathway and allows a new potential strategy for inhibiting human and plant pathogens using these unique enzymes as targets for antibacterial agents. The previous work also indicated that E92 of *X. campestris* AOTCase (the corresponding residue P90 in *B. fragilis* SOTCase) is critical for conferring specificity toward the second substrate.

In this paper, the activity and crystal structures of E92A, E92P, E92S, and E92V mutants of AOTCase from *X. campestris* and the P90E/T242L (T242L is a cloning artifact) mutant of SOTCase from *B. fragilis* have been investigated. The activity assays demonstrate that a single mutation in the active site interconverts the substrate specificity between AOTCase and SOTCase. The crystal structures of *X. campestris* AOTCase mutants complexed with CP and *N*-succinylnorvaline (SNOR) and a *B. fragilis* SOTCase mutant complexed with CP and *N*-acetylnovaline (ANOR) confirm that E92 (*X. campestris* numbering) is the key residue in determining substrate specificity.

Results and Discussion

Specificity of transcarbamylation activity

Specific activities of wild-type SOTCase from *B. fragilis*, wild-type AOTCase from *X. campestris*, and of the various mutants are listed in Table 1. These enzymatic assays clearly indicate that all four mutants of AOTCase preferentially catalyze the carbamylation of *N*-succinylornithine relative to two other potential substrates, *N*-acetylornithine and ornithine. Enzymatic activities of the mutants in the order from most to least active are: E92S > E92V > E92A > E92P. The shape and polarity of the side chains of the mutated residues also appear to significantly impact the catalytic activity since the E92P mutant shows 319-fold less activity than the E92S mutant. It is also interesting to note that the E92P mutant has 1135-fold less activity than the *B. fragilis* wild-type (WT) SOTCase in which the equivalent residue is also Pro. However, the specific activity of the P90E/T242L mutant of SOTCase, which preferentially catalyzes the carbamylation of *N*-acetylornithine, is only slightly less than the *X. campestris* wild-type

Table 1. *Specific activity of WT proteins and their mutants*

Protein	Specific activity (μmol/min/mg)			References
	ORN	AORN	SORN	
WT SOTCase	ND[a]	ND	238.4 ± 17.2[b]	Shi et al. (2006b)
T242L[c] SOTCase	ND	ND	195.5 ± 9.0	Same as above
P90E/T242L[c] SOTCase	ND	11.9 ± 0.6	1.6 ± 0.2	Same as above
WT AOTCase	ND	13.9 ± 0.4	ND	Morizono et al. (2006)
E92S AOTCase	ND	0.80 ± 0.03	66.9 ± 2.2	This work
E92A AOTCase	0.02 ± 0.01	ND	29.8 ± 1.1	This work
E92V AOTCase	ND	0.32 ± 0.02	40.5 ± 1.8	This work
E92P AOTCase	ND	0.05 ± 0.01	0.21 ± 0.01	This work

[a] ND, not detectable.
[b] Mean ± S.E. are shown ($n = 3$).
[c] T242L mutation is a cloning artifact.

AOTCase. While activity assays confirmed that Glu92 (*X. campestris* numbering) is a key residue involved in distinguishing AOTCase from SOTCase, the significantly lower catalytic efficiency of these mutants compared to the WT proteins implies that other structural differences also play a significant role in the catalytic reaction. To clarify these differences, structural information for these mutants complexed with bound substrates or substrate analogs is essential.

Structural model

The structures of all mutant proteins complexed with CP and their second substrate analog, ANOR or SNOR, have been determined. Refinement statistics of the final structural models are listed in Table 2. The stereochemistry, assessed using program PROCHECK (Laskowski et al. 1993), indicates that all structures have acceptable geometry. As in the WT protein structures, three residues, Glu144, Thr145, and Leu 295 (*X. campestris* numbering), are in energetically unfavorable conformations and there is one *cis* peptide bond between Leu 295 and Pro 296. The mutant structures are very similar to the WT structure with an RMS difference <0.25 Å when all equivalent Cα atoms are superimposed.

The electron density maps show clear and interpretable electron density in the active site corresponding to the bound ligands, CP and SNOR or CP and ANOR (Fig. 1A–E), even in the early stages of refinement. The conformation of the 80's loop where the mutation is located is well defined. The electron density around Glu92 of AOTCase and Pro90 of SOTCase verifies the presence of the engineered mutations.

Active site

One CP and one SNOR were found in the active site of *X. campestris* AOTCase E92A, E92S, E92V, and E92P mutants, while one CP and one ANOR were identified in the *B. fragilis* SOTCase P90E mutant. Almost all equivalent active site residues that interact directly with the substrates in SOTCase are also involved in binding CP and SNOR to all AOTCase mutants (Table 3). As in *B. fragilis* WT SOTCase (Shi et al. 2006a), two residues, His180 and Arg298, which hydrogen bond to the Glu92 from the adjacent subunit in the *X. campestris* WT AOTCase structure, were found to interact via hydrogen bonds with the succinyl group of SNOR in all *X. campestris* AOTCase mutant structures. Likewise, the interactions of CP and ANOR with the active site residues

Table 2. *Data collection and refinement statistics*

	AOTCase				SOTCase
	E92A	E92P	E92S	E92V	P90E/T242L
Inhibitors or substrates	CP and SNOR	CP-SNOR	CP-SNOR	CP and SNOR	CP and ANOR
Space group	$I2_13$	$I2_13$	$I2_13$	$I2_13$	$P4_3$
Resolution (Å)	2.3	2.5	2.8	2.8	2.9
Unit-cell parameters (Å)	$a = b = c = 129.4$	$a = b = c = 129.0$	$a = b = c = 129.0$	$a = b = c = 129.7$	$a = b = 156.7$ $c = 120.2$
Measurements	195,035	136,664	90,682	92,660	353,276
Unique reflections	16,186 (1608)	12,549 (1232)	8,953 (875)	9,016 (899)	64,177 (6.363)
Redundancy	12.0 (11.6)[a]	10.9 (10.6)	10.1(9.7)	10.3 (9.6)	5.5 (5.0)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	99.8 (100.0)	99.7 (100)	99.6 (99.3)
$\langle I/\sigma(I)\rangle$	31.1 (5.6)	24.5 (4.6)	15.8 (3.3)	15.3 (3.4)	9.8 (1.8)
R_{merg}[b]	7.1 (41.8)	10.1 (56.6)	15.3 (78.7)	14.3 (76.1)	14.5 (83.2)
Refinement					
Resolution range (Å)	50–2.3	50.0–2.5	50–2.8	50.0–2.8	50–2.9
No. of protein atoms	2,613	2,615	2,614	2,615	15,488
No. of water atoms	108	111	60	44	144
No. of hetero atoms	33	28	28	28	114
RMSD of bond lengths (Å)	0.006	0.006	0.006	0.007	0.006
RMSD of bond angle (°)	1.2	1.2	1.3	1.3	1.2
R_{work} (%)[c]	19.1 (23.3)	17.9 (23.7)	18.5 (26.7)	20.7 (29.1)	22.5 (36.6)
R_{free} (%)[d]	23.9 (32.4)	22.5 (32.7)	26.1 (33.7)	25.8 (32.3)	25.9 (38.5)
Average B factor (Å^2)	38.9	35.7	38.6	45.0	58.1

[a] Figures in parentheses apply to the highest-resolution shell.
[b] $R_{merg} = \Sigma_h\Sigma_i|I(h,i) - \langle I(h)\rangle|/\Sigma_h\Sigma_i I(h,i)$, where $I(h,i)$ is the intensity of the ith observation of reflection h, and $\langle I(h)\rangle$ is the average intensity of redundant measurements of reflection h.
[c] $R_{work} = \Sigma_h|F_{obs} - F_{calc}|/\Sigma_h F_{obs}$.
[d] $R_{free} = \Sigma_h|F_{obs} - F_{calc}|/\Sigma_h F_{obs}$ for 5% or 10% of the reserved reflections.

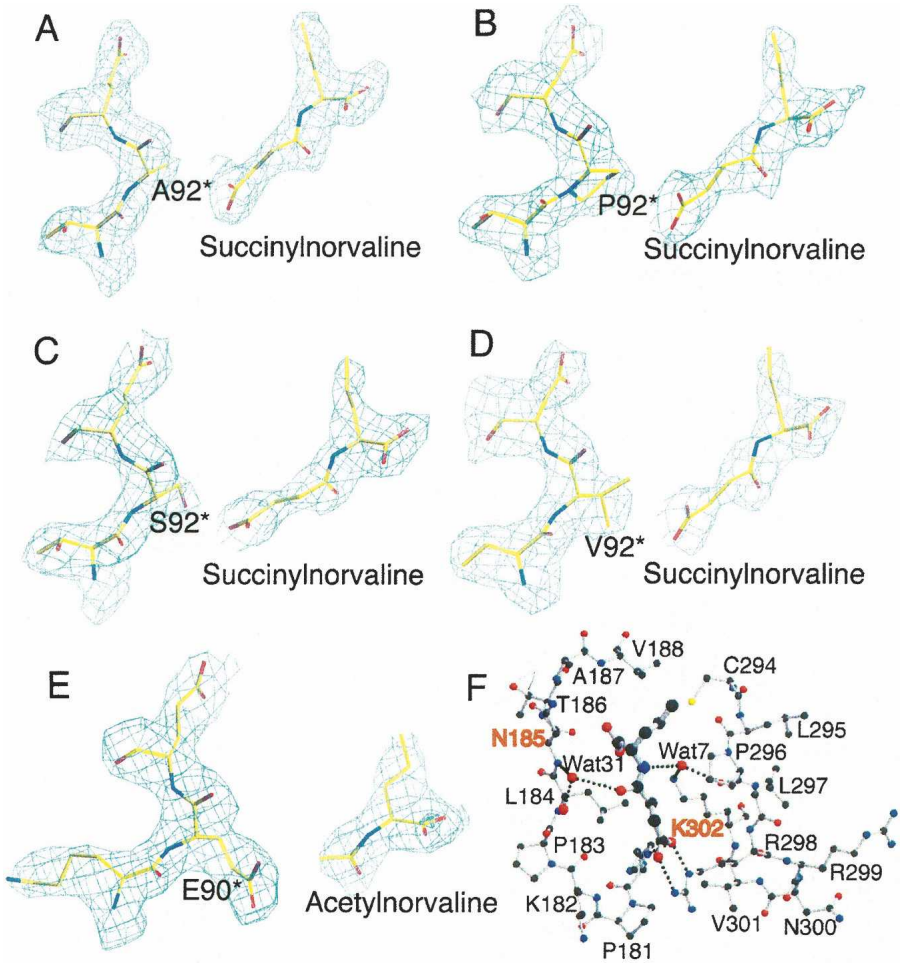

Figure 1. *N*-acetylnorvaline or *N*-succinylnorvaline binding of SOTCase or AOTCase mutants. (*A–D*) Contours of the electron density map ($2F_o - F_c$) (1.0 σ shown in blue cage) around *N*-succinylnorvaline and E92 region of *X. campestris* AOTCase mutants. (*E*) Contours of the electron density map ($2F_o - F_c$) (1.0 σ shown in blue cage) around *N*-acetylnorvaline and P90 region of *B. fragilis* SOTCase mutant. The final refined positions of the ligands and residues are represented as colored sticks. The residue indicated by * is from the adjacent subunit. The electron density clearly indicates the binding of substrate analogs and the expected amino acid substitutions. (*F*) *N*-succinylnorvaline binding site of *X. campestris* E92A mutant. SNOR is shown as thick ball-and-stick models, while the surrounding residues are shown as thin ball-and-stick models. Two water molecules, Wat7 and Wat31, form hydrogen-bonding bridges to link the ligand to protein residues. These water-mediated hydrogen-bonding networks can only be found in *X. campestris* AOTCase and its mutants, but not in *B. fragilis* SOTCase and mutants. The residues, N185 and K302 (*X. campestris* AOTCase numbering) shown in red, are proposed as the residues that distinguish SOTCase from AOTCase in addition to the key residue of E92.

in *B. fragilis* SOTCase P90E mutant are similar to those in *X. campestris* WT AOTCase (Shi et al. 2006b). The Glu90 (mutated from Pro90) interacts with His176 and Arg278 (*B. fragilis* numbering) in a manner similar to WT AOTCase, thus occupying the space that in WT SOTCase accommodates the succinyl moiety, and precluding SNOR binding to the P90E mutant of SOTCase.

Even though the data quality, especially the presence of high resolution data, will influence the number of water molecules identified from the electron density map, four water molecules involved in binding ANOR or SNOR are conserved in *X. campestris* WT AOTCase and all its mutant structures (Table 3). In contrast, only two of them can be identified in the *B. fragilis* WT SOTCase structure and none were found in the P90E SOTCase mutant structure. Two residues in the active site, Asn185 and Lys302 (*X. campestris* numbering), appear to affect the role of water in substrate binding and catalytic activity. One molecule of water links the acyl O atom of ANOR or SNOR to the main-chain N atom of Asn185 in *X. campestris* WT or mutant structures (Fig. 1F). The equivalent residue in *B. fragilis* is Pro181, which is unable to contribute its main-chain N atom for hydrogen bonding. Consistent with this observation, all known

Table 3. *Active site interactions in AOTCase, and SOTCase and their mutants*

Hetero atoms	Protein atoms or water atoms	Distances (Å)							
	SOTCase			AOTCase					
CP		WT	P90E/T242L		WT	E92A	E92S	E92V	E92P
O1P	R110 NH1	2.61	2.75	R112 NH2	2.44	2.43	2.46	2.60	2.44
				W77* NE1	2.78	2.72	2.68	2.85	2.69
O2P	S47 OG	2.58	2.54	S49 OG	2.78	2.76	2.64	2.70	2.76
	T50N	2.57	2.72	T52 N	2.84	2.91	3.03	3.02	3.02
	T50OG1	3.06	2.95	T52 OG1	2.58	2.64	2.59	2.65	2.64
O3P	L48 N	2.75	2.63	M50 N	2.79	2.68	2.75	2.60	2.64
	R49N	3.14	2.92	R51 N	3.07	2.88	2.80	2.69	2.74
	R49NH2	3.06	3.08	R51 NH2	2.88	2.92	2.98	2.84	2.95
	W75[a] NE1	2.90	3.09						
OA	R49 NE	2.88	2.94	R51 NE	2.82	2.85	2.95	2.69	2.73
O1	T50 OG1	2.66	2.58	T52 OG1	3.04	3.05	2.98	3.06	2.96
	R110 NH2	2.74	2.90	R112 NH1	3.16	3.24	2.46	3.20	3.33
	H147 NE2	3.29	3.20	H148 NE2	3.03	3.03	2.87	3.11	2.99
	R302 NH1	3.19	2.86	R322 NH1	3.09	3.16	3.20	3.45	2.95
NP	Q150 OE1	3.00	3.07	Q151 OE1	3.01	2.78	2.78	2.70	2.85
	C274 O	3.17	3.24	C294 O	3.04	2.92	3.13	3.22	2.94
	L275 O	3.03	3.17	L295 O	3.12	2.97	3.31	3.31	3.04
SNOR or ANOR									
OXT	K236 NZ	2.82	2.83	K252 NZ	2.69	2.62	2.76	2.83	2.57
				Wat O[b]			3.04	3.17	3.23
O	E142 OE1	2.66	2.71	E144 OE1	2.60	2.61	2.53	2.60	2.51
	Wat O	3.17		Wat O	2.66	2.91	3.02	2.77	2.73
N1				Wat O	3.18	3.12	3.02	3.22	3.05
O1	Wat O	2.70		Wat O	2.52	2.63	2.76	2.72	2.59
				Wat O	3.23		3.26		3.01
OD1	H180 NE2	2.94		H180 NE2		2.89	2.88	2.51	2.67
	R278 NE	2.91		R298 NE		2.62	2.87	2.98	2.85
				Wat O		2.48			2.56
OD2	R278 NH2	2.89		R298 NH2		2.82	2.60	2.94	2.83
				Wat O		2.79	2.79	2.97	2.65
				Wat O		3.01			
				(S92 OG)			2.60		

[a] Residue is from an adjacent monomer.
[b] Wat O is different water molecule in each row.

sequences of SOTCases and AOTCases classified on the basis of the residue at the Glu92 position (*X. campestris* numbering) have an Asn in the equivalent position of AOTCase and a Pro in SOTCase (Fig. 2). Similarly, Lys302 may help a water molecule hydrogen bind to the imino N atom of AORN or SORN and the main-chain carbonyl O atom of Pro296 (Fig. 1F). The presence of Glu, Val, or Ile in the equivalent position of SOTCase appears to prevent water from participating in hydrogen bonding. The different ways in which water is involved in substrate binding in AOTCase and SOTCase appear to be one reason for the significantly lower catalytic efficiency of the mutants relative to the WT protein. Additional mutants such as N185P and K302I (*X. campestris* numbering) need to be studied to clarify this prediction.

As pointed out recently by two independent groups there is a novel deep trefoil knot in AOTCase and SOTCase structures close to the active site that is not present in ATCase or OTCase structures (Fig. 3; Khatib et al. 2006; Virnau et al. 2006). It appears that the knot modifies the active sites to allow AOTCase and SOTCase to recognize acetylated or succinylated substrates and to use different amino acid motifs in substrate recognition. In ATCase and OTCase, the flexible 240's loop (termed SMG loop in OTCase) is the main recognition site for the second substrate, aspartate or ornithine (Krause et al. 1987; Gouaux and Lipscomb 1988; Ke et al. 1988; Ha et al. 1997; Shi et al. 1998, 2001), while in AOTCase and SOTCase, the unique proline-rich loop becomes part of the second substrate binding site.

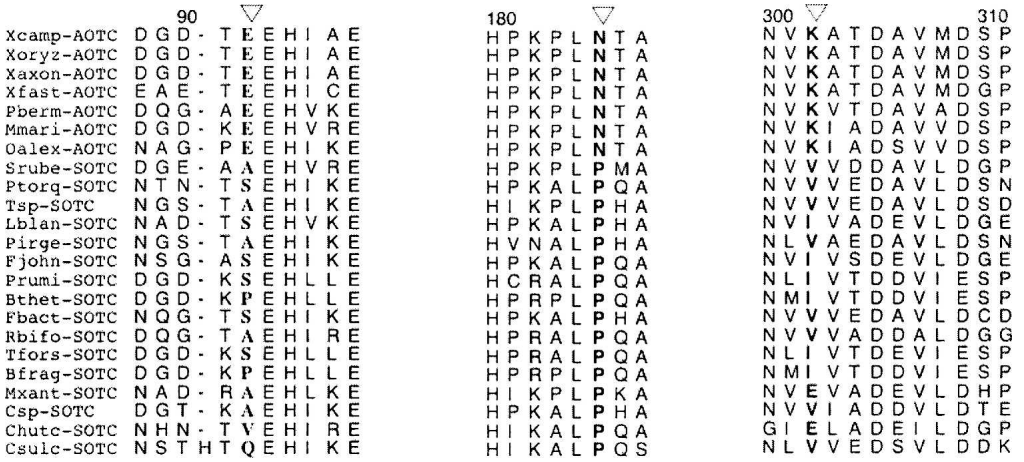

Figure 2. Sequence alignments of the regions containing E92, N185, and K302 (shown in bold and indicated by a triangle) to define the substrate specificity of AOTCase and SOTCase. The sequences are from bacteria, *Xanthomonas campestris*, *Xanthomonas oryzae*, *Xanthomonas axonopodis*, *Xylella fastidiosa, Parvularcula bermudensis*, *Maricaulis maris*, *Oceanicaulis alexandrii*, *Salinibacter ruber*, *Psychroflexus torquis*, *Tenacibaculum sp*. Med152, *Leeuwenhoekiella blandensis* MED217, *Polaribacter irgensii* 23-P, *Flavobacterium johnsoniae* UW101, *Prevotella ruminicola*, *Bacteroides thetaiotaomicron*, *Bacteroides fragilis*, *Flavobacteriales bacterium* BBFL7, *Robiginitalea biformata*, *Tanneralla forsythensis*, *Myxococcus xanthus*, *Cellulophaga sp*. MED134, *Cytophaga hutchinsonii*, and *Candidatus Sulcia Muelleri*.

Structural comparison of X. campestris AOTCase E92S, E92V, E92A, and E92P mutants

Even though the specific SOTCase activities of these four mutants range from 66.9 μmol/min/mg for E92S to 0.22 μmol/min/mg for E92P, their structures are very similar with RMS <0.25 Å for all superimposed 332 Cα atoms. The relatively high activity of E92S may be a result of a favorable hydrogen-bonding interaction between the side chain of the mutated residue and the succinyl group of substrate which is not present in the other mutants. The reason for the extremely low activity of E92P mutant needs to be investigated further.

SOTCase and AOTCase in bacteria

Although other subtle differences may affect the enzymatic activity of AOTCase and SOTCase, Glu92 (*X. campestris* numbering) plays a key role in distinguishing AOTCase from SOTCase. In addition to bacteria such as *Bacteroides fragilis*, *Bacteroides thetaiotaomicron*, *Cytophaga fastidiosa*, *Tannerella forsythensis*, *Prevotella ruminicola*, and *Myxococcus xanthus* in which SOTCase has been previously identified in the genomes (Shi et al. 2006a), other bacteria, including *Robiginitalea biformata*, *Polaribacter irgensii*, *Flavobacterium sp*., *Flavobacteriales bacterium*, *Cellulophaga sp*., and *Salinibacter ruber*,

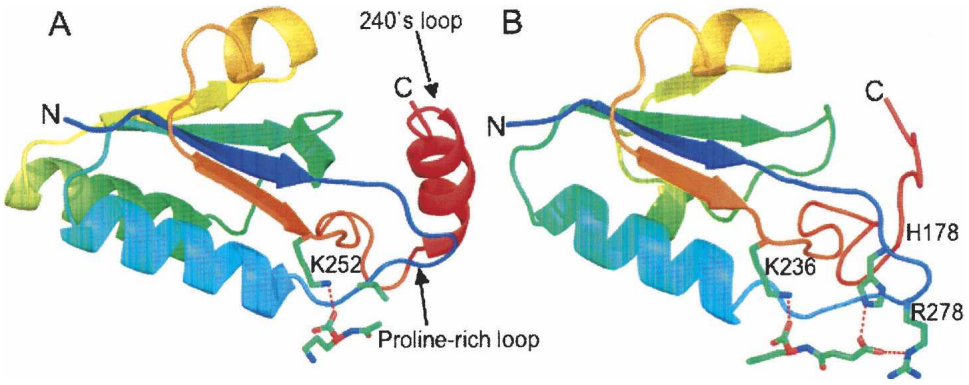

Figure 3. A ribbon diagram of the knotted region of *X. campestris* AOTCase (*A*) (residues 170–275) and *B. fragilis* SOTCase (*B*) (residues 166–255) with bound ligands shown in colored sticks. Colors change continuously from blue (first residue in the region) to red (last residue in the region). The proline-rich loop (from dark blue to light blue), in similar conformation for both AOTCase and SOTCase, appears to be critical for the knot formation. The 240's loop (from brown to red) was threaded through the proline-rich loop to form a trefoil knot.

also have transcarbamylases that lack the signature Glu–Asn–Lys triad of AOTCase, suggesting that they are almost certainly SOTCases and that their arginine biosynthetic pathways likely use succinylated intermediates. Conversely, in addition to *Xanthomonas campestris*, *Xanthomonas axonopodis*, and *Xylella fastidiosa*, which we identified as AOTCases previously, *Maricaulis maris*, *Oceanicaulis alexandrii*, and *Parvularcula bermudensis* also have the AOTCase signature and are, therefore, likely to be AOTCases. These bacteria possess a modified pathway to synthesize arginine but still use acetyl intermediates as in the canonical pathway found in most microbes (Fig. 4). More interestingly, bacteria which use AOTCase to convert acetylornithine to acetylcitrulline in arginine biosynthesis have a unique bifunctional enzyme that catalyzes the first two steps of the pathway (Qu et al. 2007). The reannotation of these transcarbamylase sequences as AOTCase and SOTCase is provided in Table 4. The function of these transcarbamylases needs to be confirmed by enzymatic assay, particularly for the transcarbamylase (NBCI accession No. ZP_01308773) from *Candidatus Sulcia muelleri*, which has Gln in the position equivalent to Glu92 but lacks the Asn and Lys residues in the other two equivalent positions.

Classification of transcarbamylases

The three-dimensional structures of four distinct members of the transcarbamylase superfamily, ATCase, OTCase, AOTCase, and SOTCase, are now available. The structures, especially for those bound with substrates or substrate analogs, provide important information about how enzymes recognize their substrates. The classification scheme is shown in Figure 5. The highest conservation across the transcarbamylase superfamily involves three regions: the SxRT motif, the HPxQ motif, and the HxLP motif (Allewell et al. 1999; Shi et al. 2001). These conserved motifs define the CP binding site characteristic of all known transcarbamylases including putrescine transcarbamylase (PTCase) (Naumoff et al. 2004) and other unannotated transcarbamylases in the genome databases. Therefore, these three motifs can be used to screen for all transcarbamylases in genomic databases.

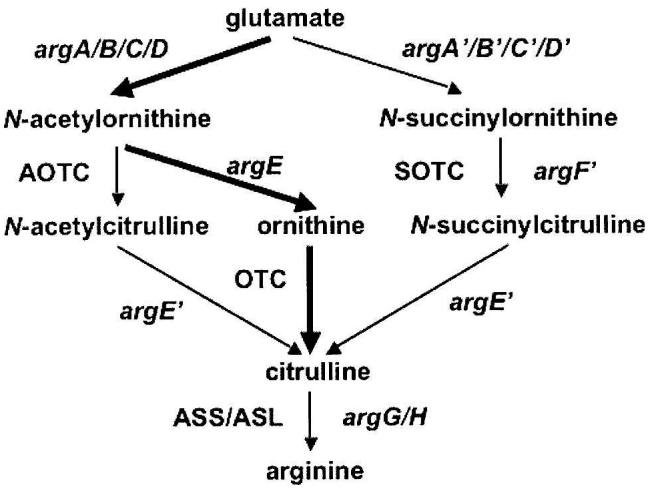

Figure 4. Comparison of the canonical and two proposed linear arginine biosynthetic pathways. The canonical pathway is shown in thick arrows.

The sequence variations in four loop regions, 80's loop, 120's loop, proline-rich loop, and 240's loop, confer specificity for second substrate binding. The presence of the extended 80's and 120's loops and the proline-rich loop seem to be characteristic of AOTCase and SOTCase. We can use these features to distinguish AOTCase and SOTCase from other transcarbamylases while AOTCase can be distinguished from SOTCase using three residues, Glu92, Asn185, and Lys302 (*X. campestris* numbering). As noted previously, the 240's loop is the main binding site for the second substrate in ATCase and OTCase. The specific features of the 240's loop, the RxQxER motif in ATCase and the DxxxSMG motif in OTCase, distinguish ATCase and OTCase from other transcarbamylases, as well as from each other. No structural information for the remaining members of the transcarbamylases, such as PTCase, is available. Sequence alignment indicates the putrescine transcarbamylases have high sequence similarity to OTCase with sequence identity ~35%–45% but do not have the DxxxSMG motif which is characteristic of OTCase.

Disordered regions in proteins evolve more rapidly than ordered regions (Brown et al. 2002), and loop regions seem to evolve faster than regions which have secondary structures. This observation is consistent with the notion that all transcarbamylases have evolved from a common ancestor, with branches defined by the sequence diversity of their loop regions.

Three possible arginine biosynthesis pathways in bacteria

Our results demonstrate that arginine biosynthesis occurs via three slightly different pathways in prokaryotes (Fig. 4). Both newly discovered pathway variants follow similar routes in which *N*-succinyl-L-ornithine or *N*-acetyl-L-ornithine is first carbamylated to form *N*-succinyl-L-citrulline or *N*-acetyl-L-citrulline, respectively, then desuccinylated or deacetylated to produce citrulline (Shi et al. 2005a, 2006a). In contrast, in the canonical arginine biosynthetic pathway, *N*-acetyl-L-ornithine is first deacetylated to ornithine, then carbamylated to form citrulline. The acetyl pathway is utilized by most microorganisms and plants, while the succinyl pathway seems to be used only by certain bacteria.

Table 4. *New annotation of putative transcarbamylase sequences as AOTCase or SOTCase*

Species name	NCBI	GI	Database annotation	Proposed annotation
Xanthomonas campestris pv.	YP_24952	66768190	OTCase	AOTCase
Campestris str. 8004	NP_637604	21231687		
Xanthomonas campestris pv. Vesicatoria str. 85–10	YP_364252	78048107	OTCase	AOTCase
Xanthomonas oryzae pv. Oryzae KACC10331	YP_201315	58582299	OTCase	AOTCase
	YP_451551	84624179		
Xanthomonas axonopodis pv. Citri str. 306	NP_642668	21243086	OTCase	AOTCase
Xylella fastidiosa Temeculal	NP_778526	28198212	OTCase	AOTCase
Xylella fastidiosa Ann-1	ZP_00680033	71897807	ATC/OTCase	AOTCase
	EAO034416.1	71732362		
	ZP_00651934	71275649		
	EAO13257	71163540		
Xylella fastidiosa 9a5c	NP_298288	15837600	OTCase	AOTCase
Oceanicaulis alexandrii HTCC2633	ZP_00952755	83859234	OTCase	AOTCase
	EAP90534	83852681		
Parvularcula bermudensis HTCC2503	ZP_01018069	84704569	OTCase	AOTCase
	EAQ15247	84689405		
Stenotrophomonas maltophilia R551–3	ZP_01643361	119876311	ATC/OTCase	AOTCase
Maricaulis maris[a]	YP_755595	114568915	OTCase	AOTCase
Salinibacter ruber DSM 13855	YP_446323	83814628	ATC/OTCase	SOTCase
Myxococcus xanthus DK 1622	YP_633258	108758773	OTCase	SOTCase
Cytophaga hutchinsonii	ZP_00309405	48855246	OTCase	SOTCase
	YP_679667	110639458		
	ABG60325	110282139		
Flavobacterium johnsoniae UW101	ZP_01243099	90587446	ATC/OTCase	SOTCase
	EAS61751	90436395		
Flavobacteria bacterium BBFL7	ZP_01202604	89891096	OTCase[b]	SOTCase
	EAS19399	89516740		
Flavobacteriales bacterium HTCC2170	ZP_01107420	88713337	OTCase[b]	SOTCase
	EAR00484	88708247		
Cellulophaga sp. MED134	ZP_01050687	86132091	OTCase[b]	SOTCase
	EAQ38605	85817425		
Leeuwenhoekiella blandensis MED217	ZP_01060680	86142170	OTCase[b]	SOTCase
	EAQ49379	85830922		
Tenacibaculum sp. Med152	ZP_01052148	86133566	OTCase[b]	SOTCase
	EAQ41576	85820429		
Psychroflexus torquis ATCC 700755	ZP_01253266	91216299	OTCase	SOTCase
	EAS71813	91185437		
Bacteroides fragilis NCTC 9343	YP_210161	60680017	OTCase[b]	SOTCase
Bacteroides fragilis YCH46	YP_097777	53711785	OTCase	SOTCase
Bacteroides thetaiotaomicron VPI-5482	NP_812628	29349125	OTCase	SOTCase
Candidatus Sulcia muelleri Str. Hc	ZP_01308773	94502294	OTCase	SOTCase[b]
	EAT14092	94451146		
Polaribacter irgensii 23-P	ZP_01118859	88803333	OTCase[b]	SOTCase
	EAR12078	88780899		
Robiginitalea biformata HTCC2501	ZP_01120437	88804917	OTCase	SOTCase
	EAR16965	88785796		
Tannerella forsythensis	AAO33834[c]	28274159	ATC/OTCase	SOTCase
Algoriphagus sp. PR1	ZP_01717877	126645333	OTCase[b]	SOTCase
	EAZ82908	126578744		
Gramella forsetii KT0803	YP_862135	120436449	AOTCase	SOTCase
	CAL67068	117578599		
Microscilla marina ATCC 23134	ZP_01694046	124009368	OTCase	SOTCase
Polaribacter dokdonensis MED152	EAD41576	85820429	OTCase[b]	SOTCase
Stigmatella aurantiaca DW4/3–1	ZP_01466631	115379540	OTCase	SOTCase
Microscilla marina ATCC 23134	EAY24981	123985030	OTCase	SOTCase
	ZP_01694046	124009368		
Gramella dorsetii	YP_862135	120436449	AOTCase	SOTCase
	CAL67068	117578599		
Dordonia donghaensis MED 134	EAQ38605	85817425	OTCase[b]	SOTCase
Prevotella ruminicola[d]				SOTCase

[a] The complete sequence can also be found in Computer Biology at ORNL (http://genome.ornl.gov/microbial/mari_mar, gene Mmar100364 at Contig13).
[b] The proteins were annotated as putative OTCase, ATCase, or SOTCase.
[c] This NCBI record has only partial sequence. The complete sequence can be found in Oral Pathogen Sequence database (http://www.oralgen.lanl.gov/).
[d] The complete sequence can be found in TIGR microbial database (http://www.tigr.org/tdb/mdb/mdbinprogress.html Contig 702 from 2,650,367 to 2,653,332).

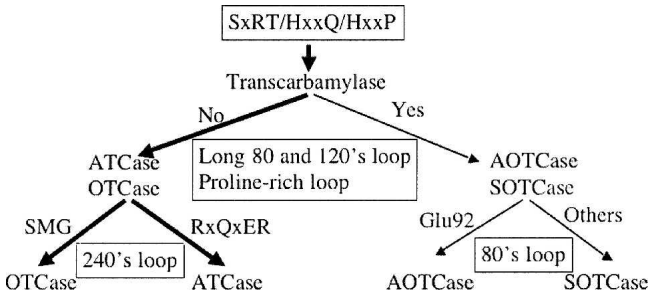

Figure 5. A schematic diagram representing the classification of four known transcarbamylases, ATCase, OTCase, AOTCase, and SOTCase, based on structure-derived amino acid signature motifs.

Variant pathways have been observed in *meso*-diaminopimelate (*m*-DAP)/lysine biosynthesis (Born and Blanchard 1999). In prokaryotes, three slightly different pathways, which diverge after the production of tetrahydrodipicolinate, lead to the formation of *m*-DAP and lysine. Two follow a similar route, using succinylated intermediates in most bacteria but acetylated intermediates in certain *Bacillus* species. The third route, found in a few members of the firmicutes, uses *m*-DAP dehydrogenase to convert tetrahydrodipicolinate to *m*-DAP directly, circumventing the use of acyl intermediates and shortening the pathway. In contrast to the arginine biosynthetic pathway in which the acetyl pathway dominates, the succinyl pathway for *m*-DAP/lysine biosynthesis is used by all Gram-negative and many Gram-positive bacteria while the acetyl pathway is limited to certain *Bacillus* species.

Materials and Methods

Site-directed mutagenesis, protein expression, and purification

Site-directed mutagenesis was carried out using the ''QuikChange'' mutagenesis kit (Stratagene) according to the manufacturer's protocol. The primers used to create E92P, E92V, E92A, and E92S mutants of *X. campestris* AOTCase and the P90E mutant of *B. fragilis* SOTCase are listed in Table 5. Preparation of the parents plasmids of *X. campestris* AOTCase and *B. fragilis* SOTCase was described previously (Shi et al. 2002, 2005a, 2006b). The correct mutants were confirmed by DNA sequencing. Recombinant proteins were expressed and purified as described previously (Shi et al. 2005a, 2006b). The protein purity was verified by SDS-PAGE (12% polyacrylamide gel) followed by Coomassie staining; a single band of the expected molecular mass was observed for each purified protein. Protein concentration was determined by the Bradford method using the BioRad protein assay dye reagent with bovine serum albumin as a standard (Bradford 1976).

Enzymatic assays

The modified colorimetric assay, which detects ureido groups, was used to test for enzymatic activity (Pastra-Landis et al. 1981). The method has been successfully used for the enzymatic assay of ATCase (Pastra-Landis et al. 1981), OTCase (Morizono et al. 1997), AOTCase (Morizono et al. 2006), and SOTCase (Shi et al. 2006a). With the exception of the E92P activity assay, the reaction mixtures contained in a final volume of 1 mL were 50 mM Tris-HCl buffer pH 8.3, 1 mM carbamyl phosphate (dilithium salt), 10 mM succinylornithine or acetylornithine or ornithine, and 0.2 μg of enzyme. Activity assay for E92P were performed by the same method with the exception that 0.12 mg of enzyme was used since its activity is much lower than that of the other proteins. All enzymes were diluted prior to the reaction with filtered storage buffer containing 20 mL Tris-HCl, 250 mM NaCl, and 1 mM EDTA pH 7.4. Reactions were initiated by adding the second substrate, run for 5 min at 25°C, and the reactions were quenched by adding 1 mL of quench reagent containing antipyrine in 50% H_2SO_4 and butadione monoxime in 2:1 ratio. Standard curves were prepared by adding varying amounts of succinylcitrulline or acetylcitrulline or citrulline in 50 mM Tris-HCl pH 8.3 to a total volume of 1 mL followed by the addition of 1 mL quench reagent. All reaction tubes were covered with foil and shaken in a dark chamber overnight before developing color at 42°C hot bath for 25 min. The absorbance was then measured at 466 nm.

Crystallization and data collection

The purified protein was concentrated to ~10 mg/mL with an Amicon-Y30 membrane concentrator (Millipore). The E92P, E92A, E92V, and E92S mutants of *X. campestris* AOTCase complexed with CP and SNOR were crystallized in similar condition as wild-type AOTCase (Shi et al. 2005a), while the P90E mutant of *B. fragilis* SOTCase complexed with CP and ANOR was crystallized as described previously (Shi et al. 2006a).

During data collection, a crystal was dipped quickly into a mother solution containing 20% (v/v) ethylene glycerol for cryoprotection. The cryoprotected crystals were frozen by plunging crystals directly into liquid nitrogen. Data were collected

Table 5. *Primers used for mutagenesis*

Mutants	Primer sequence
E92S AOTCase	5′-TGGACGGCGATACC**TCG**GAGCACATCGCCGAGG-3′
E92A AOTCase	5′-TGGACGGCGATACC**GCG**GAGCACATCGCCGAGG-3′
E92V AOTCase	5′-TGGACGGCGATACC**GTG**GAGCACATCGCCGAGG-3′
E92P AOTCase	5′-TGGACGGCGATACC**CCG**GAGCACATCGCCGAGG-3′
P90E SOTCase	5′-TGGACGGTGACAAA**GAG**GAACATCTGCTGGAAG-3′

The mutated nucleotides are highlighted.

on an R-axis IV image-plate diffractometer mounted on a Rigaku RU-200 rotating-anode generator with a Cu target (λ = 1.54178 Å). All data were processed using the HKL2000 package (Otwinowski and Minor 1997) and reduced using the program TRUNCATE in the CCP4 suite (Collaborative Computational Project, Number 4, 1994). The data collection statistics are summarized in Table 2.

Structure solution and refinement

The five structures reported herein were refined by slight variations of a basic protocol. The starting model for AOTCase mutants taken from the structure of wild-type AOTCase complexed with CP and ANOR (PDB code 2ZQ6) were stripped of the solvent atoms and hetero compounds. The starting model for the SOTCase P90E mutant was taken from the SOTCase structure complexed with CP and SNOR (PDB code 2FG7). The models were originally treated as single rigid bodies and refined with CNS 1.1 (Brünger et al. 1998) to 3 Å resolution. Further refinements involved simulated annealing, individual *B*-factor refinement manual with CNS 1.1, and model building with O (Jones et al. 1991). During the refinement process, 5% or 10% of the data, randomly selected, were used to monitor the free *R*-value (Brünger 1992). Water molecules were added using the WATERPICK protocol with CNS (Brünger et al. 1998). The final refinement statistics are listed in Table 2. The model quality was checked with the program PROCHECK (Laskowski et al. 1993), which indicated good stereochemistry for all structural models.

Figures were drawn using the programs MOLSCRIPT (Kraulis 1991), Raster3D (Merrit and Murphy 1994), PyMOL (DeLano 2002), and O (Jones et al. 1991). The coordinates have been deposited with the RCSB Protein Data Bank as entries 2G7M, 2G6C, 2G6A, 2G65, and 2G68.

Acknowledgments

This work was supported by Public Health Service grant DK-47870 (M.T.) and DK-067935 (D.S.) from the National Institute of Diabetes and Digestive and Kidney Diseases. We thank Dr. David Davies for facilitating our use of the diffraction equipment in the Molecular Structure Section of the National Institute of Health and Dr. Fred Dyda for help in data collection and processing.

References

Allewell, N.M., Shi, D., Morizono, H., and Tuchman, M. 1999. Molecular recognition by ornithine and aspartate transcarbamoylases. *Acc. Chem. Res.* **32:** 885–894.

Born, T.L. and Blanchard, J.S. 1999. Structure/function studies on enzymes in the diaminopimelate pathway of bacterial cell wall biosynthesis. *Curr. Opin. Chem. Biol.* **3:** 607–613.

Bradford, M. 1976. Quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72:** 1219–1223.

Brown, C.J., Takayama, S., Campen, A.M., Vise, P., Marshall, T.W., Oldfield, C.J., Williams, C.J., and Dunker, A.K. 2002. Evolutionary rate heterogeneity in proteins with long disordered regions. *J. Mol. Evol.* **55:** 104–110.

Brünger, A.T. 1992. Free R value: A novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355:** 472–475.

Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Crosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr.* **D54:** 905–921.

Collaborative Computational Project, Number 4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.* **D50:** 760–763.

Cunin, R., Glansdorff, N., Pierard, A., and Stalon, V. 1986. Biosynthesis and metabolism of arginine in bacteria. *Microbiol. Rev.* **50:** 314–352.

Davis, R.H. 1986. Compartmental and regulatory mechanisms in arginine pathways of *Neurospora crassa* and *Saccharomyces cerevisiae*. *Microbiol. Rev.* **50:** 280–313.

DeLano, W.L. 2002. The PyMOL molecular graphics system. http://www.pymol.org.

Gouaux, J.E. and Lipscomb, W.N. 1988. Three-dimensional structure of carbamoyl phosphate and succinate bound to aspartate carbamoyltransferase. *Proc. Natl. Acad. Sci.* **85:** 4205–4208.

Ha, Y., McCann, M.T., Tuchman, M., and Allewell, N.M. 1997. Substrate-induced conformational change in a trimeric ornithine transcarbamylase. *Proc. Natl. Acad. Sci.* **94:** 9550–9555.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. 1991. Improved methods for building protein models in electron density maps and location of errors in these model. *Acta Crystallogr.* **A47:** 110–119.

Ke, H.M., Lipscomb, W.N., Cho, Y.J., and Honzatko, R.B. 1988. Complex of N-phosphonacetyl-L-aspartate with aspartate carbamoyltransferase. X-ray refinement, analysis of conformational changes and catalytic and allosteric mechanisms. *J. Mol. Biol.* **204:** 725–747.

Khatib, F., Weirauch, M.T., and Rohl, C.A. 2006. Rapid knot detection and application to protein structure prediction. *Bioinformatics* **22:** e252–e259. doi: 10.1093/bioinformatics/btl236.

Kraulis, P.J. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24:** 946–950.

Krause, K.L., Volz, K.W., and Lipscomb, W.N. 1987. 2.5 Å structure of aspartate carbamoyltransferase complexed with the bisubstrate analog N-(phosphonacetyl)-L-aspartate. *J. Mol. Biol.* **193:** 527–553.

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

Merritt, E.A. and Murphy, M.E. 1994. Raster3D Version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.* **D50:** 869–873.

Morizono, H., Tuchman, M., Rajagopal, B.S., McCann, M.T., Listrom, C.D., Yuan, X., Venugopal, D., Barany, G., and Allewell, N.M. 1997. Expression, purification and kinetic characterization of wild-type human ornithine transcarbamylase and a recurrent mutant that produces ''late onset'' hyperammonaemia. *Biochem. J.* **322:** 625–631.

Morizono, H., Cabrera-Luque, J., Shi, D., Gallegos, R., Yamaguchi, S., Yu, X., Allewell, N.M., Malamy, M.H., and Tuchman, M. 2006. Acetylornithine transcarbamylase: A novel enzyme in arginine biosynthesis. *J. Bacteriol.* **188:** 2974–2982.

Naumoff, D.G., Xu, Y., Glansdorff, N., and Labedan, B. 2004. Retrieving sequences of enzymes experimentally characterized but erroneously annotated: The case of the putrescine carbamoyltransferase. *BMC Genomics* **5:** 52. doi: 10.1186/1471-2164-5-52.

Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276:** 307–326.

Pastra-Landis, S.C., Foote, J., and Kantrowitz, E.R. 1981. An improved colorimetric assay for aspartate and ornithine transcarbamylases. *Anal. Biochem.* **118:** 358–363.

Qu, Q., Morizono, H., Shi, D., Tuchman, M., and Caldovic, L. 2007. A novel bifunctional N-acetylglutamate synthase-kinase from *Xanthomonas campestris* that is closely related to mammalian N-acetylglutamate synthase. *BMC Biochem.* **8:** 4. doi: 10.1186/1471-2091-8-4.

Shi, D., Morizono, H., Ha, Y., Aoyagi, M., Tuchman, M., and Allewell, N.M. 1998. 1.85 Å resolution crystal structure of human ornithine transcarbamylase complexed with *N*-phosphonacetyl-L-ornithine. *J. Biol. Chem.* **273:** 34247–34254.

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

Shi, D., Gallegos, R., DePonte III, J., Morizono, H., Yu, X., Allewell, N.M., Malamy, M., and Tuchman, M. 2002. Crystal structure of a transcarbamoylase-like protein from the anaerobic bacterium *Bacteroides fragilis* at 2.0 Å resolution. *J. Mol. Biol.* **320:** 899–908.

Shi, D., Morizono, H., Yu, X., Roth, L., Caldovic, L., Allewell, N.M., Malamy, M.H., and Tuchman, M. 2005a. Crystal structure of *N*-acetylornithine transcarbamoylase from *Xanthomonas campestris*: A novel enzyme in a new arginine biosynthetic pathway found in several eubacteria. *J. Biol. Chem.* **280:** 14366–14369.

Shi, D., Yu, X., Roth, L., Morizono, H., Hathout, Y., Allewell, N.M., and Tuchman, M. 2005b. Expression, purification, crystallization and preliminary X-ray crystallographic studies of a novel acetylcitrulline deacetylase from *Xanthomonas campestris*. *Acta Crystallogr.* **F61:** 676–679.

Shi, D., Morizono, H., Cabrera-Luque, J., Yu, X., Roth, L., Malamy, M.H., Allewell, N.M., and Tuchman, M. 2006a. Structure and catalytic mechanism of a novel *N*-succinyl-L-ornithine transcarbamylase in arginine biosynthesis of *Bacteroides fragilis*. *J. Biol. Chem.* **281:** 20623–20631.

Shi, D., Yu, X., Roth, L., Morizono, H., Tuchman, M., and Allewell, N.M. 2006b. Structures of *N*-acetylornithine transcarbamoylase from *Xanthomonas campestris* complexed with substrates and substrate analogues imply mechanisms for substrate binding and catalysis. *Proteins* **64:** 532–542.

Shi, D., Yu, X., Roth, L., Tuchman, M., and Allewell, N.M. 2007. Structure of a novel *N*-acetyl-L-citrulline deacetylase from *Xanthomonas campestris*. *Biophys. Chem.* **126:** 86–93.

Slocum, R.D. 2005. Genes, enzymes and regulation of arginine biosynthesis in plants. *Plant Physiol. Biochem.* **43:** 729–745.

Virnau, P., Mirny, L.A., and Kardar, M. 2006. Intricate knots in proteins: Function and evolution. *PLoS Comput. Biol.* **2:** e122. doi: 10.1371/journal.pcbi.0020122.